https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


# Statistical Assessment of the Representative Elementary Area for Areal Fracture Intensity (P21) in Digital Outcrop Models

Stefano Casiraghi[1], Daniela Bertacchi[2], Gabriele Benedetti[1], Silvia Mittempergher[1], Federico Agliardi[1], Mattia Martinelli[3], Francesco Bigoni[3], Cristian Albertini[3], Andrea Bistacchi[1]

[1]Dipartimento di Scienze dell'Ambiente e della Terra, Università degli Studi di Milano-Bicocca, Milan, 20126, Italy
[2]Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, Milan, 20125, Italy
[3]Eni S.p.A, Global Natural Resources, San Donato Milanese, Italy

*Correspondence to*: Stefano Casiraghi (s.casiraghi21@campus.unimib.it)

**Abstract.** The definition of a representative elementary volume (REV) or area (REA) for a target parameter is a fundamental step toward the upscaling of fracture network properties generated by discrete fracture network models (DFN) to an equivalent continuous medium for in situ applications in engineering geology, hydrogeology, and structural geology. The target parameter of this work is the areal fracture intensity ($P_{21}$), a key metric often used as a stopping criterion in stochastic DFN simulations, that is derived directly from surface data collected at natural outcrops. We propose a novel approach to define the REA as a range bounded by a lower and an upper limit. The upper limit, often overlooked but nonetheless theorized, identifies the largest representative domain, which is crucial for optimizing computational efficiency. We evaluate the REA range based on three statistical parameters, namely: the shape, mean, and variance of the $P_{21}$ distributions obtained with progressively increasing scan area sizes. Each statistical parameter is assessed by combining formal statistical tests and diagnostic plots. Within a multi-parametric framework, the method enables a detailed analysis of the statistical behaviour of the dataset, supporting informed decisions in defining the REA range. The methodology is tested on two fractured limestone outcrops with markedly different characteristics: (i) an abandoned quarry in the Murge Plateau (Puglia, Italy) and (ii) the Lilstock Benches in the southern Bristol Channel basin.

## 1 Introduction

The Geometrical, mechanical and hydraulic properties of fractured rock masses are influenced by the occurrence, geometry and spatial distribution of discontinuities at every observation scale. Increasing the size of the sampling element results in a variation in almost all measured parameters of a rock mass, due to (i) the progressive inclusion of larger discontinuities or heterogeneities, and (ii) the increasing size of the sampling domain over which these parameters are averaged. Representative elementary volume (REV) or area (REA), depending on the dimensionality, represent the area/volume above which a certain parameter value becomes independent of the position and of sampling element size with which it is calculated, and thus the value can be used to constrain wider models (Bear, 1975). In other words, REV/REA mark the transition from an inhomogeneous and discontinuous behaviour of the response variable to an homogeneous and continuous

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


one (Hudson and Harrison, 1997; Franzosi et al., 2023a, b). Continuing to increase the size of the sampling element, theoretically speaking, another transition to a heterogeneous behaviour can be found due to the progressive inclusion of large-scale features. Both the lower and upper limits of REA/REV are fundamental because they define the range of stability of the target parameter (Fig.1, Hudson and Harrison, 1997).

The REV/REA concept finds application in different fields of research (Huang et al., 2020; Sari, 2021). Its robust definition is crucial for numerical modelling applications involving large-scale problems (i.e. fluid reservoirs, tunnelling, mining, large rock slope instability), based on both the "equivalent continuum" approximation (Hoek and Brown, 1997; Hudson and Harrison, 1997) and the discontinuum one, where an explicit account of all fractures is not feasible and outcrop-scale rock mass properties must be incorporated in the modelled "rock matrix" (Agliardi et al., 2013; Gerstner et al., 2025).

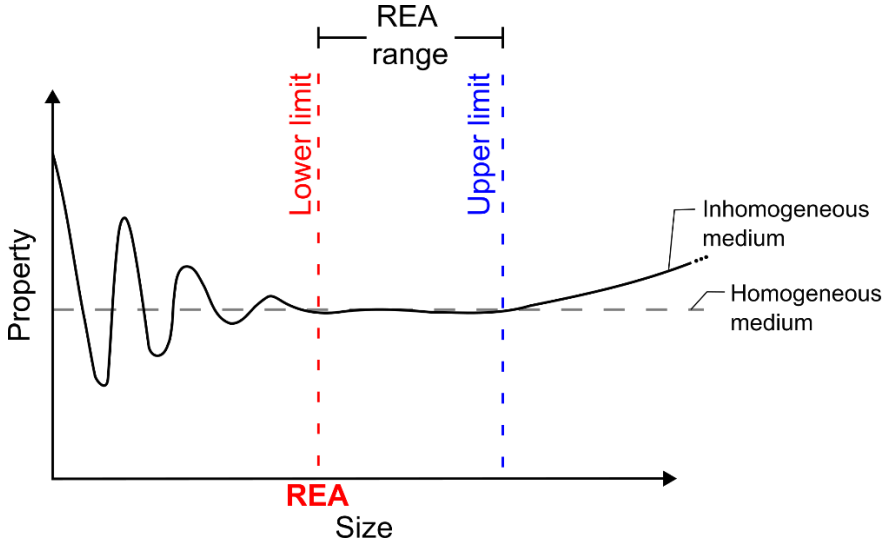



**Figure 1 REV/REA theoretical behaviour, modified after Hudson and Harrison (1997).**

It has been demonstrated that there is not necessarily a unique REV for a given fracture network, but every specific parameter is associated to its own specific REV (Martinelli et al., 2020). Particularly, in engineering geology literature, the REV has been determined for many parameters of interest, for example: Geological Strength Index (Marinos et al., 2005; Huang et al., 2020), damage coefficient (Ni et al., 2017), blockiness (Xia et al., 2016), fracture persistence (Song et al., 2015), among many others (Sari, 2021). In hydrogeology applications, the REV has been associated with parameters like the average block permeability (Kulatilake and Panda, 2000) and the hydraulic conductivity (Wang et al., 2002). Considering the parameters describing the geometry of fracture networks, methods for REV lower limit calculations have been proposed for the average block size (Rohrbaugh Jr. et al., 2002), mean spacing (Pariseau et al., 2008; Zeeb et al., 2013), and has been associated with the mean trace length through a scalar multiplier that goes from 3 to 8 depending on the network connectivity (Li and Zhang, 2011).

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


In the realm of fracture network stochastic modelling, one of the most important parameters is the volumetric fracture intensity $P_{32}$, defined as the ratio between the total fracture area and the sampling volume, or its surface equivalent - $P_{21}$, defined as the total fracture trace length per unit area (Dershowitz and Herda, 1992). These parameters are critical because they serve as stopping criteria in marked point process stochastic DFN (Discrete Fracture Network) modelling, meaning that fracture generation ceases once the target intensity is reached. REV calculations for $P_{32}$ have been reported, for example, by Zhang et al., (2012) and Esmaieli et al., (2010). In such applications, REV definitions are supported by stochastic models to upscale geometrical parameters derived from outcrops into three-dimensional volumes. Alternative approaches addressing the same parameters rely directly on features mapped from outcrops (e.g., digitalized fracture traces; Martinelli et al., 2020; Casiraghi et al., 2025b). In these cases, the evaluated parameter is $P_{21}$, with the focus placed on REA (2D) rather than REV (3D). This work focuses on the REA rather than the REV because most fracture parameters used to constrain stochastic DFN models are derived from field analogues and obtained from the analysis of outcrop surfaces (either vertical or horizontal), where fracture traces are digitized manually (Forstner and Laubach, 2022) or through automated approaches (Prabhakaran et al., 2021). Since these parameters are two-dimensional, a representative area must first be defined to ensure that parameters such as $P_{21}$ are meaningfully calculated before being used to constrain their volumetric counterparts.

While several methodologies have been proposed for REV/REA calculation, they consistently focus on identifying the lower limit of the REV. However, the upper limit—beyond which the assumption of homogeneity is no longer valid—has received much less consideration. However, the upper limit of the REV is of particular importance in modelling applications because, for example, in numerical methods, such as finite element or finite differences, it represents the largest cell size that can be used without losing the possibility to describe the spatial variability of this property, and hence the cell size associated with the lowest attainable computational power (the larger the cell size, the lower the computational cost). Nevertheless, a univocal definition of the upper limit of REV/REA is still elusive.

$P_{21}$ is computed by placing sampling areas of arbitrary shape (e.g., squares, hexagons, or circles) and size within the outcrop boundary, measuring the total cumulative length of fractures (or fracture segments when only partially contained), and dividing this by the area of the sampling window. By progressively increasing the scan area size (for example, the radius in the case of circular windows), it is possible to evaluate how sampling scale influences the resulting $P_{21}$ values (Martinelli et al., 2020; Casiraghi et al., 2025b). Within the REV/REA range, following the REV definition (Bear, 1975), similar statistics are expected among groups of scan areas with the same radius, and comparable statistical distributions of $P_{21}$ value among scan areas with different radii, provided they fall within the REV range.

In this study, we present a new methodology for quantifying the Representative Elementary Area (REA) range, based on the evaluation of three statistical parameters of the $P_{21}$ value distribution: the mean, the variance, and the distribution shape.

Each parameter is assessed using a different statistical test and an associated diagnostic plot. The shape of the distribution is evaluated using the Shapiro–Wilk test (Shapiro and Wilk, 1965), under the assumption that $P_{21}$ values follow a normal distribution within the REA range. Test results are checked with normal probability plots. Variance equality among groups of $P_{21}$ values obtained with different scan area sizes is tested using Levene's test (Brown and Forsythe, 1974), and results are

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


visually inspected with residual plots (Kozak and Piepho, 2018). Differences in mean values between groups are assessed through one-way ANOVA test (Stahle and Wold, 1989).

We demonstrate and discuss the applicability of this approach on two different case studies: (i) Pontrelli quarry (Panza et al., 2019; Casiraghi et al., 2025b), and (ii) the Lilstock outcrop in Bristol channel (Prabhakaran et al., 2021, and references therein). Each of these case studies exhibits distinct characteristics, enabling the method behaviour to be evaluated under different conditions and at varying scales. The Pontrelli quarry is characterized by multiple fracture sets, although only one can be clearly identified across the whole quarry pavement. The Lilstock outcrop, by contrast, is particularly clean and contains several fracture sets that together produce an extremely dense fracture network.

## 2 Case studies

### 2.1 Pontrelli quarry pavement

The Pontrelli abandoned quarry is an outcrop of fractured limestone in the Murge Plateau near Altamura (Puglia, Italy), in the outer Apulian platform, in front of the Southern Apennines fold and thrust belt (Panza et al., 2019). The quarry is carved into the shallow marine intertidal limestones of the Calcare di Altamura Formation (Coniacian to Early Campanian, Panza et al., 2016). The quarry has been the subject of several studies since it offers a wide and clean pavement of subhorizontal bedding surfaces (around 18.000 $m^2$) where fractures are well exposed and vertical walls, thus enabling the characterization of both horizontal and vertical fracture traces and the integration of information obtained from the two complementary types of exposures.

The outcrop presents three main fracture sets, with well-defined topological relationships (Casiraghi et al., 2025b). Among the three sets, Set 1 is the most persistent and the only one detectable across the entire pavement. The fracture traces of Set 2 and Set 3 on the pavement are partly hidden by artificial fractures from quarrying activities, although they can be observed on the adjacent walls. This effect cannot be resolved simply by isolating the no-data zones, since in areas where Sets 2 and 3 are not identifiable, Set 1 is still well visible. Therefore, we limited our analysis to Set1, the most persistent and clearly visible over the largest outcrop area (Fig. 2).

### 2.2 Lilstock outcrop – Bristol channel

The Lilstock outcrop is located in the southern coast of the Bristol Channel in West Somerset (UK), near the hamlet of Lilstock (Prabhakaran et al., 2021, and references therein). The outcrop is composed of weakly deformed, shallowly dipping Jurassic sediments, and it is characterized by large-scale folds, faults, veins and fractures (Passchier et al., 2021). The fracture network consists of eight distinct fracture sets (Passchier et al., 2021), and the exceptional cleanliness of the pavement allows for the mapping of every single fracture, with virtually no gaps in data coverage (Fig. 2). The analysis is performed on the Area E of the Bench IV (Passchier et al., 2021), a portion of the outcrop with a size comparable to the part of Pontrelli quarry pavement free of no-data zones. The fracture trace shapefile is provided as a supplementary material by

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Prabhakaran et al., (2021), and obtained with an automatic digitalization method. The interpretation boundary, not provided in the dataset, is instead digitalized as a "convex hull", joining the fracture trace tip lines. This would be an issue if the goal of the analysis was, for instance, to determine the fracture length distribution, as the assumption of independence between the censoring mechanism and the phenomenon would no longer hold (Benedetti et al., 2025). In the case of REA estimation, however, the interpretation boundary simply delineates the area where data are observable from the area where they are not.

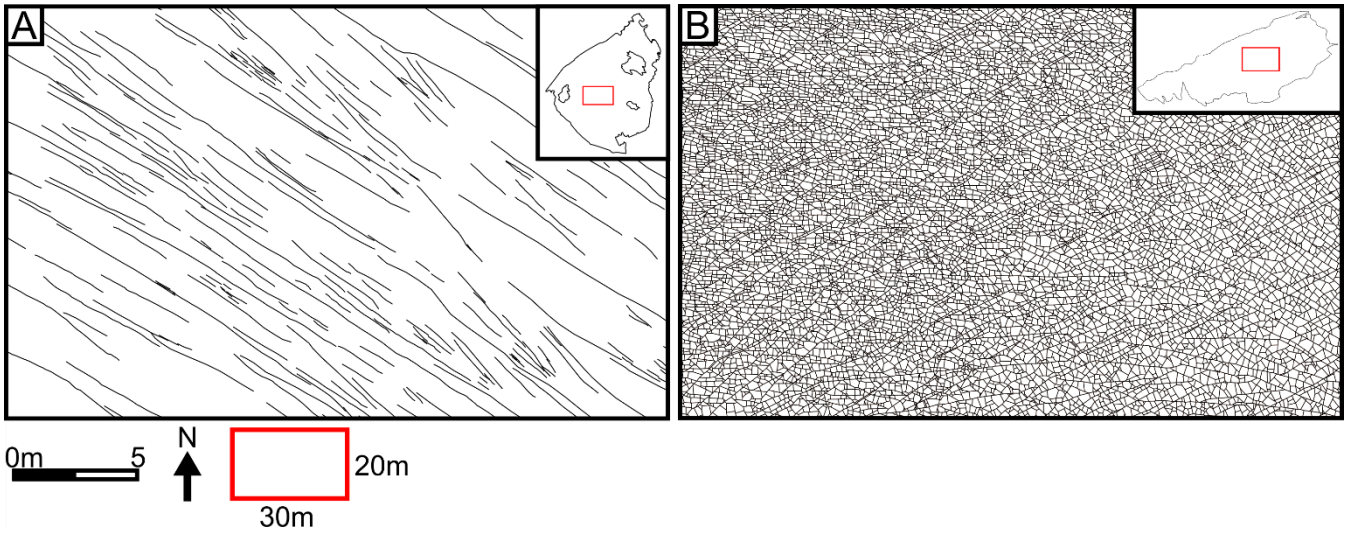



**Figure 2 Close-up view of the digitized fracture network for the Pontrelli quarry (A) and the Lilstock outcrop (B). Only the fracture sets considered in the analysis are shown: Set 1 for the Pontrelli quarry and all eight recognized fracture sets for the Lilstock outcrop (Passchier et al., 2021).**

## 3 Methods

Analysis of Variance (ANOVA), first introduced by Ronald Fisher, is a well-established statistical methodology that finds applications in a wide range of research fields and experimental design (Stahle and Wold, 1989; Mooi et al., 2018; Rajaraman et al., 2019). It aims to evaluate the effect of different factors on a response variable. Aligning this concept to our problem, the response variable is $P_{21}$, while the factor under investigation is the size of the scan area radius. Therefore, the number of factors in the proposed experimental design will be equal to the number of scan area radii needed collect $P_{21}$ values from the outcrop.

The upper limit of scan area radius range is defined by the outcrop size (the scan area cannot be larger than the outcrop itself), while the lower limit is qualitatively defined by checking preliminary results, for example with the Tukey (1977) box plot method. Decreasing too much the scan area radius when the number of outliers is already significant will result in a waste of computational power. As a heuristic initial estimate, the minimum scan radius can be defined by the minimum spacing for a single fracture set, or the minimum total spacing when multiple sets are present. Having defined the range,

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


scan area radii are here defined through equal-spacing sampling between the maximum and minimum values (but different spacing could be used in theory).

ANOVA requires different underlying assumptions to be validated. The basic assumptions that allow working in the ANOVA framework are: (i) the continuity of the response variable and (ii) the independence of each measurement. The continuity assumption is supported by the nature of $P_{21}$, which is a numerical parameter ranging from 0 to infinity and can theoretically be measured at infinitely many values within this range.

The independence of observations within each factor is ensured by employing a random sampling with replacement strategy. This approach guarantees that the selection of one sampling unit does not influence another, thereby satisfying the statistical requirement for independence regardless of the underlying spatial correlations of the fracture network. In this approach, scan area centres are randomly positioned within the interpretation boundary. If a scan area intersects the interpretation boundary, it is rejected and randomly repositioned until the conditions are met, that are: (i) complete inclusion into the interpretation boundary, or (ii) tangency, that is permitted to ensure equal sampling probability across the outcrop. This step represents the most computationally demanding part of the workflow. Increasing the scan area radius leads to more frequent rejections due to boundary intersections, and scan areas larger than the interpretation boundary are systematically rejected. This process, and the following testing procedures are repeated one hundred times to account for the variability of the random sampling.

With regard to grid sampling strategy, for example that employed in Casiraghi et al. (2025b), or Martinelli et al., (2020) random sampling allows the same number of $P_{21}$ data to be collected for every scan area radius. Even if it is not an underlying assumption, equal sample size is a desirable condition when it comes to testing different groups of data. Sample size is related to the power of a statistical test and comparing the test results of samples with different sample sizes will produce misleading results (Kozak and Piepho, 2018).

Aside from continuity and independence, that are assumptions related to the nature of the response variable (continuous, discrete, categorical etc.) and sampling strategy (e.g. random sampling with replacement, grid sampling), ANOVA test requires two additional assumptions that are: (i) the normality of the error factor in the ANOVA linear model (Eq. [4]), and (ii) the homogeneity of variances between the tested factors. In other words, it is first necessary to assess whether the $P_{21}$ values collected using a given scan area radius (e.g., 1 m; Fig. 3) can be reasonably described by a normal distribution, and subsequently to evaluate whether datasets obtained at different radii (e.g., 1 m, 2 m, 3 m) exhibit comparable variances.

In order to test these two assumptions (Stahle and Wold, 1989), we propose integrating formal significance tests with diagnostic plots. Normality assumption is tested with the Shapiro-Wilk (Shapiro and Wilk, 1965) test and normal probability plots (Kozak and Piepho, 2018), while the homogeneity of variances is tested with the Levene test (Brown and Forsythe, 1974) and residual plots (Kozak and Piepho, 2018). The Shapiro-Wilk test, Levene test, ANOVA test and normal probability plots are performed using the SciPy stats package (https://docs.scipy.org/doc/scipy/reference/stats.html).

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


### 3.1 Data collection

The datasets are composed by an interpretation boundary (polygon shapefile) and a series of fracture traces (polyline shapefile). $P_{21}$ is calculated on digitalized fracture traces from TS-DOMs (Textured Surface Digital Outcrop Models, (Bistacchi et al., 2022; Casiraghi et al., 2025b). $P_{21}$ values are straightforwardly obtained by calculating the total sum of length of the fracture traces (complete or partial) inside the circular scan area and dividing it by the scan area surface area. $P_{21}$ is not calculated using estimators (e.g. Mauldon et al., 2001). For any fixed scan area radius, we collect a random sample of size $n$ of such areas and compute the variable $P_{21}$ (Fig. 3). The number of scan area radii depends on the interpretation boundary size and the radii discretization strategy, that means the number of radii collected between the maximum and minimum radius, and the stepping strategy (e.g. equally spaced). The procedure is repeated $m$ times, each time with a new and independent random sample. Consequently, the statistical tests, which will be described in the following sections, are applied $m$ times. In the present simulation $n = m = 100$. This kind of "meta-analysis" allows for consideration of the variability inherent in random sampling while ensuring robustness to the statistical analysis.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

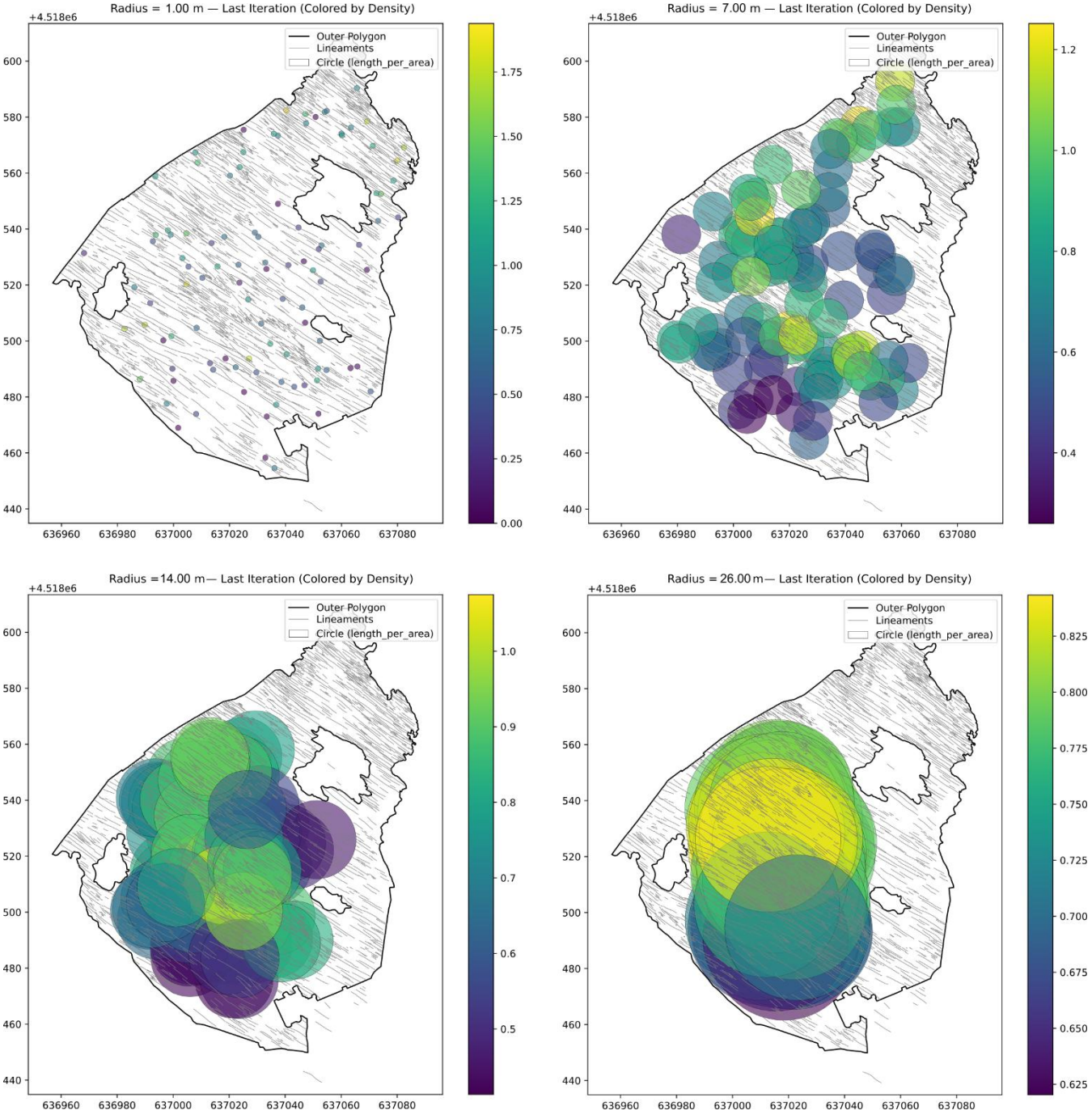




**Figure 3 Example of $P_{21}$ data collection at progressively increasing scan area radius for the Pontrelli quarry outcrop. One hundred scan area are randomly distributed inside the interpretation boundary for every scan area radius. At small scan area radii, the scan areas can be placed even in the most convex part of the outcrop. At increasing radii, the position of the scan area is progressively constrained by the shape of the interpretation boundary and by the presence of no-data zones inside the outcrop. This image represents one realization.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


### 3.2 Exploratory data analysis

Before performing the testing procedure, data collected from one or more realizations (if necessary) are first analysed using graphical tools (Fig. 4, 5, 6) to identify potential outliers or anomalous patterns. This preliminary step also provides an initial understanding of how $P_{21}$ evolves with increasing scan area radius, thereby supporting a more critical assessment of the statistical tests and diagnostic plots presented in the following sections. The graphical tools employed include: (i) variance plots, showing the relationship between sampling variance and scan area radius, (ii) mean plot showing the mean variation at increasing scan area radius, and (iii) boxplots of $P_{21}$ values for each scan area radius, constructed following the method of Tukey (1977).

### 3.3 ANOVA mathematical model

The ANOVA test aims at assessing the effect of different factors on the response variable behavior (Stahle and Wold, 1989). In other words, the ANOVA test is aimed at detecting significant differences in mean between groups of data. The easiest configuration is therefore a situation in which every factor has the same mean and every observed difference is due to random errors. Given a certain number of factors the model would be:

$$x_{i,j} = \mu + e_{ij}, \quad [1]$$

where $x_{i,j}$ is the observation $i$ belonging to factor $j$, $\mu$ is the mean of the considered variable populations and $e_{ij}$ is the random error associated with observation $i$, belonging to factor $j$. Adding a level of complexity, a configuration in which differences are actually related to the effect of different factors:

$$x_{i,j} = \mu_j + e_{ij}, \quad [2]$$

where $\mu_j$ is the mean of the group of observations related to factor $j$. It is possible to explicitly express the effect of factor $j$ as:

$$\alpha_j = \mu_j - \mu, \quad [3]$$

where $\alpha_j$ is the effect of the factor j on the response variable. Adding [2] into [3], it yields:

$$x_{i,j} = \mu + \alpha_j + e_{ij}. \quad [4]$$

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Equation [4] is the so-called ANOVA linear model, an additive model that consist of the sum of three effects: (i) $\mu$, the common effect, (ii) $\alpha_j$, the effect of the j-th factor, and (iii) $e_{ij}$ the random error associated to each observation of each factor (Stahle and Wold, 1989).

In our analysis the factors $j$ are the scan-area radiuses, the observations $i$ are the individual scan-areas and the response variable is $P_{21}$, that is affected by both $i$ and $j$. In ANOVA we are testing the effect of $j$ on $P_{21}$, and for this purpose we design our sampling strategy in order to randomize the effect of $i$.

### 3.4 Normality assumption

The normality assumption of the ANOVA test involves the error term of the linear model ($e_{ij}$), described in equation [4] (Kozak and Piepho, 2018). The theoretical distribution of the errors is unknown, but errors can be estimated with the residuals, defined as:

$$\hat{\varepsilon}_{i,j} = x_{i,j} - \mu_j, \quad [5]$$

where $\hat{\varepsilon}_{i,j}$ is the residual associated with the $i$-th observation of the $j$-th factor, $x_{i,j}$ is the $i$-th observation of the $j$-th factor and $\mu_j$ is the mean of the group of observations related to factor $j$, as already explicated in equation [3].

Normality assumption is tested with the Shapiro-Wilk test (Shapiro and Wilk, 1965). The test is based on the W statistic that measures the goodness-of-fit to a normal distribution. The values of this statistic range from 0 to 1.0, where values close to 1.0 indicate a normal distribution of the residuals. Significance level for the test is set at 0.05.

Associated with the statistical test, normal probability plots for every tested factor are also provided. Normal probability plots are a graphical procedure to check if data are approximately normally distributed, and they are a specific case of Q-Q plots, where the ordered sample percentiles are plotted against the theoretical percentiles of a normal distribution. If data are normally distributed, points should approximately plot along a straight line. From the shape of the distribution in a normal probability plot, it is possible to detect other characteristics of the dataset, for example the presence of outliers, left or right skewed residuals, or heavy-tailed residuals (presence of both extreme positive and negative values). According to the expected qualitative behaviour of the REA, at small scan area radii the $P_{21}$ distribution is expected to be characterized by extreme values, with a significant amount of zeros on one side and high $P_{21}$ values on the other. This asymmetrical configuration typically corresponds to a right-skewed distribution. Within the REA range, where the heterogeneous medium can be approximated as homogeneous, a more symmetrical distribution is expected, with $P_{21}$ values distributed approximately around the mean. At the upper limit of the REA range, the progressive inclusion of large-scale features leads to a loss of symmetry, producing a different data distribution (it varies on a case-by-case basis), in which zeros are not to be expected.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


### 3.5 Homogeneity of variances assumption, REA upper limit definition

Homogeneity of variances or homoscedasticity is a condition in which the variances of the different factors are comparable. Homoscedasticity is tested with the Levene test (Brown and Forsythe, 1974), performed on residuals (Eq. [5]). While Shapiro-Wilk test is straightforwardly applied to every factor, application of Levene test is controlled by two variables. On one hand there is the number of factors to be tested. The minimum number of factors is equal to three, that is also the minimum number of groups that can be tested with the ANOVA test. The upper limit is not defined, but qualitatively it is possible to say that the higher the number of tested factors, the higher the probability of finding one group that is significantly different from the others. On the other hand, there is the "location". Considering a sequence of scan area radii of increasing size, the Levene test can be conceptualized as a moving window, applied progressively from smaller to larger scan areas. According to REA theoretical behaviour, small scan areas exhibit high variability due to the heterogeneous behaviour of the response variable, leading to rapidly changing group variances and frequent test rejections. This is followed by a stability range (the REA range), where the variance across factors becomes qualitatively stable, resulting in an increased rate of test acceptance. Beyond the upper limit of the REA range, acceptance rates are expected to decrease again as large-scale features introduce additional variability. To capture this theoretical behaviour, Levene test is applied to every consequent group of scan area radius, in an ordered way, from the smallest windows to the largest ones. Group size is iteratively increased starting from groups of three factors, until acceptances are reduced to zero.

Levene test results are complemented by residual plots, which serve as diagnostic tools to assess differences in residual variance across factors. In this framework, standardized residuals—calculated by dividing each residual by the standard deviation of the group of residuals—are plotted against factors (scan area radii), treated here as categorical variables. Standardization centres the residuals around zero, facilitating comparison among groups. Homoscedasticity (i.e. homogeneity of variance) is assumed when the random scatter around the zero line is qualitatively similar across groups. Following Kozak and Piepho, (2018), a highlighted band corresponding to ±1.96 of the estimated standard deviation of the standardized residuals is added to aid visual inspection. Funnel- or cone-shaped patterns in the plots indicate substantial variance heterogeneity (Kozak and Piepho, 2018).

### 3.6 Testing procedure and ANOVA test

Having described the core features of the presented statistical method, it is now possible to describe the testing procedure for REA range definition. The normality of the $n$ $P_{21}$ values collected for any fixed scan area radius is tested with the Shapiro-Wilk test. Acceptances and rejections, for the $m$ realizations, are collected and represented by a bar (see Fig. 7). This means that for any fixed scan area radius we have a bar representing the outcome of the $m$ test: orange bars represent acceptance, grey ones represent rejection. The normality assumption is considered plausible when the acceptance ratio is around 80%. The normality of $P_{21}$ of these scan area radii, that meets the previous condition is also checked with the probability plots. Similarly, the Levene test is iteratively applied from the smallest radius to the largest, progressively increasing the number of

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


tested radii. The number of stacked bar plots for the Levene test is equal to the number of times the group size has increased. The largest group of factors maintaining a significant level of acceptance will be defined as the candidate REA range to be tested with ANOVA. In the same way as for the normality assumption, Levene test results are also checked with the residual plots. The accepted group is cross-validated using the Shapiro–Wilk test results, as each factor accepted by the Levene test must also satisfy the normality assumption. If any factor within the group fails to meet the normality criterion, the group is progressively reduced until a group of consecutive factors fulfils both the homoscedasticity and normality assumptions.

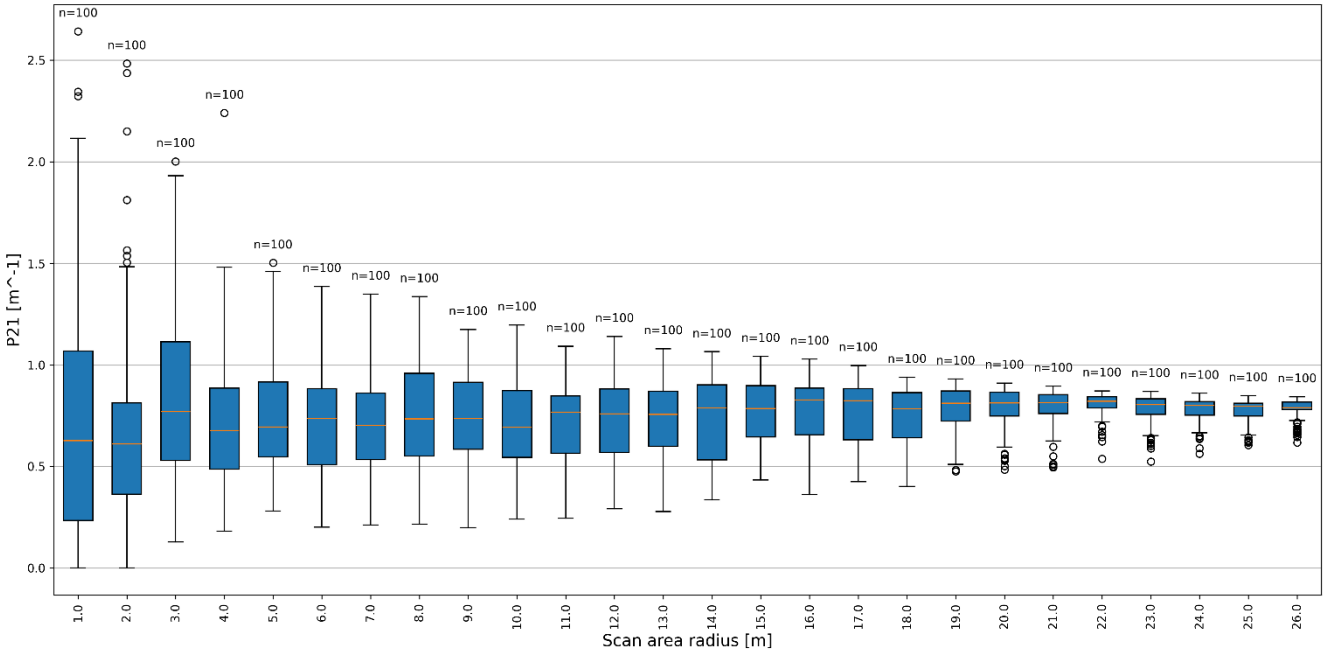
The ANOVA test is applied to the accepted group of data. The ANOVA core test is also called the F-test, which is built such that the value of the F-statistics is large if significant differences are detected between the tested factors (Stahle and Wold, 1989). If the ANOVA test is accepted with a significance level of 0.05, then it is possible to assume the tested groups of factors as the range of stability of the $P_{21}$ parameter, i.e. the REA range.

## 4 Results

### 4.1 Pontrelli quarry

The Pontrelli quarry pavement extends approximately 120 m NW–SE and 150 m NE–SW. In principle, this would allow for the use of relatively large scan areas, with an upper radius limit of about 60 m. However, the presence of no-data zones within the pavement restricts the maximum scan area radius to ~26 m (Fig. 3). The lower limit is set to 1 m, which, as shown by the boxplot results (Fig. 4), is sufficiently small to yield both a considerable number of empty scan areas and outliers. Based on the boxplots, the variance plot (Fig. 5) and the mean (Fig. 6), it is possible to qualitatively define three zones. The first zone is characterized by strong variance and mean fluctuations, and asymmetrical distributions. This zone extends to a radius of approximately 7 m. This is followed by a zone of relative stability, that goes from 7m in radius to 13-15m (depending on the iteration). In this zone the mean $P_{21}$ stabilizes around 0.71-0.73 $m^{-1}$. Finally, as the scan area radius approaches the maximum size allowed by the interpretation boundary, scan areas can no longer be freely positioned across the outcrop (Fig. 3), and the $P_{21}$ distribution collapses toward a single value, corresponding to the only feasible position. In this final stage, variance tends toward zero and most $P_{21}$ values are classified as outliers (Fig. 4).

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


305

**Figure 4 Boxplot of $P_{21}$ data collected at increasing scan area radius, for Pontrelli quarry outcrop.**

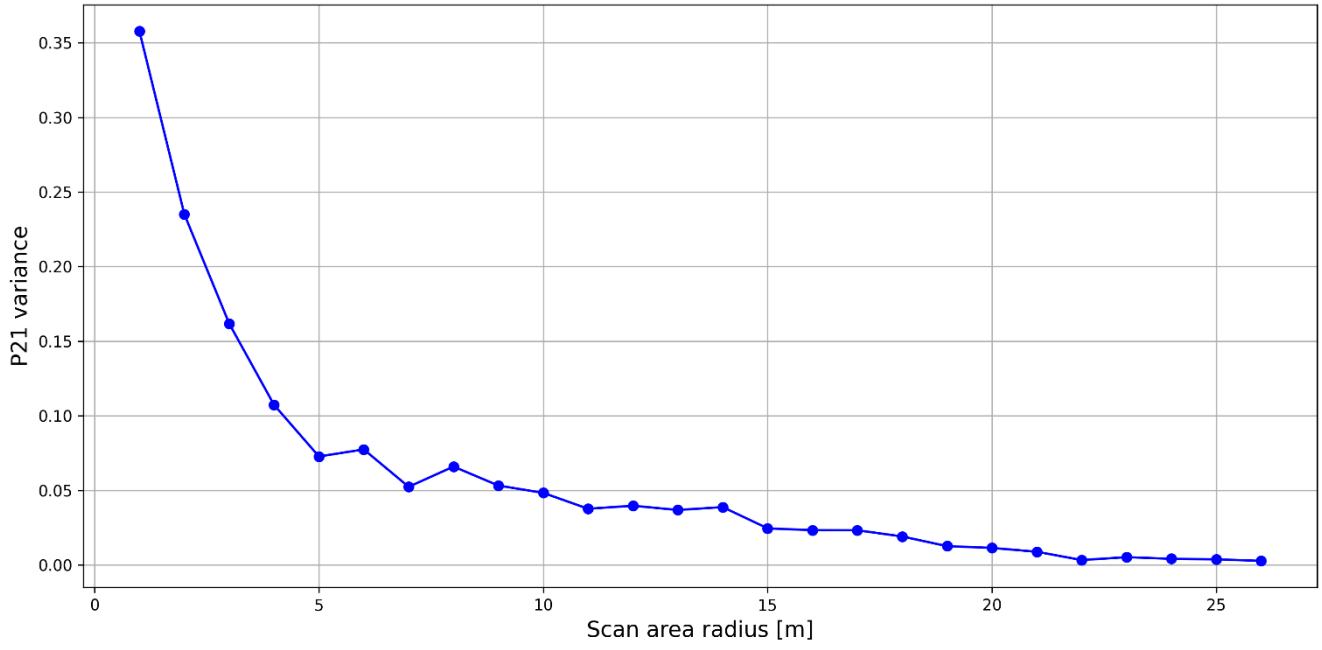



**Figure 5 Variance plot of $P_{21}$ statistical distribution for Pontrelli quarry outcrop.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

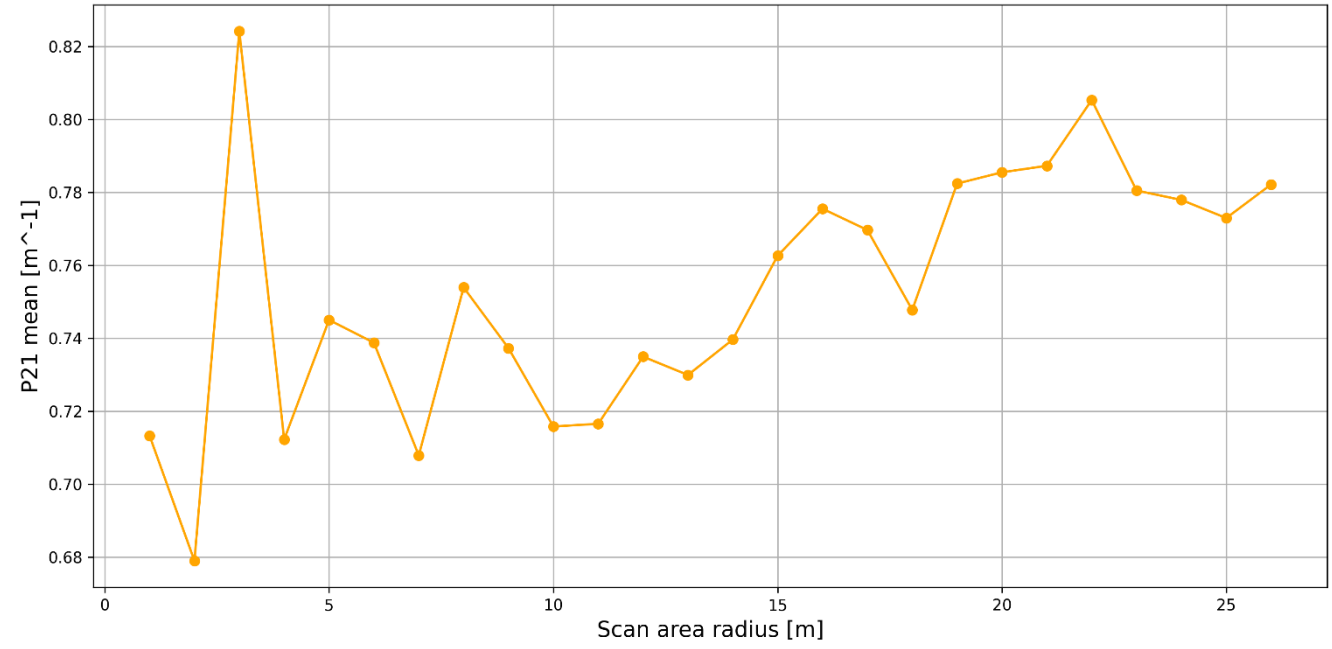




**Figure 6 Mean plot of $P_{21}$ statistical distribution for Pontrelli quarry outcrop.**

The Shapiro-Wilk test results show that the best performing scan area radius are 8m and 9m with acceptances above the 80% threshold (Fig. 7). 7 m and 10 m are few acceptances away from the threshold, and at 6 m and 11 m a huge drop in acceptances is detected, falling below 50%. This result well matches the qualitative assessment made with the boxplots, the variance and the mean plot (Fig. 4, 5, 6). Scan area radius from 7 m to 10 m are checked with the respective probability plots (Fig. 8A, B). From the probability plots strong deviation from the normal model (e.g. heavy tails, left or right skewed distributions) cannot be detected and therefore the normality assumption is confirmed.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

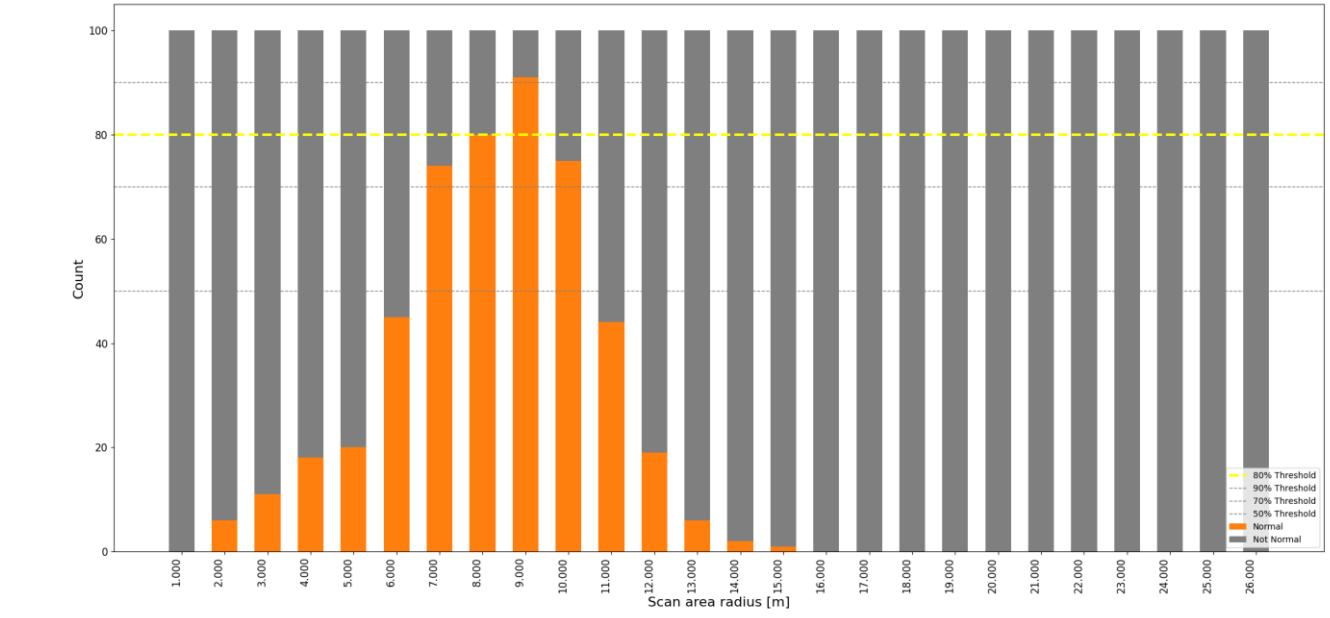




**Figure 7 Shapiro-Wilk test results for Pontrelli quarry outcrop. The orange bars represent acceptances, while the corresponding grey bars indicate rejections. Their sum always equals 100, corresponding to m, the total number of realizations.**

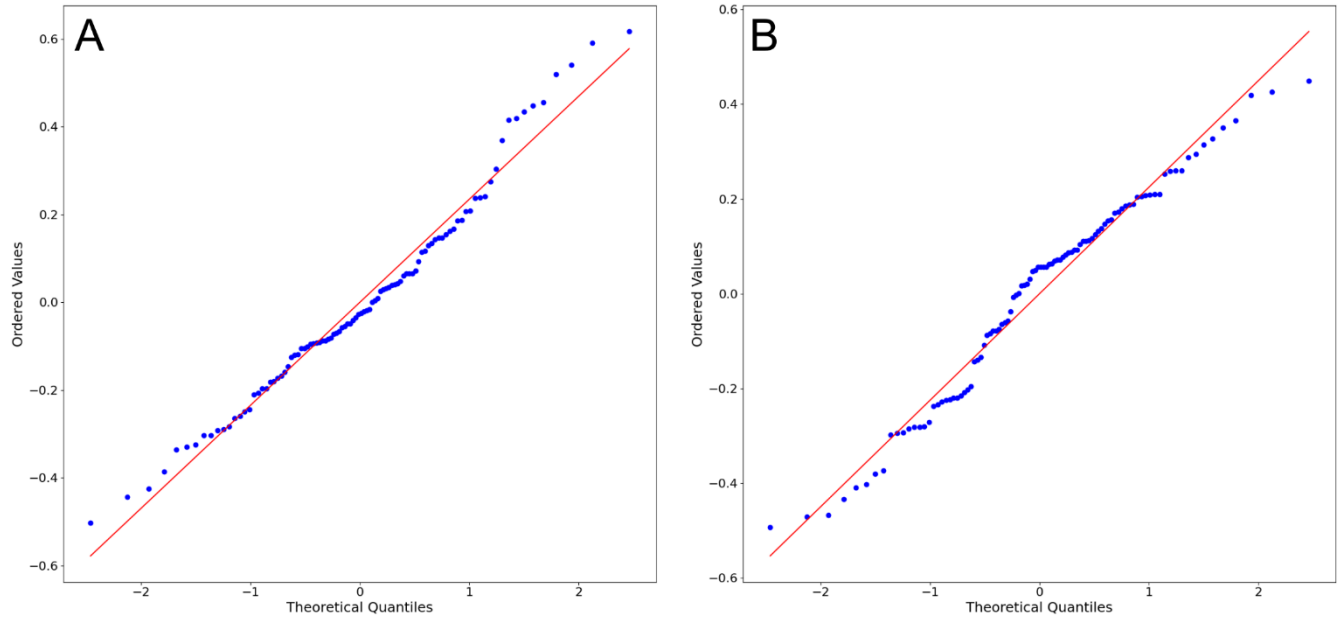



**Figure 8 Normal probability plots for scan area radius equal to 7 m (A) and 10 m (B) in Pontrelli quarry outcrop.**

Levene test results show that the tested groups remain representative up to a range of five factors (Fig. 9C). From six factors onward acceptances decrease by a significant amount (Fig. 9D). Between the size-five groups, the candidates that keep an

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


acceptance ratio around 80% threshold are: 7 m to 11 m (close to 80% threshold), 8 m to 12 m, and 9 m to 13 m. Homoscedasticity is supported by the residual plot, as each accepted group displays a region with almost straight boundaries (Fig. 10).

By cross-validating these candidate groups with the Shapiro-Wilk test results, it is possible to see that each group contains at least one factor that does not match the normality assumption (11 m for the first group, 11 m and 12 m for the second group and 11 m 12 m and 13 m for the last group). In the size-four groups (Fig. 9B), the 7 m to 10 m group both keep an acceptance ratio above the 80% threshold, and every factor is accepted for normality.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

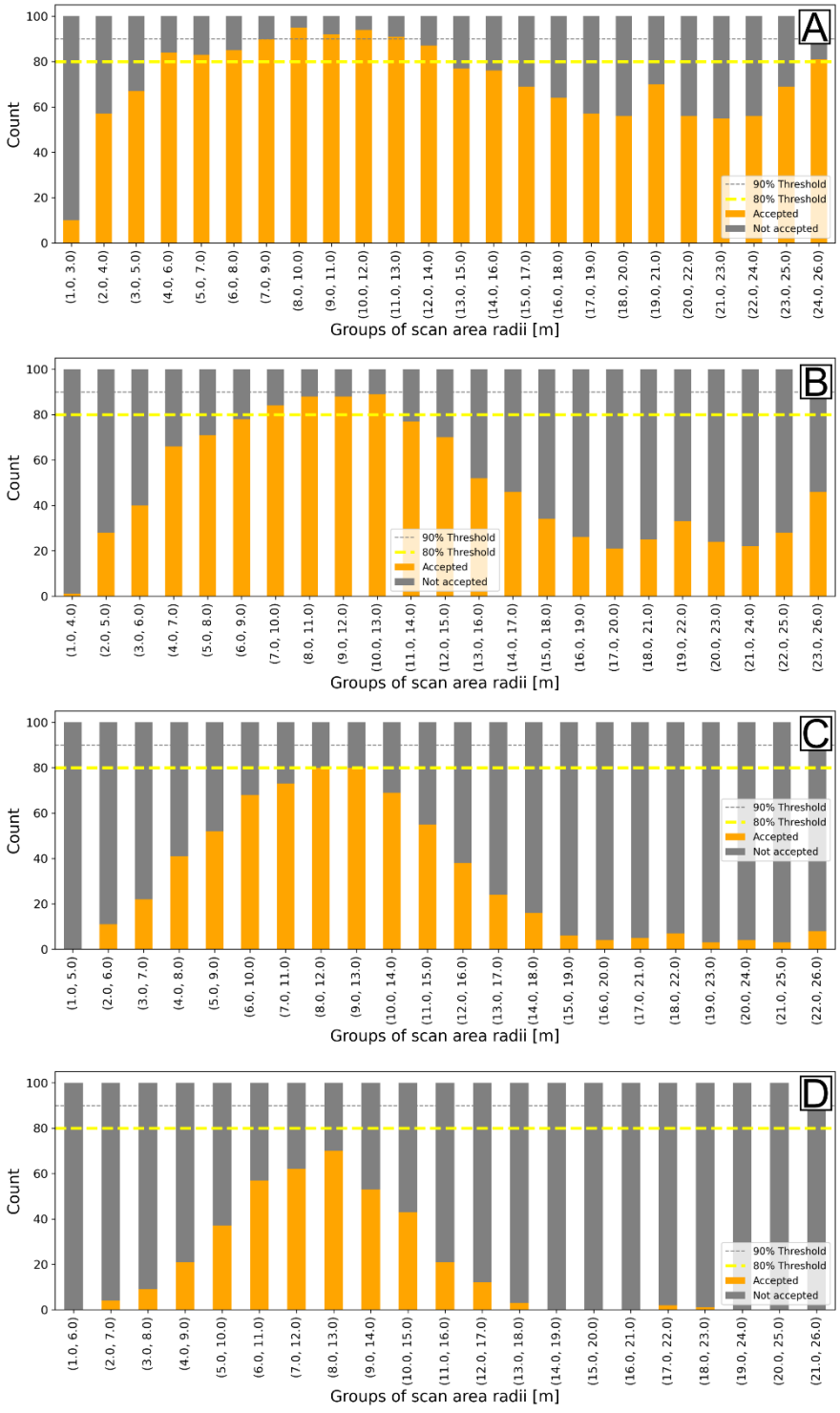




**Figure 9 Levene test results for Pontrelli quarry outcrop for different group sizes. (A) Group size = 3, (B) Group size = 4, (C) Group size = 5, (D) Group size = 6.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

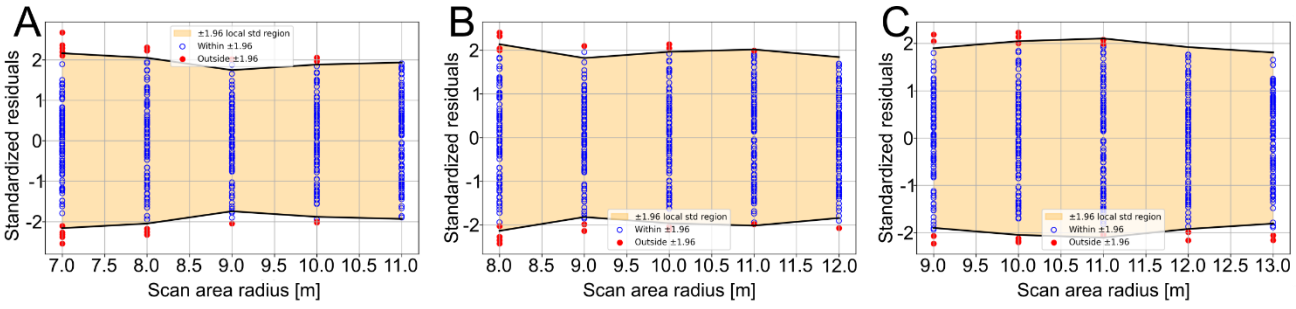




**Figure 10 Residual plots for Pontrelli quarry outcrop. (A) Scan area radii from 7 m to 11 m, (B) Scan area radii from 8 m to 12 m, (C) Scan area radii from 9 m to 13 m.**

The ANOVA test on the 7 m to 10 m group is accepted with an acceptance ratio way above 80% (Tab. 1). Finally, it is possible to conclude that, given the convergence between the Shapiro-Wilk, Levene and ANOVA test results, and the validation through the diagnostic plots and exploratory data analysis, the REA range for the Pontrelli quarry pavement is between scan areas that go from 7 m to 10 m in radius.

| | ANOVA acceptances | ANOVA rejections | Test result |
|---|---|---|---|
| 7 m to 10 m group | 95 | 5 | accepted |

**Table 1 ANOVA test result for Pontrelli quarry outcrop.**

### 4.2 Lilstock outcrop - Bristol channel

The Lilstock outcrop in Bristol channel is characterized by a surface area of the same order of magnitude as the Pontrelli quarry, but it presents a much denser fracture network, composed of multiple fracture sets. As an initial try, the lower limit of the sampling window radius is set to 0.005m (qualitatively equal to the average total spacing), while the upper limit is set to 18m. Given the higher fracture density, we tentatively used 50 scan area radii, to try to catch every change in the $P_{21}$ distribution behavior.

With this stepping, exploratory data analysis shows that, after an initial sharp decrease between the first and second scan area radii, the mean is stable around 8.5 $m^{-1} \pm 0.25$ $m^{-1}$ (Fig. 11, 13). The sharp drop is also captured in the variance plot (Fig. 11, 12), after which the variance decreases steadily with a gentle slope. This trend is likely related to the progressive increase in scan area radius.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

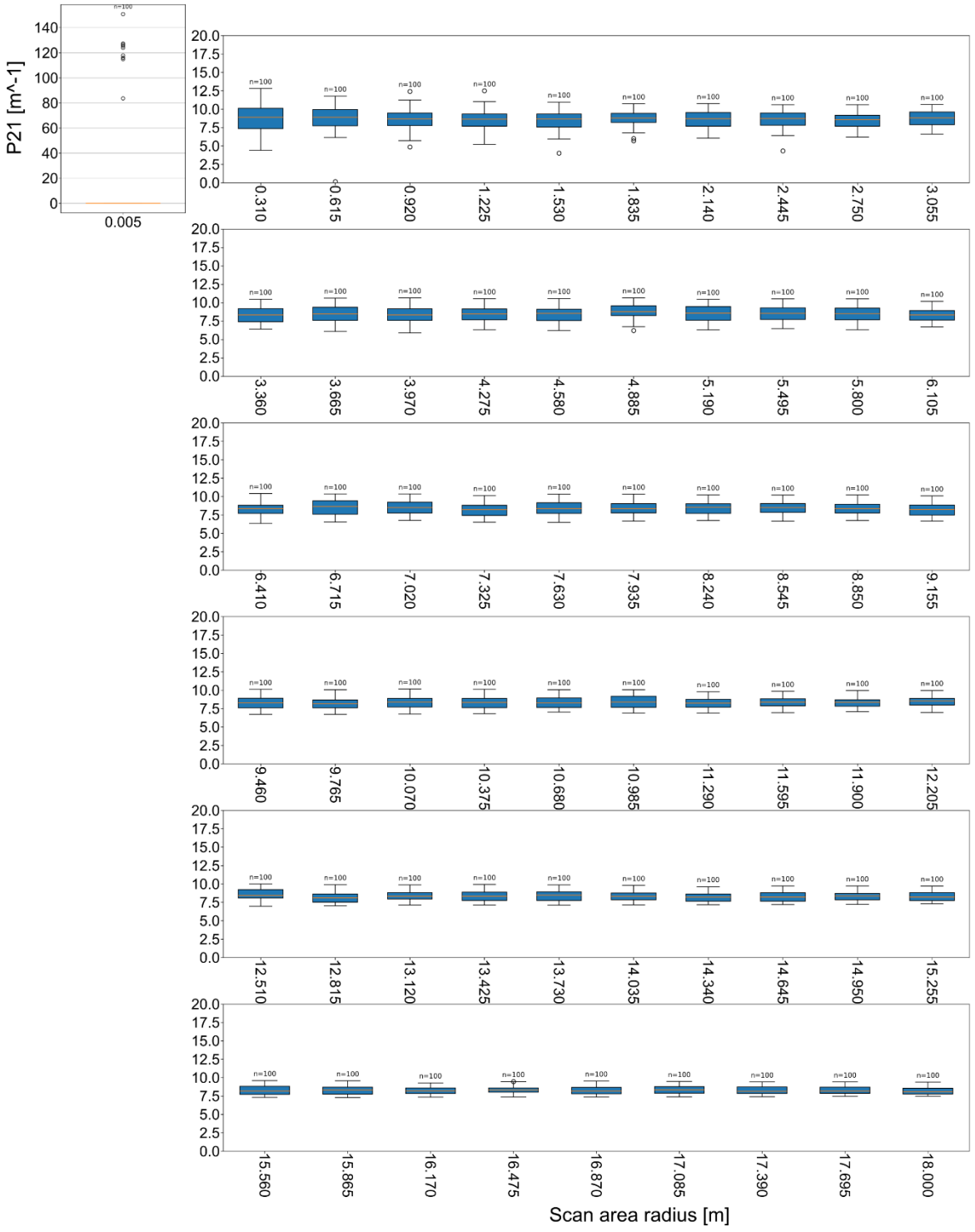




**Figure 11 Boxplot of $P_{21}$ data collected at increasing scan area radius, for Lilstock outcrop. The first boxplot is shown separately so that its larger scale does not compress the remaining boxplots and hinder readability.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

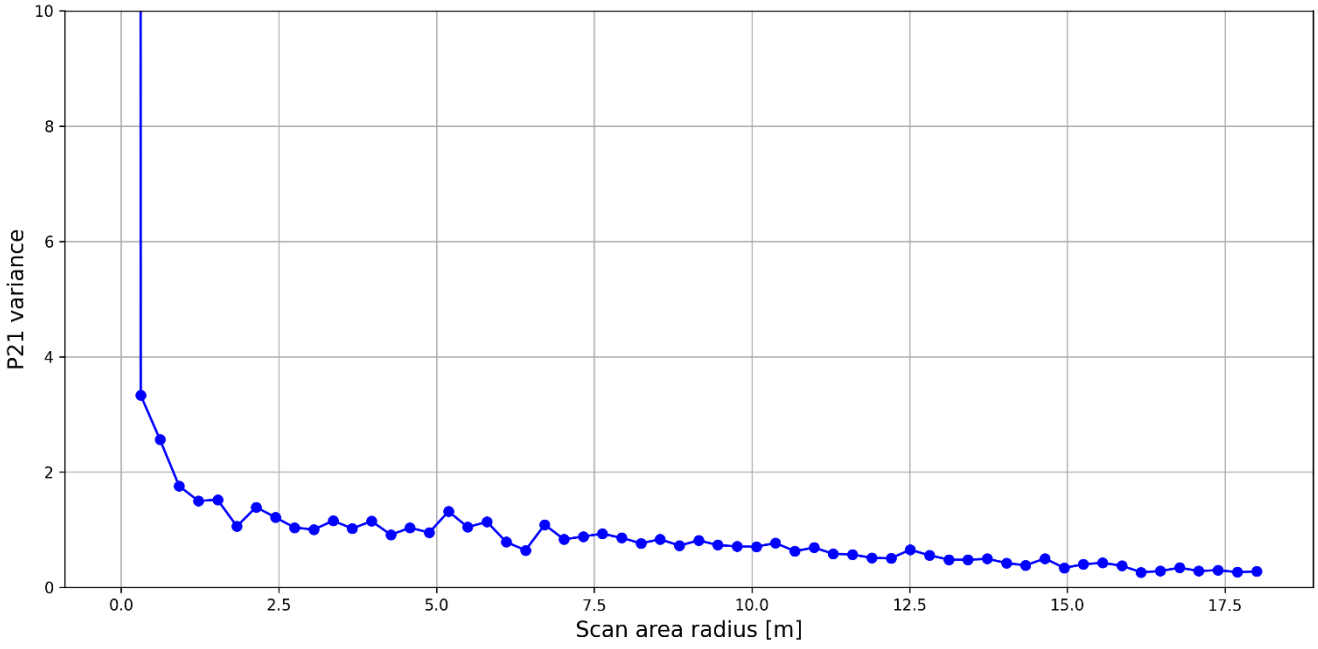




**Figure 12 Variance plot of $P_{21}$ statistical distribution for Lilstock outcrop. The y axis is limited to y = 10, to better highlight the variance trend. For scan area radius equal to 0.005 m the variance is equal to ≈1200.**

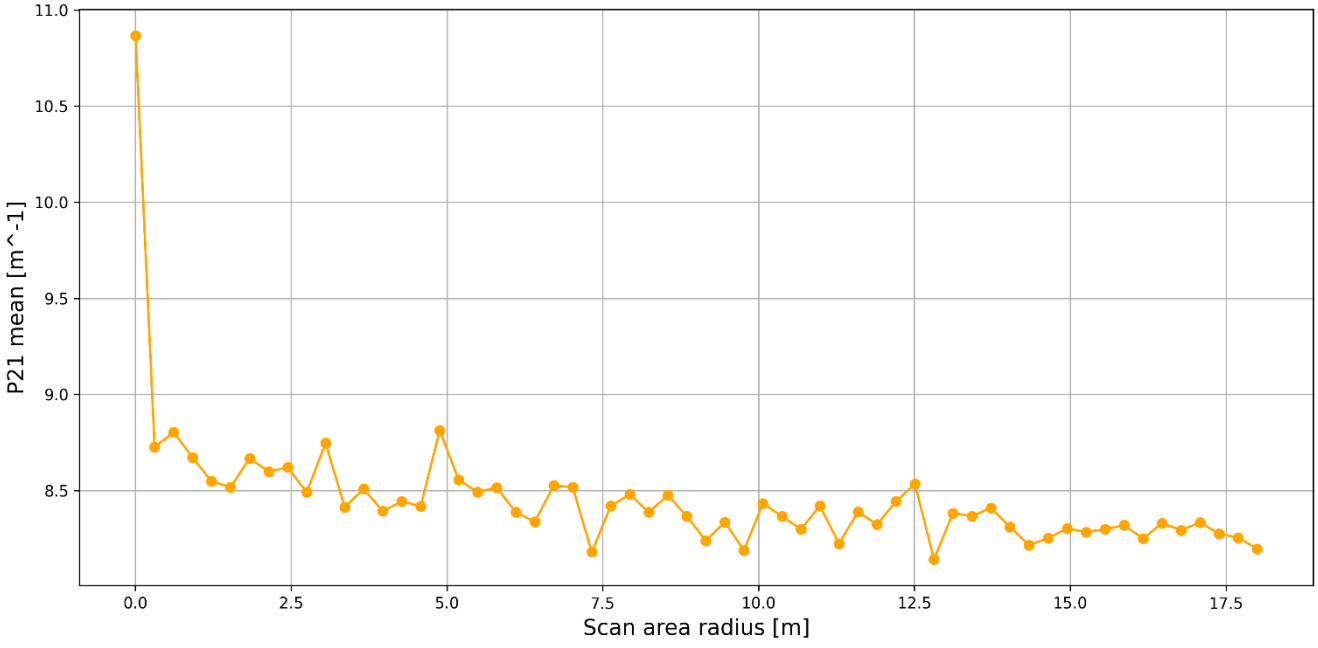



**Figure 13 Mean plot of $P_{21}$ statistical distribution for Lilstock outcrop.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Using this stepping strategy, normality is never fully achieved, as indicated by the Shapiro–Wilk test results, with the highest acceptance ratio (around 70%) corresponding to the second smallest scan area radius (Fig. 14). Inspection of the normal probability plots reveals that deviations from normality arise from local overestimation or underestimation zones, or from areas where $P_{21}$ values are clustered or align along an almost straight line, suggesting that the scan areas are sampling nearly identical $P_{21}$ values (Fig. 15). No strong departures from normality, such as pronounced left- or right-tailed distributions, are observed as, for instance, in the largest scan area radius at Pontrelli quarry. As results by inspecting the normal probability plots, although the $P_{21}$ distributions are not perfectly normal, they are approximately symmetrical independently on the scale of observation (Fig. 15).

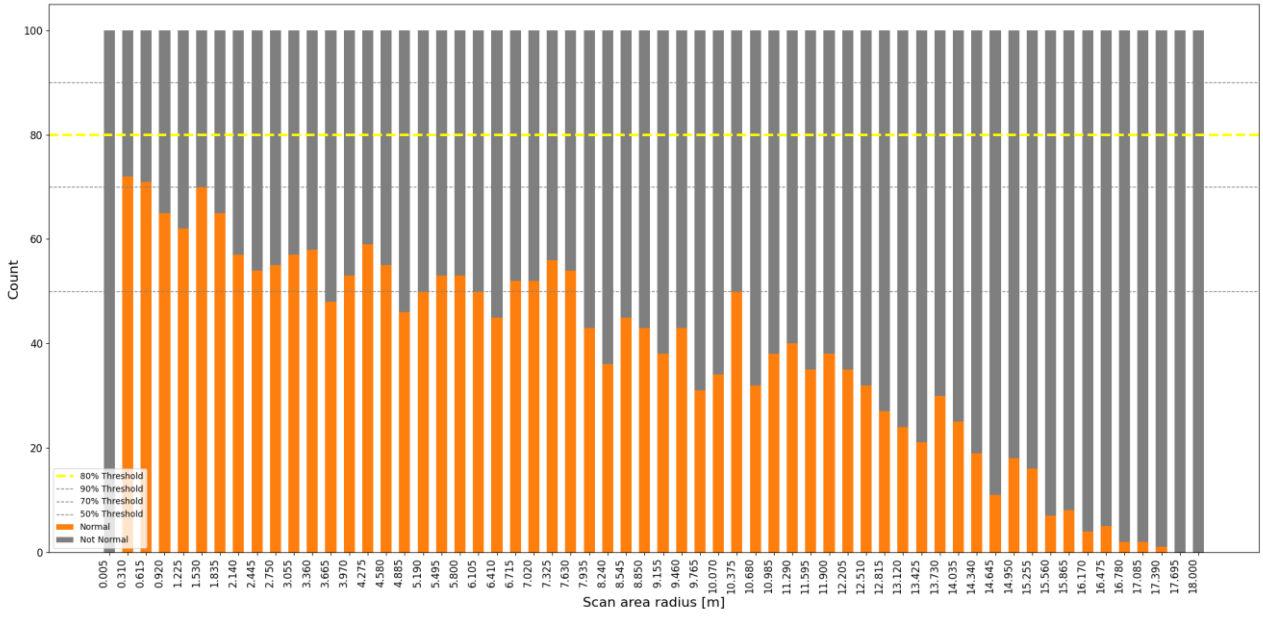
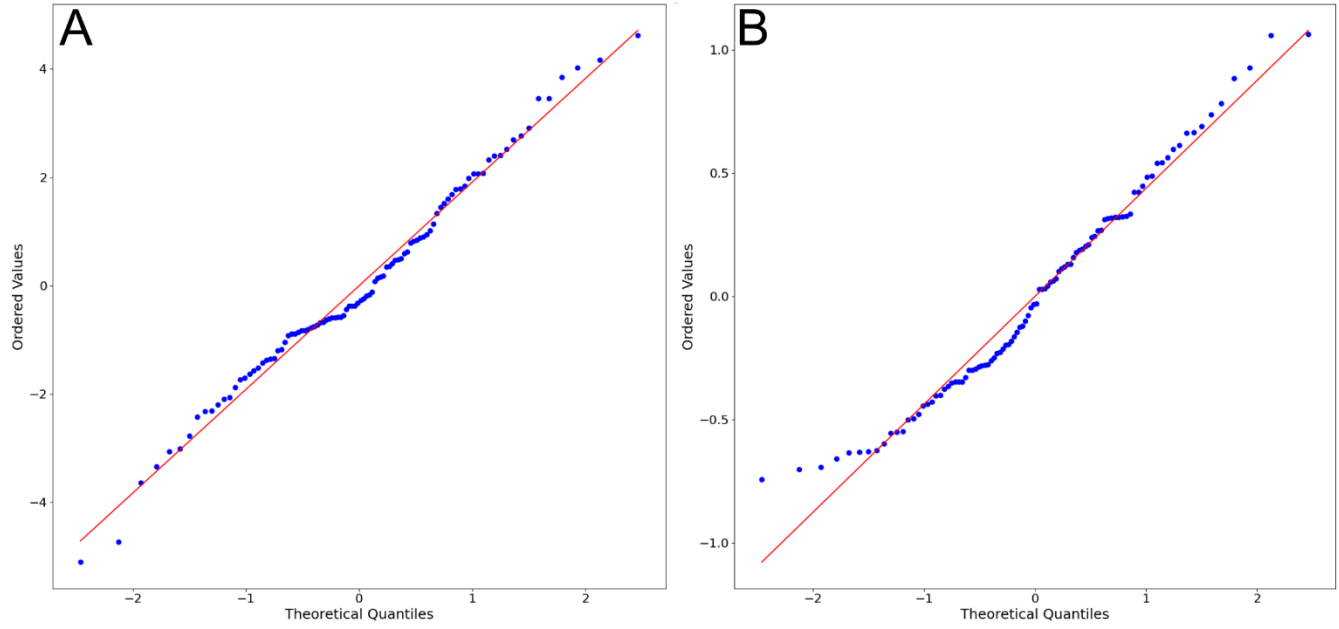



**Figure 14 Shapiro-Wilk test results for Lilstock outcrop.**

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


375 **Figure 15 Normal probability plots for scan area radius equal to 0.31 m (A) and 18 m (B) in Lilstock outcrop.**

Levene test results show that homoscedasticity maintains a significant level of acceptance up to size-fourteen groups, after which acceptance rates progressively decrease from size-fifteen groups onward, likely reflecting the variance reduction observed in the variance plot (Fig.16A, B). This high degree of acceptance contrasts with the behaviour observed at the Pontrelli quarry, where acceptance rates sharply decline beyond the REA range (Fig. 9).

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

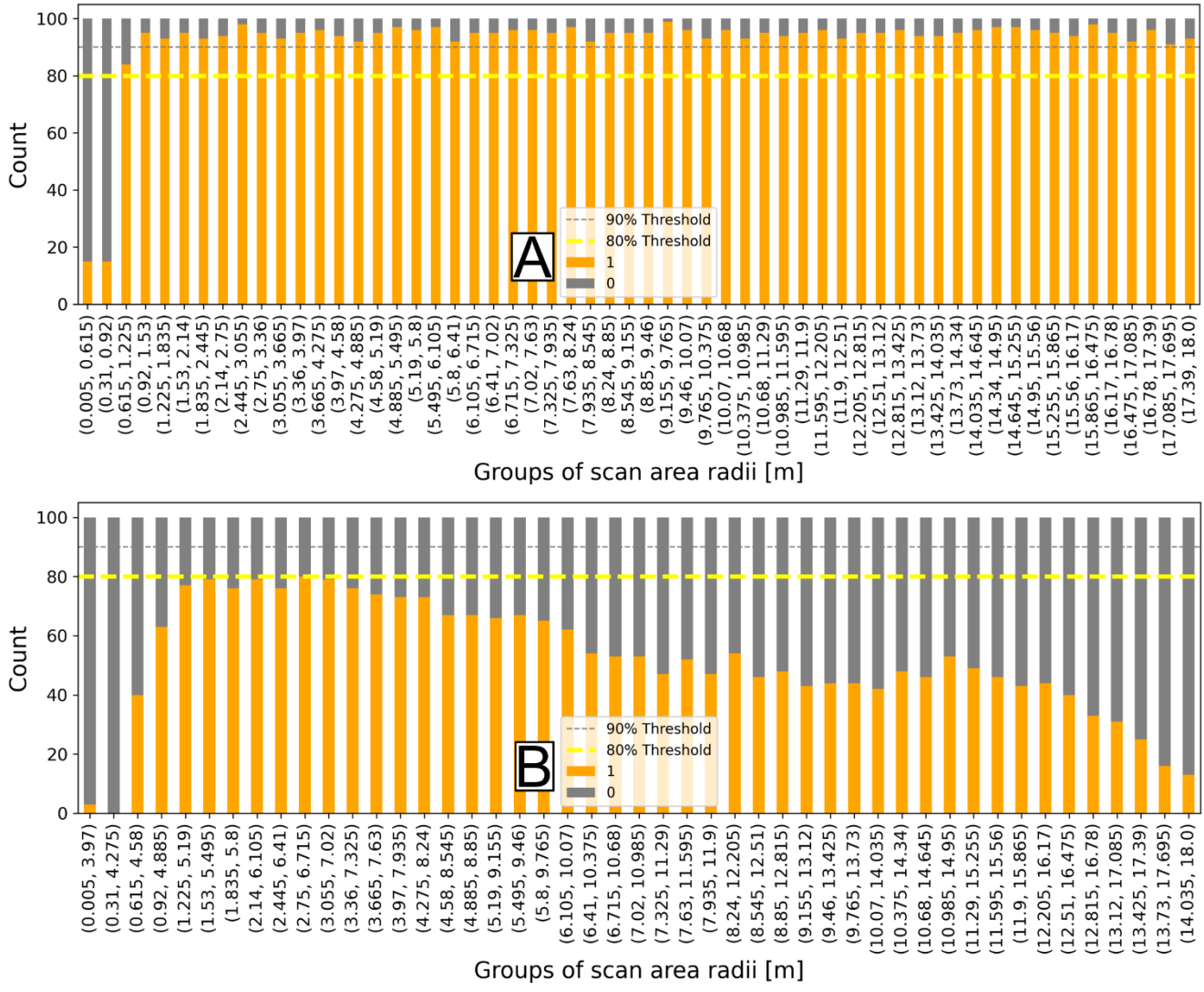




**Figure 16 Levene test results for Lilstock outcrop for different group sizes. (A) Group size = 3, (B) Group size = 14.**

As a REA range could not be established within the initial intervals, the list of scan area radii is further expanded to include ten intermediate values between 0.005 m and 0.31 m. This expansion allows for a more detailed analysis of the abrupt transitions observed in the mean, variance, and statistical tests within this specific range. In particular, before the expansion, it can be observed a sharp drop in both the $P_{21}$ mean (Fig. 13) and variance (Fig. 12). Moreover, the Shapiro–Wilk test shows an increase from 0% acceptances at 0.005 m to 70% at 0.31 m (Fig. 14), while the Levene test rises from nearly 0% to over 90% acceptances for size-three groups (Fig. 16).

The new range of scan area radii reveals an acceptance pattern very similar to that observed at the Pontrelli quarry, but at a much smaller scale (Fig. 17,18).

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

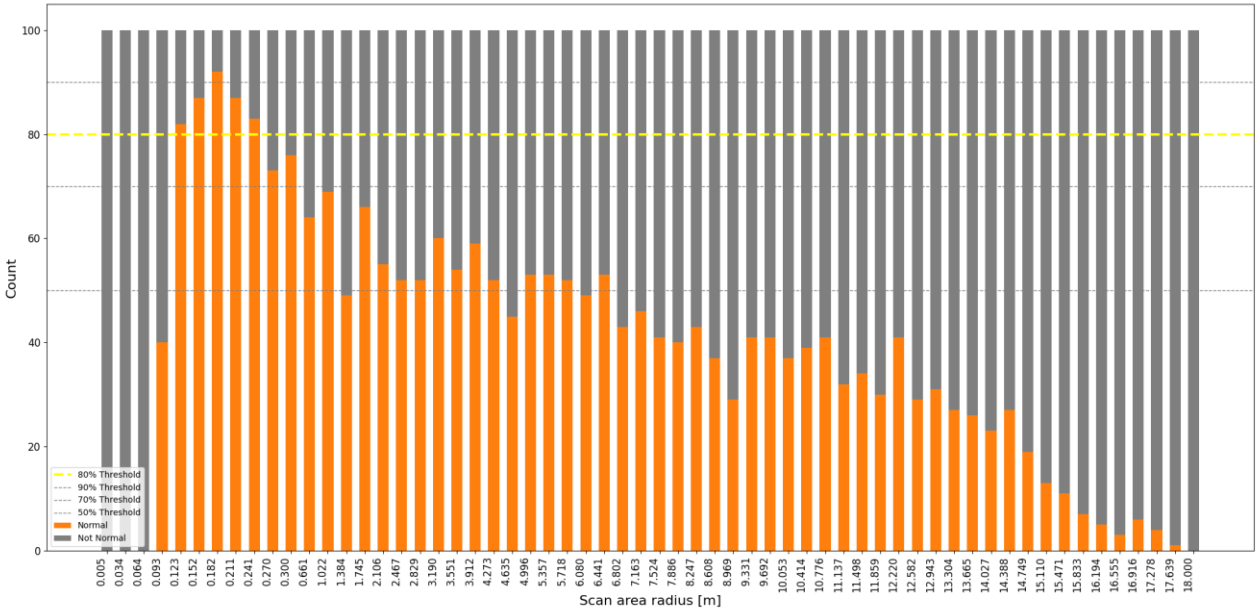




**Figure 17 Shapiro-Wilk test results for Lilstock outcrop, after the radii expansion between 0.005 m and 0.31 m.**

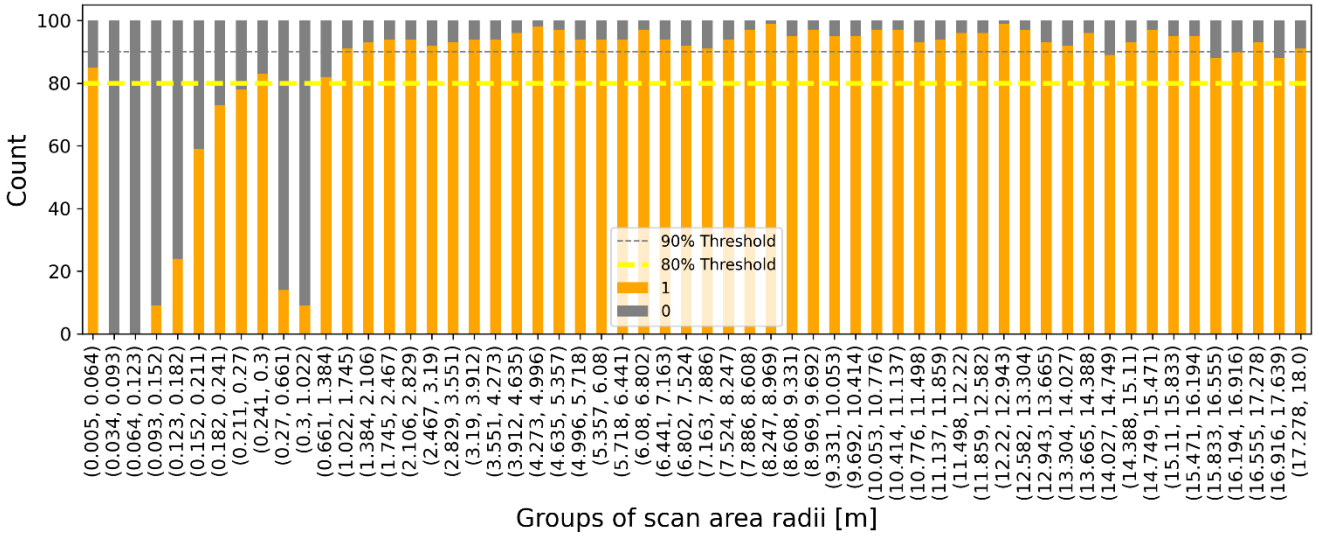



**Figure 18 Levene test results for Lilstock outcrop, after the radii expansion between 0.005 m and 0.31 m. Group size = 3.**

The size-three groups 0.211–0.27 m and 0.241–0.3 m can both be considered as the REA range (Fig. 17, 18, 19, Tab. 2). This is followed by a sharp drop in acceptances and the subsequent onset of the "homogeneous" behaviour described earlier. The drop in acceptances observed around 0.3 m may be related to the change in sampling step size. Prior to 0.3 m, the sampling step is densified and remains close to 0.03 m, whereas beyond 0.3 m it increases to 0.3 m (Fig. 17, 19). This pattern

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


is particularly evident in the residual plot shown in Figure 19C. In contrast, Figure 19D illustrates another rejection interval in which the scan area radii are equally spaced. In Figure 19E, which corresponds to the first interval excluding the 0.3 m radius, an increase in acceptances above the 80% threshold is observed. In this context, 0.3 m appears to mark the upper limit of the REA and represents the boundary between the classical "REA behaviour" discussed in this work and a regime characterized by qualitatively consistent distributions.

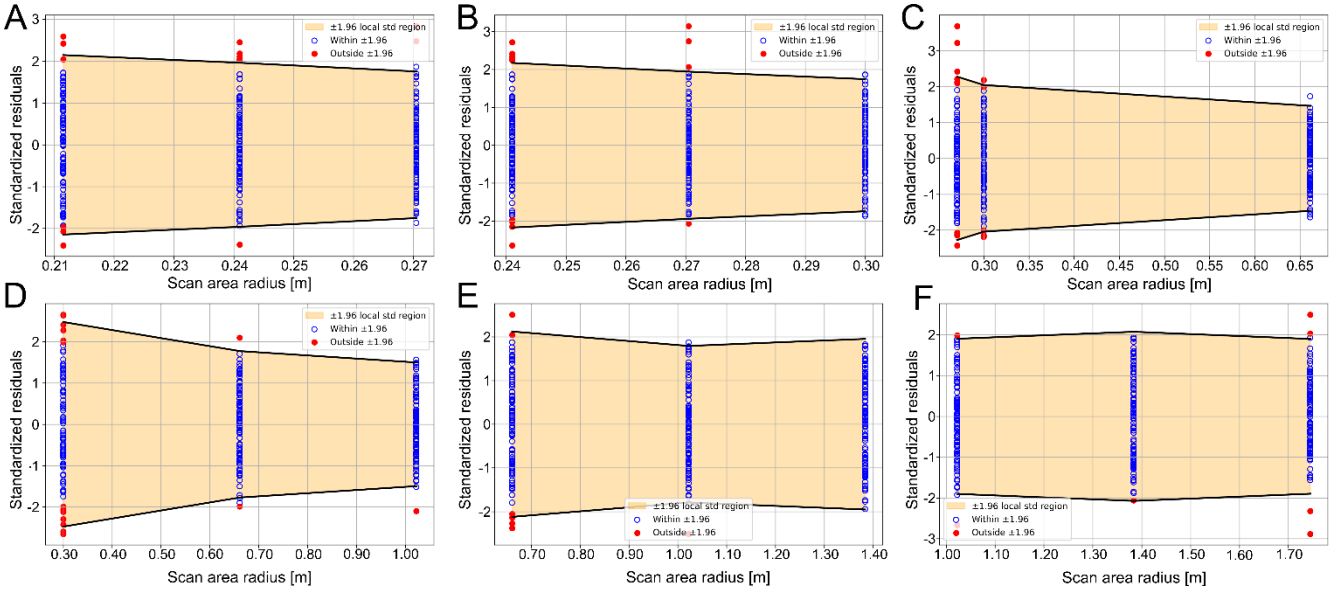



**Figure 19 Residual plots for Lilstock outcrop. (A) Scan area radii from 0.211 m to 0.27 m, (B) Scan area radii from 0.241 m to 0.3 m, (C) Scan area radii from 0.27 m to 0.661 m, (D) Scan area radii from 0.3 m to 1.022 m, (E) Scan area radii from 0.661 m to 1.384 m, (F) Scan area radii from 1.022 m to 1.745 m. Figures (A) and (B) correspond to the identified REA range; Figures (C) and (D) capture the drop in acceptances; and Figures (E) and (F) represent the beginning of the "homogeneous" zone.**

| | ANOVA acceptances | ANOVA rejections | Test result |
|---|---|---|---|
| 0.211 to 0.27 m group | 95 | 5 | accepted |
| 0.241 to 0.3 m group | 96 | 4 | accepted |

**Table 2 ANOVA test results for Lilstock outcrop.**

## 5 Discussions

We have presented a sequence of statistical tests that can be used to define the lower and upper boundaries of the REA for $P_{21}$ and applied them to two case studies of large and continuous natural outcrops, demonstrating the practicability and efficacy of this approach. In the following we discuss more in detail how we can set the acceptance thresholds for the tests, the effect of sample size, the effect of choosing different scan area radiuses, the applicability of the presented methodology considering the problem of the spatial autocorrelation of fractures, and the problem of the finite size of outcrops.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


### 5.1 Acceptance thresholds

Both Shapiro-Wilk, Levene and ANOVA test are repeated for one hundred different realizations, with a significance level set at 0.05.

In case of a theorical normal distribution – that we recall does not necessarily apply to data from real-world phenomena – the outcome of the repeated Shapiro–Wilk test should reflect this significance level. In other words, around 95 out of 100 realisations should fail to reject the null hypothesis, while the remaining ~5 out of 100 are expected to incorrectly reject it due to type I error. Following this line of thought, it would be natural to think of setting the acceptance threshold at 95%. From another perspective, the dominance of the normal model among real-world phenomena has been hugely debated (Micceri, 1989). In most cases, the so-called Data Generating Process (DGP), that is the underlying process that produce the observed data (or the errors), is unknown. Moreover, the test assesses whether a 'sample', defined as a fixed set of realizations, appears to be drawn from a normal distribution, rather than evaluating an actual continuous random variable (Rochon et al., 2012). These considerations suggest that applying the testing procedure with the expectation of a precise outcome (e.g., a strict 95% significance threshold) may lead to misinterpretation. Although no prior information is available regarding the error distribution of $P_{21}$ within the REA range, both exploratory data analysis and REA theory indicate that a symmetrical distribution can be expected (Fig. 4, 5, 6). Normality is therefore not an inherent property of the $P_{21}$ values but a working hypothesis; the data should be 'normal enough' for the assumption to hold. In this study, an 80% acceptance threshold is adopted to consider $P_{21}$ values asymptotically normal. At this threshold, the distribution closely resembles, though does not fully conform to, a normal distribution, allowing reasonable assumptions to be made about its overall shape.

This is also true for homoscedasticity and ANOVA analysis. These tests are built such that they are considered accepted if two groups have the exact same variance (Levene) or mean (ANOVA). However, in the context of natural phenomena, expecting this condition to be consistently met across realizations is unrealistic.

### 5.2 Sample size effect on statistical test results

The power of a statistical test, that is the probability of a correct rejection of the null hypothesis (H0), is directly related to the sample size (Kozak and Piepho, 2018). This means that at small sample size there is a greater probability of accepting a false null hypothesis (type II errors), which translates into a higher acceptance rate. Increasing the sample size will progressively reduce the probability of making type II errors, to the point where, at sufficiently large sample sizes, even the smallest departure from the null hypothesis will be detected and consequently H0 will be rejected.

This is questionable especially when testing the possibility to fit a natural dataset to some theoretical distribution. For instance, given a distribution $\rho$ of some variable measured in a natural dataset, which is only slightly more or less peaked than the normal distribution, for most practical purposes we can use the normal distribution as a model for any of its samples. However, if we draw a sufficiently large sample from $\rho$, and test its normality, we are bound to reject this hypothesis, because natural phenomena do not obey exactly to some theoretical distribution.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


This behaviour makes the outcome of statistical tests strongly dependent on the sample size, meaning that by adjusting the sample size it becomes possible to accept or reject a null hypothesis which is only "slightly" false (and therefore acceptable for practical use). This has led to a decline in confidence in the use of formal statistical tests and a renewed emphasis on graphical approaches (e.g., probability plots), which, although more qualitative, allow the data to be assessed more explicitly (Kozak and Piepho, 2018). At the same time, when dealing with large datasets and multiple realizations, checking every diagnostic plot becomes time-consuming, and identifying variations in data behaviour or setting boundaries (such as defining the REA range) is not straightforward. For example, considering the Pontrelli quarry — the smallest of the case studies in terms of the number of scan area radii — there are 26 unique radii. This results in 26 normal probability plots per realization, and therefore 2,600 plots when 100 realizations are considered. For the Levene test, 23 residual plots are produced for size-three groups, 22 for size-four groups, and so on, for a total of 171 residual plots if the analysis stops at size-eleven groups. This number must then be multiplied by the number of realizations. Therefore, checking every single plot, albeit possible, is inefficient.

In this situation, we cannot exclude the possibility of using formal statistical tests, at least as screening tools, to narrow the range of candidate scan area radii that will be eventually validated with diagnostic plots.

The advantage of $P_{21}$ data is that the sample size can be arbitrarily modified to qualitatively assess the effect of sample size on statistical tests results. We performed this test on the Pontrelli case study, which, given the reduced computational time compared to the Lilstock outcrop, allowed us to increase the sample size up to n = 500. Results of the Shapiro–Wilk test are shown in Figure 20 for sample sizes of n = 30 (Fig. 20A), n = 100 (Fig. 7), and n = 500 (Fig. 20B). For n = 30, acceptance rates are relatively high across all scan area radii, but start to decline markedly (below 50%) at radii larger than approximately 14 m. For n = 500, almost all tests are rejected, with only a few isolated acceptances around 7 m, 8 m, and 9 m. At n = 100, the test produces a well-defined response, with distinct zones of acceptance and rejection that are consistent with both the diagnostic plots and the exploratory data analysis. In support of this choice of sample size, the effect of sample size on the Shapiro–Wilk W statistic has been investigated by Souza et al., (2023). Although their study was conducted on a different dataset, it employed the same sampling strategy used in this work. Their results indicate that the W statistic fully stabilizes for sample sizes around n = 75–100.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

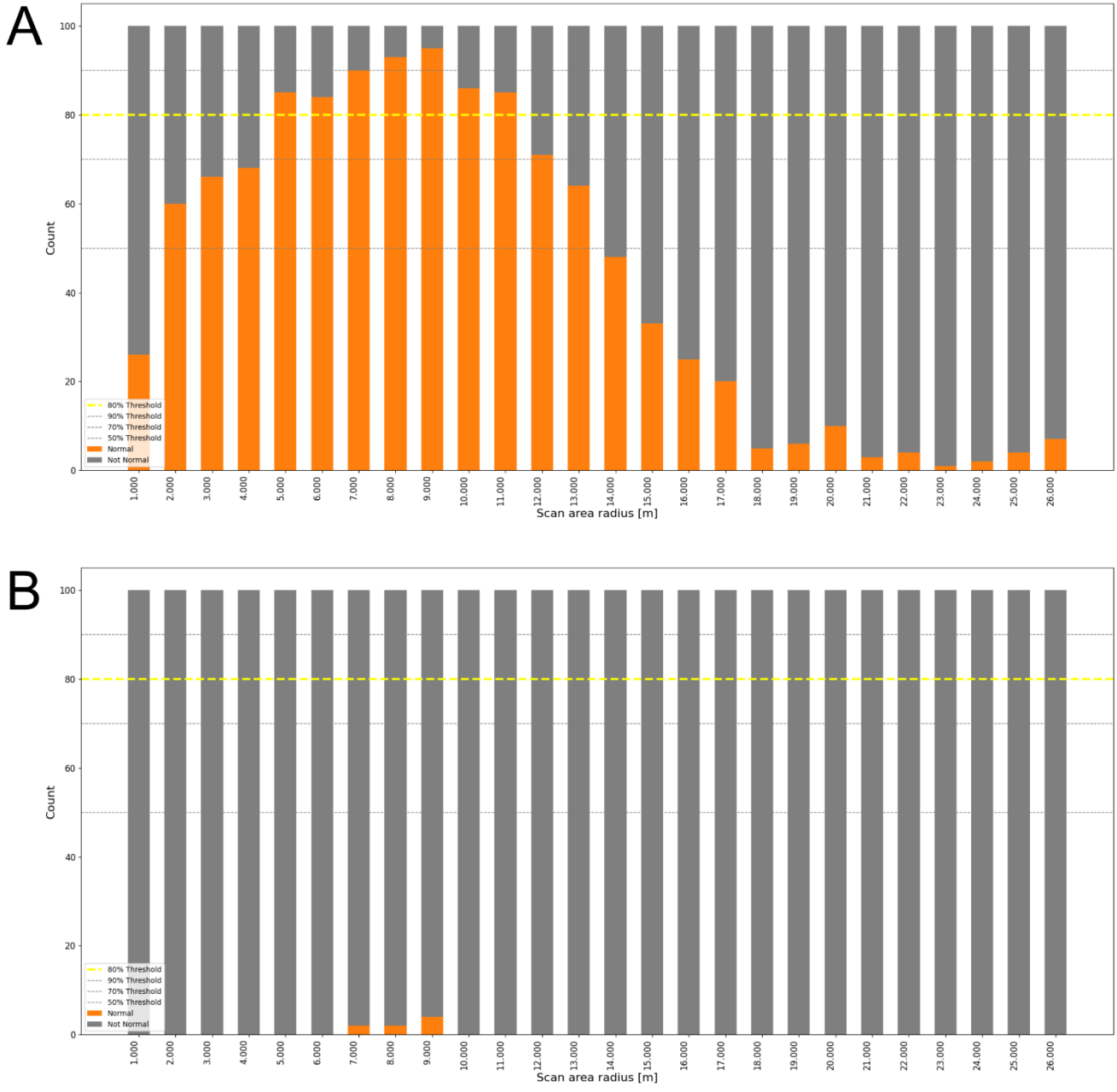




**Figure 20 Shapiro-Wilk test results for Pontrelli quarry outcrop. (A) n = 30, (B) n = 500.**

We do not have a decisive answer or a quantitative method for determining an ideal sample size. Moreover, even the definition of "ideal sample size" is not straightforward. One interpretation could be the sample size that maximizes the number of acceptances; however, reducing the sample size also increases the likelihood of type II errors. Another

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


interpretation might be the sample size that maximizes test power. Yet, as we have shown, beyond a certain sample size this approach no longer yields meaningful results. It is also questionable to pursue perfect agreement with the assumptions of normality or homogeneity of variances, as already discussed in the previous section. In two different case studies, which have markedly different characteristics, it has been seen that the response produced by the tests at n = 100 corresponds to statistically significant peaks of acceptance and well-defined rejection zones, consistent with the diagnostic plots. Particular attention should be paid to radii with acceptance rates only slightly below the 80% threshold. The statistical power of normality tests depends on sample size, and for large samples it becomes possible to reject the null hypothesis even when the data are still reasonably consistent with a normal model. In such cases, a formal rejection does not necessarily imply that the dataset should be discarded, as the outcome may simply reflect the sensitivity of the test rather than a substantive deviation from normality. This is precisely where diagnostic plots become essential, as they provide a visual benchmark to contextualize the statistical result.

### 5.3 Effect of radii discretization strategy

Another variable left to the user's choice is the number and stepping strategy for scan area radii. The maximum radius simply depends on the outcrop size, while the minimum radius should be small enough to capture the discontinuous behavior of the fracture network being analyzed. How intermediate radii are collected in between the maximum and minimum radii is a trial-and-error procedure that starts from an initial guess, and it is very important since, if the resolution of the analysis is not enough, the REA range could not be defined, or could be defined with an insufficient precision.

In the Pontrelli quarry, as an initial guess, $P_{21}$ values were collected with scan area radii spaced at 1m intervals (min. radius = 1 m, max. radius = 26 m). A reasonable question arises as to whether decreasing the spacing between radii would lead to significant changes in the test results. By halving the sampling step (Fig. 21), it appears that the overall trend of acceptances remains unchanged, and no anomalous behaviour emerges from the denser sampling. This means that keeping a sampling step of 1m or halving it does not produce significant results aside from increasing the computational time.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

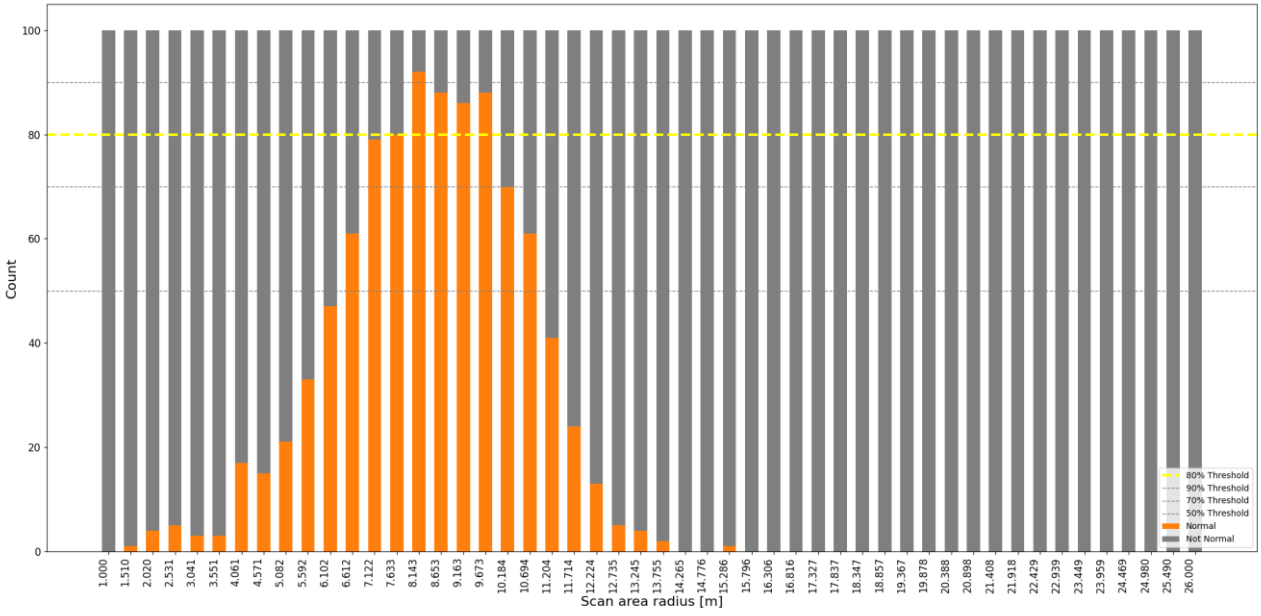




**Figure 21 Shapiro-Wilk test results for Pontrelli quarry outcrop with 50 different scan area radii.**

In contrast, for the Lilstock outcrop case study, the initial sampling consisted of 50 different scan area radii (min. radius = 0.005 m, max. radius = 18 m, sampling step = 0.305 m), as the higher fracture density suggested that $P_{21}$ distributions might vary more substantially. Using these 50 radii, an anomalous behaviour at 0.005 m was detected, prompting a local densification of radii to investigate it further. As a cross-check, the analysis was repeated using an equal spacing of 1 m, as done for the Pontrelli quarry (Fig. 22). In this case, although the overall trend of acceptances remained unchanged, the anomaly could not be detected.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026

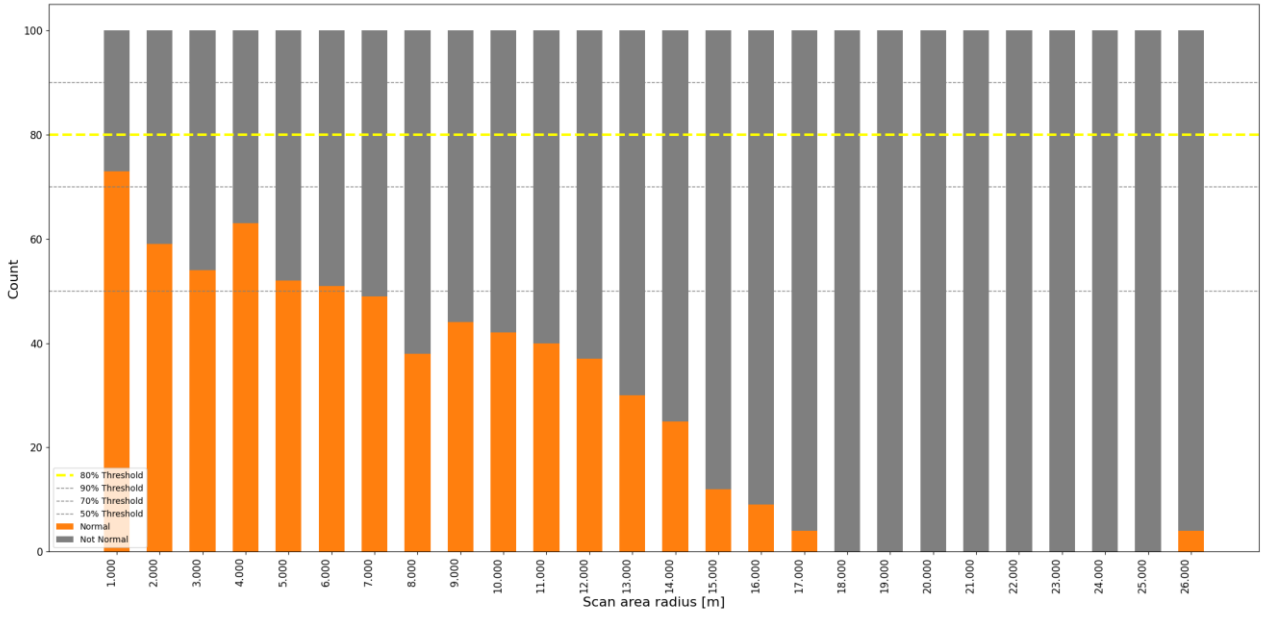




**Figure 22 Shapiro-Wilk test results for Lilstock outcrop for scan area radii with an equal spacing of 1m.**

In our case studies, we observed that starting with a relatively high number of radii is generally preferable: in the worst-case scenario, as for the Pontrelli quarry, the additional radii are simply redundant and do not affect the results, while in other situations, such as the Lilstock outcrop, a higher initial number allows detection of anomalies that might otherwise be missed. Alternative strategies, such as sampling radii according to a logarithmic or geometric series, could also be explored depending on the fracture network characteristics and the objectives of the analysis. In any case, if proper upper and lower limits to the REA range do not emerge in the first stages of the analysis, increasing the resolution could be a reasonable strategy.

### 5.4 Spatial autocorrelation of $P_{21}$ data and applicability of ANOVA analysis

In section 3 it is explained that the requirement for an independent dataset, that is a necessary condition to avoid Type I (false positive) errors in ANOVA, is addressed through the implementation of a random sampling with replacement strategy. However, this solution only works partially for spatially autocorrelated data (Griffith, 1978, 1992).

Spatial autocorrelation is a property of a random spatial variable where, taking values at pairs, it can be observed that their correlation is correlated with their proximity in the spatial domain, i.e. values of nearby data points are more related than distant ones (Legendre, 1993). In the context of fracture networks and their related parameters, autocorrelation can be produced by the underlying mechanical laws that guided the fracture generating process, and at a bigger scale by the geological and structural context (i.e. mechanical stratigraphy, deformation gradients, etc.). Therefore, quite often in a fully

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


developed fracture network, the position of a certain fracture is not independent from the others surrounding it, at least at some scales of observation. Depending on the interplay between the scale of spatial autocorrelation (if any) and the size of scan areas, this could lead to a partial predictability of the $P_{21}$ value collected with two adjacent scan areas, and consequently to a possible violation of the independence hypothesis of the ANOVA test (Griffith, 1978, 1992). It is therefore important to contextualise the methodology presented in this contribution within the problem of spatial autocorrelation and understand the conditions under which our methodology can be applied.

The ANOVA independence assumption requires both independence between groups (e.g., $P_{21}$ data from 1 m vs. 2 m scan radii) and independence of observations within each group (e.g. data collected with scan areas with radius of 1 m). Regarding independence between groups, spatial autocorrelation would pose a significant concern if the study factor was the geographic position of the scan areas rather than the scan area radius (Legendre, 1993). This would have meant checking whether the mean of $P_{21}$ variable would undergo significant changes when the position of the scan areas changes, for instance due to the presence of clusters, anticlusters, or spatial gradients (as e.g. in Bistacchi et al., 2020).

However, in the methodology presented here, the factor taken into consideration is the scan area radius, that is systematically changed, and every group of data is composed of scan areas sampled in random positions from the whole region considered. This implies that the position is not taken into consideration as a factor of the analysis.

Moving to the in-group independence, there is always the possibility that two randomly positioned scan areas will fall close together and sometimes partly overlap, consequently, P21 collected in these two scan areas will not be independent. As shown in Figure 3, the probability for two scan areas to be positioned close together or overlap increases as the radius of the scan areas increases, and for instance in Cava Pontrelli scan areas with a radius of 26 m are much more overlapped with respect to scan areas with a radius of 1 m. This means that the possibility of obtaining an independent sample conflicts with the need to obtain a similar sample size for the various radiuses as discussed in 5.2. In this situation, it is necessary to accept the possibility that part of the collected scan area are not spatially independent, possibly increasing the Type 1 error (false positive), resulting in a slightly higher number of acceptances. Obviously, this effect becomes progressively more important as the radius of the scan area increases when compared to the size of the outcrop, so for instance it will be more pronounced at Cava Pontrelli with respect to the Lilstock outcrop, and it will become unavoidable where the REA approaches the outcrop size. Noteworthy, this effect is partially mitigated by the fact that our approach is based on several parameters, and ANOVA is only one of the parameters that contribute to the definition of REA range, therefore performing the whole analysis workflow outlined in this contribution and using ANOVA results per se is insufficient.

### 5.5 REA upper limit vs representativity limit

In the introduction section, REA has been defined as a range with a lower limit corresponding to the discontinuum-continuum transition and an upper limit that represents another transition to a large-scale heterogeneity. However, in the context of a natural outcrop with limited extent, the detected upper limit could be constrained by the overall outcrop size and the geometry of the interpretation boundary.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


For the Pontrelli quarry, when the scan area radius exceeds approximately 14 m (Fig. 3), a strong overestimation emerges relative to the normal model, producing left-tailed distributions characterized by a few small values. At this stage, the variance approaches zero and the $P_{21}$ values converge toward a single value, the only one that can be sampled, as the scan areas become increasingly restricted by the boundary size and shape and thus strongly overlapped (Fig. 3, larger radii). The observed decrease in acceptance rates for the Shapiro–Wilk and Levene tests is therefore linked to a loss of measurement representativity rather than to the detection of the true upper limit of the REA range. Nevertheless, this representativity limit provides essential information, as it defines the largest representative scan area that can be extracted from the dataset. Additionally, this behavior could be considered as a significant clue to a situation already depicted in the previous section 5.4, where spatial autocorrelation becomes too much important as the scan areas overlap too much.

In the Lilstock outcrop case study, we observed a somewhat different situation. Although REA conditions are achieved in a range limited to small scan area radii (0.211 – 0.3 m), homoscedasticity acceptances increase again at larger radii after a sill between 0.27 – 1.022 m (Figure 18). In this case, the importance of the normality test in defining the REA is highlighted, and we propose that the acceptances at large radii in the Levene test are due to this test losing power when the samples have non-normal distributions.

### 5.6 Relevance of a multiparametric methodology for REA definition

In this paper, we presented a multiparametric methodology based on the mean, variance and shape of the $P_{21}$ distribution to detect the REA range, starting from data collected from outcrops large enough to distribute scan areas with radius that span some order of magnitude in radius. Compared to several studies in the scientific literature, where the REV is generally discussed, we want to highlight two distinctive features of our methodology: (i) considering REA instead of REV, and (ii) using a multiparametric approach.

The REA is a parameter that can be properly measured on an outcrop, a 2D slice provided by erosion in a 3D fractured volume, while any attempt at defining a property of a REV based on field data will involve some assumption, and therefore must be considered a model. For instance some authors model the REV with a stochastic approach (Esmaieli et al., 2010; Zhang et al., 2012; Huang et al., 2020), but the size of the REV depends not only on size and orientation parameters of the fractures, but also on their spatial distribution. Using stochastic models to generate a volume for REV estimation actually replaces the natural fracture-generating process with the spatial point process assumed in the stochastic model, which in most commercial and open-source stochastic DFN algorithms is a Poisson point process, where the position of each fracture centroid is independent of those already generated. Consequently, calculating the REV from a stochastic DFN corresponds to evaluating the REV of a different, model-derived fracture network, with a completely different spatial distribution (and probably a very different connectivity) with respect to the natural network.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Other studies propose a simple definition of the REV based on a multiplier of either the mean trace length (Li and Zhang, 2011), the average spacing (Zeeb et al., 2013), the average block size (Rohrbaugh Jr. et al., 2002) or the maximum spacing (Pariseau et al., 2008).

Applying these methods to our case studies, the multipliers proposed by Li and Zhang (2011) can only be applied to connected networks. Consequently, they cannot be applied to Pontrelli quarry, as the REA there is calculated on a single fracture set. Assuming Lilstock outcrop represents a highly connected network, the proposed multiplier is 2.8. Using the mean trace length of the fracture sets reported by Passchier et al. (2021), the average fracture length in Lilstock outcrop is approximately 7 m, resulting in a circle with a diameter of 19.6m, yielding an equivalent REA radius of about 10 m, at least one order of magnitude larger than the upper value detected in this study.

Rohrbaugh Jr. et al. (2002) suggested that the REV should exceed the average block size. Using the polygon shapefile provided by Prabhakaran et al. (2021), the average block size in Lilstock outcrop is 0.05 m², corresponding to a circle with an equivalent radius of 0.13 m. In this case, the REA identified with our method aligns well with this criterion.

Zeeb et al. (2013) proposed that the REV should exceed the average fracture spacing. For Pontrelli quarry Set 1, Benedetti et al. (2025) report an average spacing of 1.55 m. Here, our method identifies a considerably larger REA, ranging from 7 m to 10 m of scan area radius.

Pariseau et al., (2008) proposed that the REV is reached at roughly twice the maximum total fracture spacing for networks with multiple sets. For Lilstock outcrop, the largest block size is 0.5 m², yielding an equivalent REA radius of 0.399 m, which approximately corresponds to the onset of homogeneous behaviour in the $P_{21}$ values.

While these approaches have been tested in multiple case studies, they remain less informative than multi-parametric methods and do not allow defining the upper limit of the REA (or REV) range, that with our approach was detected in both the Pontrelli and Lilstock case studies. Also, regarding the lower limit of the range, the outcome of our analysis may be similar, but it is achieved through a more informed and approach rather than a 'black box' one.

### 5.7 Comparison of REA range between case studies

We applied the methodology to two markedly different natural case studies (Fig. 2): the Pontrelli quarry outcrop, where a single fracture set is consistently detectable across the entire surface, and the Lilstock outcrop, which instead exhibits an extremely dense network composed of eight distinct fracture sets (Passchier et al., 2021). This difference is reflected in the results.

In Pontrelli quarry we identified a REA range between 7 m and 10 m. The upper limit is influenced by the limited size of the outcrop, thus what is identified as an "upper limit" is instead a representativity limit, that is nonetheless an important information and can be seen as a sort of "lower bound to the upper limit". In the Lilstock outcrop case study, the detected REA range is between one and two orders of magnitude smaller than that of the Pontrelli quarry. This difference can be due to several factors. First, the Lilstock network is significantly denser — even when compared with those portions of the Pontrelli quarry where all fracture sets are visible (Casiraghi et al., 2025b) Lilstock still exhibits a much higher $P_{21}$. Second,

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


since scan areas are circular windows to avoid orientation bias. larger radii are required to reach representativity in a network dominated by a single, consistently oriented fracture set, as occurs at Pontrelli. In contrast, highly connected and multi-directional networks, such as at Lilstock, stabilize much earlier. At Pontrelli we would likely have obtained a smaller lower bound for the REA if appropriately oriented anisotropic sampling windows were used, but this would have meant introducing a subjective bias in the analysis.

Another difference is observed in the behaviour of the test results beyond the REA range upper limit. In the Pontrelli quarry fracture network, acceptances exhibit a progressive decline, eventually reaching zero. This applies to both the Shapiro–Wilk and Levene tests. The effect arises because, as the scan radius increases, the placement of sampling windows becomes increasingly constrained by the interpretation boundary. Consequently, the largest scan areas tend to overlap, reducing spatial variability as discussed in section 5.4. The effect is further accentuated by the presence of no-data zones, which limit positioning of the scan areas even more (Fig. 3). In Lilstock, the Levene test shows a very different trend (Fig. 18): after an abrupt decrease in acceptances at small radii, the equality of the variance is consistently accepted again for larger scan areas but this can be an effect of this test losing power since distributions are not normal for large radii.

## 6 Conclusions

In this work, we addressed the problem of REV/REA determination for the $P_{21}$ parameter in natural fractured outcrops. We proposed defining the REA as a range, characterized by both a lower and an upper limit. This definition is particularly relevant in the context of reservoir modelling, as it allows us to define a range of cell size resolutions, from the one associated with the lowest computational cost (upper limit) to the lower limit that defines the smallest representative cell size. To calculate the REA range, we applied the ANOVA test, where the underlying assumptions are verified both with formal statistical tests and diagnostic plots.

The applicability of the method was demonstrated on two case studies with distinct characteristics. In the Pontrelli quarry, the analysis was performed on a single fracture set. The REA range was defined following the hypothesized definition. It was observed that its upper limit does not correspond to a true REA upper limit but is related to a loss of representativity due to the scan area radii size approaching the interpretation boundary size. Beyond this limit, the $P_{21}$ distributions become left-tailed. In the Lilstock outcrop case study, a different behavior is observed since the REA range is confined to smaller radii and the outcrop is much larger, hence also the right limit of the REA can be defined with confidence.

The proposed method combines multiple statistical tests with diagnostic plots, offering a multi-parameter framework for REA estimation. Rather than relying on a single numerical output, it enables the user to explore the reasons behind different statistical behaviors, enhancing interpretability and reducing the “black box” aspect often associated with statistical tests.

Some aspects of the procedure remain user-dependent, notably the sample size and the discretization strategy for the scan area radii. Both parameters directly affect the outcome of formal statistical tests. However, given that it is always possible to extract additional scan areas it is possible, at least, to qualitatively evaluate the sample size effect on the outcome of

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


statistical tests. As for the radii discretization, although there is not a quantitative rule, our results indicate that using a larger number of radii is preferable. This approach increases computational cost but does not introduce significant drawbacks, while it may reveal anomalous behaviors that would otherwise remain undetected.

## 7 Code and Data availability

Codes and data are available at the following GitHub repositories owned by the Gecos-lab group of the University of Milano-Bicocca (https://github.com/gecos-lab, Last access date April 24 2026):

- Pontrelli quarry dataset: Available at the following repository: https://doi.org/10.5281/zenodo.17483588 (Casiraghi et al., 2025a)
- FracArea: Python code to run the statistical analysis described in this contribution https://doi.org/10.5281/zenodo.19735298 (Bistacchi and Casiraghi, 2026).

## 8 Author contributions

SC, DB, AB and SM conceived the research. SC wrote the paper. SC, AB and GB wrote the code. AB, SM, FA, DB heavily contributed to the revision process and in defining the final structure of the paper. MM, FB, CA participated in fundamental discussions. AB led the project and provided funding.

## 9 Competing interests

The authors declare that there are no competing interests.

## 10 Acknowledgments

We thank Eni SpA for funding, reviewing and agreeing to publish the results of the study. We thank SABAP BA (Soprintendenza Archeologia, Belle Arti e Paesaggio per la Città Metropolitana di Bari) for granting the access to the Pontrelli quarry, making this work possible. We also warmly thanks prof. Sophie Viseur for the valuable discussion regarding the spatial autocorrelation of P21 data.

## 11 Financial support

The project is funded by Eni S.p.A, Global Natural Resources, San Donato Milanese, Italy through the Joint Research Agreement 5210001818 between Eni and the University of Milano-Bicocca.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


## 12 References

Agliardi, F., Crosta, G. B., Meloni, F., Valle, C., and Rivolta, C.: Structurally-controlled instability, damage and slope failure in a porphyry rock mass, Tectonophysics, 605, 34–47, https://doi.org/10.1016/j.tecto.2013.05.033, 2013.

Bear, J.: Dynamics of Fluids in Porous Media, Soil Sci., 120, 162, https://doi.org/10.1097/00010694-19750800000022, 1975.

Benedetti, G., Casiraghi, S., Bertacchi, D., and Bistacchi, A.: Unbiased statistical length analysis of linear features: adapting survival analysis to geological applications, Solid Earth, 16, 367–390, https://doi.org/10.5194/se-16-367-2025, 2025.

Bistacchi, A., Casiraghi, S. gecos-lab/FracArea: v1.0.0 (v1.0.0), Zenodo [code]. https://doi.org/10.5281/zenodo.19735298, 2026.

Bistacchi, A., Mittempergher, S., Martinelli, M., and Storti, F.: On a new robust workflow for the statistical and spatial analysis of fracture data collected with scanlines (or the importance of stationarity), Solid Earth, 11, 2535–2547, https://doi.org/10.5194/se-11-2535-2020, 2020.

Bistacchi, A., Mittempergher, S., and Martinelli, M.: Digital Outcrop Model Reconstruction and Interpretation, in: 3D Digital Geological Models, John Wiley & Sons, Ltd, 11–32, https://doi.org/10.1002/9781119313922.ch2, 2022.

Brown, M. B. and Forsythe, A. B.: Robust Tests for the Equality of Variances, J. Am. Stat. Assoc., 69, 364–367, https://doi.org/10.2307/2285659, 1974.

Casiraghi, S.: Pontrelli quarry dataset, Zenodo [data set], https://doi.org/10.5281/zenodo.17483588, 2025a.

Casiraghi, S., Benedetti, G., Bertacchi, D., Mittempergher, S., Agliardi, F., Monopoli, B., La Valle, F., Martinelli, M., Bigoni, F., Albertini, C., and Bistacchi, A.: An integrated workflow for parametrization of fracture network geometry in digital outcrop models, Solid Earth, 16, 1351–1382, https://doi.org/10.5194/se-16-1351-2025, 2025b.

Dershowitz, W. S. and Herda, H. H.: Interpretation of fracture spacing and intensity, The 33rd U.S. Symposium on Rock Mechanics (USRMS), ARMA-92-0757, 1992.

Esmaieli, K., Hadjigeorgiou, J., and Grenon, M.: Estimating geometrical and mechanical REV based on synthetic rock mass models at Brunswick Mine, Int. J. Rock Mech. Min. Sci., 47, 915–926, https://doi.org/10.1016/j.ijrmms.2010.05.010, 2010.

Forstner, S. R. and Laubach, S. E.: Scale-dependent fracture networks, J. Struct. Geol., 165, 104748, https://doi.org/10.1016/j.jsg.2022.104748, 2022.

Franzosi, F., Casiraghi, S., Colombo, R., Crippa, C., and Agliardi, F.: Quantitative Evaluation of the Fracturing State of Crystalline Rocks Using Infrared Thermography, Rock Mech. Rock Eng., 56, 6337–6355, https://doi.org/10.1007/s00603-023-03389-x, 2023a.

Franzosi, F., Crippa, C., Derron, M.-H., Jaboyedoff, M., and Agliardi, F.: Slope-Scale Remote Mapping of Rock Mass Fracturing by Modeling Cooling Trends Derived from Infrared Thermography, Remote Sens., 15, 4525, https://doi.org/10.3390/rs15184525, 2023b.

Gerstner, R., Maschler, A., Schneider-Muntau, B., Agliardi, F., Avian, M., Frießenbichler, M., and Zangerl, C.: The critical role of fracture propagation in the evolution of extensive, structurally preconditioned rockslides, Eng. Geol., 358, 108359, https://doi.org/10.1016/j.enggeo.2025.108359, 2025.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Griffith, D. A.: A Spatially Adjusted ANOVA Model, Geogr. Anal., 10, 296–301, https://doi.org/10.1111/j.1538-4632.1978.tb00661.x, 1978.

Griffith, D. A.: A spatially adjusted N-way ANOVA model, Reg. Sci. Urban Econ., 22, 347–369, https://doi.org/10.1016/0166-0462(92)90034-X, 1992.

Hoek, E. and Brown, E. T.: Practical estimates of rock mass strength, Int. J. Rock Mech. Min. Sci., 34, 1165–1186, https://doi.org/10.1016/S1365-1609(97)80069-X, 1997.

Huang, H., Shen, J., Chen, Q., and Karakus, M.: Estimation of REV for fractured rock masses based on Geological Strength Index, Int. J. Rock Mech. Min. Sci., 126, 104179, https://doi.org/10.1016/j.ijrmms.2019.104179, 2020.

Hudson, J. A. and Harrison, J. P.: Chapter 4. In Situ Stress, in: Engineering Rock Mechanics, Pergamon, Amsterdam [etc.], 1997.

Kozak, M. and Piepho, H.-P.: What's normal anyway? Residual plots are more telling than significance tests when checking ANOVA assumptions, J. Agron. Crop Sci., 204, 86–98, https://doi.org/10.1111/jac.12220, 2018.

Kulatilake, P. H. S. W. and Panda, B. B.: Effect of Block Size and Joint Geometry on Jointed Rock Hydraulics and REV, J. Eng. Mech., 126, 850–858, https://doi.org/10.1061/(ASCE)0733-9399(2000)126:8(850), 2000.

Legendre, P.: Spatial Autocorrelation: Trouble or New Paradigm?, Ecology, 74, 1659–1673, https://doi.org/10.2307/1939924, 1993.

Li, J. H. and Zhang, L. M.: Connectivity of a network of random discontinuities, Comput. Geotech., 38, 217–226, https://doi.org/10.1016/j.compgeo.2010.11.010, 2011.

Marinos, V., Marinos, P., and Hoek, E.: The geological strength index: applications and limitations, Bull. Eng. Geol. Environ., 64, 55–65, https://doi.org/10.1007/s10064-004-0270-5, 2005.

Martinelli, M., Bistacchi, A., Mittempergher, S., Bonneau, F., Balsamo, F., Caumon, G., and Meda, M.: Damage zone characterization combining scan-line and scan-area analysis on a km-scale Digital Outcrop Model: The Qala Fault (Gozo), J. Struct. Geol., 140, 104144, https://doi.org/10.1016/j.jsg.2020.104144, 2020.

Mauldon, M., Dunne, W. M., and Rohrbaugh, M. B.: Circular scanlines and circular windows: new tools for characterizing the geometry of fracture traces, J. Struct. Geol., 23, 247–258, https://doi.org/10.1016/S0191-8141(00)00094-8, 2001.

Micceri, T.: The Unicorn, The Normal Curve, and Other Improbable Creatures, Psychol. Bull., 105, 156–166, 1989.

Mooi, E., Sarstedt, M., and Mooi-Reci, I.: Market Research, Springer Singapore, Singapore, https://doi.org/10.1007/978-981-10-5218-7, 2018.

Ni, P., Wang, S., Wang, C., and Zhang, S.: Estimation of REV Size for Fractured Rock Mass Based on Damage Coefficient, Rock Mech. Rock Eng., 50, 555–570, https://doi.org/10.1007/s00603-016-1122-x, 2017.

Panza, E., Agosta, F., Rustichelli, A., Zambrano, M., Tondi, E., Prosser, G., Giorgioni, M., and Janiseck, J. M.: Fracture stratigraphy and fluid flow properties of shallow-water, tight carbonates: The case study of the Murge Plateau (southern Italy), Mar. Pet. Geol., 73, 350–370, https://doi.org/10.1016/j.marpetgeo.2016.03.022, 2016.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Panza, E., Agosta, F., Rustichelli, A., Vinciguerra, S. C., Ougier-Simonin, A., Dobbs, M., and Prosser, G.: Meso-to-microscale fracture porosity in tight limestones, results of an integrated field and laboratory study, Mar. Pet. Geol., 103, 581–595, https://doi.org/10.1016/j.marpetgeo.2019.01.043, 2019.

Pariseau, W. G., Puri, S., and Schmelter, S. C.: A new model for effects of impersistent joint sets on rock slope stability, Int. J. Rock Mech. Min. Sci., 45, 122–131, https://doi.org/10.1016/j.ijrmms.2007.05.001, 2008.

Passchier, M., Passchier, C. W., Weismüller, C., and Urai, J. L.: The joint sets on the Lilstock Benches, UK. Observations based on mapping a full resolution UAV-based image, J. Struct. Geol., 147, 104332, https://doi.org/10.1016/j.jsg.2021.104332, 2021.

Prabhakaran, R., Urai, J. L., Bertotti, G., Weismüller, C., and Smeulders, D. M. J.: Large-scale natural fracture network patterns: Insights from automated mapping in the Lilstock (Bristol Channel) limestone outcrops, J. Struct. Geol., 150, 104405, https://doi.org/10.1016/j.jsg.2021.104405, 2021.

Rajaraman, S., Jaeger, S., and Antani, S. K.: Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images, PeerJ, 7, e6977, https://doi.org/10.7717/peerj.6977, 2019.

Rochon, J., Gondan, M., and Kieser, M.: To test or not to test: Preliminary assessment of normality when comparing two independent samples, BMC Med. Res. Methodol., 12, 81, https://doi.org/10.1186/1471-2288-12-81, 2012.

Rohrbaugh Jr., M. B., Dunne, W. M., and Mauldon, M.: Estimating fracture trace intensity, density, and mean length using circular scan lines and windows, AAPG Bull., 86, https://doi.org/10.1306/61EEDE0E-173E-11D7-8645000102C1865D, 2002.

Sari, M.: Determination of representative elementary volume (REV) for jointed rock masses exhibiting scale-dependent behavior: a numerical investigation, Int. J. Geo-Eng., 12, 34, https://doi.org/10.1186/s40703-021-00164-1, 2021.

Shapiro, S. S. and Wilk, M. B.: An Analysis of Variance Test for Normality (Complete Samples), Biometrika, 52, 591–611, https://doi.org/10.2307/2333709, 1965.

Song, B., Zheng, N., Li, D., Chen, R., and Li, L.: Reconstructing DEM using TLS point cloud data and NURBS surface, Trans. Nonferrous Met. Soc. China, 25, 3165–3172, https://doi.org/10.1016/S1003-6326(15)63947-4, 2015.

Souza, R. R. D., Toebe, M., Mello, A. C., and Bittencourt, K. C.: Sample size and Shapiro-Wilk test: An analysis for soybean grain yield, Eur. J. Agron., 142, 126666, https://doi.org/10.1016/j.eja.2022.126666, 2023.

Stahle, L. and Wold, S.: Analysis of variance (ANOVA), Chemom. Intell. Lab. Syst., 6, 259–272, https://doi.org/10.1016/0169-7439(89)80095-4, 1989.

Tukey, J. W.: Exploratory Data Analysis, Addison-Wesley Publ. Co. Read., https://doi.org/10.1002/bimj.4710230408, 1977.

Wang, M., Kulatilake, P. H. S. W., Um, J., and Narvaiz, J.: Estimation of REV size and three-dimensional hydraulic conductivity tensor for a fractured rock mass through a single well packer test and discrete fracture fluid flow modeling, Int. J. Rock Mech. Min. Sci., 39, 887–904, https://doi.org/10.1016/S1365-1609(02)00067-9, 2002.

Xia, L., Zheng, Y., and Yu, Q.: Estimation of the REV size for blockiness of fractured rock masses, Comput. Geotech., 76, 83–92, https://doi.org/10.1016/j.compgeo.2016.02.016, 2016.

https://doi.org/10.5194/egusphere-2026-2409
Preprint. Discussion started: 12 May 2026


Numerical modelling of brittle fracture and step-path failure: from laboratory to rock slope scale: https://summit.sfu.ca/item/9076, last access: 27 October 2025.

Zeeb, C., Gomez-Rivas, E., Bons, P. D., and Blum, P.: Evaluation of sampling methods for fracture network characterization using outcrops, AAPG Bull., 97, 1545–1566, https://doi.org/10.1306/02131312042, 2013.

Zhang, W., Chen, J., Liu, C., Huang, R., Li, M., and Zhang, Y.: Determination of Geometrical and Structural Representative Volume Elements at the Baihetan Dam Site, Rock Mech. Rock Eng., 45, 409–419, https://doi.org/10.1007/s00603-011-0191-0, 2012.