



# Alternative methods for semi-automatic clusterization and extraction of discontinuity sets from 3D point clouds

Stefano Cardia<sup>1</sup> · Biagio Palma<sup>2</sup> · Francesco Langella<sup>2</sup> · Marco Pagano<sup>2</sup> · Mario Parise<sup>1</sup>

Received: 1 March 2023 / Accepted: 4 June 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

Recent advances in the use of remote sensing techniques allow the acquisition of dense 3D information helpful for the characterization of the rock mass joints. This implies the necessity of having robust and reliable methods to evaluate and extract the primitive geometries representing discontinuities on a rock outcrop. Moreover, these methods have to be easy to use, fast and accurate, which leads nowadays to the tendency of developing automated methods, often having limitations as concerns processing time, definition of parameters and, especially, accuracy. We present here an alternative approach based on two new semi-automatic algorithms, the Iterative Pole Density Estimation (IPDE) and the Supervised Set Extraction (SSE), used in combination with well-known and suitable clustering and density estimation methods. The IPDE performs an analysis based on a threshold value, within which it searches for coplanar points in a range of tolerance, automatically eliminating those below the established threshold, and then finding principal orientations by Kernel Density Estimation (KDE) and identifying clusters by a manual evaluation or through automated clustering methods. The SSE is a tool that allows to extract discontinuity sets from point clouds through an approach aimed to combine observations made in situ with digital results, taking into account the crucial importance of traditional analysis by an expert user. The method was tested in Campania (Italy) at the Cocceio Cave and at the Cetara Tower cliff: at the cave, we were able to recognize an additional set, not identified during previous digital analysis. In the second case, a fully automatic technique, with little or no human intervention on the point cloud, was compared with a previously made supervised method to perform the semi-automatic approach, eventually checking both results with those from traditional surveys, which led the whole analysis to shift the focus on the combined method proposed.

**Keywords** Data science · Algorithms · Geomechanics · Discontinuity · 3D point cloud · Rock Mass · Laser scanner

## Introduction

Geological hazards associated with rock mass stability are widespread, and several natural factors act in predisposing many types of geological environments to landslides and rock failures, covering a wide range of size and volume. A rock mass is defined as formed by blocks of rock material characterized by systems of discontinuities that significantly affect its mechanical behavior, predisposing peculiar types

of kinematics (Bieniawski 1989). For this reason, the issues related to the study of rock masses have always been of primary importance. For protection of the territory, safeguard of human lives and urban areas, and of natural and cultural heritage as well, it is essential to define the susceptibility to rock instabilities, identifying the potentially affected areas. The orientation of discontinuities, as well as other properties (i.e. spacing, persistence, roughness, infilling, weathering), have a capital importance on the geo-mechanical behavior of the rock mass (Bieniawski 1973; Calcaterra and Parise 2010; Piteau 1970). These characteristics play a crucial role and contribute to the geo-structural and geo-mechanical characterization of the rock mass and are defined as the key features for a complete survey according to the 1978 ISRM recommendations (ISRM 1978). In case of soluble rocks such as carbonates and evaporites, the presence of voids related to karst dissolution should also be taken into account (Andriani

---

Communicated by: H. Babaie

✉ Stefano Cardia  
stefano.cardia@uniba.it

<sup>1</sup> Earth and Environmental Sciences Department, Aldo Moro University, Bari, Italy

<sup>2</sup> IdroGeo Srl, Vico Equense, Italy

and Parise 2015, 2017; Palmer 2007), since these features strongly control the water flow, with significant effects as concerns stability (Parise 2022).

To quantitatively describe the structural set-up of rock masses, new methods have been proposed in the last decades, with particular reference to LiDAR (Light Detection and Ranging) and Photogrammetry, both often coupled with RPAS (Remotely Piloted Aircraft Systems), as essential tools for the geomechanical analysis (Abellán et al. 2014; Barnobi et al. 2009; Jaboyedoff et al. 2012; Oppikofer et al. 2009; Tomás et al. 2020; Viero et al. 2010). Application of these methods allows the acquisition of high resolution geo-localized 3D point clouds over large areas, in a relatively short time. The general theme, represented by the use of remote systems for the definition of a quantitative assessment of rock mass instability, is today a subject of great interest: the recent scientific literature demonstrates that these methodologies still require the development of innovative survey techniques and data analysis, in order to fully define their potential, advantages, but above all to ensure their reliability (Abellan et al. 2016; Cardia et al. 2022; Ferrero et al. 2009; Galgaro et al. 2004; Liotine et al. 2021a, b; Pagano et al. 2020; Riquelme et al. 2014, 2017; Slob 2010; Slob et al. 2005). For several years researchers made efforts to implement new algorithms to standardize the extraction of primitive geometries from 3D point clouds, in order to identify detailed geometric characteristics on scanned structures of the real world and, therefore, to identify discontinuities and blocks on rock masses (Borrmann et al. 2011; Dewez et al. 2016; Hammah & Curran 1998; Jaboyedoff et al. 2007; Li et al. 2019; Lombardi et al. 2011; Menegoni et al. 2021; Roncella & Forlani 2005; Schnabel et al. 2007; Tran et al. 2015; Xia et al. 2020). New algorithms have to be reproducible, ready to use and, possibly, fast, automatic, and reliable. The problem with fully automatic methods is that they have several limitations, such as incorrect computation of the normals, or uncertainty in the automatic input of parameters, which can greatly vary from case to case. Results can often be misleading, due to the geometrical peculiarities normally presented by a rock outcrop. Even with the huge progresses done in the last decades for softwares able to manage and analyze point clouds, acquiring the geo-mechanical data requires necessarily the control by experts in order to select the planes of geological significance, and delete all those related to anthropogenic works. This is a crucial point to highlight, since our firm belief is that it is not possible yet to entirely rely on automated systems.

Therefore, a correct and complete characterization of a rock mass cannot disregard the importance of an analysis carried out by an expert user, together with the precision of unbiased input parameters. Thus, a previous analysis conducted by means of traditional methods, or in situ observations, is crucially important to standardize the subsequent process for a reliable definition of the principal characteristics of both rock blocks and discontinuity systems. The traditional methods of surveying rock outcrops require the work of specialized geologists to carry out accurate and time-consuming surveys on sites that, especially in steep mountains and underground environments, are logistically difficult, often requiring intervention of rock climbers. This operation, which is definitely not simple due to logistics, requires multi-disciplinary skills. Consequently, there is the risk of making sampling errors, or to collect insufficient data for a complete geo-structural characterization. For this reason, in situ observations should be coupled with those acquired through new technologies, to have more robust and reliable data to perform geostatistical analysis. Moreover, rock outcrops always present complex geometrical relations among joint, faults and fractures; it is therefore necessary to detect and evaluate ranges of angular values for the discontinuities, rather than extracting a single value.

Starting from the above considerations, this paper proposes a different approach for the semi-automatic evaluation and extraction of discontinuity sets affecting rock mass stability, using 3D point cloud data. The work is structured as follows: in this section, we have summarized the main objectives and a brief state of the art regarding rock mass stability analysis; the main novel contributions of the proposed method is presented in Section "Methodology", which includes the methodology for the discontinuities evaluation/extraction, and a description of the material used: it illustrates (a) the user-supervised method of reducing the point cloud density following a minimum threshold and the coplanarities using the Iterative Pole Density Estimation (IPDE), (b) the semi-automatic identification of discontinuity sets using a Kernel Density Estimation (KDE) analysis coupled with a manual sets selection or automated clustering, and (c) the user-defined extraction of the main sets with the Supervised Set Extraction (SSE). Section "Case studies" summarizes the results of the method applied to two case studies. In the second case, this is also done by comparing a density estimation of the various set of discontinuities made with little or no previous manual intervention on the whole point cloud, with a supervised method made before performing the density estimation, linking at last both the outcomes with the results from

traditional analysis. Section "[Discussion](#)" discusses the most important matters of the results, and, eventually, the last section explores the future perspective of the research. This paper includes the public availability of the complete programming code used, along with the mention of all used modules, in order to provide the method validation for other researchers.

## Methodology

The proposed method aims to evaluate and extract structural discontinuities from 3D digital point clouds of rock outcrops. Starting from a set of raw data points ( $D$ ), that could be represented by the point coordinates  $X$ ,  $Y$ , and  $Z$ , the surface described by the points is limited by discontinuities, which can be classified into sets, each defining a single plane, with a more or less wide angle of tolerance in terms of orientation. The proposed methodology is developed through the following 4 main steps: (1) determination of the approximate orientation of every normal of the points in  $D$ ; (2) statistical analysis of  $D$  with IPDE, which consists in detecting all relevant poles of planes on the rock mass, representing the different sets of discontinuities; (3) manual or automatic clusters identification: localization of the points defining different clusters in space; and (4) extraction of the identified sets with SSE.

## Materials

The surveys were carried out with a RIEGL VZ400 laser scanner, and integrated with a photogrammetric survey from RPAS, performed with the aircraft Italdron 4HSE, equipped with a Sony Alpha 7 (24.7 MP sensor) digital camera. The scanner is a class 1 laser scanner, with a detection range  $> 500$  m, equipped with a high-definition digital camera Nikon D700 (12.1 MP sensor) and an integrated inclinometer, lead laser, GPS system and compass. It performs high velocity acquisition, precisely more than 120.000 pts/sec with 3 mm precision. The drone is assembled for critical scenarios, it reaches a maximum flying height of 150 m and has a maximum range of 1,5 km. The sensor size is  $35.8 \times 23.9$  mm (42 MP). It is equipped with a parachute, and a 3 axes gimbal. Further, it has an autonomy of 30 min and a wind resistance up to 50 km/h. The whole analysis of acquired data was conducted on an Intel i5-8250U 8th generation 1.60 GHz processor, 16 GB DDR4 RAM, integrated GPU and Windows 10 operating system. The georeferencing of the point clouds was done using a Leica GS14 dual-frequency GNSS and a Leica TS16 total station. The

adopted system is the WGS84 UTM33. For the visualization of the point clouds, the internal software of the laser scanner RiSCAN PRO 2.0 was used, and the results were visualized by means of the open GPL software CloudCompare v2.11beta.

## Procedure

Before performing the main steps of the method, in the pre-processing phase it is necessary to clean the point cloud by manually removing all spurious elements without geological meaning (e.g. vegetation, floor points, walls, etc.) as these areas could lead to unnecessary computation. A point sub-sampling is also carried out to reduce the large density of points in excess, especially for small areas, but above all to lighten the cloud and limit long computation times, not leading to any further benefit for the recognition of discontinuities. This step was carried out using the sub-sampling algorithm in CloudCompare, a method that removes double points (points with equal values present around a specified radius), and does not affect the resolution of the three-dimensional cloud.

### Step 1: Determination of the orientation of the points

For this first step the values of the normal of every point in the initial dataset  $D$  have to be calculated. To compute normals on a point cloud, the local surface represented by a point and its neighbors must be estimated. This can be done by a Principal Component Analysis (PCA), that specifically computes eigenvectors of the covariance matrix  $C$  based on the local neighborhoods of each point, the equation of which can be written as:

$$C = \frac{1}{N} \sum_{i=1}^N (p_i - \mu)(p_i - \mu)^T \quad (1)$$

where  $p_1 \dots p_N$  are the points in a given neighborhood,  $\mu$  is the centroid of the points and  $T$  in the apex indicates the transpose of a matrix formed by exchanging rows into columns. The resulting eigenvector for each neighborhood of points, that can be chosen a priori, therefore represents the vector relative to the normal for each point ( $a, b, c$ ). The decomposition of the eigenvectors on  $x, y$  and  $z$  for each point gives the component of the normal. Given that  $(a, b, c) = n_{(x,y,z)}$  is the vector of a single plane derived by the equation  $a_x + b_y + c_z + d = 0$ , if then the condition  $n_x + n_y + n_z = 1$  is satisfied for every point, the procedure goes on, otherwise the result must be divided by the normal vector. Then, normals need to be adjusted in terms of orientation, because this method computes only the direction of the vector. It is

therefore a common procedure to determine a heuristic preferred orientation ( $\pm X, \pm Y, \pm Z$ ); in most of the cases this will be  $+Z$ . Such operation can be done by placing the condition  $|n_x|$  on every value present in the third column of the resultant 3-dimensional normal matrix. A fast and accurate way to compute normals is to use the computation plug-in available in CloudCompare: it performs the analysis after the user has chosen a local surface model (plane, 2d triangulation, quadric); then, the default neighbors extraction process relies on an octree-cell structure. The user must only choose the neighborhood radius, and all points relying in that radius will be used for the PCA analysis. Eventually, it proceeds with the user-specified heuristic orientation or with the Minimum Spanning Tree (MST) algorithm, which allows the user to determine again a radius by choosing a number of nearest normal neighbors, useful to compute the orientation. In this case a sort of region growing process attempts to re-orient all normals in a consistent way: it starts from a random point, and then propagates the normal orientation from one neighbor to the other. We calculated normals with the PCA method, and then relied on MST to have uniform orientations. However, the level of the initial noise and the number/distance of neighbors will change the surface of the 3D structure, so that the actions above cannot be done disregarding the previous pre-processing point cloud cleaning and a case sensitive parameters input. After this sub-step, it is necessary to convert the normal values into dip and dip direction values through a standardized method of conversion, lastly revised in Kemeny et al. (2006) and generally available on line (Girardeau-Montaut 2016). The dip ( $\delta$ ) value expressed in radians for every point is determined by the equation:

$$\delta = \arctan\left(\frac{|n_x|}{|n_z|}\right) \quad (2)$$

Where  $n_x$  and  $n_z$  are the components of the normal values, respectively along the X and Z axes. Since the result of this equation is in radians, it is sufficient to convert it in degrees by multiplying every value for 180 and dividing for  $\pi$ . The same occurs for dip direction ( $\gamma$ ) values, by means of the equation to express it in radians starting from every point normal triplet:

$$\gamma = \arctan\left(\frac{|n_x|}{|n_y|}\right) \quad (3)$$

Where  $n_y$  is the component of the normal value expressed along the Y axis. Conversion to degrees is done in the same way as for dip values. The only difference is that, in order to compute correctly the  $\gamma$  values before the conversion, a mask has to be made for every quadrant,

**Table 1** Structure of the masks made to compute the right azimuthal values along the Y axis

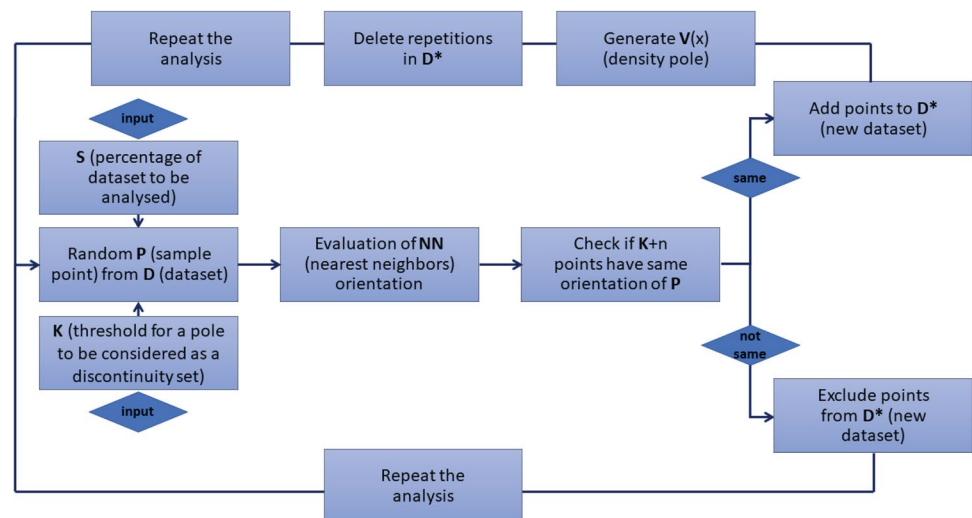
Quadrant	$N_x$	$N_y$	$N_z$	Mask
First	$N > 0$	$N > 0$	$N > 0$	None
Second	"	$N < 0$	"	$\gamma = \pi - \gamma$
Third	$N < 0$	"	"	$\gamma = \pi + \gamma$
Fourth	"	$N > 0$	"	$\gamma = 2\pi + \gamma$

except for the first, to re-orient the values with the 0 along the  $+Y$  axis and to compute the right azimuthal values (0–360). The structure of the masks is summarized in Table 1.

#### Step 2: Statistical analysis of datasets with the IPDE

This step is fundamentally based on recognition of the parallel orientation associated with the normals to the points. A certain number of given sample points ( $S$ ) are considered to perform the analysis. Starting from a random point, the algorithm iterates  $S$  times the analysis over  $D$ , searching for every point with the same orientation as the sample point, in a given range of tolerance for both dip and dip direction. This is very straightforward, since for every point an orientation matrix with these two values has been previously generated. Given the tolerance angles, very planar or almost planar features present in  $D$  are evaluated, and each time one of these has at least a density greater than a given threshold ( $K$ ), the system keeps and stores it in a new dataset ( $D^1$ ), that will be used to perform the subsequent analysis and plot the data. Values below the threshold will be discarded, as they probably represent minor discontinuities or transition edges between discontinuities, most likely non-geological planes. An overall look at the functionality of the algorithm is provided in alg. 1 and in Fig. 1. Next, the analysis is performed by means of the stereographic projection of the plane poles for data plotting. Every point orientation matrix is converted to stereographic projection, and the density of the poles for each region of the projection is calculated. The statistical analysis calculates the distribution by means of the KDE technique, a non-parametric way to estimate the probability density function of a random variable (Silverman 1986). This implementation was done through Gaussian kernels, useful for visualization purposes. Kamb and Schmidt density estimation methods were also used; these are instead parametric methods very similar to KDE, as they process Gaussian curves by discarding all points below a certain given threshold (Vollmer 1995).

**Fig. 1** Schematic flow chart of the IPDE



**Algorithm 1** Iterative Pole Density Estimation in pseudocode

D = Point cloud dataset

K = Pole density threshold

S = Covering density sampling

R = Random sampled value

V = Random interval as a function of Dip & Dip direction

T = Orientation tolerance angle

Input: D, K, S, T

for n = 1, ..., S do

1. Choose R | R ∈ D

2. Compute  $V \leftarrow \{R - T \leq R \leq R + T\}$

if V > K do

3. Generate density pole  $V_n$

end for

4. Generate  $D_1 \ni V_n, V_{n+1}, \dots, V_{n=S}$

5. Delete repeated values in  $D_1$

Output:  $D_1$

**Step 3: Manual or automatic clusters identification**

At this step it is possible to visualize the results on a stereographic projection, and the user can select cluster points manually or go for an automatic clustering that will give as output the values of the centroids of each identified cluster. As KDE allows calculation of the width of the kernels (e.g. bandwidth) and computation of their density, it is possible to see the plot ranges of values with local maxima; the user in practice can hover the mouse and click on the plot while this is interactively showing dip and dip direction values. The selected values are printed as output when the plot is closed. The same happens if the automatic clustering is chosen, with the difference that no point can be selected, but the user can compare his/her observations on the stereographic projection along with the output values. In any case, every plot is stored in temporary file, which the system is able to reload again for further observations, without the need to re-perform the analysis. The types of clustering implemented are the K-means and the Gaussian Mixture. The first one is implemented with its “+ +” version (Arthur & Vassilvitskii 2007), that uses a smart procedure to initialize the cluster centers before proceeding with the standard k-means optimization: the first cluster center is chosen uniformly at random from the data points that are being clustered, after which each subsequent cluster center is chosen from the remaining data points with probability proportional to its squared distance from the point's closest existing cluster center. The second type of clustering relies on a method that computes a finite-dimensional model, performing a hierarchical evaluation consisting of a set of parameters, each specifying



**Table 2** Comparison of the clustering algorithms implemented in the workflow

Method	Parameters	Scalability	Use cases	Geometry (metric used)
K-Means	Number of clusters	Very large $n_{\text{samples}}$ ; medium $n_{\text{clusters}}$ with MiniBatch code	General-purpose, even cluster size, flat geometry, lots of clusters, inductive	Distances between points
Gaussian mixtures	Branching factor, threshold, number of clusters, optional global clusterer	Not scalable	Flat geometry, good for density estimation, induc- tive	Mahalanobis distances to centers

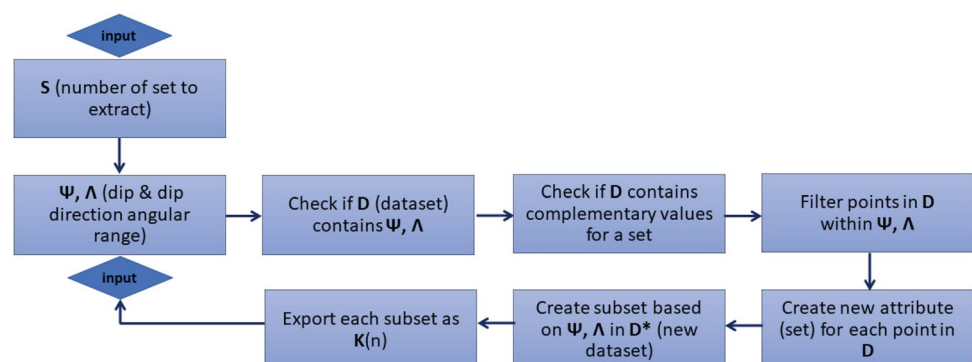
the degree of belonging of a component (point) to the corresponding mixture (Figueiredo and Jain 2002). If the mixture components are Gaussian distributions, as in this type of process, there will be a mean and variance for each component. The two described techniques are generally good clustering approaches for a high number of sample points, and are very similar also in processing and results (Table 2). The main difference relies on the metric technique to evaluate the geometrical relations between clusters: K-means uses distances between points, whilst the Gaussian Mixture takes into account the Mahalanobis distances (1936). The main issue is that the number of clusters to evaluate has to be known previously, and must be specified manually. It can be also evaluated with common indexes estimators such as the Calinski-Harabasz Index (CHI), Silhouette or Davies-Bouldin, all of which could be used independently to evaluate the optimal number of clusters at the expense of processing time (Singh et al. 2021). At the moment, no such estimator has been yet implemented in our method, as these constraints could lead to misinterpretation of data and provide in some cases biased results.

#### Step 4: Extraction of the identified sets with SSE

This step is intended to isolate the clusters of discontinuities (sets) to the point cloud, giving  $D$  as input

and resulting in various minor datasets ( $k_n, \dots, k_{n+1}$ ), stored by the system as new files. Again, the procedure is very straightforward, since  $D$  has orientation values already computed; in practice, it only needs to be specified how many sets one wants to extract and, subsequently, an interval of tolerance for both dip and dip direction values for each set. In addition, it must be specified if a set has complementary values to highlight on the point cloud, and which could be useful for later analysis. The meaning of specifying complementary values is that very often poles of interest with high dip angle (e.g.  $> 80^\circ$ ), could be included in the cloud points; thus, they will be considered as part of separate clusters, but in fact they are not. In these cases, the system generates a single set instead of splitting it into two. In practice, SSE is a tool that filters the original point cloud as the user controls the process step by step, not leaving any choice to the machine. In this way the user can compare the results of the other steps, and decide if they are useful to the last evaluation, or repeat the analysis tweaking the parameters, or, alternatively, extract the sets identified with other methods. In this sense, SSE is a stand-alone tool that can be used independently from the algorithms and methods seen above, and can be defined as a fully-supervised method of extracting clusters (Alg. 2; Fig. 2).

**Fig. 2** Schematic flow chart of the SSE



**Algorithm 2** Supervised Set Extraction in pseudocode

$D$  = Point cloud dataset

$N$  = Number of set to extract

$\Psi$  = Dip angular range of the set  $k_n$

$\Lambda$  = Dip direction angular range of the set  $k_n$

$\Lambda_c$  = complementary values of  $\Lambda$

Input:  $D, N, \Psi, \Lambda, \Lambda_c$

for  $n = 1, \dots, N$  do

  If  $\Psi, \Lambda, \Lambda_c \subset D$  do

1. Create new attribute  $k_n \ni \Psi, \Lambda, \Lambda_c$  in  $D$
2. Filter  $D$  by locating  $\Psi, \Lambda$  generating  $D(k_n)$
3. Export  $D(k_n)$
4. Filter  $D$  by locating  $\Psi, \Lambda_c$  generating  $D(k_{nc})$
5. Export  $D(k_{nc})$

  If  $\Psi, \Lambda \subset D$  and  $\Lambda_c \not\subset D$  do

6. Create new attribute  $k_n \ni \Psi, \Lambda$  in  $D$
7. Filter  $D$  by locating  $\Psi, \Lambda$  generating  $D(k_n)$
8. Export  $D(k_n)$

end for

9. Use  $(k_n, k_{n+1}, \dots, k_{n=N})$  to create new attribute  $\delta$  in  $D \rightarrow$  column specifying discontinuity set

10. Use  $(k_n, k_{n+1}, \dots, k_{n=N})$  to filter  $D$  generating  $D_2 \rightarrow$  dataset with extracted sets

Output:  $D, D_2$

## Case studies

### Description

The first case study is the Cocceio cave, located in the Phlegrean Fields of Campania, in southern Italy (Fig. 3).

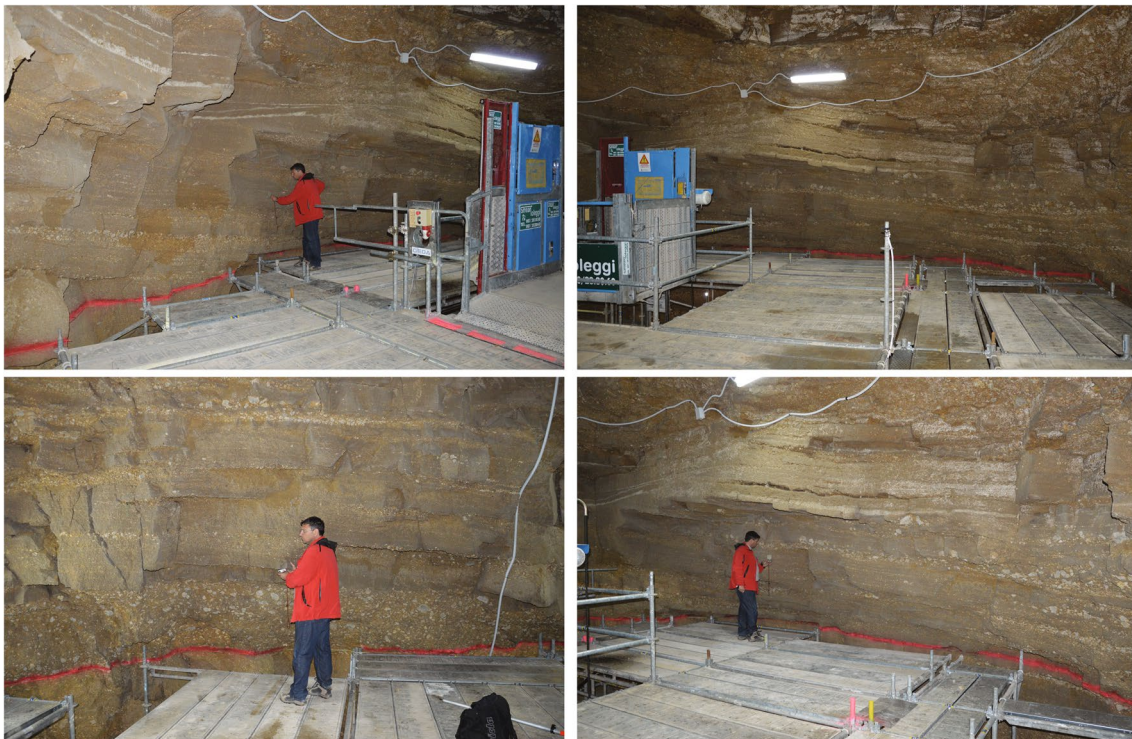
This artificial cave belongs to the category of military works in the classification of artificial cavities proposed by the dedicated UIS (International Union of Speleology) Commission (Parise et al. 2013). Built in Roman time, around 37 B.C. (Beloch 1989) by Lucio Cocceio (from whom it takes its name), it was entirely dug into



**Fig. 3** Geographical location of the two case studies

the tuff, with 6 vertical wells crossing the above hill to reach the ground surface (Pagano et al. 2018), in order to connect Cuma (fortification and lookout point on the

Domitian-Phlegrean coast) to Portus Iulius, an important military infrastructure reaching, through a series of canals and basins, the Gulf of Pozzuoli. During World



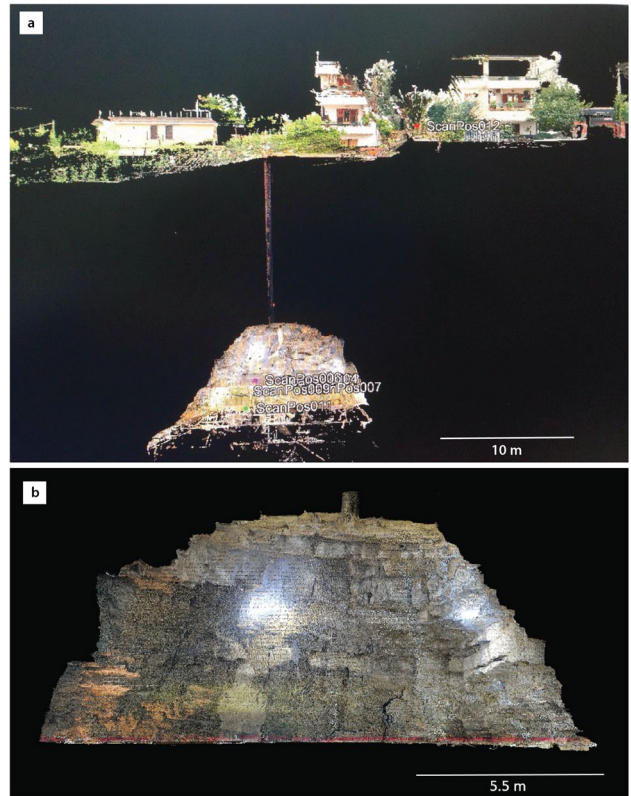
**Fig. 4** Photographs inside the Cocceio cave, showing one of the authors taking measurements of the discontinuities in the traditional way, with compass and inclinometer. The red line drawn with a spray

represents the layer where the 3D point cloud was subsequently cropped to start the analysis



War II the cave was used as an ammunition depot and suffered severe damage following blastings that generated an explosion cap, strongly susceptible to collapses of rock blocks. The cave, large enough to allow the passage of two wagons, extends for 950 m, has trapezoidal section and straight, slightly uphill, course. Since the post-war period, it underwent reclamation works without ever being fully consolidated. The Cocceo Cave insists in the western sector of the Phlegrean Fields (Fig. 3, 4), a large volcanic field whose origin is connected to the tectonic events related to the opening of the Tyrrhenian basin, as the result of two large caldera collapses linked to the eruptions of the Campanian Ignimbrite (39,000 years ago) and of the Neapolitan Yellow Tuff (15,000 years ago) (De Riso et al. 2004; Di Girolamo et al. 1984; Rolandi et al. 2003). Within this latter the cave was entirely excavated, through a sequence of massive to pseudo-stratified ashes, with abundant accretionary and lithic lapilli, and subordinate pumice. Surge deposits are distributed over an area of about 34 km<sup>2</sup>, while fall deposits are found only on a restricted area to the north (Di Vito et al. 2011; Lirer et al. 2011). After being put in place, these materials were affected by zeolitisation processes which led to the formation of lithoid facies characterized by intense fracturing.

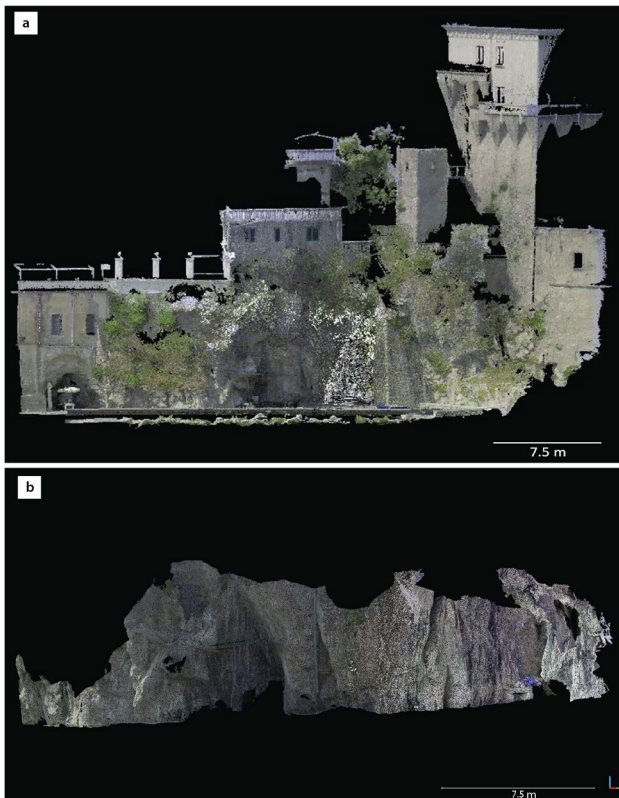
The second case study is part of the slope, along with a natural cave, standing below the medieval watchtower of Cetara, on the Amalfi Coast in Campania (Fig. 3–5), which later became a fort and a prison, and nowadays houses a civic museum. The total scanned area is about 150 m-long and



**Fig. 6** Cocceo cave 3D point cloud. **a** Point cloud resulting from the combined scans. **b** Portion of the cave vault (seen from one of the sides of the cap) used for the analysis

**Fig. 5** Photographs of the scanned portion of the Cetara cliff, with details of the two main areas. To the right, it is also visible the entrance to the cave and part of the above buildings





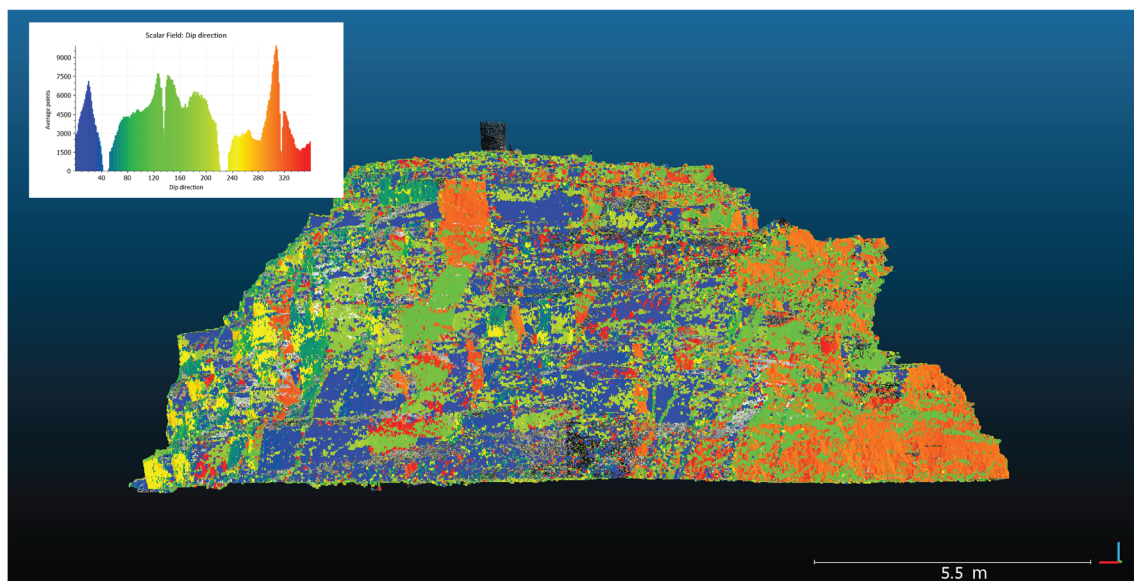
**Fig. 7** 3D point cloud at the Cetara cliff. **a** Point cloud of the cliff, resulting from the combined scans. **b** Portion of the point cloud used for the analysis

60 m-high, with Mesozoic limestones and dolostones as the main lithotypes. The Cetara cliff refers to the lower Jurassic Upper Dolomite formation, more precisely to a laminated bioclastic dolomite (Pappone et al. 2009), representing the

bedrock, locally covered by loose pyroclastic and lithoid deposits of the Campanian volcanic systems.

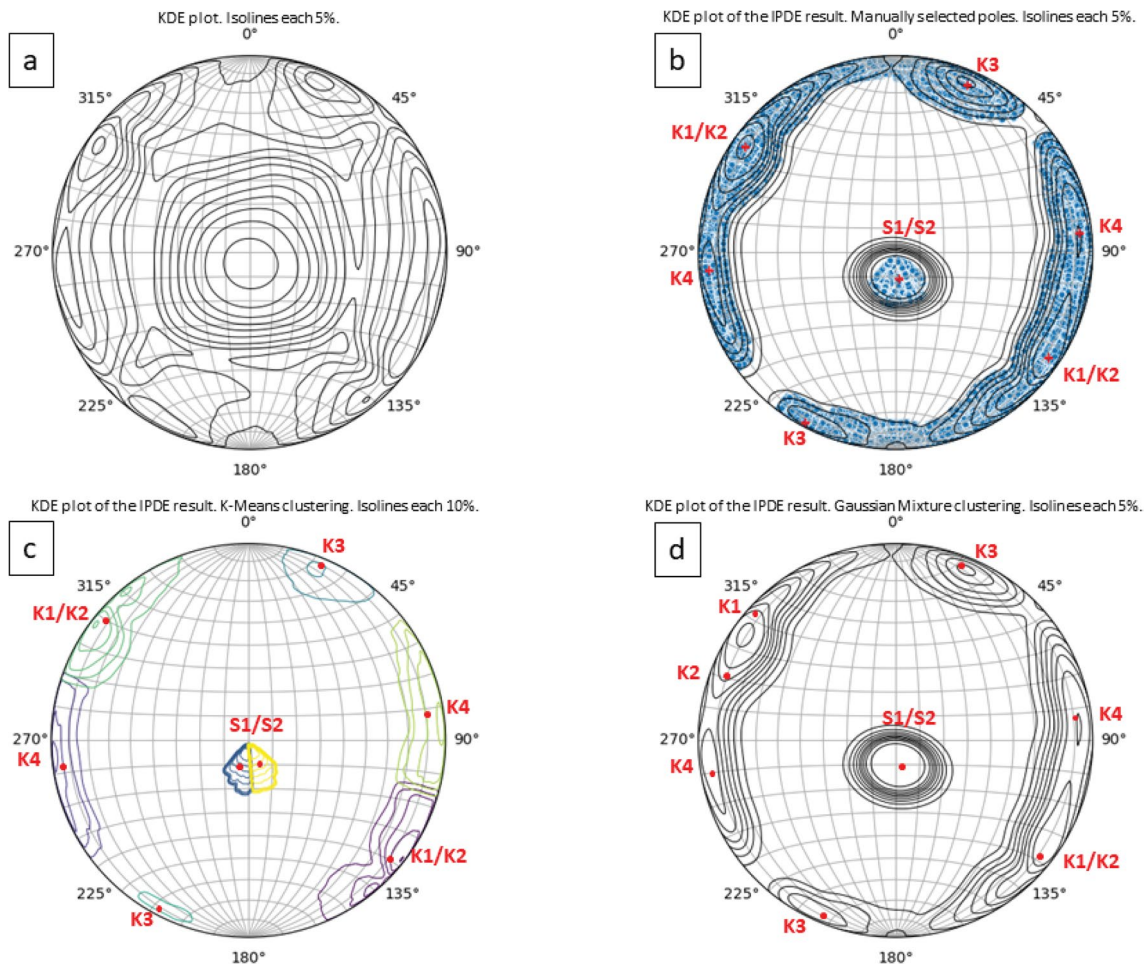
In the SE portion of the Lattari Mts., where Cetara falls, the dolomitic rocks have been intensively modeled by fluvial incisions, which almost completely canceled the morphological relics of the ancient base levels. The geo-morphological structure is characterized by a significantly faulted bedrock, due to various Miocene to Quaternary tectonic phases. Over the carbonate rocks, Quaternary detrital deposits (breccias on the slope, gravels and conglomerates of alluvial deposits) and pyroclastic deposits rest discontinuously. The alluvial deposits (gravels and sands) are found in the main valley bottoms where most of the inhabited centers arise. In particular, the town of Cetara rests on a large fan built through numerous debris-flow and alluvial events, highly frequent in the Campanian geological settings (Calcaterra et al. 1999, 2000, 2003; Vennari et al. 2016).

The survey at the Cocceio cave (Fig. 4) concerned a portion of the cave vault and the ground level over it. It was carried out with laser scanner, which acquired a georeferenced point cloud consisting of 349 million points, with a density of 350,000 points per square meter (Fig. 6). The survey consisted of 10 scan positions (9 inside the cave, 1 outside), integrated with a photogrammetric survey from RPAS. Each single scan indicates a sufficient number (minimum of 4) of georeferenced targets (control points). The registration error between different scans varies from 2 to maximum 4 mm. The control points were georeferenced by survey celerimeter with total station and GPS. The survey was carried out through a polygonal survey line starting from the outside to within the cave.



**Fig. 8** Cocceio Cave: point cloud resulting from the IPDE, highlighted on the original cloud and colored with scalar values indicative of the dip direction





**Fig. 9** Results at the Cocceio Cave: **a** KDE performed on the whole point cloud. **b** Stereographic projection of the IPDE and subsequent KDE – points of the point cloud are spread on the plot in blue color; red crosses are the manually selected centers of the poles. **c** K-Means clustering performed on the result of the IPDE. Note that the isolines

in this type of classification have different colors for each cluster, and are smoothed compared to the classical KDE. **d** Gaussian Mixture clustering performed on the result of the IPDE. The red dots in both (c) and (d) are the automatic centroids of the clusters resulting from the analysis

**Table 3** Cocceio Cave: resulting numerical values of the clustering on the IPDE result

SET	Manual clustering		K-means clustering		G. Mixture clustering	
	Dip	Dip dir	Dip	Dip dir	Dip	Dip dir
K1	/	/	/	/	84.6	313.0
K2	/	/	/	/	80.1	290.1
K1/K2	83.0	304.3	81.7	310.2	/	/
K1/K2*	85.8	125.4	81.1	130.0	81.4	129.3
K3	87.8	208.0	83.7	208.8	83.7	201.7
K3*	79.7	27.4	81.4	24.3	81.3	19.0
K4	83.5	82.9	81.3	82.7	81.0	83.8
K4*	85.9	265.0	83.8	261.6	84.0	260.0
S1	/	/	11.5	205.2	/	/
S2	/	/	11.7	154.6	/	/
S1/S2	11.5	179.0	/	/	11.6	174.8

\* complementary values

**Table 4** Cocceio Cave: discontinuity sets recognized by traditional techniques. Being in presence of a tuff formation, the trend of the layers is the consequence of its emplacement, or pyroclastic flows of the surge or flow type. This originates a kind of folded layers, likely due to the curvilinear trend that characterizes pyroclastic flows

SET	Traditional techniques acquisition	
	Dip	Dip direction
K1	85	309
K2	88	287
K3	90	208
K4	79	82
S1	10	228
S2	7	50

The survey at Cetara (Fig. 5) was performed without the drone but with the same laser scanner, acquiring a point cloud of about 570 million points, with a density of 700,000 points per square meter on 9 scan positions (Fig. 7). The targets, positioned on the ground, were georeferenced with GPS. Georeferencing of the entire group of targets by means of GPS recorded an average error of about 4 mm. The scans were merged with the C2C (cloud to cloud) method, i.e. the best possible overlap between two neighboring scans which results in the smallest error (registration error of the entire group was 1 mm).

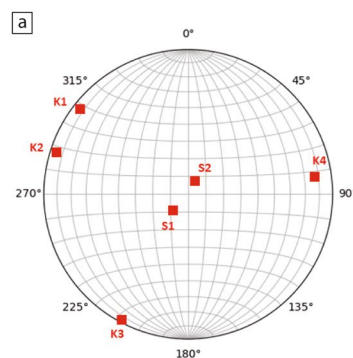
## Results

After manual cleaning and sub-sampling, the point cloud of the Cocceio cave, only referring to the rock outcrop, turned out to be about 2 million points, with average density of 20,000 points/m<sup>2</sup>, and that of Cetara about 760,000 points, with average density of 8,500 points/m<sup>2</sup>. IPDE was used in its basic version to estimate its effectiveness on the first case study point cloud, where the functioning of the basic version has a fixed tolerance range of  $\pm 5^\circ$  for both dip and dip direction. Parameters used for the starting cloud were  $S = 100,000$

(~5% of total points) and  $K = 30,000$  (density for the coplanarity pole of ~1.5% of total points) for the first case study. The result is a cloud consisting in about half the starting points, with average density of 11.300 points/m<sup>2</sup>, where the main systems of discontinuity are highlighted (Fig. 8). With this method the four main sets recognized with the traditional analysis have been identified, even though, due to high density and their vicinity, k1 and k2 sets are merged in the resulting plot. They can be split for extraction by relying on the observation made previously on the rock outcrop. Interestingly, the k4 set, not identified in previous studies (Pagano et al. 2018), came out clearly as a separate set. The two sets referring to bedding planes are merged in the resulting plot, with most of the points lying approximately towards south, at about 180°, revealing a slightly higher density of s1 in a kind of weighted average between the two poles identified on-site. This outcome confirms the previous analysis made on the point cloud by Pagano and coworkers (2018). For the automated clustering evaluation, we gave as input for the number of clusters a total of eight, six of which referring to k1-k4 (i.e. two taking into account k1/k2 merging stage and their complementary value, two for k3 and k4, two for complementary values of these sets), one to the bedding plane, plus another as a bonus to detect either the extra bedding plane that we had from the traditional acquisition or another set. Both K-Means and Gaussian Mixture clustering have been performed on the IPDE resulting point cloud (graphical results in Fig. 9, output centroids of clustering methods and values of points selected manually in Table 3). Eventually, combining the results from traditional techniques (Table 4; Fig. 10) and the new approach, five sets were extracted with the previously given tolerance angle (Table 8).

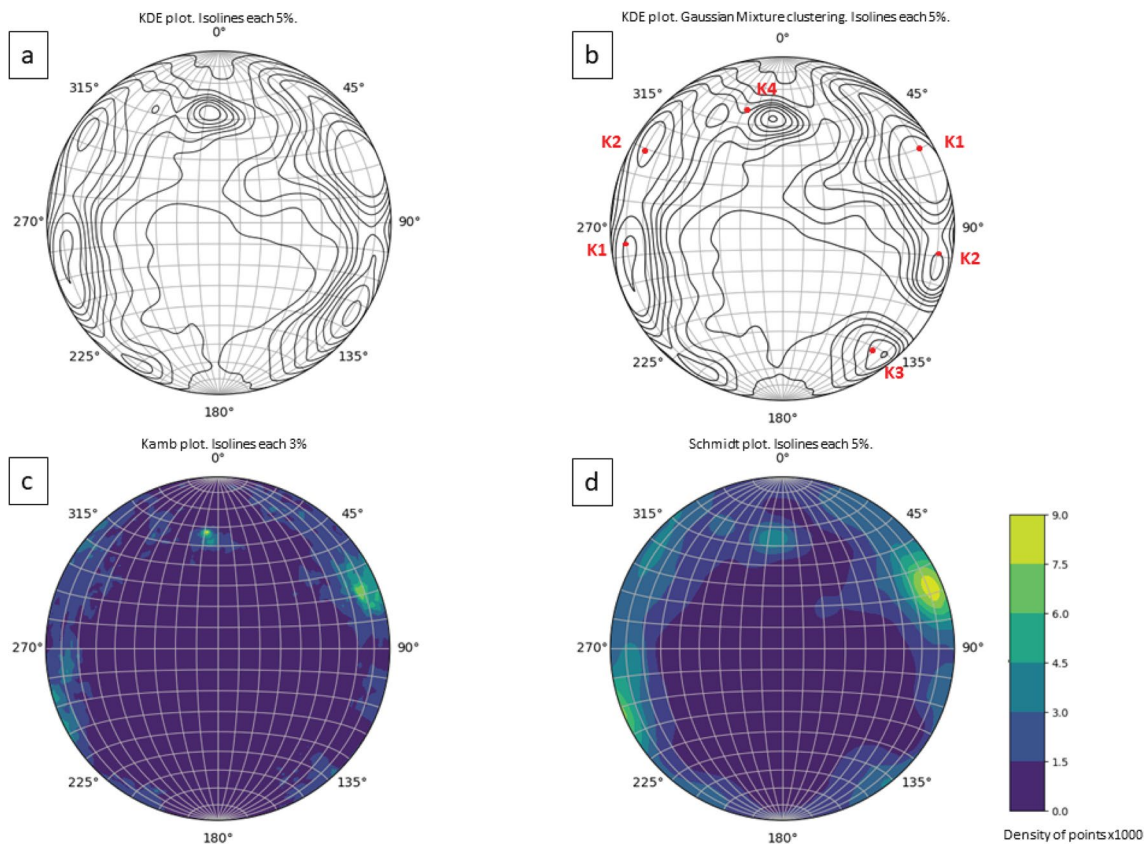
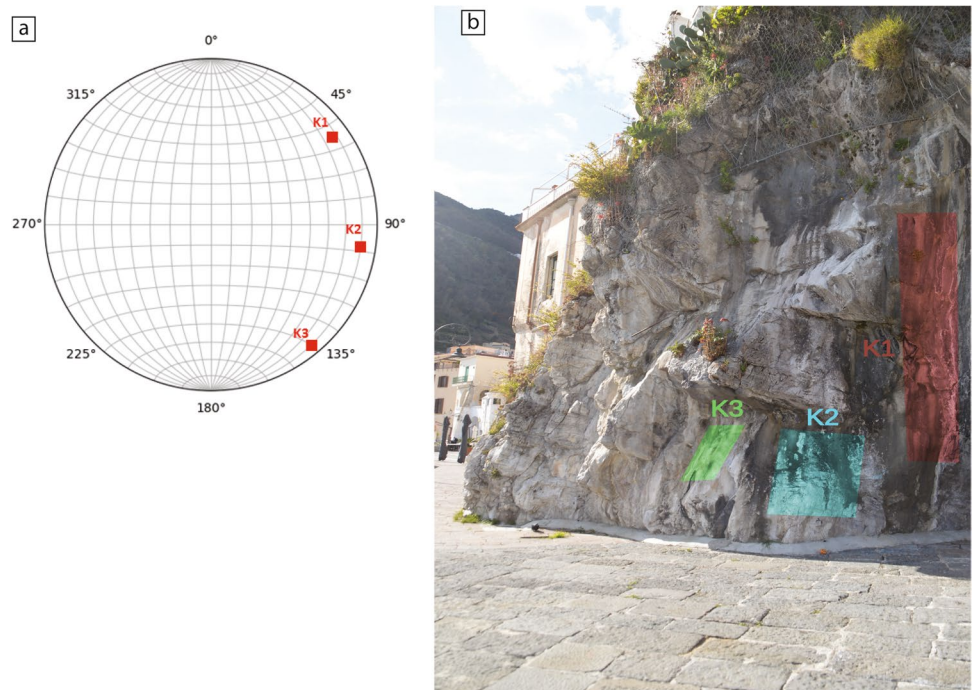
At Cetara, on the graphs resulting from the KDE analysis performed on the cloud, density poles not compatible with the outcomes of traditional methods went to mask the sets recognized before (Fig. 12a). It was thus realized that some of these poles (two in particular, with dip direction ~125°, and ~350°), were planes of anthropogenic origin, not considered in the

**Fig. 10** Cocceio cave. **a** Stereographic projection with the sets identified through traditional methods. **b** Detail of the rock outcrop, highlighting some of the major sets identified during the in situ survey. Note that both K4 and S2 are not shown in the picture, due to their relatively scarce presence





**Fig. 11** Cetara cliff. **a** Stereographic projection with the sets identified through traditional methods. **b** Detail of the rock outcrop, highlighting the major sets identified during the in situ survey



**Fig. 12** Cetara case study: **a** KDE on the whole point cloud. **b** KDE with Gaussian Mixture clustering on the point cloud resulting by the manual intervention; the more biasing influent pole at  $\sim 125^\circ$  of dip direction was eliminated – the red dots are the centroids of the clus-

ters resulting from the automatic analysis. **c** Kamb evaluation on the partially cleaned point cloud. **d** Schmidt evaluation on the partially cleaned point cloud. For these two types of evaluations the density of points per pole is highlighted with a clear color ramp

on-site analysis. By manually extracting such planes, filtering the cloud by eliminating them, and performing again the digital density analysis (i.e., KDE and Kamb/Schmidt) and the Gaussian Mixture clustering, the planes recognized in the traditional surveys were actually identified (Fig. 11, 12b, Table 5, 6). Subsequently, IPDE was also performed on the total cloud. This made possible to lighten it, eliminating minor discontinuities and non-geological planes between discontinuities (Fig. 13). The parameters utilized in this case were  $S = 70,000$  ( $\sim 10\%$  of total points) and  $K = 20,000$  (density for the coplanarity pole of  $\sim 2.5\%$  of total points), and the cloud passed from  $\sim 760,000$  to  $\sim 580,000$  points, with average density of  $7,000$  points/m<sup>2</sup>. Manual filtering was required to eliminate the anthropogenic planes which, having abundant densities, were still identified by the algorithm. The final result ( $\sim 520,000$  points) was relatively in agreement with that of the traditional survey: the main three sets were identified, with the addition of two others, which were not considered significant during the in-situ surveys (Fig. 14; Table 7). The latter refer to complementary values of 2 out of the 3 sets recognized by the digital analysis. The midpoints for the construction of the poles to extract have an orientation of  $+5^\circ$  for dip direction of k1 and  $+4^\circ$  for k3, when compared to those identified in the traditional analysis. Taking into account the nature of the rock mass, composed mainly by tuff (see previous chapter), that in this area is typically formed by blocks with surfaces showing evident undulations, sets with an angular tolerance of  $\pm 10^\circ$  for dip direction and  $\pm 5^\circ$  of dip have been subsequently extracted, being the points of the poles very scattered in the stereographic projection (Table 9).

## Discussion

At the Cocceio Cave the fully automated method gave reliable results, compared with the analysis performed through traditional techniques; these latter served as ground truth to have an analytical basis to integrate and

compare the digital analysis. In this case it also proved unnecessary to proceed with any manipulation for manually eliminating non-geological features before starting the whole evaluation. Both clustering methods are relatively precise and correlate well with the in-situ observations on the point cloud when executed on the IPDE result. This algorithm, in fact, provides a standalone solution to mitigate common errors coming from performing density and clustering analysis directly on the whole point cloud resulting from the pre-processing stage. By eliminating points lying below a certain case-sensitive threshold given by the user (the parameter  $K$  of the algorithm), that in this case was of  $30,000$  points, corresponding to  $\sim 1.5\%$  of the total of the pre-processed cloud, it is possible to automatically take into account for the analysis only features having geological meaning, i.e., planar or almost planar rocky surfaces, leaving aside many useless points in the cloud, not to be considered for the analysis. This process, in addition to speeding up the computation, tends to minimize the errors due to the evaluation of curved (e.g., passages from one discontinuity to another) and small surfaces (e.g. tiny fractures), thus leading the results towards the expectations of the user's observation conducted by traditional method, and improving it. This important feature is rarely taken into account by the scientific literature on detection of discontinuity sets (Dewez et al. 2016; Jaboyedoff et al. 2007; Hammah and Curran 1998; Menegoni et al. 2022; Pagano et al. 2020; Riquelme et al. 2014; Singh et al. 2021; Tomás et al. 2020); a relatively new method explores the concept of getting rid of the rock surfaces with too pronounced curvature (Tsui et al. 2021) and, in this sense, a possible future perspective for this type of work could be the application of machine/deep learning algorithms to the classification of 3D point clouds, to improve the accuracy of object segmentation and identification process. Several approaches of this type can be found nowadays in the literature, even though most of them are not directly connected to geology; among these, it is worth mentioning the works by Li et al. (2019), representing

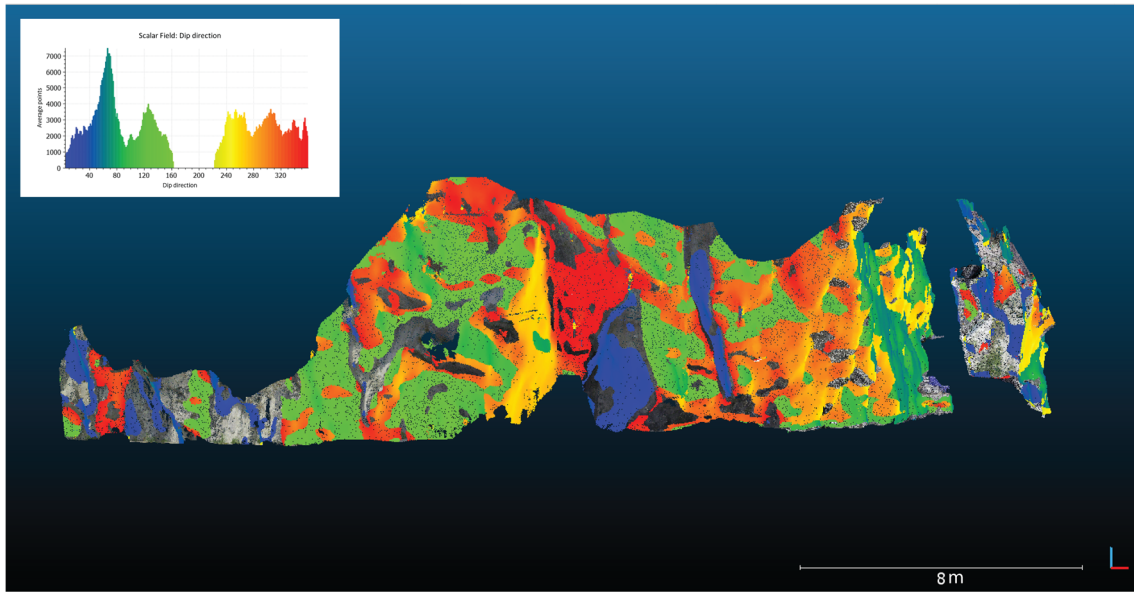
**Table 5** Cetara case study: discontinuity sets recognized by traditional techniques

SET	Traditional techniques acquisition	
	Dip	Dip direction
K1	78	55
K2	80	99
K3	82	140

**Table 6** Cetara case study: resulting numerical values of the Gaussian Mixture clustering on the cleaned dataset

SET	Gaussian Mixture clustering	
	Dip	Dip direction
K1	79.4	61.5
K1*	81.9	263.6
K2	81.3	101.2
K2*	79.1	297.0
K3	73.7	144.5
K4**	59.4	342.8

\*complementary values. \*\*additional set

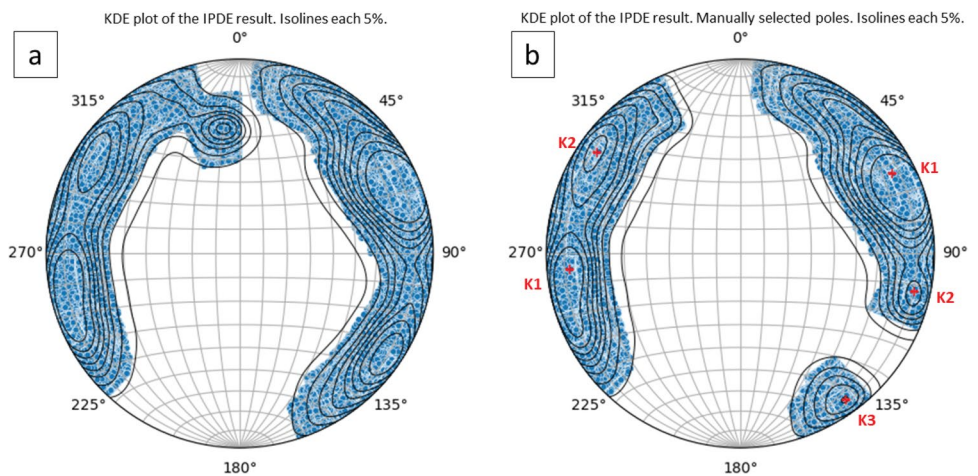


**Fig. 13** Cetara case study: point cloud resulting from the IPDE, highlighted on the original cloud and colored with scalar values indicative of the dip direction

one of the first use of machine learning methods to recognize primitive shapes within 3D point clouds, and that by Mammoliti et al. (2022), who have been the first to explore the possibility of using such systems for the recognition of geological features on the results of digital surveys. The K-means clustering gave reliable outputs, since the isolines are eventually smoothed, and thus the results are even more defined. This type of clustering is considered in fact reliable and useful when the clusters to detect on a rock mass are consistent, for example more than four/five sets, in rather complex cases. It relies on point distances and the centroid of each cluster would be a sort of weighted mean of the cluster itself, resulting in

good outcomes when evaluating many sets, i.e. the more the clusters, the more precise the analysis on a rock mass would be. In this case the surface layer pole was split into two, giving as output two very close centroids, one of which can be precisely correlated with the result of the digital evaluation previously performed (Pagano et al. 2018). The Gaussian Mixture, on the other hand, splits k1 and k2 on their main orientations, while the complementary values of the two poles are unified in a single one; it also leaves the bedding plane in a unified pole cluster, the centroid of which is pointing approximately towards south (Fig. 9d). This method is more suitable for cases with low number of clusters, since it strongly relies on

**Fig. 14** Cetara case study results of the IPDE. Points of the point cloud are spread on the plot in blue color (a) KDE performed on the point cloud without manual intervention. (b) KDE performed on the point cloud where two poles referring to planes of anthropogenic origin (dip direction = 125°, 350°) were eliminated. Manual points selection





**Table 7** Resulting numerical values of the manual clustering on the IPDE result

SET	Manual clustering	
	Dip	Dip direction
K1	81.9	62.5
K1*	78.5	267.4
K2	80.4	102.0
K2*	79.0	307.1
K3	79.8	141.5

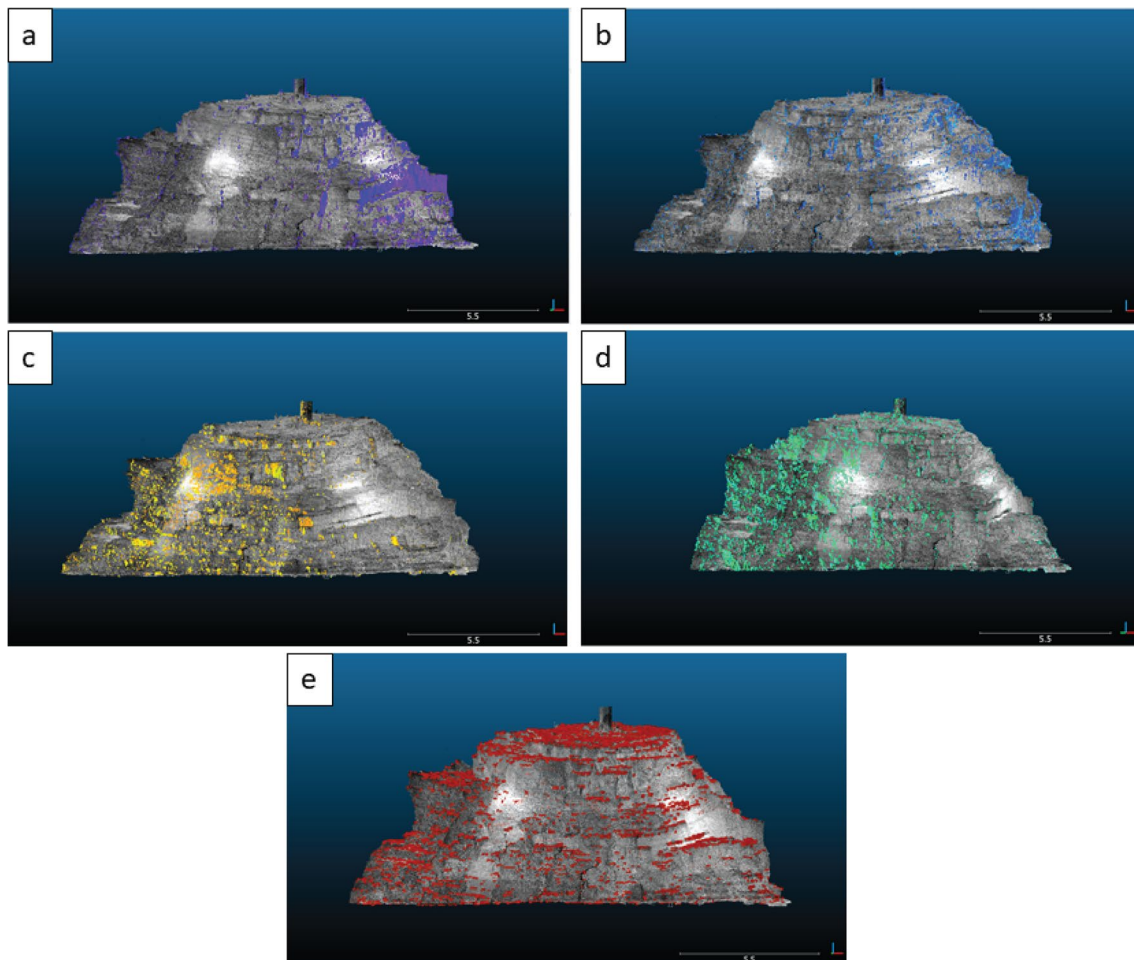
\*complementary values

the point density, then giving outputs that closely match the traditional observations or the manual clustering on digital results when the sets to detect are few (from two to a maximum of five/six, as we were able to verify in the second case study). Finally, the manual selection of points of interest is also a precise and effective method to deal

**Table 8** Numerical values of the extracted sets at the Cocceio Cave

SET	Supervised Set Extraction	
	Dip	Dip direction
K1	$85 \pm 5$	$309 \pm 5$
K2	$85 \pm 5$	$287 \pm 5$
K3	$85 \pm 5$	$208 \pm 5$
K4	$80 \pm 5$	$82 \pm 5$
S1	$10 \pm 5$	$210 \pm 5$

with the stereographic projections. This method is simple and intuitive and instead of giving automatic outputs to be validated later, by adding or deleting what the program recognizes (Dewez et al. 2016; Jaboyedoff et al. 2007; Riquelme et al. 2014), it gives to the user the possibility to choose the most significant poles directly on the stereoplot. Finally, the extracted sets through the SSE, are a sort of mean among the analyses performed (Fig. 15; Table 8).

**Fig. 15** Extracted sets highlighted on the original point cloud at the Cocceio Cave: **a** k1, **b** k2, **c** k3, **d** k4, and **e** s1. The software scale is in meters



**Table 9** Numerical values of the extracted sets at Cetara

SET	Supervised Set Extraction	
	Dip	Dip direction
K1	$77 \pm 5$	$60 \pm 10$
K2	$80 \pm 5$	$100 \pm 10$
K3	$82 \pm 5$	$144 \pm 10$

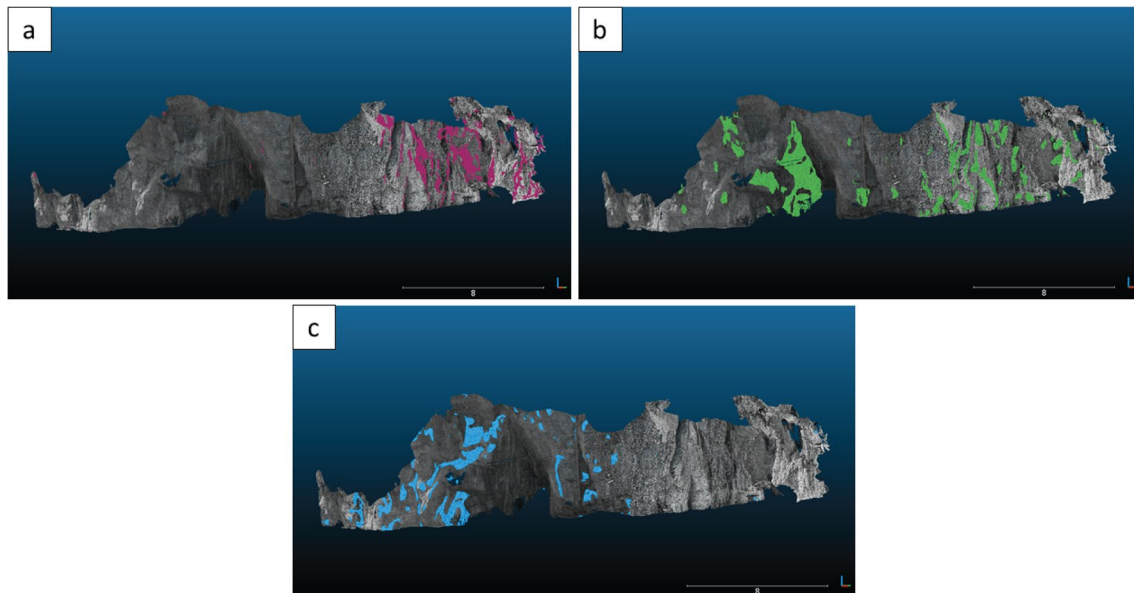
At Cetara, instead, a fully automatic procedure would have led to detect also planes not relevant for geological purposes, and the density of these would have masked completely some sets recognized in situ. Therefore, starting an automated clustering, or even performing IPDE, we would have had biased outcomes. This is because these planes, namely two walls, had an high mole of points density to be eliminated by the IPDE without resulting in a significant loss of geological poles. Nonetheless for the Gaussian Mixture it was chosen not to eliminate one of these, i.e. the pole at dip direction  $350^\circ$ , as in this case we realized after a density analysis that, unlike that at  $\sim 125^\circ$ , it was not going to hide any main set. Both were anyway eliminated before starting IPDE to further lighten the computation and have more defined results. In this case, too, the sets were extracted by wisely mixing the results from the various methods, including the traditional one (Fig. 16; Table 9). Comparing the KDE/Kamb-Schmidt evaluation timing, to perform the KDE and plot the resulting data, the computer takes about 10 min for the Cetara case, compared to about 3 for Kamb's analysis and even about 1 and a half minute for Schmidt's. These results point to the implementation of the last two

as a standard. However, the fact that the accuracy of traditional KDE is greater, makes it preferable not to completely abandon this method, and induces, on the contrary, to use the approaches in a combined manner, aimed at comparing the results in terms of accuracy. In case quick preliminary analysis is needed before undertaking more in-depth investigations, the Kamb/Schmidt system could be used from the beginning.

## Conclusions

Stability assessment of rock masses plays a crucial role in the mitigation of the related risk. Any rock mass characterization requires the acquisition of information on discontinuities at the outcrops, which is an essential step for the fully comprehension of their kinematic behavior.

With this work we presented an alternative perspective on the analysis of data acquired by means of 3D scanning techniques. In detail, we highlighted and stressed the importance of the role by the expert, without entirely relying upon the use of the machine; even though a great improvement in workflow automation is obtained using the proposed methodology, a solid background in structural geology and rock mechanics is always needed to make the right choice throughout the different steps. This must be coupled and integrated with field reports from in-situ surveys, aimed at gaining the best visual recognition of the results. Further, we pointed out the importance of considering ranges of values to describe clusters of poles and to extract them, instead of fixed orientations values, as it is nearly impossible to



**Fig. 16** Extracted sets highlighted on the original point cloud at Cetara: **a** k1, **b** k2, and **c** k3. The software scale is in meters

**Table 10** List of the main python packages utilized

PYTHON PACKAGE	UTILIZATION
NUMPY (HARRIS ET AL. 2020)	Matrix reading and calculation
PANDAS (MCKINNEY 2010)	Reading and organizing data
OPEN3D (ZHOU ET AL. 2018)	Geometrical calculations
MATPLOTLIB (HUNTER 2007); SEABORN (WASKOM 2021); MPLSTEREONET (KINGTON 2019)	Plotting
SCIKIT-LEARN (PEDREGOSA ET AL. 2011)	Statistical calculations and clustering

quantify the orientation of a discontinuity set by using a single number in most rock types.

In its first part, the method presented is useful for highlighting the main systems of discontinuity and does not take into account numerous minor planes of transition among sets or smaller fractures, thus lightening the point cloud to avoid loss of information on the density of the major poles. This type of approach facilitates both dimensional and statistical analysis, and speeds up the generation of stereographic projections by increasing the accuracy in the recognition of density poles through the use of well-established methods and algorithms for their visualization and detection. Eventually, the packages utilized to develop the code implemented in this workflow are listed in Table 10.

The second part of this method focuses on the utility of a fully supervised function for the extraction of discontinuity sets. After comparing the various outputs that different methods can provide, and having established the relationships among digital and traditional methods, it is strongly advisable to decide which intervals referring to individual sets to extract, without leaving the machine to perform automatically the isolation of the main families present on the rock mass. This could enhance the overall workflow, leaving to the user the possibility to filter and extract as many sets as possible, to repeat the procedure also in order to widen or narrow the given interval depending on the case, or to isolate and eliminate those portions of the 3D point cloud potentially inducing to errors or to unexpected peaks when performing a simple density analysis.

**Acknowledgements** The authors would like to thank all members of IdroGeo S.R.L., in particular Angela Caccia and Anna Ruocco for valuable discussions and for producing relevant data available to the project. Further, the authors thank Domenico Diacono (PhD) of IRIS (Institutional Research Information System, Università degli Studi di Bari Aldo Moro), for his valuable advices on how to set up different parts of the code used in this work.

**Author contributions** Stefano Cardia: conceptualization; methodology; software; formal analysis; data curation; writing—original draft; visualization. Biagio Palma: conceptualization; validation; resources; supervision. Francesco Langella: validation; investigation; visualization. Marco Pagano: validation; investigation; visualization. Mario Parise: writing—review & editing; supervision; project administration. All authors reviewed the manuscript.

**Data Availability** The data that support the findings of this study may be available upon request. Restrictions apply to the availability of these data, which were used under license for this study. Data may be available with the permission of Idrogeo S.R.L. (Vico Equense – Naples – ITA).

**Code availability** Name of the code/library: GeoDS (Geological Data Science).

Contact: stefano.cardia@uniba.it.

Program language: Python 3.x

Code size: 30 KB.

The source codes are available for downloading at the link: <https://github.com/StewTheBrew/geods>

## Declarations

**Competing interests** The author declare no competing interests in the work done to writing this manuscript.

## References

- Abellán A, Oppikofer T, Jaboyedoff M, Rosser NJ, Lim M, Lato MJ (2014) Terrestrial laser scanning of rock slope instabilities. *Earth Surf Proc Land* 39:80–97
- Abellan A, Derron MH, Jaboyedoff M (2016) Use of 3D Point Clouds in Geohazards. *Remote Sens* 8:130. <https://doi.org/10.3390/rs8020130>
- Andriani GF, Parise M (2015) On the applicability of geomechanical models for carbonate rock masses interested by karst processes. *Environ Earth Sci* 74:7813–7821. <https://doi.org/10.1007/s12665-015-4596-z>
- Andriani GF, Parise M (2017) Applying rock mass classifications to carbonate rocks for engineering purposes with a new approach using the rock engineering system. *J Rock Mech Geotech Eng* 9:364–369
- Arthur D, Vassilvitskii S (2007) K-means++: the advantages of careful seeding. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. 1027–1035
- Barnobi L, La Rosa F, Leotta A, Paratore M (2009) Analisi geomeccanica e di caduta massi tramite rilievo geostrutturale con geologi rocciatori e Laser Scanner 3D. *GdiS* 4(2009):33–40
- Beloch J (1989) Campania – Storia e topografia della Napoli antica e dei suoi dintorni. Reprint by Ferone C. and Pugliese Carratelli F., Bibliopolis, Napoli, pp 461
- Bieniawski ZT (1973) Engineering classification of jointed rock masses. *T S Afr Inst Civ Eng* 15:335–344
- Bieniawski ZT (1989) *Engineering Rock Mass Classifications: A Complete Manual for Engineers and Geologists in Mining, Civil, and Petroleum Engineering*. John Wiley & Sons (New York City), Technology & Engineering, p 272

- Borrmann D, Elseberg J, Lingemann K, Nüchter A (2011) The 3D hough transform for plane detection in point clouds: A review and a new accumulator design. *3DR express* 2:2–14
- Calcaterra D, Parise M, Palma B (2003) Combining historical and geological data for the assessment of the landslide hazard: a case study from Campania, Italy. *Nat Hazard* 3(1/2):3–16
- Calcaterra D, Parise M (Eds.) (2010) Weathering as a predisposing factor to slope movements. Geological Society of London, Engineering Geology Special Publication no. 23, 233
- Calcaterra D, Parise M, Palma B, Pelella L (1999) The May 5<sup>th</sup>, 1998, landsliding event in Campania (Southern Italy): inventory of slope movements in the Quindici area. *Proc. International Symposium on Slope Stability Engineering*, Shikoku, Japan, November 8–11, 1999, 2, 1361–1366
- Calcaterra D, Parise M, Palma B, Pelella L (2000) Multiple debris flows in volcanoclastic materials mantling carbonate slopes. *Proceedings 2<sup>nd</sup> International Conference on “Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment”*, edited by Wieczorek G.F. and Naeser N.D., Taiwan, August 16–18, 2000, 99–107
- Cardia S, Langella F, Pagano M, Palma B, Parise M (2022) 3D structural analysis of the cave of Saint Michael at Minervino Murge, Bari (Italy) – a typical case of karst environment in Puglia, EGU General Assembly, Vienna (Austria), 23–27 May 2022, EGU22-2329. <https://doi.org/10.5194/egusphere-egu22-2329>
- De Riso R, Budetta P, Calcaterra D, Santo A, Del Prete S, De Luca C, Di Crescenzo G, Guarino PM, Mele R, Palma B, Sgambati D (2004) Fenomeni di instabilità dei Monti Lattari e dell’area flegrea (Campania): scenari di suscettibilità da frana in aree campione. *Quaderni Di Geologia Applicata* 11:1–26
- Dewez TJB, Girardeau-Montaut D, Allanic C, Rohmer J (2016) Facets: a CloudCompare plugin to extract geological planes from unstructured 3d point clouds. *Int Arch Photogramm Remote Sens Spat Inf Sci*. 41. <https://doi.org/10.5194/isprs-archi-2016-11-01-01>
- Di Girolamo P, Ghiara MR, Lirer L, Munno R, Rolandi G, Stanzione A (1984) Vulcanologia e petrologia dei Campi Flegrei. *Boll Soc Geol Ital* 103:349–413
- Di Vito MA, Arienzo I, Braia G, Civetta L, D’antonio M, Di Renzo V, Orsi G (2011) The Averno 2 fissure eruption: a recent small-size explosive event at the Campi Flegrei caldera (Italy). *Bull Volcanol* 73:295–320
- Ferrero AM, Forlani G, Roncella R, Voyat HI (2009) Advanced Geospatial Survey Methods Applied to Rock Mass Characterization. *Rock Mech Rock Eng* 42:631–665. <https://doi.org/10.1007/s00603008-0010-4>
- Figueiredo MAT, Jain AK (2002) Unsupervised Learning of Finite Mixture Models. *IEEE Trans Pattern Anal Mach Intell* 24:381–396. <https://doi.org/10.1109/34.990138>
- Galgaro A, Genevois R, Teza G, Conforti D, Rocca M (2004) Monitoraggio da terra di fenomeni di instabilità geologica mediante sistemi di telerilevamento laser scanner: un primo test sulle Cinque Torri (Cortina d’Ampezzo, BL). *Telerilevamento* 2004:11–15
- Girardeau-Montaut D (2016) <https://www.cloudcompare.net/forum/>
- Hammah RE, Curran JH (1998) Fuzzy Cluster Algorithm for the Automatic Identification of Joint Sets. *Int J Rock Mech Mining Sci* 35(8):889–905 (PII: S0148 9062(98)00011-4)
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Oliphant TE (2020) Array programming with NumPy. *Nature* 585:357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9(3):90–95
- ISRM (International Society for Rock Mechanics) (1978) Suggested Methods for the Quantitative Description of Discontinuities in Rock Masses. *Int J Rock Mech Mining Sci Geomech Abstr* 15:319–368
- Jaboyedoff M, Oppikofer T, Abellán A, Derron MH, Loye A, Metzger R, Pedrazzini A (2012) Use of LiDAR in landslide investigations: a review. *Nat Hazards* 61:5–28
- Jaboyedoff M, Metzger R, Oppikofer T, Couture R, Derron M-H, Locat J, Turmel D (2007) New insight techniques to analyze rock-slope relief using DEM and 3D-imaging cloud points: COLTOP-3D software. In Eberhardt, E., Stead, D and Morrison T. (Eds.): *Rock mechanics: Meeting Society’s Challenges and Demands* 1, Taylor & Francis. 61–68. <https://doi.org/10.1201/NOE0415444019-c8>.
- Kemeny J, Turner K, Norton B (2006) LIDAR for Rock Mass Characterization: Hardware, Software, Accuracy and Best-practices, Laser and Photogrammetric Methods for Rock Face Characterization. ARMA Golden, Colorado
- Kington J (2019) <https://pypi.org/project/mplstereonet/>. <https://github.com/joferkington/mplstereonet>
- Li L, Sung M, Dubrovina A, Yi L, Guibas L (2019) Supervised Fitting of Geometric Primitives to 3D Point Clouds. *IEEE Explore*. Computer Vision Foundation. *CVPR* 2019: 30-39
- Lirer L, Petrosino P, Alberico I, Armiero V (2011) Carta Geologica di Cuma, Averno e Monte Nuovo (Scala 1:10.000)
- Liotine L, Andriani GF, Jaboyedoff M, Parise M, Derron M-H (2021a) Comparison of remote sensing techniques for geostructural analysis and cliff monitoring in coastal areas of high tourist attraction: the case study of Polignano a Mare (Southern Italy). *Remote Sens* 13:5045
- Liotine L, Wolff C, Wyser C, Andriani GF, Derron M-H, Jaboyedoff M, Parise M (2021b) QDC-2D: a semi-automatic tool for 2D analysis of discontinuities for rock mass characterization. *Remote Sens* 13:5086
- Lombardi G, Rozza A, Casiraghi E, Campadelli P (2011) A Novel Approach for Geometric Clustering based on Tensor Voting Framework. *Conference: WIRN* 2011, pp. 210–217
- Mahalanobis PC (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India* 2:49–55
- Mammoliti E, Di Stefano F, Fronzi D, Mancini A, Malinverni ES, Tazioli A (2022) A Machine Learning Approach to Extract Rock Mass Discontinuity Orientation and Spacing, from Laser Scanner Point Clouds. *Remote Sens* 14:2365. <https://doi.org/10.3390/rs14102365>
- McKinney W (2010) Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* Vol. 445, pp. 51–56
- Menegoni N, Giordan D, Perotti C (2021) An Open-Source Algorithm for 3D Rock Slope Kinematic Analysis (ROKA). *Appl Sci* 11:1698. <https://doi.org/10.3390/app11041698>
- Menegoni N, Giordan D, Inama R, Perotti C (2022) DICE: An open-source MATLAB application for quantification and parametrization of digital outcrop model-based fracture datasets, *J Rock Mech Geotech Eng* 320–344. <https://doi.org/10.1016/j.jrmge.2022.09.011>
- Oppikofer T, Jaboyedoff M, Blikra L, Derron MH, Metzger R (2009) Characterization and monitoring of the Åknes rockslide using terrestrial laser scanning. *Nat Hazard* 9:1003–1019
- Pagano M, Palma B, Parise M, Ruocco A (2018) Analisi geostrutturale su nuvola di punti acquisita con laserscanner 3d: applicazione alla grotta di Cocceio, Bacoli (Campania, Italia). *Geologia Dell’ambiente Suppl* 4:87–94
- Pagano M, Palma B, Ruocco A, Parise M (2020) Discontinuity Characterization of Rock Masses through Terrestrial Laser Scanner and Unmanned Aerial Vehicle Techniques Aimed at Slope Stability Assessment. *Appl Sci* 10:2960. <https://doi.org/10.3390/app10082960>
- Palmer A (2007) *Cave geology*. Cave Books, Smerzer ed. (Dayton OH), pp 454
- Pappone G, Casciello E, Cesarano M, D’Argenio B, Conforti A (2009) Note illustrative della Carta Geologica d’Italia alla scala 1:50.000 Foglio 467 Salerno. *SystemCart*, Roma

- Parise M, Galeazzi C, Bixio R, Dixon M (2013) Classification of artificial cavities: a first contribution by the UIS Commission. Proc. 16<sup>th</sup> Int. Congress of Speleology, Brno, 21–28 July 2013, vol. 2: 230–235
- Parise M (2022) Sinkholes, Subsidence and Related Mass Movements. In: Shroder J.J.F. (Ed.), *Treatise on Geomorphology*, vol. 5. Elsevier, Academic Press, pp. 200–220. <https://doi.org/10.1016/B978-0-12-818234-5.00029-8>. ISBN: 9780128182345
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: Machine Learning in Python. *The Journal of Machine Learning Research* 12(85):2825–2830. arXiv:1201.0490. <https://doi.org/10.48550/arXiv.1201.0490>
- Piteau D (1970) Geological factors significant to the stability of slopes cut in rock. *J South Afr Inst Min Metall* 33–53
- Riquelme A, Abellan A, Tomas R, Jaboyedoff M (2014) A new approach for semi-automatic rock mass joints recognition from 3D point clouds. *Comput Geosci* 68:38–52
- Riquelme A, Cano M, Tomas R, Abellan A (2017) Identification of Rock Slope Discontinuity Sets from Laser Scanner and Photogrammetric Point Clouds: a Comparative Analysis. In: *Symposium of the International Society for Rock Mechanics*. Procedia Engineering 191:838–845
- Roland G, Bellucci F, Heizler MT, Belkin HE, De Vivo B (2003) Tectonic controls on the genesis of ignimbrites from the Campanian Volcanic Zone, southern Italy. *Mineral Petrol* 79:3–31
- Roncella R, Forlani G (2005) Extraction of planar patches from point clouds to retrieve dip and dip direction of discontinuities. ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands, September 12–14, 2005
- Schnabel R, Wahl R, Klein R (2007) Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum* 26(2):214–226
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. Boca Raton: CRC press. 176. <https://doi.org/10.1201/9781315140919>
- Singh SK, Banerjee BP, Lato M, Sammut C, Raval S (2021) Automated rock mass discontinuity set characterisation using amplitude and phase decomposition. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.3899864>
- Slob S, van Knapen B, Hack R, Turner K, Kemeny J (2005) Method for automated discontinuity analysis of rock slopes with three-dimensional laser scanning. *Transp Res Rec* 1913:187–194
- Slob S (2010) *Automated Rock Mass Characterisation Using 3-D Terrestrial Laser Scanning* (Ph.D. thesis), TU Delft, Delft University of Technology (URL: <http://www.narcis.nl/publication/RecordID/oai:tudelft.nl:uuid:c1481b1d-9b33-42e4-885a-53a6677843f6>)
- Tomás R, Riquelme A, Cano M, Pastor JL, Pagán JI, Asensio JL, Ruffo M (2020) Evaluation of the stability of rocky slopes using 3D point clouds obtained from an unmanned aerial vehicle. *Revista De Teledetección* 55:1–15. <https://doi.org/10.4995/raet.2020.13168>
- Tran TT, Cao VT, Laurendeau D (2015) Extraction of Reliable Primitives from Unorganized Point Clouds. *3DR Express*. *3D Res* 6:44. <https://doi.org/10.1007/s13319-015-0076-1>
- Tsui R, Cheung CS, Hart J, Hou W, Ng A (2021) Intersection-Based Potential Plane Failure Detection on 3D Meshes for Rock Slopes. Conference: HKIE Geotechnical Division 41st Annual Seminar: Hong Kong. <https://doi.org/10.21467/proceedings.126.27>
- Vennari C, Santangelo N, Santo A, Parise M (2016) A database on flash flood events in Campania, southern Italy, with an evaluation of their spatial and temporal distribution. *Nat Hazard* 16:2485–2500
- Viero A, Teza G, Massironi M, Jaboyedoff M, Galgaro A (2010) Laser scanning-based recognition of rotational movements on a deep-seated gravitational instability: the Cinque Torri case (north-eastern Italian Alps). *Geomorphology* 122:191–204
- Vollmer FW (1995) C Program for Automatic Contouring of Spherical Orientation Data Using a Modified Kamb Method. *Comput Geosci* 21(1):31–49
- Waskom ML (2021) Seaborn: statistical data visualization. *J Open Source Softw* 6(60):3021. <https://doi.org/10.21105/joss.03021>
- Xia S, Chen D, Wang R, Li J, Zhang X (2020) Geometric Primitives in LiDAR Point Clouds: A Review. *IEEE J-STARS* 99:1–1. <https://doi.org/10.1109/JSTARS.2020.2969119>
- Zhou QY, Park J, Koltun V (2018) Open3D: A Modern Library for 3D Data Processing. *Deep AI*. ArXiv: 1801.09847v1

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)