geosciences

MDPI

Article

AI-Assisted 3D Rock Fracture Characterization from Multi-View Smartphone Image–LiDAR Data in Underground Environments

Alessandra Spadaro * and Marco Piras

Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy; marco.piras@polito.it

* Correspondence: alessandra.spadaro@polito.it

Abstract

Three-dimensional characterization of rock fractures in underground environments is challenging because restricted access, uneven illumination and partial visibility hinder mapping and structural measurements. This study presents an AI-assisted workflow that combines text-prompted Segment Anything Model 3 segmentation with smartphone image–LiDAR reconstruction to derive three-dimensional fracture supports. Candidate masks are combined into graded image evidence, transferred to the reconstructed surface and consolidated across views. Assessment combines manual image annotations, terrestrial laser scanning (TLS) surface and fracture references, field-compass measurements and conditional bootstrap analysis. On 25 annotated images, pixel-level precision and recall were 90.4% and 73.5%. Following alignment to TLS, block-level median absolute M3C2 surface-normal distances were 2.2–5.0 mm, while independent three-dimensional checkpoint RMSEs were 1.10–2.55 cm. For four fractures in a calcarenite chamber, full-image no-threshold supports covered 83.7–100% of the independent TLS traces within 5 cm. TLS plane discrepancies ranged from 0.3–32.7° without thresholding to 0.6–8.0° after thresholding, with fracture-specific coverage trade-offs. Controlled aggregation comparisons showed no general advantage of weighting, while reduced-image characterization decreased coverage for all four fractures. The workflow links image evidence to independently assessed fracture geometry and orientation. Its structural reliability depends on the spatial distribution of the recovered support, motivating joint evaluation of recovery, plane agreement and resampling stability.

Keywords: GeoAI; rock fracture characterization; rock discontinuities; multi-view photogrammetry; image–LiDAR fusion; underground environments; structural mapping; SAM 3

Check for updates

Academic Editor: Ioannis Koukouvelas

Received: 15 September 2026
Revised: 1 November 2026
Accepted: 3 November 2026
Published: 5 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The structural characterization of fractures and discontinuities in natural and human-made underground cavities is a fundamental requirement for rock-mass assessment, stability analysis, risk mitigation and the documentation of subterranean heritage. Discontinuity orientation, persistence, spacing, aperture, roughness and surface condition are among the descriptors controlling the mechanical behavior of fractured rock masses and the possible development of local instability mechanisms [1,2]. In hypogeal cavities excavated in soft or weakly cemented rock, these descriptors are also relevant for conservation and monitoring because local fracture patterns may affect both the underground void and the built environment above it. Reliable fracture mapping in such contexts therefore requires methods able

to document fracture traces, relate them to three-dimensional geometry and quantify their structural orientation.

Acquiring these descriptors in situ is, however, intrinsically difficult. Global Navigation Satellite System (GNSS) signals are unavailable in underground environments, which are often poorly and unevenly illuminated, geometrically irregular and physically constrained, with narrow passages, limited working distances, restricted accessibility and safety limitations that reduce the number and distribution of viable survey stations. Fracture surfaces are often only partially observable, being exposed mainly as traces over undulating walls, with limited exposure of the fracture planes, while conventional instruments may be difficult to position because of limited floor space, occlusions and uneven walking surfaces. In addition, materials commonly encountered in hypogeal and cave environments, including moist or wet surfaces, porous and weakly cemented calcarenite, locally karstified carbonate surfaces, weathering crusts and variable roughness or reflectance, can degrade image texture, Light Detection and Ranging (LiDAR) returns and the visual continuity of fine discontinuities. These factors make the direct transfer of survey and detection strategies developed for well-lit, accessible and geometrically regular scenes uncertain.

Traditional geological-structural survey methods, including geological-compass measurements, scanline mapping and window mapping, remain essential because they provide direct expert interpretation and physically meaningful structural measurements [1,2]. Nevertheless, in confined underground environments, they are necessarily local, operator-dependent and limited by accessibility, visibility and safety. A compass measurement usually samples a restricted portion of a fracture, whereas the same discontinuity may be long, curved, discontinuous or only partially visible on the wall. As a result, the relationship between local field measurements and the mean attitude of an extended three-dimensional fracture entity can be difficult to establish, especially when fracture traces are irregular or when different reference methods sample different spatial supports.

Close-range photogrammetry, Structure-from-Motion (SfM) and Multi-View Stereo (MVS) have become widely adopted for digital outcrop and rock-mass characterization because they can provide dense, textured and repeatable 3D models at relatively low cost and with flexible acquisition geometries [3,4]. Terrestrial photogrammetry and terrestrial laser scanning (TLS) have been shown to support discontinuity orientation and persistence measurements on rock exposures and slopes, while high-resolution digital outcrop models facilitate the interpretation of structural features beyond the limitations of single-point measurements [5,6]. More recently, smartphone-based LiDAR and mobile devices have been investigated as low-cost tools for rapid 3D surveying and digital outcrop model generation, showing clear operational advantages in terms of portability, speed and accessibility, although their metric reliability remains scale- and acquisition-dependent [7,8]. These properties are attractive for underground surveys, where compact sensors and short acquisition times can reduce on-site occupation, but they also require careful validation before structural parameters are inferred.

TLS, LiDAR and point-cloud analysis provide a complementary route to rock-mass characterization by directly measuring 3D geometry. Dense point clouds have been extensively used in landslide, slope and rock-engineering investigations, and robust cloud-to-cloud comparison methods such as Multiscale Model-to-Model Cloud Comparison (M3C2) support quantitative assessment of geometric agreement between surveys [9,10]. Semi-automatic and automatic point-cloud procedures based on clustering, coplanarity analysis and plane fitting can extract discontinuity sets, estimate their orientations and map persistence [6,11]. Expert-guided tools such as Compass further support rapid semi-automatic fracture and contact mapping across point-cloud and image data [12]. Recent

research is also moving towards deep-learning-based characterization of rock discontinuities from large-scale point clouds, including point-based networks and CNN-based identification followed by extraction of individual discontinuity entities using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and principal component analysis (PCA) [13,14]. Deep-learning-assisted photogrammetric workflows have also begun to combine dense point-cloud reconstruction with automated extraction of individual fracture surfaces and their structural attributes from multi-angle imagery [15]. However, purely point-cloud-based methods are primarily driven by geometry: they may be sensitive to point density, roughness, occlusion and acquisition geometry, and they do not necessarily exploit the graded semantic and radiometric evidence contained in the images.

Image-based crack and fracture detection methods address this semantic component more directly. Classical image processing, CNN-based approaches and recent deep-learning models have improved automated crack and fracture segmentation in civil infrastructure and rock-mass imagery [16–19]. In geological applications, the growing availability of annotated datasets, such as GeoCrack, is facilitating data-driven segmentation of fracture traces across diverse lithologies and imaging conditions [20]. Nevertheless, many image-based methods remain confined to the 2D domain, depend on image-specific annotations or training data, and often produce binary decisions at the pixel level. In underground scenes, artificial illumination, shadows, wet or weathered surfaces and fine rock texture can cause ambiguous fracture evidence; a single image may miss portions that are visible from adjacent viewpoints, while a detected 2D trace does not by itself guarantee metric consistency in three dimensions.

Recent crack-inspection studies in civil engineering have explicitly addressed this limitation by coupling image segmentation with 3D reconstruction. U-Net-based crack segmentation has been used to improve pixel-level detection under variable image conditions [21], while stereo-vision and photogrammetry-based approaches have projected detected crack information into 3D models to reduce errors caused by perspective, surface curvature, occlusion, camera pose and scale conversion [22–24]. More recently, SAM-based crack segmentation has been coupled with visual-inertial LiDAR fusion to align multi-frame crack masks with dense 3D point clouds for automated crack modeling and measurement [25]. These works are important because they demonstrate that crack measurements become more meaningful when image evidence is spatially anchored to a 3D support. At the same time, these civil-infrastructure studies mainly target concrete surfaces, crack-width measurement and infrastructure inspection, often under controlled or semi-controlled conditions. In underground rock engineering, photogrammetric tunnel-face models have also been projected into multiple 2D attribute images for deep-learning-based discontinuity identification and subsequent reprojection to their original 3D context [26]. Their methodological lesson is therefore relevant to the present study, but the geological problem addressed here is different: the objective is to infer structurally interpretable 3D fracture entities and their orientation from multi-view evidence in a confined underground rock environment.

Foundation models for image segmentation have introduced a further shift in how image evidence can be generated. The Segment Anything Model (SAM) demonstrated promptable segmentation and strong zero-shot transfer across many image domains [27]. SAM 3 extends this line towards promptable concept segmentation, where a short noun-phrase concept can be used to detect and segment matching instances in images and videos [28]. Recent work has begun to apply SAM specifically to rock-mass discontinuity mapping, combining interactive SAM segmentation with SfM-based transfer to identify facet-type discontinuities in 3D point-cloud models [29]. Such models are attractive for geological and geomatic workflows because they can be queried with prompts and can

return candidate masks with associated scores, reducing the need for task-specific training in early exploratory applications. However, in structural-geological use, these outputs should not be interpreted as calibrated structural measurements. Candidate masks may be ambiguous, scores may be uncalibrated, and a valid engineering or geomatic interpretation requires perspective-consistent transfer to a metric 3D support, multi-view consolidation and independent validation against geometric and structural references.

Existing approaches increasingly connect image interpretation with reconstructed geometry through deep-learning-based reprojection, SAM–SfM integration and spatially constrained semantic transfer. Weighted multi-view voting has also been used to consolidate image labels on digital outcrop models [30]. For underground fracture characterization, a central challenge is to establish how image evidence distributed across views constrains both the recovered fracture geometry and its structural orientation. Fractures exposed mainly as narrow or discontinuous traces may be geometrically well recovered while providing limited information for plane estimation. Addressing this problem requires a common framework for transferring image evidence, comparing aggregation and selection rules and assessing the resulting supports against separately acquired geometric and structural references.

This paper develops and assesses an AI-assisted multi-view workflow for 3D fracture characterization from smartphone imagery and embedded LiDAR. Four fractures in a chamber of the San Francesco hypogeal complex in Gravina in Puglia (southern Italy) provide the case study. The main contributions are the following:

1. A rapid smartphone image–LiDAR acquisition and reconstruction workflow for a poorly illuminated underground chamber, with geometric assessment against survey-grade TLS and independent registration checkpoints;
2. Text-prompted segmentation that retains candidate masks as graded fracture-support and ambiguity maps, accompanied by pixel-level assessment against manual image annotations;
3. Perspective-consistent image-to-surface transfer and multi-view consolidation, with explicit observation pooling and controlled comparisons of weighted aggregation, union and majority voting against TLS fracture references;
4. Joint assessment of fracture recovery and structural orientation using TLS and field-compass references, support-geometry diagnostics, conditional point- and image-level bootstrap analysis, parameter sensitivity and reduced-image characterization.

The remainder of the paper is organized as follows. Section 2 describes the study area and the acquired datasets. Section 3 details the methodology underpinning the proposed workflow. Section 4 reports the quantitative results at the geometric, 2D, 3D-extraction and structural levels. Section 5 discusses their interpretation, limitations and generalizability, and Section 6 draws the conclusions.

2. Study Area and Dataset

2.1. Study Area

The case study is located in Gravina in Puglia, southern Italy, within the San Francesco hypogeal caves, an artificial underground cavity belonging to the wider rupestrian and subterranean system developed beneath the historical settlement. In this geological and geomorphological context, the widespread presence of the Calcarenite di Gravina favored the excavation of a dense network of artificial underground spaces directly connected with the urban fabric above ground, where the cavities are closely associated with buildings, streets, squares and service spaces excavated within the soft calcarenitic formation [31]. The combination of a soft, weakly cemented lithology and a dense network of cavities developed directly beneath the inhabited urban fabric makes these underground spaces

susceptible to fracturing and localized instability phenomena, which can in turn affect the overlying built environment [32]. This setting motivates rapid, low-cost survey methods able to document and characterize fractures without dedicated infrastructure or prolonged on-site occupation.

The analyzed site corresponds to a single chamber of the San Francesco cavity, selected as the test environment for fracture analysis within the broader underground system. As shown in Figure 1, the investigated chamber belongs to a more articulated complex composed of multiple connected spaces. Its approximate internal dimensions are 6.0 m in height, 5.5 m in width and 9.0 m in longitudinal extent, measured in plan from the access side to the rear wall. The chamber floor lies approximately 9.30 m below the local ground level above the cavity. The site is located at E 619,354 m, N 4,519,651 m, elevation 360 m, in the WGS84/UTM zone 33N reference system. It was selected because, within a single confined environment, it exposes fractures spanning a range of geometric expressions, lengths and illumination conditions and therefore provides a case study for assessing image–LiDAR fracture characterization across contrasting fracture geometries and visibility conditions. The fractures are developed in a single, relatively homogeneous lithology (soft calcarenite); the implications of this choice for the transferability of the results are addressed in Section 5.

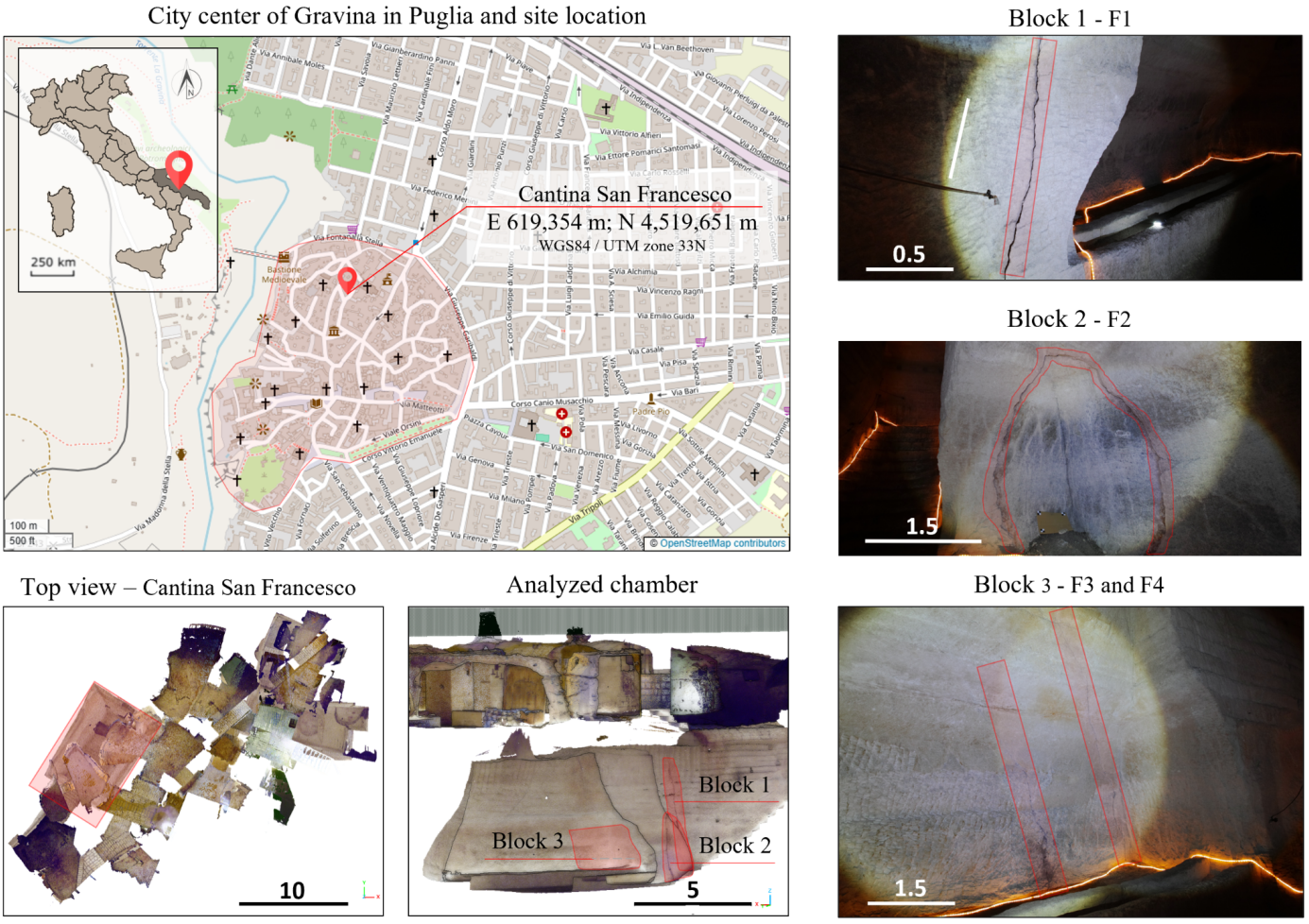


Figure 1. Study area and dataset used in this work. The figure summarizes the geographic and local context of the case study at Cantina San Francesco in Gravina in Puglia (southern Italy), together with the top view of the surveyed underground complex, the analyzed chamber and the three acquisition blocks considered in the paper. Block 1 includes Fracture 1 (F1), Block 2 includes Fracture 2 (F2) and Block 3 includes Fractures 3 and 4 (F3 and F4). The insets on the right show representative close-up views of the analyzed fractures on the reconstructed model, with the corresponding detected fracture traces highlighted (red overlay).

2.2. Data Acquisition and Reference Data

The mobile survey, based on low-cost acquisition hardware, was performed with an Apple iPhone 13 Pro operated through the PIX4Dcatch mobile application (https://www.pix4d.com/product/pix4dcatch/, accessed on 4 September 2026) (v2.14.0). The device combines a 12 MP triple rear-camera system, whose optical characteristics are summarized in Table 1, with an embedded solid-state time-of-flight LiDAR scanner [33]. The LiDAR provides short-range depth observations, with an operational range of up to approximately 5 m under suitable conditions [34], contributing to the metric scaling and geometric completion of the reconstruction. Acquisition was carried out under head-torch illumination only, with the operator walking the chamber while PIX4Dcatch recorded the RGB images together with device pose and depth observations and provided a live preview of the covered geometry. Acquisition was rapid, proceeding at a rate of approximately three seconds per image so that each block was surveyed in roughly one to three minutes (Section 3). The resulting imagery corresponds to an average ground sampling distance of approximately 1 mm or finer, consistent with the photogrammetric quality reports. Images were acquired with nominal forward and side overlaps of approximately 80% and 60%, respectively. RGB imagery and LiDAR-derived depth observations were subsequently combined in PIX4Dmatic (https://www.pix4d.com/product/pix4dmatic-large-scale-photogrammetry-software/, accessed on 4 September 2026) to reconstruct a metrically scaled hybrid image–LiDAR model of the chamber, which constitutes the geometric support for all subsequent processing.

The acquisition hardware comprises a consumer smartphone and a head torch, with an estimated combined cost on the order of EUR 1000 in 2026, more than an order of magnitude below the cost of survey-grade TLS instrumentation. This comparison refers to acquisition hardware and does not imply equivalent costs for software, licensing or processing time. This hardware-cost asymmetry is the operational motivation for assessing whether such a setup can support quantitative structural estimation.

As shown in Figure 1, the dataset comprises 108 images distributed across three acquisition blocks and four analyzed fractures. Block 1 (21 images) contains a single long and narrow fracture, denoted Fracture 1 (F1). Block 2 (54 images) contains one large curved fracture, Fracture 2 (F2). Block 3 (33 images) contains two thinner fractures, Fracture 3 (F3) and Fracture 4 (F4), with different local continuity and visibility. The block organization was maintained throughout acquisition, processing and validation and was intentionally designed to include fractures with contrasting geometric and radiometric characteristics within the same chamber.

Table 1. Summary of datasets and reference sources used in the study.

Data Source	Sensor/Software	Technical Characteristics	Outputs Used in This Study	Use in This Study
Low-cost mobile survey	Apple iPhone 13 Pro/ PIX4Dcatch; PIX4Dmatic	12 MP rear camera system (wide f/1.5, ultra-wide f/1.8 with 120° FoV, and a third camera at f/2.8); embedded short-range time-of-flight LiDAR scanner	Undistorted RGB images, calibrated internal orientation parameters, optimized exterior orientation parameters, hybrid image–LiDAR point cloud and textured mesh	Main dataset for 2D segmentation, 2D–3D transfer, multi-view fusion, and 3D fracture extraction, including the automatic estimation of fracture plane orientation
TLS survey	Leica RTC360/ Cyclone FIELD 360; Cyclone REGISTER 360	Measuring rate up to 2 million points/s; medium resolution (6 mm at 10 m); HDR imaging system	HR TLS point cloud; semi-automatic reference traces and plane orientations	Block-level geometric validation and fracture-level comparison
Field structural survey	Geological compass	In situ structural measurements on selected fractures	Dip and dip-direction reference values	Structural validation of the estimated orientations

Independent reference data were acquired in the same chamber with a Leica RTC360 terrestrial laser scanner. The RTC360 provides a measuring rate of up to two million points per second and an integrated HDR imaging system; it was selected because it yields a dense, metrically reliable reference point cloud suitable for validating both the geometry of the hybrid model and the derived fracture orientations. The scanner was operated at medium resolution (6 mm at 10 m), and the registered TLS survey provides the georeferenced frame into which the three hybrid models were subsequently aligned (Section 3.3). On-site data capture and real-time pre-registration were managed with Leica Cyclone FIELD 360 (https://leica-geosystems.com/products/laser-scanners/software/leica-cyclone/leica-cyclone-field-360, accessed on 4 September 2026) (v6.4), which aligns the individual scans in the field and provides immediate control over coverage and registration quality, while the final registration and optimization of the point cloud were carried out in Leica Cyclone REGISTER 360 PLUS (https://leica-geosystems.com/it-it/products/laser-scanners/software/leica-cyclone/leica-cyclone-register-360, accessed on 4 September 2026) (v2025.0.0). The separately acquired TLS dataset provides the reference geometry for surface comparison, fracture tracing and plane estimation. It also defines the alignment frame of the hybrid models. Reference traces are digitized with the Compass Trace tool by one operator, while field compass readings represent selected locations on the exposed fractures. The TLS and field references therefore describe complementary spatial supports. A separate set of manual annotations covers all visible fractures in 25 images distributed across the three blocks and supplies the reference masks for the pixel-level assessment. The points used to align the hybrid models and the independent checkpoints used to assess registration are distinct from these fracture references.

The hybrid reconstructions were generated independently for the three image blocks. Their calibration completeness, tie-point support, dense-cloud size and processing settings are summarized in Table 2. These quantities describe the input datasets on which the fracture-analysis method operates.

Table 2. Per-block acquisition and processing parameters of the hybrid image–LiDAR reconstruction (PIX4Dmatic). Common settings: maximum extracted keypoints 20,000; calibration image scale 1; depth-and-dense fusion distance 0.025 m; detailed mesh quality. Image dimensions are 1920 × 1440 pixels.

Block (Fractures)	Acquisition Time	Calibrated Images	Median Matches/Image	ATPs	Dense Points	Processing Time
1 (F1)	∼1 min	21/21	2875	21,564	363,015	∼6 min
2 (F2)	∼3 min	54/54	4092	72,412	902,367	∼12 min
3 (F3–F4)	∼2 min	33/33	1834	20,487	450,406	∼8 min

3. Methodology

3.1. Overview of the Proposed Workflow

The workflow combines image-based semantic evidence with calibrated camera geometry to extract and characterize selected rock fractures from a hybrid image–LiDAR reconstruction (Figure 2). Acquisition and reconstruction (A–B) provide the surface model, undistorted imagery and camera parameters. SAM 3 [28] generates candidate image masks (C), which are combined into support and ambiguity maps (D). These maps are transferred to the reconstructed surface by ray casting (E) and consolidated across views (F). Spatial clustering identifies candidate fracture supports (G), and plane fitting yields their structural orientations (H).

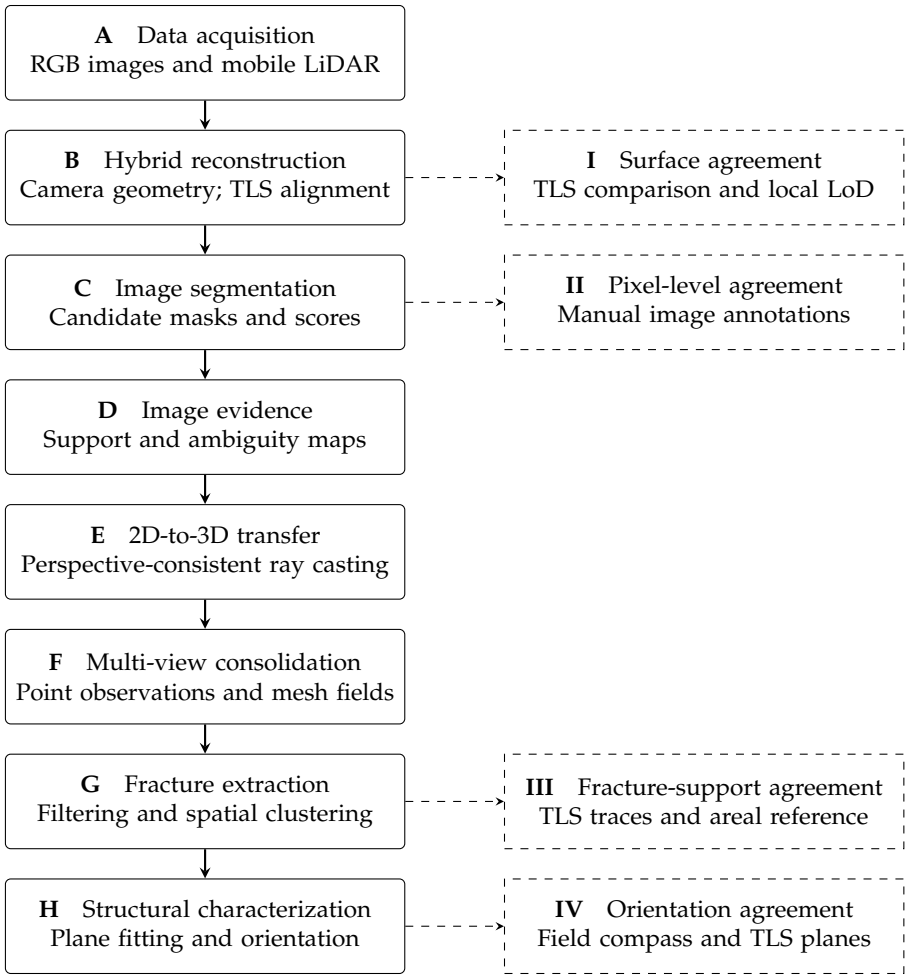


Figure 2. Processing and validation framework. Stages A–H transform the mobile survey into fracture supports and structural orientations. Validation addresses surface agreement with TLS (I), pixel-level segmentation agreement (II), fracture-support agreement (III), and structural orientation (IV). Solid arrows connect processing stages, and dashed arrows connect their outputs to the corresponding validation.

The multi-view stage retains two complementary representations. Concatenated ray–surface observations preserve the contribution of each image and form the input to the primary characterization. A separate mesh field aggregates image evidence at shared surface locations and supports comparison between weighted fusion, union and majority voting. This distinction separates changes in the representation of image evidence from changes introduced by clustering and plane fitting.

Validation first assesses the hybrid surface against TLS (I), followed by pixel-level segmentation agreement with manual annotations (II). Fracture supports are compared with TLS traces and an areal reference (III), and structural orientations are compared with field-compass and TLS measurements (IV). Parameter sensitivity and reduced-image analyses assess the response to changes in processing conditions. Numerical settings and sample sizes are reported with the results.

3.2. Hybrid Image–LiDAR 3D Reconstruction and Data Preparation

The hybrid reconstruction (B) combines photogrammetric image matching with mobile depth observations to obtain a metrically scaled surface model for each acquisition block. The resulting data comprise undistorted RGB images, calibrated internal orientation, optimized camera poses and a textured mesh and point cloud. Processing each block inde-

pendently preserves its camera geometry and provides a consistent surface for transferring the corresponding image evidence.

Following the standard projective camera formulation [35], for each image i, the internal orientation is described by the calibration matrix $\mathbf{K}_i$

$$\mathbf{K}_i = \begin{bmatrix} f_{x,i} & 0 & c_{x,i} \\ 0 & f_{y,i} & c_{y,i} \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where $f_{x,i}$ and $f_{y,i}$ are the focal lengths expressed in pixel units along the image axes, and $c_{x,i}$ and $c_{y,i}$ are the coordinates of the principal point. The camera position is described by the camera center $\mathbf{C}_i$ expressed in the local model frame, whereas the rotation matrix $\mathbf{R}_{cl,i}$ maps vectors from the local model frame l to the camera frame c (a model-to-camera rotation); its transpose $\mathbf{R}_{cl,i}^T$ therefore maps camera-frame directions back into the local model frame.

Projection and ray casting are performed in the local reconstruction frame to maintain numerical precision in the intersection calculations. A rigid transformation relates this frame to the georeferenced TLS coordinates used for structural comparison. The transformation is retained with the reconstructed model and applied to the extracted points before characterization.

3.3. Georeferencing and Registration Control

Structural comparison is performed in the WGS 84/UTM zone 33N reference frame (EPSG:32633), with east, north and vertical axes. The TLS dataset is connected to this frame through an outdoor mobile mapping system (MMS) survey georeferenced using six static GNSS control points. The outdoor survey links the cavity entrances. Shared surfaces at the San Francesco entrance provide the overlap for transferring the reference frame to the underground TLS dataset through an initial rigid alignment using six homologous points, followed by ICP refinement. The three hybrid models are subsequently registered to this georeferenced TLS dataset.

For each block, homologous features identify corresponding positions in the hybrid model and TLS cloud. An initial rigid transformation is estimated from these correspondences and refined by iterative closest point alignment over overlapping surface regions, keeping the model scale fixed. The resulting transformation maps a local point into TLS coordinates as

$$\mathbf{x}_{\mathrm{TLS}} = \mathbf{R}_{\mathrm{TLS}\leftarrow l}\,\mathbf{x}_l + \mathbf{t}_{\mathrm{TLS}\leftarrow l}, \quad (2)$$

where $\mathbf{R}_{\mathrm{TLS}\leftarrow l}$ and $\mathbf{t}_{\mathrm{TLS}\leftarrow l}$ are the rotation and translation. The translation already includes the relationship between the two coordinate origins. Any common coordinate shift used for numerical conditioning is accounted for consistently in the point clouds and reference geometries.

Registration quality is evaluated using identifiable checkpoints that are excluded from both transformation estimation and ICP refinement. Their three-dimensional residuals assess the positional agreement achieved at independent locations. The scan-network bundle error characterizes internal TLS alignment, whereas the checkpoint residuals assess hybrid-to-TLS agreement; these quantities are reported separately.

3.4. Prompt-Based Image Segmentation and Evidence Aggregation

Image segmentation and evidence aggregation (C–D) use SAM 3 [28] to identify candidate discontinuity features from undistorted RGB imagery. The text prompt `crack` is used as an image-level query; geological interpretation is assigned to the resulting three-dimensional supports. Each complete image is processed without task-specific retraining.

The model returns candidate masks and associated scores, retaining all candidates that satisfy its confidence criterion. This criterion acts before the application combines the masks and is distinct from subsequent image-support filtering.

Images are converted to RGB and passed to the model processor, which applies its required resizing and channel normalization. Returned masks are expressed on the original image grid; binary-mask resizing, when required, uses nearest-neighbor interpolation. No additional ranking step restricts the output to the highest-scoring candidate. The processor settings and model configuration are specified with the segmentation results.

For image i, the k-th candidate binary mask is denoted by $m_{i,k}(u,v) \in \{0,1\}$, where (u,v) are the pixel coordinates and $k = 1,\ldots,K_i$ indexes the candidate masks generated for that image. The score associated with the same hypothesis is denoted by $s_{i,k}$. The scores are converted into normalized weights through a softmax function [36],

$$\omega_{i,k} = \frac{\exp(\beta s_{i,k})}{\sum_{j=1}^{K_i} \exp(\beta s_{i,j})}, \tag{3}$$

where $\beta > 0$ is a sharpness parameter controlling the concentration of the weights. When β increases, the weighting becomes more selective and concentrates on the highest-scoring mask hypotheses; when β decreases, the weighting approaches a more uniform combination of the candidate masks.

The pixel-wise support map is then computed as

$$p_i(u,v) = \sum_{k=1}^{K_i} \omega_{i,k}\, m_{i,k}(u,v), \tag{4}$$

which yields a continuous support value in the interval $[0,1]$ for each pixel. Although p_i takes values in the interval $[0,1]$, it is not interpreted as a calibrated probability. It is instead used as an uncalibrated, score-weighted fracture-support map derived from the candidate-mask ensemble. A second map is derived as

$$q_i(u,v) = p_i(u,v)\,[1 - p_i(u,v)], \tag{5}$$

that is, a Bernoulli-form variance term evaluated at the support value $p_i(u,v)$ [37]. This quantity is minimal when the support is close to 0 or 1 and maximal when the candidate-mask ensemble provides intermediate support values. In the present workflow, q_i is used as a compact descriptor of pixel-wise segmentation ambiguity generated by the multi-mask ensemble.

The logical union of all candidate masks provides the binary image representation used for comparison with the manual annotations. Support and ambiguity maps preserve the spatial distribution of the candidate evidence.

Block-level summaries are calculated over images with non-empty candidate masks. Two image-level descriptors are retained. The first is $s_{\max,i} = \max_k s_{i,k}$, i.e., the highest segmentation score among the candidate masks generated for image i. The second is the Shannon entropy [38] of the normalized weights,

$$H_i = -\sum_{k=1}^{K_i} \omega_{i,k} \log \omega_{i,k}, \tag{6}$$

The entropy describes the distribution of score-derived weights among candidate masks. Low values indicate concentration on a few candidates, while high values indicate a more even allocation. Image-level descriptors are summarized over cases in which they are defined, with empty predictions counted separately. They describe the model response and do not measure reference-based segmentation accuracy.

3.5. Perspective-Consistent 2D-to-3D Transfer

The perspective-consistent 2D-to-3D transfer (stage E in Figure 2) maps the image evidence onto the 3D model by camera-based back-projection, following standard projective camera geometry [35]. For a pixel coordinate $\mathbf{u} = (u, v)^T$ in image i, the corresponding direction in camera coordinates is

$$\mathbf{d}_{c,i}(\mathbf{u}) = \frac{\mathbf{K}_i^{-1}[u, v, 1]^T}{\left\|\mathbf{K}_i^{-1}[u, v, 1]^T\right\|}, \tag{7}$$

where the denominator normalizes the direction vector to unit length. The direction in the local model frame is then obtained as

$$\mathbf{d}_{l,i}(\mathbf{u}) = \mathbf{R}_{cl,i}^T \, \mathbf{d}_{c,i}(\mathbf{u}), \tag{8}$$

and the corresponding ray equation is

$$\mathbf{r}_i(s; \mathbf{u}) = \mathbf{C}_i + s \, \mathbf{d}_{l,i}(\mathbf{u}), \qquad s > 0, \tag{9}$$

where s denotes the distance along the ray. The first valid ray–mesh intersection is retained as the 3D hit point associated with pixel $\mathbf{u}$.

To control the density and cost of back-projection, image support is sampled on a regular pixel grid. Two outputs are retained. The no-threshold output includes locations with positive image support; the thresholded output selects locations using an image-dependent quantile of positive support. This formulation accommodates differences in the distribution of support between images without imposing a common absolute value. The same sampling and threshold rules are applied to all acquisition blocks.

Every retained intersection stores its coordinates, support, ambiguity, image identifier and local surface association. These attributes preserve the link between the three-dimensional point and its originating image, allowing subsequent analyses to vary image contributions without re-estimating the camera geometry.

3.6. Multi-View Consolidation on the Reconstructed Surface

Multi-view consolidation (F) combines observations of the same scene from different camera positions. Viewing geometry is represented by a weight that decreases with obliquity and camera-to-surface distance. For a surface point $\mathbf{x}$ observed in image i, the weight is

$$\lambda_i(\mathbf{x}) = [\max(0, \cos\theta_i(\mathbf{x}))]^\gamma \exp\left(-\frac{d_i(\mathbf{x})}{\tau_d}\right), \tag{10}$$

where θ_i is the angle between the surface normal and the direction from the point toward the camera, d_i is their separation, γ controls angular selectivity and τ_d defines the distance-decay scale. Clipping the cosine before exponentiation gives back-facing observations zero weight.

In the point representation, observations from all images are concatenated while retaining their image identifiers and evidence attributes. This representation does not merge coincident coordinates or normalize the number of observations per surface region. Repeatedly imaged portions can therefore contribute multiple points to the subsequent spatial analysis. DBSCAN operates on the point coordinates, and each retained observation contributes equally to the PCA fit.

The mesh representation instead evaluates image support at shared surface vertices and computes normalized cross-view averages. A view contributes when the vertex projects in front of the camera, lies within the image and has positive geometric weight. Views

outside this domain do not contribute to the denominator. Explicit occlusion testing is not applied in the evaluated configuration. For vertices with a non-zero sum of weights, support and ambiguity are

$$P(v) = \frac{\sum_i \lambda_i(v) p_i(v)}{\sum_i \lambda_i(v)}, \qquad Q(v) = \frac{\sum_i \lambda_i(v) q_i(v)}{\sum_i \lambda_i(v)}. \tag{11}$$

The mesh fields provide a shared surface representation for the comparison of multi-view aggregation rules described below.

Comparison of Multi-View Aggregation Rules

A controlled comparison examines whether geometry-dependent weighting changes the fracture support relative to two simpler aggregation rules. Union retains a surface location when at least one valid view provides positive evidence. Majority voting requires positive evidence from more than half of the valid views and at least two valid observations. Ties and single-view observations are excluded as insufficient consensus. This comparison uses the same images, camera geometry and surface domain for all three rules.

In no-threshold mode, positivity means non-zero support. In thresholded mode, union and majority apply the image-specific support threshold before aggregation, whereas the weighted field is thresholded after aggregation. The comparison evaluates the complete aggregation rules, including the stage at which support selection is applied.

To control geometric sampling, the method outputs are represented on a common voxel grid with identical representatives for shared occupied cells. One clustering radius is estimated from the combined comparison domain within each block and used across aggregation rules and configurations. Whole-cluster correspondence is established by spatial overlap with the primary fracture supports and checked at an alternative association distance. This correspondence step identifies comparable physical supports; independent TLS references are then used to assess their geometry and orientation. The controlled comparison is kept separate from the primary concatenated-point estimates.

3.7. 3D Fracture Clustering, Plane Fitting and Structural Orientation

Fracture extraction and characterization (G–H) convert the transferred image observations into spatial supports on which plane orientation can be estimated. The points are first expressed in the TLS frame and selected using optional lower limits on image support and view weight. The inequalities are inclusive, so a zero lower limit retains zero-valued attributes. The selected points are clustered with DBSCAN [39], which partitions density-connected observations and labels isolated points as noise. A minimum retained cluster size suppresses small components.

The neighborhood radius is adapted to the spatial sampling of each primary input. A sorted k-nearest-neighbor distance curve is normalized, and its knee is identified from the largest absolute discrete curvature within an interior interval that excludes the curve ends. A specified distance quantile provides a fallback when the knee is invalid. Multiplication by a scale factor α controls neighborhood extent. Using a common estimation rule preserves the procedure across datasets while allowing the numerical radius to respond to density changes.

Each retained cluster is a candidate fracture support, not an independently established geological entity. Plane fitting is performed separately for spatially associated supports. The primary estimate is unweighted PCA [40], so the coordinates and multiplicity of the retained observations determine the fitted plane. For cluster c with N_c points, the centroid and sample covariance are

$$\mathbf{c}_c = \frac{1}{N_c}\sum_{j=1}^{N_c}\mathbf{x}_{c,j}, \qquad \mathbf{S}_c = \frac{1}{N_c-1}\sum_{j=1}^{N_c}(\mathbf{x}_{c,j}-\mathbf{c}_c)(\mathbf{x}_{c,j}-\mathbf{c}_c)^T. \tag{12}$$

The eigenvector corresponding to the smallest covariance eigenvalue gives the unit plane normal $\mathbf{n}_c$. Its sign is chosen consistently so that the vertical component is non-negative. With east, north and vertical components denoted by $n_{x,c}$, $n_{y,c}$ and $n_{z,c}$, dip and dip direction follow

$$\mathrm{dip}_c = \arccos(|n_{z,c}|)\frac{180}{\pi}, \qquad \mathrm{dipDir}_c = \left[\mathrm{atan2}(n_{x,c}, n_{y,c})\frac{180}{\pi}\right] \bmod 360. \tag{13}$$

The fit is accompanied by orthogonal residuals and covariance-based descriptors of spatial support. These distinguish a broad planar distribution from a near-linear trace, for which the least-variance direction may be weakly constrained. The support geometry and resampling stability jointly characterize the reliability of the fitted orientation.

3.8. Validation Framework and Metrics

Validation uses distinct reference products for the four branches in Figure 2. The TLS observations are independently acquired but also provide the alignment frame for the hybrid models; this dependency is retained when interpreting geometric agreement. Manual image annotations, TLS traces, TLS plane estimates and field measurements assess different spatial quantities and are not treated as interchangeable error-free references.

3.8.1. Surface Agreement with TLS

Surface agreement (I) is assessed with M3C2 [10] over overlapping regions of each registered block. TLS provides the reference cloud and core-point locations; the comparison cloud is sampled from the hybrid mesh. Spatial subsampling controls the sampling of both products. M3C2 estimates local normals and compares the projected point distributions within cylinders centered on the core points. Normal scale, projection diameter, search depth and minimum local sample size define the spatial support of each comparison.

The standard M3C2 level of detection is used with a block-specific scalar registration-error term derived from independent checkpoint residuals. Signed distance, local dispersion, point counts, distance uncertainty and significance flags are retained. Results summarize the magnitude and distribution of surface disagreement, the valid-data fraction and the proportion exceeding the local detection limit. Missing comparisons are reported separately. Significant differences identify local surface discrepancies between the TLS and hybrid representations.

Local regions surrounding the fracture traces are evaluated as subsets of the block comparison, retaining its settings and uncertainty fields. This relates block-level geometric agreement to the surface support of individual fractures. The detection limit characterizes the local surface-distance comparison.

3.8.2. Image-Level Segmentation Agreement

Pixel-level agreement (II) is evaluated on manually annotated images from each acquisition block. The annotation subset comprised 25 images selected through purposive sampling, with approximately proportional representation of the three acquisition blocks. Selection covered different viewing directions, acquisition angles and illumination conditions, including cast shadows and reduced image sharpness. Image selection and manual annotation were performed without inspecting SAM3 predictions. All visible fractures in each selected image were delineated by one operator using CVAT. The union of the predicted masks is compared with the union of the reference polygons over the complete

image. The annotation and prediction grids are brought into correspondence by the fixed orientation mapping required by their image conventions. Polygon vertices are rounded to pixel coordinates, rasterized and combined by logical OR; no morphological correction or spatial matching tolerance is introduced.

Precision, recall, intersection over union and F1 are computed from true-positive, false-positive and false-negative pixel counts. Pooled metrics are calculated after summing counts, while macro summaries average only defined per-image ratios and report their denominators. Empty predictions with non-empty references remain detection failures. Both-empty cases are identified separately because their overlap ratios are undefined. The image sample belongs to the same case study as the processing analyses and is not treated as an external held-out test.

3.8.3. Fracture-Support Agreement with TLS References

Fracture-support agreement (III) uses reference traces digitized directly on TLS with the Compass Trace tool [12]. Companion traces on the hybrid model assess consistency between the reconstructed surfaces, whereas TLS traces provide the primary reference for automatic extraction. The tracing configuration and vertex spacing describe the spatial representation of the reference. For a fracture with sufficiently broad exposed support, a delineated TLS point set additionally provides an areal reference.

Reference polylines are sampled at a uniform curvilinear interval Δs. Automatic supports are represented at a common voxel spacing for distance evaluation, independently of the points used for primary plane fitting and bootstrap. Already normalized supports from the controlled aggregation comparison are not sampled again. The same reference extent and distance tolerances are used across configurations.

Let Γ_r denote a reference trace, $\mathbf{r}_s$ its N_r sampled positions, and $\mathcal{P}_a$ the corresponding automatic support. Coverage measures the fraction of reference samples within tolerance τ of the automatic points:

$$\mathrm{Cov@}\tau = \frac{100}{N_r} \sum_{s=1}^{N_r} \mathbf{1}[d(\mathbf{r}_s, \mathcal{P}_a) \leq \tau]. \tag{14}$$

The median and 90th percentile of these reference-to-automatic distances describe the typical and upper-tail discrepancies. The reverse relationship is summarized by the in-buffer ratio,

$$\mathrm{IBR@}\tau = \frac{100}{N_a} \sum_{j=1}^{N_a} \mathbf{1}\big[d(\mathbf{p}_{a,j}, \Gamma_r) \leq \tau\big], \tag{15}$$

where point-to-trace distance is computed against continuous polyline segments. Coverage and IBR distinguish recovery of reference extent from the presence of automatic points away from that reference. Bidirectional proximity is also evaluated for the TLS areal support.

The maximum gap is the longest consecutive run of uncovered reference samples multiplied by Δs. Internal and terminal gaps are distinguished, and the complete reference is retained independently of the automatic extent. A common primary tolerance is accompanied by alternative tolerances to assess dependence on the distance criterion. The tolerances are evaluated in relation to the local surface and reference-trace discrepancies.

Spatially corresponding clusters are compared across processing configurations for each analyzed fracture. When a support splits, geometric agreement is evaluated for the combined components, and orientations are reported for the individual components.

3.8.4. Structural Agreement and Orientation Stability

Structural agreement (IV) compares the fitted orientations with field compass readings and planes estimated from the corresponding TLS supports. A field reading represents a local measurement, whereas a TLS plane integrates a finite portion of the fracture. The reference values and spatial supports are fixed across processing configurations. Field-compass azimuths acquired on 20 October 2025 were converted from magnetic north to UTM zone 33N grid north. At the site, WMM2025 [41] gives an east-positive magnetic declination of 4.496°. Subtracting the meridian convergence of 0.925° gives an azimuth correction of +3.570°. Dip is unchanged by this conversion, and the converted field orientations are used throughout the comparisons.

Dip differences are absolute. Dip-direction differences account for angular periodicity, while the acute angle between unit normals compares the planes independently of normal sign:

$$E_{\mathrm{dd}} = \min(|\mathrm{dd}_{\mathrm{est}} - \mathrm{dd}_{\mathrm{ref}}|, 360^\circ - |\mathrm{dd}_{\mathrm{est}} - \mathrm{dd}_{\mathrm{ref}}|), \qquad \theta_{\mathrm{plane}} = \arccos(|\mathbf{n}_{\mathrm{est}} \cdot \mathbf{n}_{\mathrm{ref}}|)\frac{180}{\pi}. \tag{16}$$

Orientation estimates and angular differences are calculated using unrounded values and rounded to one decimal place for presentation. Angular differences below 0.1° are reported as <0.1°.

Orientation stability is evaluated by non-parametric bootstrap [42]. Point-level resampling draws observations with replacement at the original sample size. Image-level resampling instead draws contributing image identifiers and retains all cluster points from each selected image, preserving within-image dependence. A plane is refitted for every replicate. Image-level resampling requires more than one contributing image and is the principal diagnostic of sensitivity to view contributions; cases without sufficient images are reported as unavailable.

Bootstrap summaries include dip dispersion, circular dip-direction statistics [43] and the upper percentile of the acute angle between replicate and original normals. The axial statistic complements dip direction because reversal of an almost horizontal normal can produce a large azimuth change without a comparable change in the unoriented plane.

For covariance eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$, linearity, planarity and scattering are $L = (\lambda_1 - \lambda_2)/\lambda_1$, $P = (\lambda_2 - \lambda_3)/\lambda_1$ and $S = \lambda_3/\lambda_1$. The separation between the two smallest eigenvalues describes how distinctly the normal is determined. The relative eigengap is

$$g_n = \frac{\lambda_2 - \lambda_3}{\lambda_2}, \tag{17}$$

evaluated for non-degenerate covariance. A small eigengap indicates that similar least-variance directions can fit the same support. Geometry descriptors, plane residuals, axial and circular dispersion and stability of cluster membership are interpreted jointly. The bootstrap quantifies conditional variability of the fitted plane for the reconstructed fracture support.

3.9. Parameter Sensitivity and Reduced-Image Analysis

Sensitivity analysis evaluates how the representation and characterization respond to changes in processing choices. A common baseline is perturbed one factor at a time across all acquisition blocks. The factors control image-support selection, image sampling, angular and distance weighting, concentration of candidate-mask weights and the scale of the adaptive clustering radius. The baseline settings were established for this case study, and their sensitivity is assessed on the same fractures without a separate parameter-selection dataset. The evaluated values are reported with the corresponding results. Candidate-weight sensitivity is evaluated by recombining saved masks and scores.

The comparison first distinguishes changes in point coordinates from changes restricted to stored attributes. For spatially matched fractures, the same geometric and orientation metrics are then recomputed. Changes in cluster membership, including splitting and merging, are evaluated together with geometric metrics and plane rotations. This design assesses internal robustness of the processing chain using fixed reference data.

The reduced-image analysis restricts the observations available to each fracture while retaining the original reconstruction and camera calibration. Small subsets are manually selected from views showing the fracture, with the aim of maintaining its observable extent while reducing cross-view redundancy. This analysis assesses characterization under a restricted set of observations. The number of selected images and the effective number contributing to each fracture are reported with the results.

For each subset, existing per-image ray intersections are retained and the same no-threshold characterization procedure is repeated. The adaptive-radius rule is recomputed for the reduced support instead of imposing the numerical radius obtained from the full dataset. The comparison includes the density-dependent response of clustering to the reduced image support.

4. Results

The results follow the validation branches of Figure 2: geometric agreement with TLS (I), image-level agreement with manual annotations (II), fracture-support agreement with TLS traces (III) and structural comparison (IV). Thresholded and no-threshold point supports are evaluated separately. Controlled aggregation comparisons, parameter sensitivity and restricted image subsets then assess how the processing configuration affects these quantities.

SAM 3 inference used an NVIDIA Tesla T4, PyTorch `2.10.0+cu128` (CUDA build 12.8) and float32 model parameters. Subsequent Python analyses (Python 3.14) ran on an Intel Core i7-1360P workstation with 32 GB RAM; M3C2 comparisons used CloudCompare 2.13.2. The bootstrap analyses used 2000 replicates with base seed 42. Computational settings and repeatability conditions are documented in Supplement S1. Table 3 summarizes the common processing settings. The same parameter rules were applied across blocks; the adaptive clustering radius was estimated from each input cloud.

Table 3. Baseline processing settings. Stage identifiers refer to Figure 2; parameter definitions are given in Section 3.

Stage	Configuration
C–D	Full-image segmentation; prompt `crack`; candidate-score threshold 0.5; $\beta = 8$.
C–D	Reference processor: 1008×1008 pixels; channel mean/std 0.5/0.5.
E	Sampling stride 12 pixels; support quantile 0.90 for thresholded export.
E	No-threshold export retains positive image support, $p_i > 0$.
F	Angular exponent $\gamma = 2$; distance decay $\tau_d = 3.5$ m.
G	Inclusive lower limits $p_{\min} = \lambda_{\min} = 0$; minimum samples 8; minimum cluster size 20.
G	Eighth-neighbor distance curve; central 5–95% knee search; 0.90 quantile fallback; $\alpha = 2.5$.
III	TLS-trace interval 10 mm; automatic evaluation spacing 5 mm; tolerances 3, 5 and 7 cm.

The geometric view weight is stored as an attribute in the primary point representation; it does not weight the DBSCAN or PCA calculations. Shared-coordinate mesh aggregation is evaluated separately in Section 4.5.

4.1. Geometric Agreement of the Hybrid Models with TLS

The six GNSS control points had mean horizontal and elevation RMS precisions of 1.12 and 1.89 cm, respectively. Mean control-point residuals of the outdoor MMS solution were 2.15 cm horizontally and 2.44 cm in elevation. Co-registration of the San Francesco TLS dataset to the outdoor backbone yielded a positional RMSE of 1.19 cm. The internal

TLS bundle error was 6 mm, calculated as the arithmetic mean of cloud-to-cloud link errors within the scan network.

Each hybrid block was registered to TLS using four homologous points followed by rigid ICP refinement at fixed scale. The four-point alignment RMS residuals were 2.746, 5.702 and 3.529 cm for Blocks 1–3. ICP refinement used a theoretical overlap of 30% and returned cloud-distance RMS values of 2.724, 2.342 and 2.652 cm, respectively. Five independent checkpoints per block yielded three-dimensional RMSEs of 2.55, 1.10 and 1.59 cm for Blocks 1, 2 and 3, respectively (Table 4). All 15 checkpoints were within 5 cm, and 14 were within 3 cm.

Table 4. Hybrid-to-TLS agreement at independent checkpoints excluded from registration estimation. RMSE, median and maximum refer to three-dimensional positional residuals.

Block	Checkpoints	RMSE (cm)	Median (cm)	Maximum (cm)	Within 3 cm	Within 5 cm
1	5	2.55	2.87	3.41	4/5	5/5
2	5	1.10	0.74	1.65	5/5	5/5
3	5	1.59	1.05	2.47	5/5	5/5

M3C2 compared the mesh-sampled hybrid model with the TLS reference cloud. Spatial spacings were 2 mm for the comparison cloud and 10 mm for TLS core points. The normal-estimation scale was 40 mm; projection diameters of 20, 30 and 40 mm were examined, with 30 mm used for the principal results. Maximum search depth was 100 mm, with mean-based statistics, at least five local points, and normals oriented toward $+Y$. Median/IQR statistics, the positive-half-space-only option and precision maps were disabled. The block-specific checkpoint RMSEs, expressed in meters, supplied the scalar registration-error inputs: 0.0255, 0.0110 and 0.0159.

Median absolute M3C2 distances were 2.472, 2.191 and 4.985 mm in Blocks 1–3, respectively (Table 5). The corresponding 90th percentiles were 6.271, 8.487 and 12.444 mm. Mean signed distances were 0.472, 1.247 and −2.396 mm. Block 2 combined a small median discrepancy with the highest proportion exceeding the local level of detection, while Block 3 had the largest 90th-percentile absolute distance. The spatial maps and distributions in Figure 3 show the localization of these departures within each block.

Table 5. Block-level M3C2 surface agreement. Distances and detection limits are in millimeters. Significant discrepancies satisfy $|d| > \mathrm{LoD}_{95}$; their denominator contains only points with both distance and detection limit defined.

Block	Valid Distances /Core Points	Median $\lvert d \rvert$ (mm)	P90 $\lvert d \rvert$ (mm)	Median LoD_{95} (mm)	Significant /Evaluable	Significant (%)	No Distance (%)
1	21,876/21,913	2.472	6.271	50.620	1/21,861	0.005	0.17
2	66,984/68,165	2.191	8.487	21.888	2587/66,951	3.864	1.73
3	90,141/90,341	4.985	12.444	31.543	881/90,113	0.978	0.22

An additional 15, 33 and 28 core points in Blocks 1–3 have a defined distance but no detection limit and are excluded only from the significance calculation. The 100 mm setting is a search depth, not a post-computation distance cutoff.

Median detection limits were 50.620, 21.888 and 31.543 mm. The proportions exceeding these limits were 0.005%, 3.864% and 0.978%, respectively. The exported significance flags agreed with the distance-versus-limit criterion at all 178,925 evaluable core points.

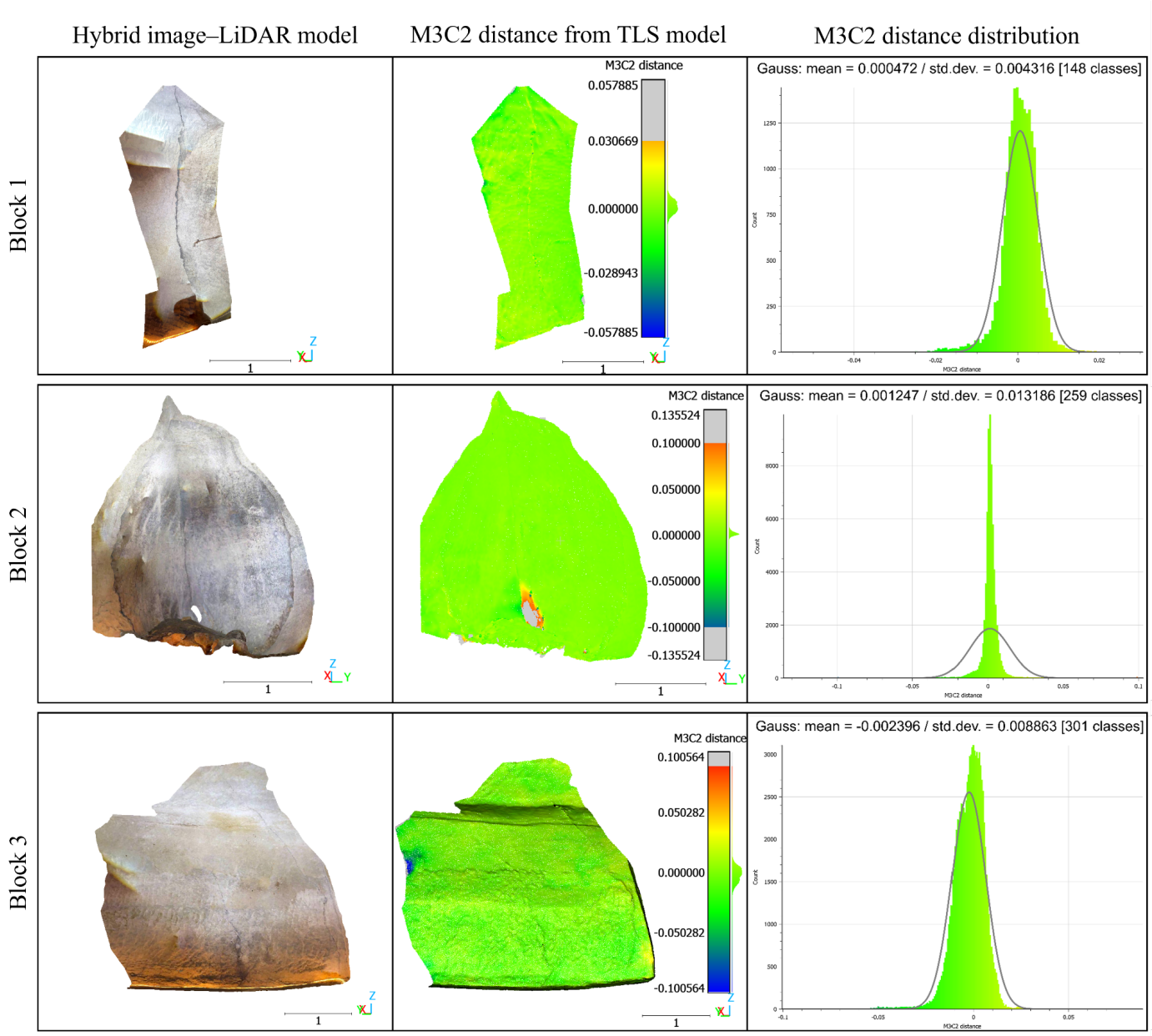


Figure 3. Block-level comparison of the hybrid image–LiDAR models with TLS. Rows show Blocks 1–3; columns show the textured hybrid model, signed M3C2 distances and their distributions. Statistics use finite distance values from the principal 30 mm projection-diameter comparison. The color-bar limits are specific to each block; missing comparisons do not enter the distance distributions.

Local subsets around the four fracture traces provide a diagnostic of the surface support without replacing the block-level assessment (Table 6). Median absolute discrepancies ranged from 2.552 to 4.844 mm. F2 had the largest local proportion of significant discrepancies (8.514%) and 478 core points without a distance, corresponding to 6.23% of its local subset; eight further points had a distance but no detection limit. The missing distances occur in the deeper fracture region represented by TLS but are absent from the hybrid reconstruction. Missing comparisons are excluded from the distance and significance statistics.

Table 6. Local M3C2 surface agreement around F1–F4, using the block-level settings. The proportion exceeding the detection limit is calculated over points with both distance and uncertainty defined.

Fracture	Median $\lvert d \rvert$ (mm)	P90 $\lvert d \rvert$ (mm)	Median LoD_{95} (mm)	Significant (%)	No Distance (%)
F1	2.552	8.243	51.334	0.067	0.00
F2	2.803	18.371	22.226	8.514	6.23
F3	3.536	9.741	31.671	0.000	0.00
F4	4.844	12.577	31.609	0.213	0.12

4.2. Image-Level Evidence and Manual Segmentation Agreement

SAM 3 returned non-empty candidate masks for 94 of 108 images: 20/21 in Block 1, 45/54 in Block 2 and 29/33 in Block 3. The mean maximum candidate scores, $\overline{s_{\max,i}}$, were 0.738, 0.702, 0.672 for Blocks 1–3, respectively. Mean weight entropies were 0.559, 0.617, 0.585. These summaries use the images with non-empty masks and describe the concentration of candidate evidence at $\beta = 8$.

Pixel-level agreement was evaluated on 25 manually annotated images, representing 23.1% of the 108-image dataset (Table 7). The sample comprised five images from Block 1 (23.8%), twelve from Block 2 (22.2%) and eight from Block 3 (24.2%). All visible fractures in each selected image were annotated. The subset was selected and annotated without inspecting SAM3 predictions, following the sampling criteria described in Section 3.8.2.

Table 7. Image availability and pixel-level agreement with manual annotations. Precision, recall, intersection over union (IoU) and F1 are pooled from pixel counts within each block; the last row pools all 25 annotated images.

Block	Total Images	Images with Masks	Annotated Images	Precision (%)	Recall (%)	IoU (%)	F1 (%)
1	21	20	5	87.89	81.64	73.39	84.65
2	54	45	12	90.65	69.75	65.07	78.84
3	33	29	8	92.05	76.80	72.02	83.73
All	108	94	25	90.44	73.47	68.18	81.08

Pooled precision was 90.44%, recall 73.47%, IoU 68.18% and F1 81.08%. Precision exceeded recall in every block, indicating that omission of annotated pixels contributed more strongly than additional predicted pixels to the aggregate disagreement. Block 2 had the lowest recall (69.75%). One annotated Block 2 image contained a reference fracture but no predicted mask and was retained as a complete omission. One Block 3 image had neither reference nor predicted fracture pixels; its overlap ratios were undefined, and it contributed only background counts to the pooled analysis. Macro-averaged precision was 90.11% over 23 defined images; macro recall, IoU and F1 were 73.11%, 67.70% and 78.77%, respectively, over 24 defined images.

Figure 4 juxtaposes manual and predicted masks with pixel-wise agreement and the continuous evidence maps. Correct detections, additional predictions and omissions are distinguished explicitly. The support map p and ambiguity map q describe the distribution of score-weighted candidate evidence; their gray levels are not calibrated probabilities of correctness. The benchmark uses views from the same case study and a single annotator, so the reported agreement does not establish performance on an external image population or quantify inter-annotator variability.

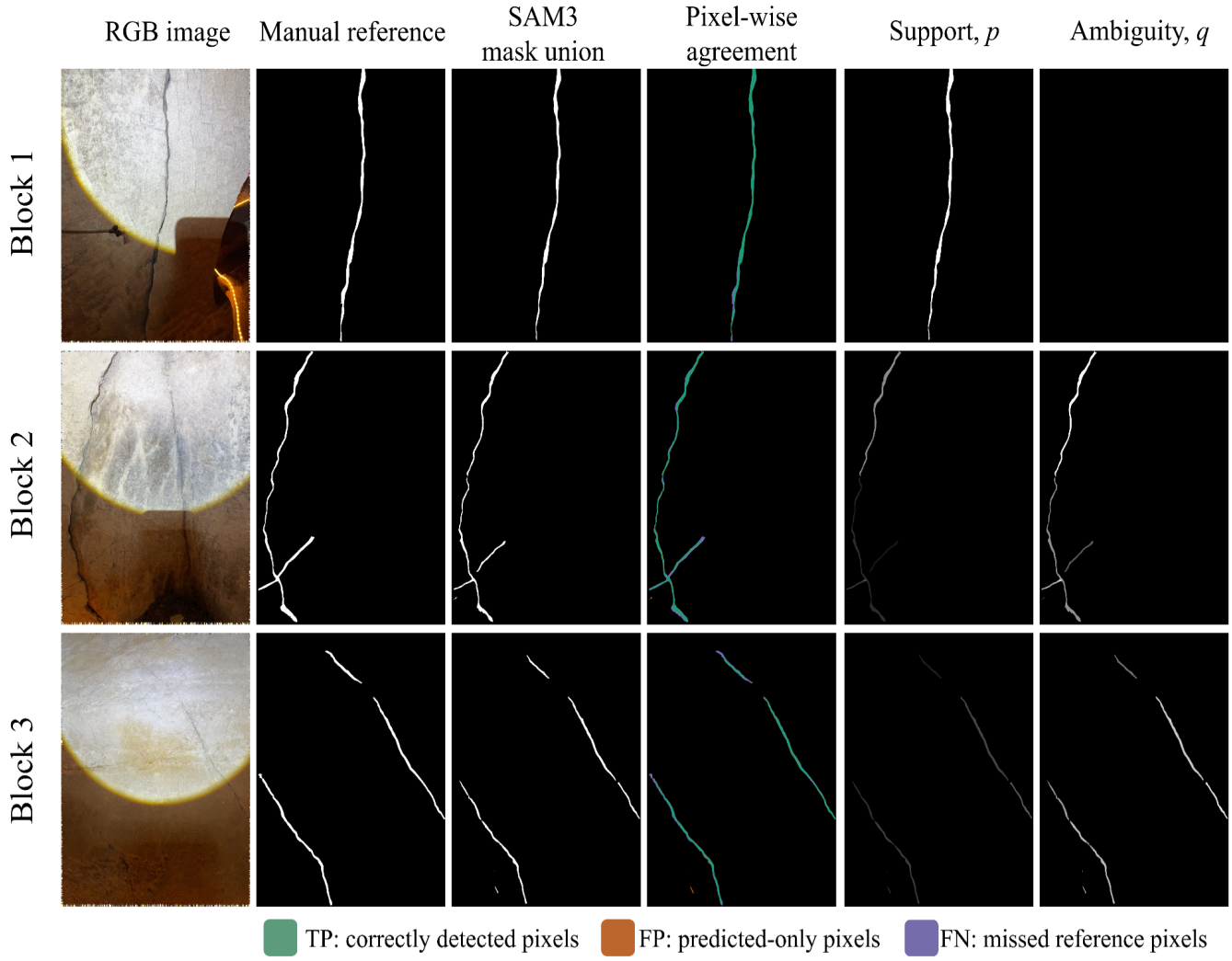


Figure 4. Representative image-level outputs for Blocks 1–3. Columns show the RGB image, manual reference, union of SAM 3 candidate masks, pixel-wise agreement, score-weighted support p and ambiguity $q = p(1 - p)$. In the agreement panels, green denotes true-positive pixels, orange false-positive pixels and purple false-negative pixels; black denotes true-negative background. Manual references include all annotated visible fractures. Support and ambiguity are qualitative grayscale displays, with lighter tones indicating higher displayed values; neither is a pixel-accuracy map.

4.3. Fracture-Support Agreement with Independent TLS Traces

The primary automatic supports comprise the ray–mesh intersections retained after clustering. Figure 5 shows the no-threshold supports on the hybrid models, colored by their image-support attribute p. Geometry and orientation are assessed for F1 in Block 1, F2 in Block 2 and F3 and F4 in Block 3.

Automatic supports were compared with traces digitized directly on TLS. The companion traces on the hybrid models provide a separate check of reference consistency: median TLS-to-hybrid-trace distances were 3.17, 2.10, 1.06 and 1.05 cm for F1–F4, with corresponding 90th percentiles of 5.70, 4.84, 2.28 and 2.35 cm. The Compass Trace settings used cost function 4 and search radii of 0.025 m on TLS and 0.0022 m on the hybrid model. Median TLS vertex spacing was approximately 2.35–2.43 cm; this describes the spacing of the vertices representing the traced geometry.

For automatic-support evaluation, traces were sampled every 1 cm and point clouds represented at 5 mm voxel spacing. The complete TLS reference extent was preserved. A common operational tolerance of 5 cm was used for the principal comparison, with 3 and 7 cm as sensitivity levels (Table A1). The primary tolerance is interpreted at the centimetric scale of the registration and cross-model trace discrepancies.

No-threshold Coverage@5 cm ranged from 83.7% to 100.0%, compared with 61.7–91.2% for thresholded supports (Table 8). Selection therefore did not improve all geometric metrics simultaneously. F3 illustrates the trade-off: thresholding increased IBR@5 cm from 63.2% to 98.9% but reduced coverage from 100.0% to 81.8%. F4 remained entirely within the 5 cm trace buffer in both modes, while coverage was 80.2% and 83.7%, respectively. High compactness of the recovered points therefore did not imply complete recovery of the reference extent.

Table 8. Independent TLS-trace validation at 5 cm. N_{raw} counts the original points in each fracture cluster, and N_{eval} its 5 mm evaluation representatives. R → A denotes reference-to-automatic distance. Thresh and no-thresh denote thresholded and no-threshold support.

Fracture	Mode	N_{raw}	N_{eval}	Coverage (%)	IBR (%)	Median R → A (cm)	P90 R → A (cm)	Max Gap (cm)
F1	thresh	2664	1420	61.7	83.7	4.05	50.69	111
F1	no-thresh	4228	2350	88.9	84.0	2.74	5.22	16
F2	thresh	7592	5517	91.2	82.3	1.90	4.77	19
F2	no-thresh	11,349	7300	91.9	80.1	1.89	4.72	19
F3	thresh	1353	1069	81.8	98.9	0.69	23.13	49
F3	no-thresh	3497	2711	100.0	63.2	0.59	1.09	0
F4	thresh	749	680	80.2	100.0	1.17	44.39	85
F4	no-thresh	1421	1152	83.7	100.0	1.03	30.03	70

Maximum gap is the number of consecutive uncovered reference samples multiplied by the 1 cm sampling interval, including terminal runs.

The largest gap in F1 thresholded support was a terminal run equivalent to 111 cm of 1 cm reference samples. The TLS trace extends beyond the represented automatic support in this portion; the maximum internal gap was 16 cm. The no-threshold result had 88.9% coverage and a maximum gap of 16 cm. Thresholded F3 and F4 likewise had terminal maximum gaps of 49 and 85 cm. These endpoint omissions account for the large upper-tail reference-to-automatic distances despite close agreement along the recovered portions. Figure 6 retains the complete profiles and distinguishes these terminal deficits from internal discontinuities.

F2 had comparable trace coverage in both modes (91.2% thresholded; 91.9% no-threshold) and a broader point support than a single line can represent. Comparison with the independently delineated TLS areal support gave automatic-to-TLS proximity within 5 cm of 99.95% and 96.25%, respectively; the reverse TLS-to-automatic proportions were 98.34% and 98.55%. At 3 cm, automatic-to-TLS proportions were 99.44% and 95.60%, whereas reverse proportions were 87.20% and 88.83%. The areal comparison complements the line-buffer IBR by assessing the broader exposed support of F2.

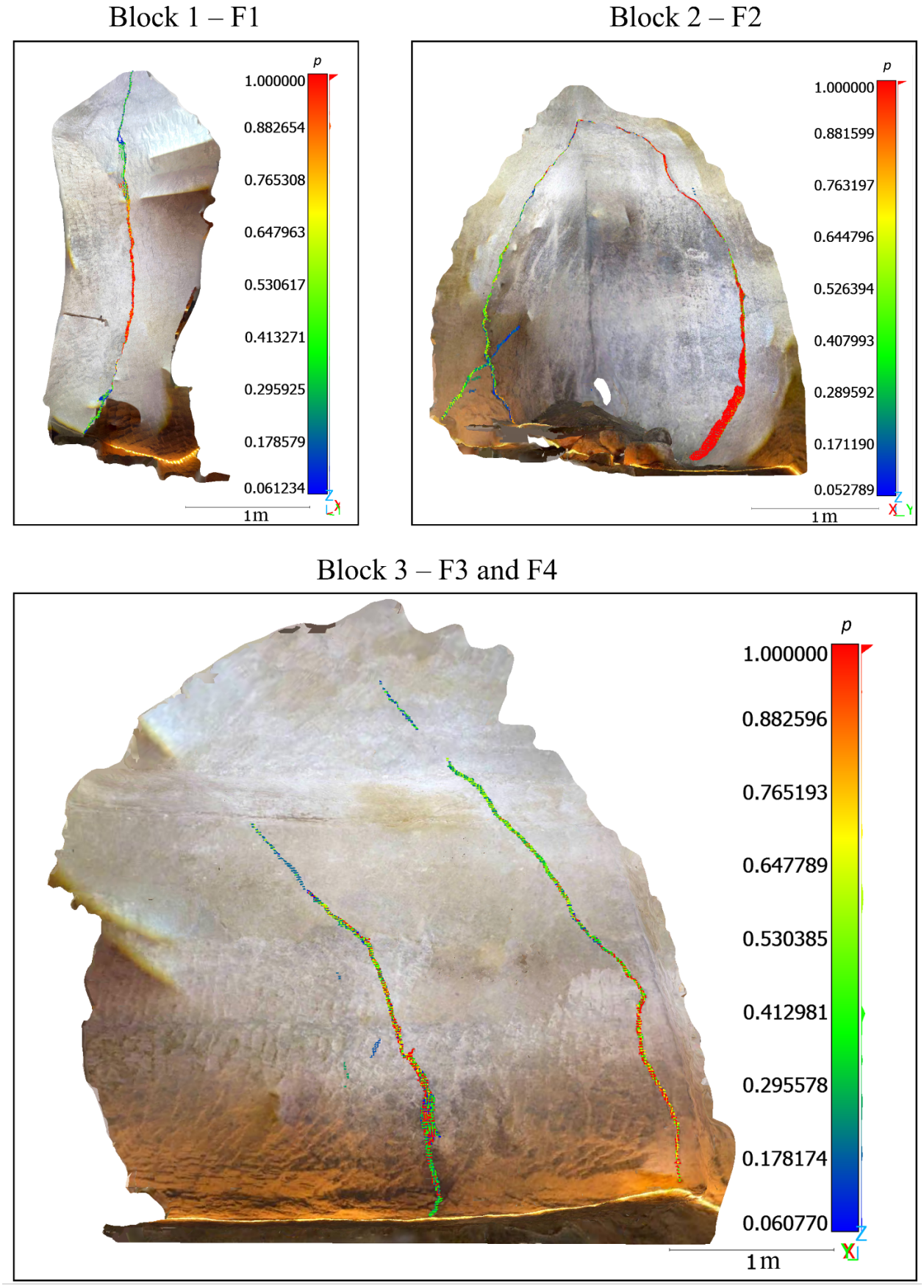


Figure 5. No-threshold fracture-support points on the textured hybrid models of Blocks 1–3. Point colors show the image-support attribute *p* associated with the projected observations.

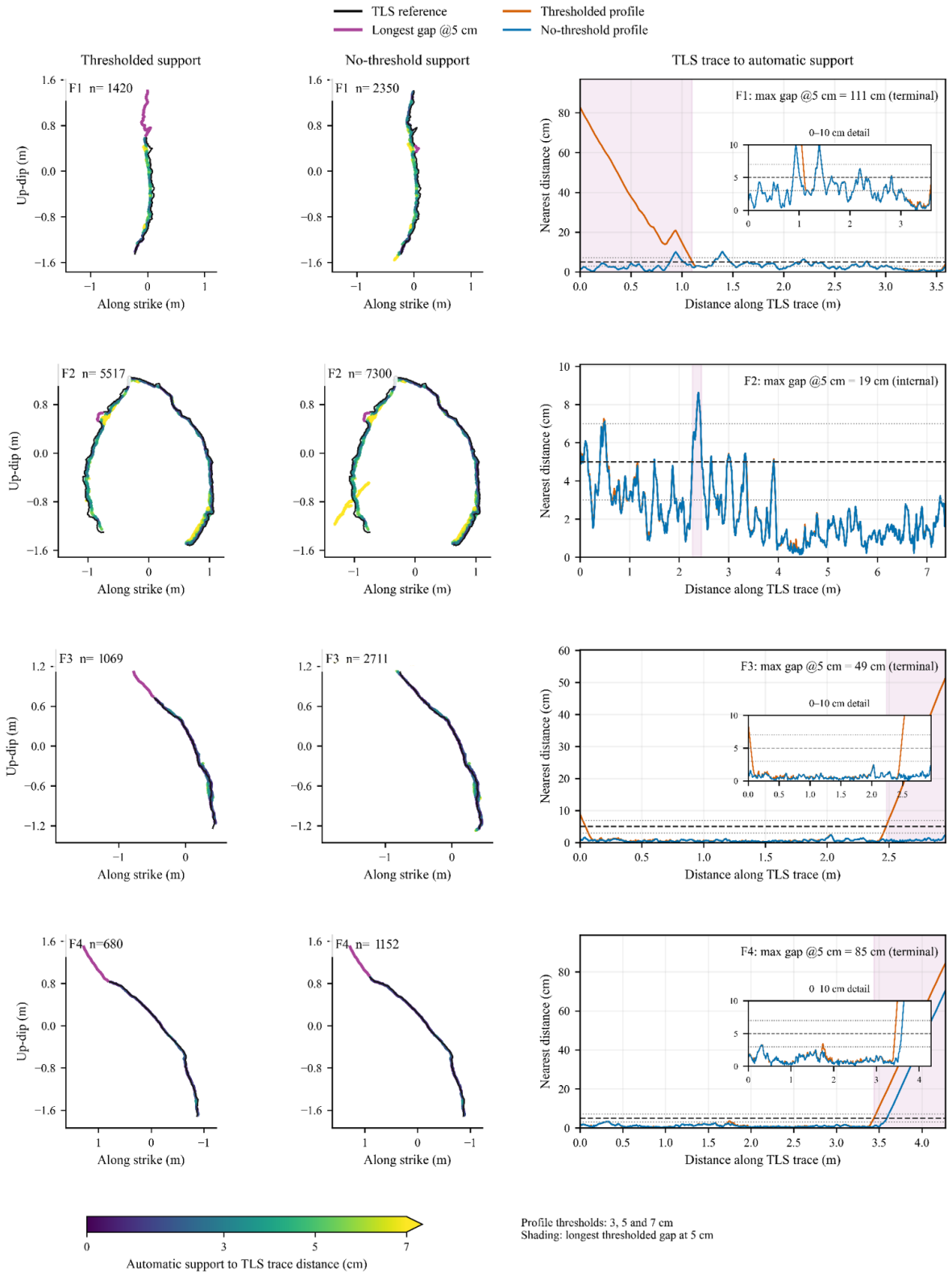


Figure 6. Validation against independent TLS traces for F1–F4. The first two columns show thresholded and no-threshold supports in fracture-oriented views; point colors indicate automatic-to-trace distance in centimetres. Black lines denote TLS references, and magenta marks each support's longest uncovered run at 5 cm. The third column shows reference-to-automatic distance profiles along the complete TLS trace. The dashed line marks 5 cm and dotted lines 3 and 7 cm; shading identifies the longest thresholded gap. Insets expand the 0–10 cm range. The displayed *n* denotes the number of evaluation points obtained using a 5 mm voxel grid.

4.4. Structural Orientations and Conditional Stability

Table 9 compares the estimated orientations with the field-compass and TLS references. Differences in dip, dip direction and plane orientation are reported in Table 10 and illustrated in Figure 7. The full-image thresholded and no-threshold results are presented here; their comparison with reduced-image estimates is developed in Section 4.7.

Table 9. Fracture orientations expressed as dip/dip direction in degrees, compared with field-compass and TLS references. Full-image estimates are reported for thresholded and no-threshold supports; reduced-image estimates use no-threshold support. Field-compass dip directions are expressed relative to UTM grid north.

Fracture	Field Reference	TLS Reference	Full-Image Thresh	Full-Image No-Thresh	Reduced-Image No-Thresh
F1	80/268.6	82/280	81.3/288.1	82.1/285.6	81.9/297.0
F2	88/267.6	82/261	82.3/260.4	82.7/260.1	82.4/260.8
F3	89/93.6	88/98	88.9/94.4	71.3/126.7	88.0/96.8
F4	86/265.6	81/258	83.8/262.3	81.3/258.0	86.2/267.7

Table 10. Orientation differences in degrees. Dip differences are absolute, dip-direction differences use circular wrapping and θ_{TLS} is the acute angle between estimated and TLS plane normals. Full and reduced indicate the image sets used for characterization. Differences are calculated before rounding the orientations reported in Table 9.

Fracture	Mode	Field $\lvert\Delta dip\rvert$	Field $\lvert\Delta DD\rvert$	TLS $\lvert\Delta dip\rvert$	TLS $\lvert\Delta DD\rvert$	θ_{TLS}
F1	Full, thresh	1.3	19.5	0.7	8.1	8.0
F1	Full, no-thresh	2.1	17.1	<0.1	5.6	5.6
F1	Reduced, no-thresh	1.9	28.4	0.1	17.0	16.8
F2	Full, thresh	5.7	7.1	0.3	0.6	0.6
F2	Full, no-thresh	5.3	7.5	0.7	0.9	1.1
F2	Reduced, no-thresh	5.6	6.7	0.4	0.2	0.4
F3	Full, thresh	<0.1	0.8	0.9	3.6	3.7
F3	Full, no-thresh	17.7	33.1	16.7	28.7	32.7
F3	Reduced, no-thresh	1.0	3.2	<0.1	1.2	1.2
F4	Full, thresh	2.2	3.3	2.8	4.3	5.1
F4	Full, no-thresh	4.7	7.5	0.3	<0.1	0.3
F4	Reduced, no-thresh	0.2	2.2	5.2	9.7	11.0

The effect of support selection was fracture-specific. For F3, the no-threshold estimate was 71.3/126.7°, differing from TLS by 16.7° in dip, 28.7° in dip direction and 32.7° in plane angle. Thresholding gave 88.9/94.4° and a plane discrepancy of 3.7°. Its increased trace-buffer compactness therefore coincided with improved orientation agreement, although coverage decreased. The converse pattern occurred for F4: the no-threshold plane differed from TLS by 0.3°, compared with 5.1° after thresholding. F1 also remained closer to TLS without thresholding (5.6° versus 8.0°), whereas F2 showed small plane differences in both modes (1.1° and 0.6°).

Table 11 separates conditional bootstrap dispersion from external agreement. Image-level resampling gave normal-axis P95 values of 14.5° for F1 thresholded and 10.0° for F4 thresholded. F2 was more stable, with P95 below 1° in both full-image configurations. Its support was also less linear: $L = 0.662$ and 0.479, with $P = 0.338$ and 0.519, respectively. The remaining supports were predominantly trace-like, although their smallest-eigenvalue separation remained non-zero. Scattering was low across the full-image supports ($S = 0.0006$–0.0038). Image-bootstrap dip standard deviations ranged from 0.15° to 2.59°. These diagnostics characterize the differing geometric constraints on each fitted plane.

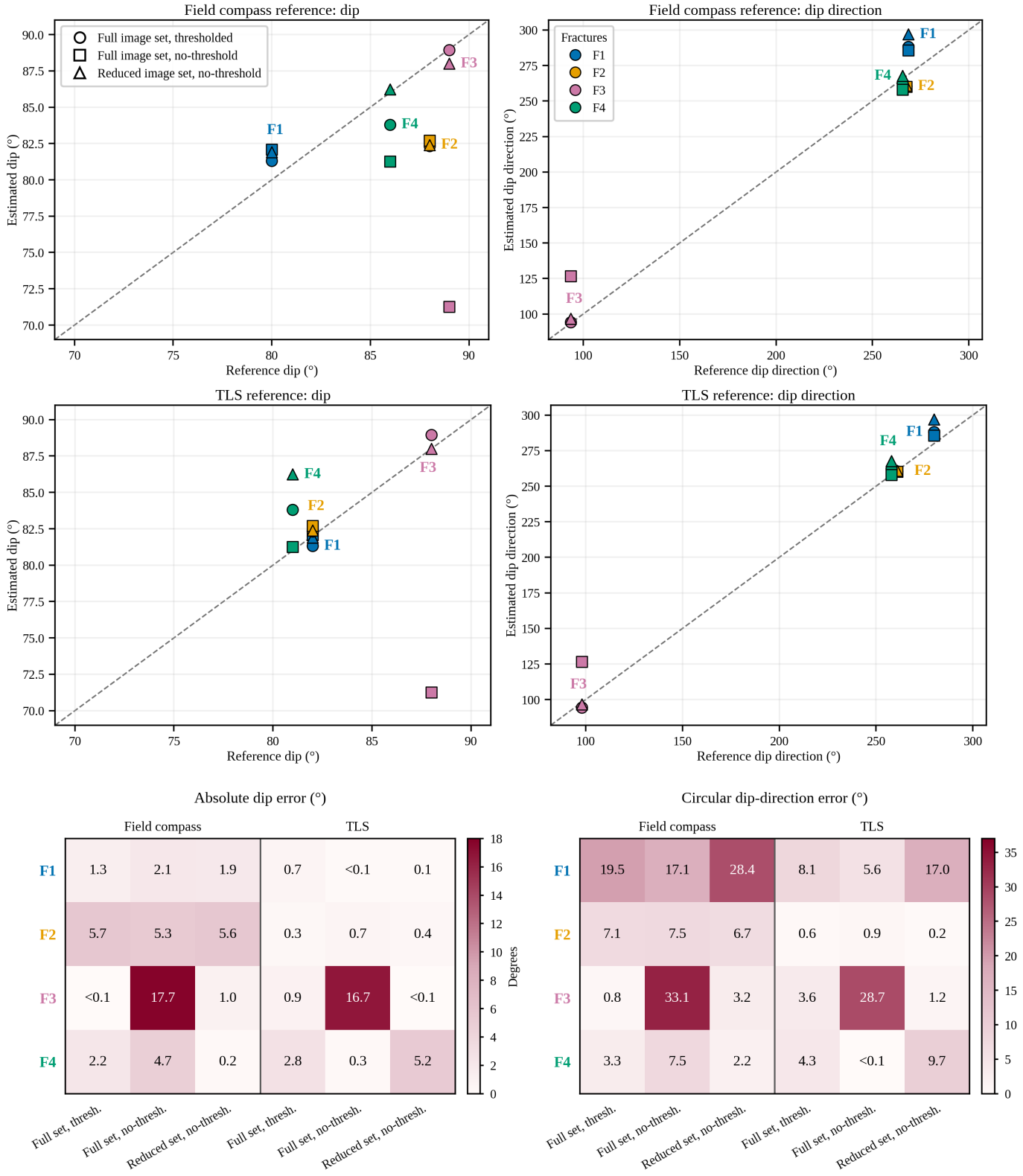


Figure 7. Structural orientation comparison for F1–F4. Scatter plots compare dip and dip direction with field-compass and TLS references for full-image thresholded, full-image no-threshold and reduced-image no-threshold configurations. Dashed lines indicate identity. Matrices report absolute dip differences and circular dip-direction differences in degrees.

Table 11. Plane-fit diagnostics and conditional bootstrap stability. RMSE is the orthogonal plane residual. L, P and g_n denote linearity, planarity and relative normal eigengap. P95 is the 95th percentile of normal-axis deviations for point-level (pt) and image-level (img) resampling. Circular dip-direction standard deviation uses image-level replicates.

Fracture	Mode	Contributing Images	RMSE (cm)	L	P	g_n	P95 pt (°)	P95 Img (°)	SD DD Img (°)
F1	Full, thresh	19	1.26	0.991	0.008	0.898	1.56	14.51	6.77
F1	Full, no-thresh	20	1.46	0.988	0.012	0.954	0.50	3.27	1.76
F1	Reduced, no-thresh	2	1.10	0.996	0.004	0.904	2.46	13.51	5.59
F2	Full, thresh	44	2.73	0.662	0.338	0.998	0.08	0.71	0.26
F2	Full, no-thresh	45	3.84	0.479	0.519	0.997	0.11	0.53	0.19
F2	Reduced, no-thresh	4	3.42	0.335	0.663	0.997	0.26	4.24	5.50
F3	Full, thresh	14	2.61	0.987	0.010	0.795	1.43	5.29	37.25
F3	Full, no-thresh	23	5.85	0.939	0.057	0.938	0.60	6.45	2.88
F3	Reduced, no-thresh	2	2.39	0.985	0.013	0.888	1.96	6.05	3.69
F4	Full, thresh	12	2.87	0.985	0.013	0.922	1.59	10.03	7.50
F4	Full, no-thresh	13	3.01	0.984	0.015	0.924	1.11	5.80	2.50
F4	Reduced, no-thresh	2	4.84	0.984	0.012	0.767	3.54	28.31	63.99

F3 thresholded had an image-bootstrap circular dip-direction SD of 37.3° but a normal-axis P95 of 5.3°. This disparity is consistent with the normal-direction convention near a vertical plane: a reversal can strongly change dip direction without a comparable change in the unoriented plane. Conversely, the no-threshold F3 plane had a 32.7° TLS discrepancy despite an image-axis P95 of 6.5°. A concentrated bootstrap distribution can therefore describe a reproducible but displaced fitted plane; it does not establish agreement with the independently measured fracture.

4.5. Controlled Comparison of Multi-View Aggregation Rules

Weighted averaging, union and majority voting were compared on a common 5 mm mesh-voxel representation, with one clustering radius per block shared across methods. These supports differ from the concatenated ray observations used in the primary characterization; their point counts and plane estimates are reported separately. Majority voting required at least two valid views and a strict majority of positive observations, excluding ties and insufficient consensus. Table 12 summarizes the fracture-level geometric and orientation results; the complete support statistics are given in Table A2.

Table 12. Controlled aggregation comparison. Each method cell reports Coverage@5 cm (%)/IBR@5 cm (%)/plane discrepancy from TLS (°). Thresh and no-thresh indicate the support-selection configurations.

Fracture	Mode	Weighted Mean	Union	Majority Vote
F1	thresh	31.1/95.2/14.2	60.0/80.9/3.2	38.6/92.2/26.6
F1	no-thresh	89.2/83.1/4.3	89.2/83.1/4.3	76.7/83.1/7.3
F2	thresh	38.6/93.0/1.9	91.5/80.1/0.9	56.0/82.5/1.5
F2	no-thresh	91.6/76.0/1.0	91.6/76.0/1.0	87.0/84.8/1.1
F3	thresh	0.0/–/–	88.6/53.9/34.1	22.9/100.0/27.9
F3	no-thresh	100.0/53.6/33.0	100.0/53.6/33.0	73.1/53.8/32.4
F4	thresh	37.5/100.0/65.4	80.2/100.0/5.1	31.0/100.0/68.1
F4	no-thresh	83.9/79.7/7.5	83.9/79.7/7.5	82.5/100.0/2.3

Thresholding follows aggregation for the weighted field and precedes aggregation for union and voting. An absent support has zero coverage; IBR and orientation are unavailable (–).

Weighted averaging and union gave identical no-threshold supports and metrics for every fracture. After thresholding, weighted averaging increased IBR relative to union for F1 (95.2% versus 80.9%) and F2 (93.0% versus 80.1%) but reduced coverage to 31.1% and 38.6%, compared with 60.0% and 91.5% for union. No F3 weighted-thresholded support

survived clustering. For F4, weighted thresholding retained a compact but incomplete support (37.5% coverage; 100% IBR) whose plane differed from TLS by 65.4°. These cases show that spatial compactness alone does not establish recovery of the intended plane.

Majority voting also produced a coverage–selectivity trade-off. In F4 no-threshold mode, it increased IBR from 79.7% for weighted averaging/union to 100%, with similar coverage (82.5% versus 83.9%) and a smaller TLS plane difference (2.3° versus 7.5°). This improvement did not extend to all fractures. The controlled results therefore do not support a general ranking in favor of geometry-dependent weighting. Because the thresholded rules also differ in the stage at which selection is applied, their differences cannot be attributed solely to the view-weight coefficients.

4.6. Sensitivity to Support Selection and Processing Parameters

The one-factor sensitivity analysis varied the support quantile (0.75 and 0.85, versus 0.90), sampling stride (6 and 24 pixels, versus 12), angular exponent ($\gamma = 1, 3$, versus 2), distance decay ($\tau_d = 1.75, 7$ m, versus 3.5 m) and removal of distance or all view weights. The same ten variants were evaluated in each of the three blocks and both export modes. Clustering multipliers $\alpha = 2, 2.5, 3$ and candidate-score concentrations $\beta = 4, 8, 16$ were examined separately. The candidate-weight analysis recombined the stored masks and scores, without running SAM 3 again. Table 13 summarizes the effects of the tested parameter variations on fracture recovery and plane orientation.

Of the 60 point-cloud comparisons from the ten input variants, 44 had exactly the baseline XYZ support and 16 changed coordinates or membership: twelve sampling-stride cases and four threshold-quantile cases. Changes to angular or distance weights altered attributes without changing the primary spatial support under the inclusive zero cutoffs. As the subsequent DBSCAN and PCA are unweighted, their unchanged orientations do not demonstrate insensitivity of a weighted estimator; those weights were not active in the plane fit.

Table 13. Parameter robustness at 5 cm. Each entry is the maximum absolute departure from its corresponding baseline across the four fractures and tested levels. Coverage and IBR differences are percentage points (pp), not relative percentages.

Parameter	Mode	Fracture Cases	Max |ΔCoverage| (pp)	Max |ΔIBR| (pp)	Max Plane Rotation (°)	Split Cases
Support quantile	thresh	8	3.5	0.8	4.6	0
Support quantile	no-thresh	8	0.0	0.0	0.0	0
Sampling stride	thresh	8	5.3	2.9	4.0	0
Sampling stride	no-thresh	8	1.1	6.2	0.4	1
Angular exponent γ	thresh	8	0.0	0.0	0.0	0
Angular exponent γ	no-thresh	8	0.0	0.0	0.0	0
Distance weighting	thresh	12	0.0	0.0	0.0	0
Distance weighting	no-thresh	12	0.0	0.0	0.0	0
View weights	thresh	4	0.0	0.0	0.0	0
View weights	no-thresh	4	0.0	0.0	0.0	0
Clustering multiplier α	thresh	12	2.8	0.1	2.4	0
Clustering multiplier α	no-thresh	12	0.0	0.1	0.0	0
Candidate concentration β	thresh	12	0.0	1.3	0.6	0
Candidate concentration β	no-thresh	12	0.0	0.0	0.0	0

Maxima summarize the analyzed fractures and parameter levels and can occur in different cases. Plane rotations concern single-cluster supports; split cases are counted separately.

The largest coverage departure was 5.3 percentage points for F1 thresholded at stride 24. The largest rotation among single-cluster cases was 4.6° for F4 thresholded at support quantile 0.85. Changing α produced a maximum thresholded rotation of 2.4° for F1 and no associated split over the tested range. Candidate concentration β left no-threshold geometry

unchanged; for thresholded supports, coverage was unchanged and the largest plane rotation was 0.6° for F3, with maximum IBR change of 1.3 percentage points across fractures.

F3 no-threshold at stride 6 split into two associated whole clusters containing 9713 and 4339 points. The first followed the TLS trace (100% coverage; 96.9% IBR at 5 cm) and had a 2.2° plane discrepancy from TLS. The second had zero coverage and IBR and a plane discrepancy of 67.3°. Their orientations were 87.0/99.9° and 27.9/138.6°, respectively. The combined support was used only for geometric evaluation; no single plane was assigned to its union. This case demonstrates a change in support structure that would be obscured by reporting only the small rotations of the remaining single-cluster cases.

The reported variations characterize robustness within the analyzed parameter ranges. F3 shows that a change in cluster structure can accompany limited changes in aggregate coverage.

4.7. Reduced-Image Characterization

The full-image and reduced-image configurations are compared using no-threshold support in both cases. The reduced sets comprised two images for Block 1, four for Block 2 and four for Block 3, with two contributing images for each of F3 and F4. Views were selected to retain visible fracture extent while reducing image redundancy. The existing reconstruction and camera calibration were retained, and characterization was repeated with the same adaptive clustering procedure.

Full-image coverage exceeded reduced-image coverage for all four fractures, by 9.4, 8.1, 13.5 and 5.6 percentage points for F1–F4, respectively (Table 14(a)). The associated changes in IBR differed among fractures: reduced-image support increased IBR for F2 and F3 and decreased it for F4. F1 retained a similar IBR in both configurations.

Table 14. Full-image versus reduced-image no-threshold characterization. Paired entries report full/reduced values. Panel (**a**) summarizes contributing images, original cluster points and geometric agreement at 5 cm. Panel (**b**) compares the plane orientations with field-compass and TLS references and reports image-bootstrap normal-axis P95.

(a) Geometric agreement				
Fracture	**Contributing images**	**Original points**	**Coverage (%)**	**IBR (%)**
F1	20/2	4228/384	88.9/79.4	84.0/83.9
F2	45/4	11,349/1029	91.9/83.8	80.1/87.0
F3	23/2	3497/436	100.0/86.5	63.2/97.3
F4	13/2	1421/353	83.7/78.1	100.0/88.2
(b) Orientation agreement and stability				
Fracture	**Field plane angle (°)**	**TLS plane angle (°)**	**Image-axis P95 (°)**	
F1	17.0/28.1	5.6/16.8	3.3/13.5	
F2	9.1/8.7	1.1/0.4	0.5/4.2	
F3	37.0/3.3	32.7/1.2	6.5/6.0	
F4	8.9/2.2	0.3/11.0	5.8/28.3	

Orientation changes are evaluated against both reference datasets (Table 14(b)). For F1, the reduced-image plane discrepancy increased from 17.0° to 28.1° relative to the field compass and from 5.6° to 16.8° relative to TLS. F2 remained close to TLS in both configurations (1.1° full-image; 0.4° reduced-image), with field-compass plane differences of 9.1° and 8.7°. For F3, reduced-image characterization improved agreement with both references: the plane discrepancy decreased from 37.0° to 3.3° against the compass and from 32.7° to 1.2° against TLS. F4 showed a larger TLS discrepancy under reduced support

(0.3° to 11.0°), while field-compass differences were 8.9° and 2.2°. Component differences in dip and dip direction are reported in Table 10 and Figure 7.

Image-bootstrap normal-axis P95 increased from 3.3° to 13.5° for F1, from 0.5° to 4.2° for F2 and from 5.8° to 28.3° for F4. F3 retained similar dispersion (6.5° full-image; 6.0° reduced-image). The larger full-image coverage therefore coincided with fracture-specific changes in orientation agreement and stability. The manually selected subsets characterize the analyzed view configurations.

5. Discussion

5.1. Image Evidence and Multi-View Consolidation

The workflow connects text-prompt segmentation to a geometric representation on which fracture orientation can be estimated and checked. Retaining candidate masks as graded support preserves their relative contribution through image-to-surface transfer, while the image identifiers allow the resulting spatial evidence to be examined across views. This is useful in the surveyed chamber, where artificial illumination, shadows and partial visibility produce heterogeneous fracture appearances. The manual annotation-based evaluation indicates that missed detections remain an important limitation despite relatively high precision (Table 7). Candidate scores and their weight entropy describe the model response, and the reference masks establish its observed pixel-level agreement. Pixel-level agreement evaluates the candidate union; its translation into 3D fracture recovery also depends on support selection and image-to-surface transfer. A missed segment can receive evidence from another view, but repeated observations can also preserve a consistently misidentified feature. Geological interpretation therefore requires examination of the reconstructed support and its relation to independently observed fractures.

Related approaches connect image interpretation with reconstructed geometry. Wang et al. [29] integrated interactively prompted SAM segmentation with SfM for discontinuity identification, while Dong et al. [30] used spatial projection and weighted voting for semantic mapping of digital outcrops. Photogrammetric tunnel-face mapping also provides a geometric basis for discontinuity identification [26]. The present study extends this application context to smartphone image–LiDAR fracture supports in an underground chamber, combining image agreement, independent TLS comparisons and structural diagnostics within one evaluation.

The controlled aggregation comparison identifies a coverage–selectivity trade-off (Table 12). Weighted averaging and union coincide without thresholding; after selection, compact supports can lose substantial reference extent or provide an unsuitable plane. Majority voting also has fracture-specific effects. Together, these results do not establish a generally superior weighting rule. Threshold placement differs among the compared rules, and the primary characterization fits concatenated ray observations with equal contribution per point. Its estimates consequently depend on the distribution of retained observations, including repeated sampling by different views. The methodological contribution lies in linking these representations to independently evaluated fracture geometry and structural estimates.

5.2. Geometric Agreement and Fracture-Support Recovery

Surface agreement provides the geometric context for interpreting the transferred image evidence. The small median M3C2 distances indicate close correspondence over most of the compared surfaces, with localized departures and different block-specific detection limits (Table 5). These limits include the registration contribution and help explain the low proportion of significant discrepancies. In F2, TLS captures deeper fracture

portions that are absent from the hybrid reconstruction, revealing a local limitation in geometric completeness.

Trace and areal references provide complementary information about extraction. Terminal gaps in Figure 6 record differences between the TLS reference extent and the represented automatic support; close agreement along the recovered portion can coexist with incomplete coverage. F1 illustrates this dependence on extent, while F2 demonstrates the limits of reducing a wider fracture support to a single reference line. Its high bidirectional proximity to the TLS areal reference supports the geological relevance of much of the lateral point distribution. Coverage, IBR and distance profiles should therefore be interpreted together. Increasing the tolerance can admit more laterally distributed points while substantial terminal omissions persist. The 3/5/7 cm comparison characterizes their dependence on the correspondence tolerance, while the surface comparison supplies a separate account of local geometric disagreement.

5.3. Support Geometry and Structural Orientation

A recovered fracture locus can constrain a plane to very different degrees depending on its spatial distribution. A broad planar support restricts the normal through two independent in-plane directions. A narrow trace provides weaker information about rotation around its dominant direction, particularly when local curvature, surface relief or uneven sampling affects the transverse distribution. F2's greater planarity and low image-bootstrap dispersion are consistent with a more strongly constrained orientation (Table 11). For the predominantly trace-like supports, linearity, normal eigengap, plane residuals and resampling variability describe different aspects of the fit. A unique least-variance eigenvector can remain sensitive to which views contribute points. The fitted plane averages the retained support, so a change in its extent or transverse distribution can rotate the estimate even when most points remain close to the visible fracture. Low plane residuals can coexist with an orientation discrepancy, particularly for narrow or curved supports.

F3 shows why trace recovery and structural agreement require separate assessment. Its full-image no-threshold support reaches complete coverage at 5 cm, while the fitted plane differs from TLS by 32.7°. Thresholding reduces the plane discrepancy to 3.7° and increases concentration around the reference trace, at the cost of coverage. The retained spatial distribution therefore changes the fitted orientation substantially. F4 shows a different response, with closer TLS agreement in the full-image no-threshold configuration (Table 10). These cases indicate that support selection must be assessed for each fracture using both spatial and structural evidence. They also preclude treating dip as uniformly well constrained: the F3 discrepancy involves a substantial dip change as well as a dip-direction change.

Bootstrap results help identify sensitivity within an extracted support. Image-level resampling preserves the grouping of observations from the same view, which share segmentation and projection conditions. It consequently examines a different source of variability from resampling individual points. The axial plane statistic also clarifies the interpretation of steep fractures: for F3 thresholded, circular dip-direction dispersion is large while normal-axis dispersion remains much smaller. A change in the adopted normal direction near a vertical plane can strongly affect azimuth while leaving the unoriented plane similar. Conversely, a concentrated bootstrap distribution can surround an estimate displaced from the reference. Agreement with compass and TLS measurements, together with conditional dispersion, therefore provides a fuller assessment than either quantity alone.

Taken together, support geometry, resampling stability and reference comparisons provide the basis for reporting an estimated orientation. An orientation should be labeled provisional and accompanied by the reason for expert verification when: a predominantly trace-like support leaves rotation of the fitted plane around its dominant direction weakly constrained; image resampling or processing variations yield competing plane-normal orientations; or clustering separates the support into components with different plane estimates or uncertain geological correspondence. In split cases, component orientations should be reported individually, without assigning a single orientation to the combined support before geological verification. The provisional designation should identify whether the concern arises from geometric constraint, orientation stability or component correspondence. High linearity, support discontinuity or dip-direction dispersion alone does not justify rejecting an estimate. These are qualitative reporting criteria, not validated numerical acceptance thresholds.

In this study, full-image no-threshold F3 warrants provisional reporting because finer sampling separates its support into differently oriented components, raising questions about point membership and geological correspondence; its disagreement with the TLS plane provides additional evidence for expert verification. Reduced-image F4 warrants the same designation because its predominantly trace-like geometry is accompanied by substantial plane-normal variability under image resampling. For thresholded F3, the reporting concern is specifically dip-direction sensitivity near verticality: its high azimuthal dispersion coexists with a more concentrated plane-normal distribution and does not, by itself, establish instability of the plane.

5.4. Parameter Sensitivity and Reduced-Image Characterization

The sensitivity analysis describes robustness over the examined processing ranges (Table 13). Most single-cluster cases retain similar orientations, but the F3 split at finer image sampling shows that cluster structure can change while aggregate coverage remains high. The individual components describe markedly different geometries and planes, making membership stability essential to interpretation. Changes confined to stored view-weight attributes leave the primary characterization unchanged because clustering and plane fitting use unweighted point coordinates.

The reduced-image comparison reveals a coverage benefit from the full image set for all four fractures, accompanied by variable structural responses (Table 14). F3 improves its orientation agreement with both references under the selected subset, whereas reduced F4 develops a normal-axis P95 of 28.3°. The usefulness of additional views therefore depends on the spatial evidence they contribute, including the distribution of points across the fracture. These subsets retain the reconstruction and camera calibration from the full dataset, and the adaptive clustering rule responds to their altered point density. The results concern restricted observations during characterization; they do not establish the image count required to reconstruct the scene. Acquisition guidance should emphasize visibility of critical fracture portions and complementary viewing geometry, with further work needed to quantify those requirements.

5.5. Reference Representativeness, Limitations and Transferability

Field-compass readings, TLS planes and automatically fitted planes represent different spatial supports. A local field attitude and an orientation averaged over a longer or curved fracture can differ without identifying a single cause for the discrepancy. The references themselves differ by 11.4° in dip direction for F1 and by 6° in dip for F2. Close agreement with one measurement does not resolve the contribution of local surface variation, measurement scale or reference uncertainty to disagreement with the other. TLS acquisition

is independent of the smartphone survey, while its coordinate frame also supports registration of the hybrid models. The validation thus combines complementary references with a shared alignment dependency. Manual image annotations and TLS interpretations were produced by one operator, leaving inter-operator variability unmeasured. Registration, reconstruction and reference-orientation uncertainties were not propagated into the angular estimates; bootstrap dispersion remains conditional on the reconstructed and selected support.

The study concerns four fractures in one calcarenite chamber. Parameter sensitivity and reduced-view comparisons characterize internal consistency within this setting, and the annotated images belong to the same survey. They do not provide an external test of detection performance across lithologies, illumination regimes or underground geometries. The controlled union and voting comparisons establish a useful internal benchmark, while evaluation against independent automated pipelines remains open. Extension to additional sites and operators, with independently documented reference supports and systematic variation of acquisition geometry, would test the transferability of the proposed characterization strategy and its practical measurement limits.

6. Conclusions

This study developed an AI-assisted workflow for characterizing rock fractures from smartphone image–LiDAR data in an underground chamber. Text-prompted segmentation, graded image evidence and image-to-surface transfer produced three-dimensional supports that could be evaluated against manual annotations, independent TLS observations and field-compass measurements. The main contribution is the integration of image interpretation, quantitative fracture recovery and structural assessment within a common spatial framework.

The four full-image no-threshold supports achieved 83.7–100% coverage of the TLS reference traces at 5 cm. Thresholding changed coverage and orientation differently among fractures: it improved the F3 plane estimate substantially, while the full-image no-threshold configuration provided closer TLS agreement for F4. These outcomes show that high trace coverage alone does not establish a reliable structural orientation. Support extent, transverse distribution and planarity influence the fitted plane and must be considered alongside reference agreement and conditional bootstrap dispersion.

Controlled comparisons with union and majority voting identified fracture-specific differences in support selection, without demonstrating a general advantage of weighted aggregation. Sensitivity analysis indicated modest orientation changes for most single-cluster cases, with a split in F3 revealing a change in cluster structure. Full-image characterization provided greater trace coverage than the selected reduced-image subsets for every fracture, although additional views did not consistently improve orientation agreement. These findings favor evaluating the spatial contribution of observations when assessing multi-view redundancy.

The evidence is limited to four fractures in one calcarenite chamber. Future work should assess additional sites and lithologies, quantify inter-operator reference variability and propagate registration and reference-orientation uncertainty into structural estimates. Systematic variation of viewing geometry would also help define acquisition requirements. Within the investigated setting, the workflow supports quantitative fracture characterization while making incomplete recovery and orientation sensitivity explicit.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/geosciences16110460/s1, Supplement S1: Computational reproducibility, including the SAM 3 application interface, archived model identifiers, software environments and conditions of repeatability.

Author Contributions: Conceptualization, A.S.; methodology, A.S.; software, A.S.; validation, A.S.; formal analysis, A.S.; investigation, A.S.; resources, A.S. and M.P.; data curation, A.S.; writing—original draft preparation, A.S.; writing—review and editing, A.S. and M.P.; visualization, A.S.; supervision, M.P.; project administration, M.P.; funding acquisition, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This study was carried out within the RETURN Extended Partnership and received funding from the European Union NextGenerationEU under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.3, D.D. 1243 2/8/2022, PE0000005, SPOKE 6—TS2. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The survey datasets and application code are not publicly available because they form part of an ongoing doctoral research project. Requests for access should be sent to the corresponding author (alessandra.spadaro@polito.it), specifying the requested materials, the verification or research purpose and the intended use. The corresponding author will assess the scope of access and, where agreed, provide the specified materials by private transfer. Computational settings and repeatability conditions are documented in Supplement S1.

Acknowledgments: The authors thank Nives Grasso (Politecnico di Torino) and Daniele Giordan (Research Institute for Geo-Hydrological Protection, Italian National Research Council, CNR-IRPI) for their support during data acquisition. The authors also thank the cultural association "Gravina Sotterranea" and, in particular, Alessandro Parisi, for granting access to the case-study site and supporting the survey activities.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
ATP	Automatic tie point
CRS	Coordinate reference system
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
FoV	Field of view
GNSS	Global Navigation Satellite System
GSD	Ground sampling distance
HDR	High dynamic range
IBR	In-buffer ratio
LiDAR	Light Detection and Ranging
LoD	Level of detection
M3C2	Multiscale Model-to-Model Cloud Comparison
PCA	Principal component analysis
RGB	Red, green and blue
SfM	Structure-from-Motion
MMS	Mobile mapping system
MVS	Multi-View Stereo
RMSE	Root-mean-square error
SAM	Segment Anything Model
TLS	Terrestrial laser scanning
UTM	Universal Transverse Mercator
WGS84	World Geodetic System 1984

Appendix A. Additional Validation Results

The following tables retain the alternative distance tolerances and the complete fracture-level summaries of the controlled aggregation comparison. They use the same fixed TLS references and support correspondences as the principal results.

Table A1. Sensitivity of geometric agreement to the distance tolerance. Columns report Coverage and IBR in percent for the full-image thresholded, full-image no-threshold and reduced-image no-threshold supports.

Fracture	Mode	Coverage 3 cm	Coverage 5 cm	Coverage 7 cm	IBR 3 cm	IBR 5 cm	IBR 7 cm
F1	Full, thresh	35.3	61.7	67.2	33.8	83.7	98.1
F1	Full, no-thresh	56.4	88.9	95.3	41.6	84.0	96.3
F1	Reduced, no-thresh	41.1	79.4	88.3	33.2	83.9	99.5
F2	Full, thresh	71.6	91.2	98.4	53.6	82.3	92.9
F2	Full, no-thresh	72.1	91.9	98.4	52.2	80.1	89.6
F2	Reduced, no-thresh	63.6	83.8	89.2	70.4	87.0	91.5
F3	Full, thresh	80.5	81.8	83.2	90.6	98.9	100.0
F3	Full, no-thresh	100.0	100.0	100.0	57.5	63.2	64.8
F3	Reduced, no-thresh	84.2	86.5	88.6	91.2	97.3	100.0
F4	Full, thresh	77.6	80.2	80.7	97.5	100.0	100.0
F4	Full, no-thresh	81.6	83.7	84.4	97.6	100.0	100.0
F4	Reduced, no-thresh	74.4	78.1	78.6	85.3	88.2	88.2

Table A2. Controlled aggregation supports on the common 5 mm mesh-voxel grid. R → A denotes reference-to-automatic distance. Orientations are dip/dip direction in degrees, and θ_{TLS} is the plane-normal discrepancy from TLS.

Fracture	Mode	Method	*N*	Median R → A (cm)	P90 R → A (cm)	Max Gap (cm)	Dip/DD (°)	θ_{TLS} (°)
F1	thresh	Weighted	231	25.07	106.44	178	81.1/294.3	14.2
F1	thresh	Union	1168	4.15	56.70	117	81.4/283.2	3.2
F1	thresh	Vote	537	14.06	76.32	149	81.0/306.9	26.6
F1	no-thresh	Weighted	1650	2.69	5.18	16	82.0/284.3	4.3
F1	no-thresh	Union	1650	2.69	5.18	16	82.0/284.3	4.3
F1	no-thresh	Vote	1073	3.26	10.46	41	81.8/287.4	7.3
F2	thresh	Weighted	302	17.86	156.33	271	83.5/259.9	1.9
F2	thresh	Union	1899	1.94	4.68	19	82.3/260.1	0.9
F2	thresh	Vote	550	3.62	111.58	198	82.8/259.8	1.5
F2	no-thresh	Weighted	2265	1.91	4.66	19	82.3/260.1	1.0
F2	no-thresh	Union	2265	1.91	4.66	19	82.3/260.1	1.0
F2	no-thresh	Vote	1144	2.00	6.24	59	82.6/260.1	1.1
F3	thresh	Weighted	0	–	–	297	–	–
F3	thresh	Union	469	1.16	6.31	28	70.8/128.0	34.1
F3	thresh	Vote	22	38.54	93.64	121	88.1/250.4	27.9
F3	no-thresh	Weighted	597	1.02	1.79	0	71.9/127.3	33.0
F3	no-thresh	Union	597	1.02	1.79	0	71.9/127.3	33.0
F3	no-thresh	Vote	277	1.57	21.11	63	71.7/126.6	32.4
F4	thresh	Weighted	85	52.87	213.22	266	78.0/140.3	65.4
F4	thresh	Union	195	1.56	44.45	85	83.9/262.3	5.1
F4	thresh	Vote	66	69.15	229.33	282	78.4/143.3	68.1
F4	no-thresh	Weighted	350	1.35	29.49	69	84.5/264.6	7.5
F4	no-thresh	Union	350	1.35	29.49	69	84.5/264.6	7.5
F4	no-thresh	Vote	206	1.49	34.56	74	82.5/259.7	2.3

For absent F3 weighted-thresholded support, coverage is zero and the 297 cm gap count spans the uncovered TLS reference. Distances, IBR and orientation are unavailable (–).

References

1. International Society for Rock Mechanics Commission on Standardization of Laboratory and Field Tests. Suggested Methods for the Quantitative Description of Discontinuities in Rock Masses. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1978**, *15*, 319–368. https://doi.org/10.1016/0148-9062(78)91472-9.
2. Priest, S.D. *Discontinuity Analysis for Rock Engineering*; Chapman & Hall: London, UK, 1993.

3. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. Structure-from-Motion Photogrammetry: A Low-Cost, Effective Tool for Geoscience Applications. *Geomorphology* **2012**, *179*, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021.
4. Sturzenegger, M.; Stead, D. Close-Range Terrestrial Digital Photogrammetry and Terrestrial Laser Scanning for Discontinuity Characterization on Rock Cuts. *Eng. Geol.* **2009**, *106*, 163–182. https://doi.org/10.1016/j.enggeo.2009.03.004.
5. Sturzenegger, M.; Stead, D. Quantifying Discontinuity Orientation and Persistence on High Mountain Rock Slopes and Large Landslides Using Terrestrial Remote Sensing Techniques. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 267–287. https://doi.org/10.5194/nhess-9-267-2009.
6. Riquelme, A.J.; Tomás, R.; Cano, M.; Pastor, J.L.; Abellán, A. Automatic Mapping of Discontinuity Persistence on Rock Masses Using 3D Point Clouds. *Rock Mech. Rock Eng.* **2018**, *51*, 3005–3028. https://doi.org/10.1007/s00603-018-1519-9.
7. Luetzenburg, G.; Kroon, A.; Kjeldsen, K.K.; Splinter, K.D.; Bjørk, A.A. High-Resolution Topographic Surveying and Change Detection with the iPhone LiDAR. *Nat. Protoc.* **2024**, *19*, 3520–3541. https://doi.org/10.1038/s41596-024-01024-9.
8. Furlan, L.M.; Piazentim, E.G. Smartphone-Based LiDAR for Generating Digital Outcrop Models (DOMs) with Field Validation. *Discov. Geosci.* **2025**, *3*, 137. https://doi.org/10.1007/s44288-025-00253-z.
9. Jaboyedoff, M.; Oppikofer, T.; Abellán, A.; Derron, M.-H.; Loye, A.; Metzger, R.; Pedrazzini, A. Use of LIDAR in Landslide Investigations: A Review. *Nat. Hazards* **2012**, *61*, 5–28. https://doi.org/10.1007/s11069-010-9634-2.
10. Lague, D.; Brodu, N.; Leroux, J. Accurate 3D Comparison of Complex Topography with Terrestrial Laser Scanner: Application to the Rangitikei Canyon (N-Z). *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 10–26. https://doi.org/10.1016/j.isprsjprs.2013.04.009.
11. Riquelme, A.J.; Abellán, A.; Tomás, R.; Jaboyedoff, M. A New Approach for Semi-Automatic Rock Mass Joints Recognition from 3D Point Clouds. *Comput. Geosci.* **2014**, *68*, 38–52. https://doi.org/10.1016/j.cageo.2014.03.014.
12. Thiele, S.T.; Grose, L.; Samsu, A.; Micklethwaite, S.; Vollgger, S.A.; Cruden, A.R. Rapid, Semi-Automatic Fracture and Contact Mapping for Point Clouds, Images and Geophysical Data. *Solid Earth* **2017**, *8*, 1241–1253. https://doi.org/10.5194/se-8-1241-2017.
13. Chen, Q.; Ge, Y.; Tang, H. Rock Discontinuities Characterization from Large-Scale Point Clouds Using a Point-Based Deep Learning Method. *Eng. Geol.* **2024**, *337*, 107585. https://doi.org/10.1016/j.enggeo.2024.107585.
14. Ge, Y.; Wang, H.; Liu, G.; Chen, Q.; Tang, H. Automated Identification of Rock Discontinuities from 3D Point Clouds Using a Convolutional Neural Network. *Rock Mech. Rock Eng.* **2025**, *58*, 3683–3700. https://doi.org/10.1007/s00603-024-04351-1.
15. Dong, S.; Fu, K.; Wang, L.; Zeng, L. Intelligent Characterization of Fractures Using 3D Point Clouds of Outcrops Generated from 2D Photos. *Comput. Geosci.* **2026**, *210*, 106136. https://doi.org/10.1016/j.cageo.2026.106136.
16. Cha, Y.-J.; Choi, W.; Büyüköztürk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. https://doi.org/10.1111/mice.12263.
17. Dorafshan, S.; Thomas, R.J.; Maguire, M. Comparison of Deep Convolutional Neural Networks and Edge Detectors for Image-Based Crack Detection in Concrete. *Constr. Build. Mater.* **2018**, *186*, 1031–1045. https://doi.org/10.1016/j.conbuildmat.2018.08.011.
18. Zhou, S.; Canchila, C.; Song, W. Deep Learning-Based Crack Segmentation for Civil Infrastructure: Data Types, Architectures, and Benchmarked Performance. *Autom. Constr.* **2023**, *146*, 104678. https://doi.org/10.1016/j.autcon.2022.104678.
19. Esmaeili, M.; Beni, T.; Gigli, G.; Tofani, V. Rock Mass Exposure Fracture Detection through 2D Close-Range Images Using Image Processing Techniques: A Review. *Earth Sci. Inform.* **2025**, *18*, 494. https://doi.org/10.1007/s12145-025-01980-0.
20. Yaqoob, M.; Ishaq, M.; Ansari, M.Y.; Konagandla, V.R.S.; Al Tamimi, T.; Tavani, S.; Corradetti, A.; Seers, T.D. GeoCrack: A High-Resolution Dataset for Segmentation of Fracture Edges in Geological Outcrops. *Sci. Data* **2024**, *11*, 1318. https://doi.org/10.1038/s41597-024-04107-0.
21. Liu, Z.; Cao, Y.; Wang, Y.; Wang, W. Computer Vision-Based Concrete Crack Detection Using U-Net Fully Convolutional Networks. *Autom. Constr.* **2019**, *104*, 129–139. https://doi.org/10.1016/j.autcon.2019.04.005.
22. Kim, H.; Sim, S.-H.; Spencer, B.F. Automated Concrete Crack Evaluation Using Stereo Vision with Two Different Focal Lengths. *Autom. Constr.* **2022**, *135*, 104136. https://doi.org/10.1016/j.autcon.2022.104136.
23. Ozturk, H.Y.; Zappa, E. Automated Crack Width Measurement in 3D Models: A Photogrammetric Approach with Image Selection. *Information* **2025**, *16*, 448. https://doi.org/10.3390/info16060448.
24. Xiao, D.; Wang, G.; Wang, K.; Liu, S.; Shang, G.; Wang, Q.-A.; Fan, X.; Hu, M.; Liu, R.; Chen, G.; et al. Intelligent Measurement of Concrete Crack Width Based on U-Net Deep Learning and Binocular Vision 3D Reconstruction. *Appl. Sci.* **2026**, *16*, 2355. https://doi.org/10.3390/app16052355.
25. Deng, P.; Yao, J.; Li, C.; Wang, S.; Li, X.; Ojha, V.; He, X. 3D Modeling and Automated Measurement of Concrete Cracks via Segment Anything Refinement and Visual Inertial LiDAR Fusion. *Comput.-Aided Civ. Infrastruct. Eng.* **2026**, *45*, 100019. https://doi.org/10.1016/j.cacaie.2026.100019.
26. Pham, C.; Kim, B.-C.; Shin, H.-S. Deep Learning-Based Identification of Rock Discontinuities on 3D Model of Tunnel Face. *Tunn. Undergr. Space Technol.* **2025**, *158*, 106403. https://doi.org/10.1016/j.tust.2025.106403.
27. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; Dollár, P.; Girshick, R. Segment Anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 2–6 October 2023; pp. 4015–4026.

28. Carion, N.; Gustafson, L.; Hu, Y.-T.; Debnath, S.; Hu, R.; Suris Coll-Vinent, D.; Ryali, C.; Alwala, K.V.; Khedr, H.; Huang, A.; et al. SAM 3: Segment Anything with Concepts. In Proceedings of the Fourteenth International Conference on Learning Representations (ICLR 2026), Rio de Janeiro, Brazil, 23–27 April 2026. Available online: https://proceedings.iclr.cc/paper_files/paper/2026/hash/e0982cbc81401df3430ee1ff780dc7a2-Abstract-Conference.html (accessed on 3 September 2026).
29. Wang, J.; Zheng, J.; Hu, J.; Gong, X.; Lü, Q.; Han, J.; Sun, J. An Interactive Framework Integrating Segment Anything Model and Structure-from-Motion for Three-Dimensional Discontinuity Identification in Rock Masses. *Int. J. Min. Sci. Technol.* **2025**, *35*, 1695–1711. https://doi.org/10.1016/j.ijmst.2025.09.005.
30. Dong, Z.; Zhang, H.; Qu, Y.; Yang, S.; Yang, H.; Tang, P. Three-Dimensional Spatial Representation Method for Semantic Boundaries in Digital Outcrop Models. *Sci. Rep.* **2026**, *16*, 15170. https://doi.org/10.1038/s41598-026-45820-7.
31. Parisi, A.; Fidelibus, M.D.; Monno, V.; Parisi, M.; Parisi, N.; Specchio, V.; Spilotro, G. Le cavità antropiche di Gravina in Puglia (Bari, Puglia): Aspetti storici e geotecnici. *Opera Ipogea* **2020**, *1–2*, 263–272.
32. D'Angeli, I.M.; Lacalamita, M.; Sasso, C.; Schingaro, E.; Parise, M. Morphological and Mineralogical Characterization of Surficial Weathering on Calcarenite Rocks in the Rupestrian System of "San Michele delle Grotte" at Gravina in Puglia (Bari, Apulia). *Catena* **2022**, *216*, 106382. https://doi.org/10.1016/j.catena.2022.106382.
33. Apple Inc. iPhone 13 Pro–Technical Specifications. Available online: https://support.apple.com/en-us/111871 (accessed on 17 July 2026).
34. OpenTopography. iPhone LiDAR with Applications for the Geosciences. Available online: https://opentopography.org/blog/iphone-lidar-applications-geosciences (accessed on 17 July 2026).
35. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004.
36. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
37. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Duxbury: Pacific Grove, CA, USA, 2002.
38. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
39. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*; AAAI Press: Portland, OR, USA, 1996; pp. 226–231.
40. Jolliffe, I.T. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, USA, 2002.
41. NOAA NCEI Geomagnetic Modeling Team; British Geological Survey. World Magnetic Model 2025; NOAA National Centers for Environmental Information. 2024. Available online: https://www.ncei.noaa.gov/products/world-magnetic-model (accessed on 2 November 2026).
42. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall/CRC: New York, NY, USA, 1993.
43. Fisher, N.I. *Statistical Analysis of Circular Data*; Cambridge University Press: Cambridge, UK, 1993.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.