This work presents a valuable and timely contribution to the field of karst studies. The acquisition campaign is extensive, methodologically sound, and applied across a wide variety of cave systems in Western and Central Europe. This diversity strengthens the scientific and practical relevance of the dataset. The authors rightly highlight the potential of such data to support a range of applications, including morphological analysis, geomorphological mapping, and numerical modelling.

One of the persistent limitations of existing 3D cave surveys is that they are often conducted for a specific purpose, with limited reusability. Many datasets are eventually lost, poorly archived, or remain inaccessible to the broader community. Several studies have demonstrated that certain geometric parameters (such as network tortuosity, wall roughness…) can strongly influence physical processes, particularly flow behaviour (Albright & Springer, 2022; Kordilla et al., 2025; Peterson & Wicks, 2006). Yet these parameters are often substituted with empirical values that lack natural extrapolation. Initiatives like the one presented here offer the potential to better constrain such parameters using real-world data, and to support more robust modelling approaches in a variety of fields.

Overall, I believe this dataset represents a significant contribution and should be made available to the community. I would suggest minor revisions to improve its visibility, usability, and long-term value. The following section outlines some general criticisms and suggestions for improvement, particularly regarding the accessibility, documentation, and contextualization of the dataset.

1. Explanation of the applications and scientific value of the dataset

While the introduction provides a detailed overview of technical aspects related to data acquisition, it could benefit from a more explicit presentation of the scientific utility and relevance of the dataset. For instance, the authors could further develop how such 3D datasets can support geomorphological interpretations, hydrological modelling, or speleogenetic reconstructions, including references to recent studies that have leveraged or emphasized the need for similar data in karst science.

2. Data visibility and accessibility

Despite the quality and richness of the dataset, its visibility and long-term impact are at risk if more effective dissemination strategies are not implemented. While the complete dataset may be suitable for large-scale studies (e.g., hydrodynamic and flow regime numerical modelling), it is more realistic to expect its broader use in studies focused on a single cave. However, this presupposes that the availability and accessibility of the data are clearly highlighted and actively promoted. For that reason, two improvements seem essential:

(1) Make it possible to download data by individual cave, rather than requiring the full 200 GB dataset. <u>As it requires more free storage than I currently have, I was unable to explore the data myself, which prevented me from analyzing it properly.</u>

(2) Provide a geographically organized and interactive visualization interface. For example, low resolution .ply 3D previews or therion files could be archived in the Karst3D database (Karst 3D Team, 2019), which already includes an overview localisation map, access to Therion and/or low-resolution .ply 3D previews and links to full-resolution downloads and metadatas could be specified in Karst3D database.

3. Limited contextual information on surveyed sites

While the acquisition workflow is well described, the rationale behind the selection of specific caves remains unclear. Providing more details about the choice of surveyed sites would enhance the scientific value and reusability of the dataset. An extended version of Table A1 could include additional descriptors for each cave, such as geographical location, geological context (e.g., lithology, structural setting), conduit morphology (e.g., phreatic, vadose..), surveyed length, and notable features. This would also help users select suitable sites for comparative or targeted analyses.

4. SLAM acquisition – usage, limits.

The use of mobile SLAM-based scanning is a particularly valuable aspect of this work, given the growing adoption of this technology in underground environments. However, the manuscript lacks sufficient detail on how SLAM performance and limitations were assessed. The comparison provided for a single cave is informative, but it remains unclear whether such comparisons were systematically repeated across all sites. I strongly recommend including statistics on the residual co-registration errors between DistoX measurements and the SLAM-based scans, using the target centers as a reference.

In addition, several practical aspects of the data acquisition process are missing: What were the typical acquisition times within the caves? Were the caves fully surveyed? If not, what were the limiting factors — time constraints, passage length, physical obstacles, etc.?

While the dataset is commendably diverse in terms of morphological and geological settings, it should be noted that, as I understand without having explored the data, it primarily consists of relatively horizontal cave passages (the author mention stopping the acquisition before a 10m shaft at the test cave). As such, the inherent limitations of both static and SLAM-based scanning in more complex cave sections — such as vertical shafts, wet areas, or narrow passages — should be clearly acknowledged.

Finally, to allow the reader to closely inspect the data and visually appreciate the differences between the two sources, it would be highly beneficial to include a figure showing a side-by-side comparison of a selected area (zoomed-in portion of the point clouds), along with the corresponding mesh reconstructions from TLS and PLS.

5. Figures

Overall, the figures are of good quality and mostly self-explanatory. However, I feel that at least one additional figure is needed to allow the reader to clearly 'see' the source data. As it stands, most illustrations are zoomed-out views of entire scanned passages, which makes it difficult to appreciate the level of detail and point cloud quality at a finer scale.

I now move on to more specific comments and minor corrections:

Line 16: The acronym *LIDAR* is used here for the first time but is only defined at its second occurrence at line 35.

Line 35: ''liDAR .. is suited to the underground as it overcomes many 35 challenges inherent to light-based techniques for the acquisition of three-dimensional point clouds (Giordan et al., 2021)''

Although the argument is valid, and could even be strengthened by mentioning the faster acquisition and post-processing times compared to visual methods, the use of this citation appears somewhat inconsistent with the conclusions of the referenced paper. In fact, Giordan et al., (2021) highlight that visual methods, particularly Structure from Motion (SfM), offer a favourable compromise in terms of accuracy, feasibility, and cost-effectiveness for 3D surveys of complex natural caves. They emphasize that SfM constitutes a strong alternative to LiDAR, rather than being subordinate to it.

Line 65: The survey method (TLS, PLS) could be added in the annex table

Line 71: ''as well as the methods we used for scanning the cave and **processing and post-processing** the point cloud dataset''.

Seams redundant.

Section 2.1.1 For the TLS acquisition, what was the registration method used (spheres or best fit?)

Line 90: ''To achieve this, the algorithm uses regular updates to the scanner position by 1) using the device's Internal Motion Unit (IMU) and 2) by triangulating between recognisable point features **(Figure 3)''**

Figure 3 shows a SLAM unit being used in a cave passage but is not really an illustration of that particular sentence regarding the SLAM method.

Line 94: The authors split the conduit into several overlapping acquisitions (scenes) acquired separately. It would be helpful if they could specify the criteria used for this splitting, such as whether it is based on time, length, or other factors. Furthermore, an explanation of the necessity/difficluty to perform loops and back-and-forth acquisitions would strengthen the understanding of the methodology.

Line 98: The sentence states that the conduit sections were scanned with a 15–35% overlap for subsequent co-registration. I understand this to mean that there is a return path of several meters or even tens of meters overlapping the previous section. However, the exact meaning and precision of these percentages in the context of field acquisition remain unclear—are these overlap values measured in real time during acquisition or calculated afterward?

Line 114-115: The two sentences could be combined into a single, clearer sentence to improve readability.

Line 116, 2.2 Georeferencing: I understand that the georeferencing was performed for all the caves, with laminated scan target as show in Fig 3 a, measured with a DistoX. It is not clear to me if the authors selected the closest point to each target center (by using the intensisty/illuminance return to clearly see the target black and white pattern?) or used another method that is less dependent to the scanning density on those targets (=acquisition distance). The authors later give some statistics about the rigid transformation for the test cave based on the splay shots but it would be helpful for the reader to give additional stats (Therion loop closure error and at least min, max and average DistoX/laser residual errors on targets), for the test cave but maybe even for the overall dataset.

Regarding the georeferencing itself, unless I missed something, it is not mentioned whether the data are shared as georeferenced point clouds and meshes, or in local coordinates with the transformation to real world parameters provided separately (e.g., in metadata). This distinction is important, as many 3D software tools do not handle large coordinate values well, and georeferenced files can be significantly larger. Clarifying this aspect in the manuscript would be useful for potential users of the dataset. My personal opinion about this is that providing data in local coordinates with georeferencing in metadata is best.

Line 136: The noise and related cleanings are well explained in this section but, again, a figure with a zoomed portion of the point cloud would help the reader to visualise the raw data for both techniques, as well as the noise cleaning and meshing.

Line 159: ''The point cloud generated by the BLK2GO device has a specific 3D structure **made of criss-crossing point trails** which originates from the scanner movement during a survey.'' Same remark, I would have appreciated a visual example of this characteristic pattern within the paper itself, rather than having to download over 200 GB of data to observe it.

Line 204: ''Intuitively, a non-rigid cloth is draped over the upturned point cloud, and points touching the cloth are labelled as ground category''.

This approach may work well for relatively simple topographies, but it could lead to misclassification in cases where the geometry is more complex or multivalued. Ex: In the presence of a big boulder with lower face overhanging the ''true ground'', parts of the boulder will not be labelled as ground. If the authors have considered such limitations or implemented specific strategies to address them, it would be helpful to mention it.

Line 213: The centreline extraction protocol is clearly presented, but the motivation for producing such data should be more explicitly stated, ideally in the introduction, and supported by relevant references (e.g., Collon et al., 2017; Jouves et al., 2017). This relates to my general comment (1), suggesting that the introduction would benefit from additional bibliography on the use and scientific value of such cave survey datasets.

Figure 6: Subfigures a) and b) appear somewhat redundant. Applying normal shading to subfigure a) could improve the visualization by better conveying surface orientation. In contrast, the illuminance-coloured point cloud in b) does not seem to add substantial additional information. Again, I would instead suggest replacing it with a zoomed-in view of the point cloud or mesh to provide more detailed insight into the data quality and geometry.

Line 276: parenthesis missing

Line 293: ''The scan was carried out in May 2024, in 24 different acquisitions assembled together, totalling approximately 400 linear metres of passage, from the entrance inwards, and stopping at a 10 m pit.''

Including acquisition time would help the reader assess the efficiency of the SLAM method in such setting.

Table 3: ''Visual Archeology Terrain blend parameters for the cave terrain shading**''**. I suggest to add the citation here too: Relief Visualization toolbox (Kokalj et al., 2016)

Line 303 305: ''The splay shots provide an independent way to check that no drift or distortion has occurred during the point cloud assembly. After georeferencing the cave point cloud using the pair-wise registration method of Arun et al. (1987) on specific targets, we used CloudCompare to…''

One could argue that analysing the residual errors on the targets after alignment — by comparing DistoX stations to the closest corresponding target centers in the scan — would provide a more reliable basis for dataset comparison than a cloud-to-cloud comparison with the splay shot ''point cloud''. The latter is extremely low-resolution, and a splay point may lie near the laser scan purely by coincidence, without reflecting an actual spatial match. At the very least, providing statistics on the residuals at the target locations would allow verification of whether the same ~12 cm error is observed. (see same remark above for line 116).

Figure 7: Same remark as above regarding the reliability of the comparison method. Additionally, the colourbar is missing for the splay shot-coloured points, which makes interpretation difficult.

Line 308: ''on Fig.7'' could be replaced by (Fig. 7).

Line 311 312: ''We conclude that for the example Markov Spodmol, both survey techniques yield consistent results with respect to cave geometry at the decimetre to metre scale.''

Same remark as for line 116: if you provide statistics on the DistoX-to-laser alignment at the target centers for all caves, this would support extending the validity of the statement to the entire dataset.

Another general remark here: It is not clearly stated which method — TLS, SLAM, or DistoX — is the most accurate in terms of absolute positioning. One would intuitively expect TLS to be the most precise, followed by DistoX and then SLAM, but some clarification or reference would help support this assumption.

Line 317: "using the mesh sculpting tool **Blender** to remove…".

Consider rephrasing to 'using the mesh sculpting tool *in* Blender' to avoid suggesting that Blender is solely a sculpting tool.

Figure 9: The image appears to show a colour-coded and segmented (ground) point cloud in orthographic view, rather than a true DEM. The legend and the naming of the station symbols are somewhat unclear: What is the distinction between 'scan target' and 'marked'? Additionally, the red circles mentioned in the legend do not seem to be visible in the figure itself.

Line 369: ''Finally, this work shows that the ease of use of mobile scanners allows for **fast acquisition** of large datasets.''

However, the text does not provide any quantitative or comparative information to support how fast the acquisition actually is.