



EGUsphere, referee comment RC2  
<https://doi.org/10.5194/egusphere-2022-633-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on egusphere-2022-633**

Thomas Blenkinsop (Referee)

---

Referee comment on "Clustering has a meaning: optimization of angular similarity to detect 3D geometric anomalies in geological terrains" by Michał P. Michalak et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-633-RC2>, 2022

---

### **Clustering has a meaning: optimization of angular similarity to detect 3D geometric anomalies in geological terrains**

This is a difficult paper to read, because it contains a lot of jargon about geometry, and because of vague general statements, some of which are unnecessary (e.g. "dip angle is not capable of showing the dip direction of faults and vice-versa" and "Geology is considered to be a subjective science (Curtis, 2012)"). A further problem for understanding the paper is that some of the methods section is couched in the technical language of the CGAL library. This is unhelpful to the general reader, and needs to be explained in simple terms.

One of the main conclusions, that applying clustering methods to normal vectors and dip direction vectors from the same data set results in different interpretations of the structure (Fig. 15), seems unlikely to be correct. There is no material difference between the geometrical significance and information contained in a normal vector compared to a dip direction vector. If there is a difference in the outcome of the clustering methods, that must be an artefact of the way the methods have been applied to each data set.

Another main conclusion is that optimisation methods must be applied to investigate clustering. This is relatively trivial: any clustering algorithm requires a similarity index, and the one used here (cosine distance) is a standard metric for assessing orientation differences. Further to the previous point, this metric should not result in significant differences between normal and dip direction vectors, because the cosine distance between two normal vectors must be the same as the cosine difference between the two dip vectors of the same surface.

There is some discussion about anomalous results:

“The above effect could be explained by several competitive hypotheses. For example, the fault plane could have been drilled, 365 thus broadening the zone of triangles genetically related to the fault (Michalak et al., 2021). Assuming the tectonic origin of the related structures, it can be hypothesized that fault drags on the hanging wall contribute to subsidiary elevation differences that must be consumed by nearby triangles. It could also be argued that an unusual lowering of the contact surface is due to a deformation zone composed of many smaller faults. Another hypothesis could be that the related feature is not a fault but rather a sedimentary slope, which would explain the gradual lowering of the contact surface.”

Such hypotheses are useful, but would be better illustrated with specific examples and some reasoning about which is the preferred hypothesis.

The determination of the optimum number of clusters is explained in Figure 7, but the results sections shows results from 2, 3 and 4 numbers of clusters. This is unnecessary: only the optimum results should be shown.

The figures could be substantially improved. The use of such a dark background does not help (e.g. Fig. 6c). In most cases the grid is the most dominant and least important aspect of the maps, obscuring the detail of the clustering. The stereoplots are not explained in the figure captions.

Please also note the supplement to this comment:

<https://egusphere.copernicus.org/preprints/2022/egusphere-2022-633/egusphere-2022-633-RC2-supplement.pdf>