This paper presents a method using support vector machines to identify faults cutting stratigraphic horizons when those horizons are represented by a triangular mesh created from scattered borehole data. The method is demonstrated on a case study and is compared to a clustering-based method developed in a previous paper by the same first author. The paper has undergone a previous round of review, in which unclear language emerged as a major issue, which the present revision seeks to address.

The method is novel to the best of my knowledge and is likely to be of interest to the geoscientific modelling community. The previous revisions have improved the clarity of the manuscript, in particular in regard to the usage of the word "terrains." However, I think that the paper still requires major revisions before it is ready for publication. Most of my comments regard the clarity of the text and figures, but I also have some questions about the use of seemingly redundant features in the classification. My detailed comments are below.

**Title**

The choice of "slopes" to replace "terrains" here and in the text is an improvement, but it still sounds odd to me. I am more likely to think of a slope as a topographic feature. I think a word such as "interfaces," "contacts," or "horizons" would sound better. I see that all three of these words are already used in some places in the text. (The example seems to be with a stratigraphic horizon, although I can see that the method could be applied to any planar interface.) If it is necessary to convey the "preferred orientation" part, perhaps this could be done by including a word such as "planar," "homoclinal," or "dipping." (For instance, one could say "homoclinal interfaces".)

**1 Introduction**

To help show the value of the study, I suggest including a sentence about the kinds of practical applications that this kind of fault identification would be useful for.

This section would benefit from a clearer statement of the problem. I think it could be something like this: Geological models are often created by interpolation of scattered borehole data. Because of the localized nature of boreholes, faults not intersected by them will be missed, and interpolation will create horizons that appear continuous across the faults. This paper proposes a way to identify the presence of faults in this situation.

Line 29: The comparison to seismic data is good, but I don't think it is clear what "subsurface topographic data such as subsurface slopes" means. I think the key difference is that this method works with scattered data, so I suggest phrasing this as a comparison of scattered data vs. seismic images.

Figure 2: The meaning of the line labeled "fault" and referred to as a "fault line" is unclear. This appears to be a 3D model, and a fault is a plane (or surface) in three dimensions. If this is meant to show the fault, it should be shown as a plane. If this is meant to show the line of intersection between the fault and the "slope," then it should be labeled and described as such.

**2 Background**

Since the paper involves significant comparisons to Michalak et al. (2022) and uses the same example as them, I think it would be helpful to divide this section into two paragraphs: One discussing other machine learning applications more generally, and the other specifically presenting Michalak et al. (2022) as a preceding work that this paper will build on.

**3 Methods**

Figure 4: I find the subfigures on the left of this figure confusing. I assume they correspond to the steps of the flow chart on the right, but it would help to state that explicitly in the caption and perhaps connect them within the figure. Also, it would help to give them letters and refer to those letters in the caption.

Also in Figure 4: Would training the model also be a step in the workflow? Or is that meant to be included in the "evaluation for synthetic data" step? If the latter, perhaps it could be "training and evaluation with synthetic data."

Lines 151-152: The use of normal and dip vectors seems redundant. Either one can be used alone to define the orientation of the plane, and either one can be derived from the other. Please either explain why both are needed or try doing the analysis using just one of them.

Lines 151-159: The different distance types also seem somewhat redundant. While they will give different numbers, they are all closely related. For instance, the square of the Euclidean distance is proportional to the cosine distance as discussed by Michalak et al. (2022). Please explain why three different distance measures are used, rather than just one as in Michalak et al. (2022).

Line 153: It would be helpful to give the variable names after the distance types: "angular distance ($d_a$), Euclidean distance ($d_e$), and cosine distance ($d_c$).

Line 161: Instead of "middle part of Fig. 4", I think it would be clearer to give this subfigure a letter (such as Fig. 4b) and refer to that.

Line 167-172: This is, indeed, a good reason to use vector representations of orientations, rather than dip and dip direction. However, this discussion doesn't seem to fit with the preceding part of the paragraph. I would suggest either making it a separate paragraph or moving it to where you first talk about using vectors.

*Tables 2 and 3*: Dip angle and dip direction in these tables appear to be in degrees, so they should have the degree symbol (°).

*Figure 5:* What is the meaning of the green box and letter ε in part c? Also, in the caption for part f, I think that "fault line" should be "fault plane" (see also my comment on Figure 2).

**4 Geological Setting**

I don't think it is necessary to divide this section into two subsections when they are only one paragraph each.

Lines 232-233: Saying the KSH is a “geological unit” sounds to me like it is a specific rock unit within the stratigraphy. But given that it is a homocline, I assume it is a structure. If that is correct, I suggest changing this to “a geological structure.” Also, if the homocline forms one side of the synclinorium, I would suggest changing “a slope” to “a limb” in Line 232.

Line 234: It would be helpful to give a range of values to quantify “low angles.”

**5 Results**

Lines 272-276: For the grid search optimization, it would be helpful to state the range of values tested and the grid spacing for each parameter.

Lines 291-293: I suggest moving the information about the specific horizon to Section 4 to go with the rest of the discussion about the case study.

**6 Discussion**

Line 329: Is the “unsupervised method” the one from Michalak et al. (2022)? It would be clearer if that were stated explicitly, such as by saying “the unsupervised method of Michalak et al. (2022)…”

Section 6.4: It appears to me that one assumption of the method is that the stratigraphy is homoclinal. That is a major modelling assumption that should be discussed.

Line 373: If it is “not clear” how the model would classify the structure in Figure 8, couldn’t the model be tested on the example in Figure 8 to find out?