Dear Editor,

Please find below my report on the manuscript “Anatomy of a fumarole field; drone remote sensing and petrological approaches reveal the degassing and alteration structure at La Fossa cone, Vulcano Island, Italy” (egusphere-2023-1692) by Daniel Müller et al.

The present manuscript presents a combination of visible and thermal imagery, mineral and geochemical analyses of rock samples and CO2 soil degassing data with the aim of classifying and quantifying the alteration and degassing structures at the Vulcano Fossa volcano. The manuscript uses a number of novel techniques, with the main emphasis on the classification of thermal and visible images (orthophotos), similar to previously published work (Müller et al., 2021), with the addition of mineralogical and geochemical analyses of hydrothermally altered samples. Whilst this approach is promising, and the results given are interesting, I think that some of the methods employed need much more extensive descriptions. Some of the techniques employed may also be flawed or, at best, misunderstood. Lastly, I do not feel that the authorship is currently up to standard for publication, though the written English is fine in itself. Hence it is my opinion that the manuscript requires major revisions before it could be considered for publication. I provide some suggestions that the authors may chose to follow. Given that the revisions may be substatial, a review of the revised manuscript may be necessary. In this case I would be happy to act as referee should you require my services.

Best regards
David Jessop

PCA and image classification

Probably my major concern regarding the manuscript concerns the lack of description of the “Principal Component Analysis”, (PCA) and the image classifications. My reasons are two fold:

1. The PCA and image classification processes, provided within the ArcGIS propitiatory software, are “black boxes” with no description of what is happening under the hood. As written, it is not easy to understand how these process work, or what data they provide. The interested (and reasonably competent) reader should be able to reproduce the results of this study and I don’t think that is currently the case.
2. In order to reproduce these results, the interested reader would require access to the proprietary ArgGIS software which means paying 500-700€. I don’t think that this is fair. Whilst this may not be an issue for wealthy universities, it could block researchers with less access to funding from utilising the approaches employed in the present study. A proper description of the methods could allow such researchers to look for cheaper of even open-source alternatives.

PCA consists of taking a multi-dimensional dataset and finding the orthonormal basis vector space that describes the data whilst minimising the variance along each vector basis (component). This is achieved by calculating by projection of the data onto a set of orthogonal axes where the variance of each data set is represented by the eigenvalues of the data. These eigenvalues are the principal components. This is usually achieved using a reduced singular value decomposition (SVD) which produces the eigenvalues of the dataset. PCA then takes the list of ordered eigenvalues which are typically used to perform dimensional reduction in high-dimensional data sets. This consists of rejecting any components that do not contribute significantly to the overall variance of the data. In the present study this is applied to an RGB image (i.e. 3D) and the authors take 3 principal components, so there is no dimensional reduction and hence the "PCA" is kind of redundant.

Regardless of how we name this process, it is unclear what the ArcGIS algorithm produces in the "Principal Component" band images (cf. L220-223) – PCA gives only the variance. Please indicate how this information is used to transform the RGB image.

The image classification, named as "unsupervised classification", process is poorly documented. Indeed unsupervised classification is a blanket term for a multitude of different families of algorithms so, to be able to reproduce these results, one would have to know which algorithm was chosen and why. The choice of 32 classes seems to be completely arbitrary (50 were used in Müller et al., 2021) and, furthermore, the individual classes are then regrouped (see fig. A1 for example). Could a smaller set of classes (say 4) not have been used to obtain similar results? I strongly urge the authors to justify their choice.

Whilst the authors refer to previous work (Müller et al., 2021) as a source for their methods, but neither the PCA nor the classification strategies are sufficiently well described in that work either. Without having to detail the algorithms in their entirety, the authors should please sufficiently explain their methods in the present manuscript so that they can be followed with the aim of reproducing their results.

Thermal image processing

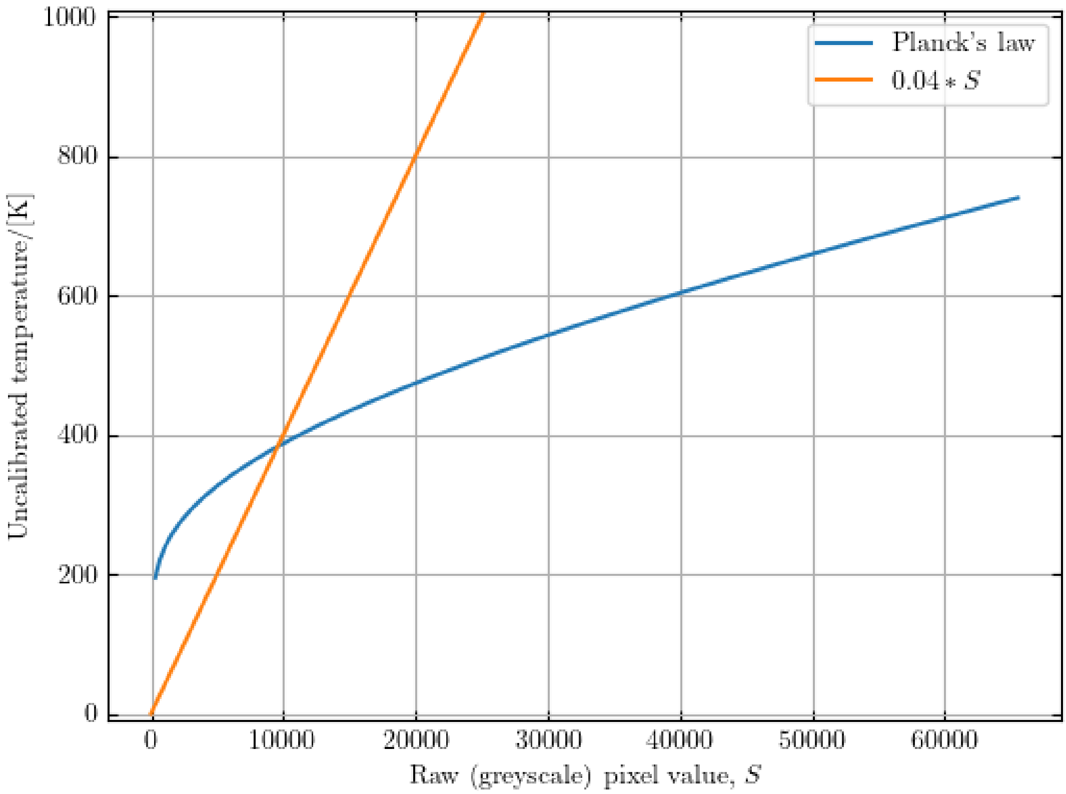

Fig. 1: "Brightness" temperature predicted by Planck's law compared to the scaling proposed by the authors (eq. 1). The two curves cross at a temperature of 383 K (about 90°C).

The authors state that they produce temperature maps from the 16-bit radiometric greyscale orthophoto using a linear mapping given by their eq. 1 using the radiometric resolution as a scaling factor. Owing to the non-linear behaviour of IR sensors this would seem unlikely to hold for more than a very limited range of greyscale values. Furthermore, it is typically necessary to use Planck's law which predicts the "brightness" temperature of an object from the intensity of incoming radiation registered by the sensor (i.e. radiometric value). By way of illustration, I have produced the above graphic (Fig. 1). Here we see that there is only one point of intersection for the two curves and the predicted temperatures can be drastically different. That said, I am not familiar with the FLIR Tau camera as used in this study and do not have access to the radiometric conversion factors necessary to correctly plot the curve for this camera. If it is like many other FLIR and other IR cameras that I have used, this information can be found in the EXIF (image metadata) which can be readily extracted using the ExifTool software, for example. However, the authors should check their data and any calculations that depend on the temperature. Some of the stated temperatures and thresholds (e.g. 40 °C for the "high-temperature fumaroles") are rather low given the vent temperatures recorded in other works (vent temperatures are well in excess of 100°C, cf. Diliberto, 2017; Mannini et al., 2019; Diliberto, 2021).

Authorship and increasing the scope of the paper

A quick survey of the the first dozen or so references showed that they were cited once in the introduction and nowhere else in the manuscript. The discussion contains only six references, and is often a rehash of the results section rather than a forum for putting these results into a fuller context and comparing them to previous works on Vulcano and, potentially, other volcanoes. Several references (e.g. Chiodini et al., 1996; Chiodini et al., 2005; Harris et al., 2009; Mannini et al., 2019) have made estimations of degassing and/or heat budgets with Harris et al. (2009), in particular, having made detailed descriptions of the fumarole field. I find it strange that the authors have not chosen to make the comparison with these works, particularly given the ongoing and unrest at Vulcano with the recent well-documented paroxysms. Fig. 11 identifies a structure with an increased radiant density and labels it as "new fumarole complex in 2021", but this is not discussed anywhere. This would be very important information for assessing the activity at this volcano. Curiously, Fig. 10 hints at a heat budget having being calculated, but this is not discussed in the manuscript. I note also that the area of the alteration zone (ALTZ), that is the area affected by hydrothermal activity, is given as 70 000 m^2 (note typo "770 000 m^2" on L301) which is very close to the "diffuse heated area" of 63 000 m^2 calculated by Mannini et al. (2019) using approximately contemporaneous data. Of course, I have my own professional biases in mentioning this, but the authors have already done the work so it is suprising that it is only mentioned in passing.

Concerning alteration of the edifice, there have been a number of studies in recent years trying to ascribe thermal properties of volcanic rocks to hydrothermally altered samples (typically andesite). These results may be interesting to discuss. See Heap et al., (2022) in particular. Section 5.3 briefly mentions the role of alteration and permeability:

"Relative gas flux values measured within unit c are lower than observed for units a and b, for instance. This might be a consequence of the dynamics of hydrothermal alteration and indicate permeability reduction or sealing processes due to the advanced state of alteration like proposed by Heap et al., 2019." (L592-594)

Furthermore, a sequence of alternating high and low permeability zones are identified in Fig. 11, but each result is only discussed independently and there is no real synthesis of the large and important set of results. This is one discussion point, in particular, that is really important for assessing volcanic hazard and I find it frustrating that this point has not been fully persued.

A minima, the figure captions should allow the reader to understand the figure in isolation and, currently, this is not the case for several of the figures. The main cuprits are

Fig 2 – give locations of each photo, also show these locations on Fig 1B.

3 – describe the grey blocks. Why are there two blocks for Remote Sensing?

8 – what are "RGB values", as only one value is given here. Is this for one band in particular or an average? Why are RGB values being used rather than one of the "PCA" image bands?

9 – What is the absicssa in this figure? How is it that Transect A blends into Transect B? I think it would be worthwhile to combine this figure with Fig. 7, particularly as one requires the locations of the transects to understand what is going on in this figure.

Other suggestions for the figure captions are:

4 – D, please give the thermal thresholds for each class.

5 – Please label your axes. The titles "Red channel values" should probably be axes labels.

6 – The bars are very confusing and there is no scale given for the gas concentrations. Instead maybe use a colour map where colour/intensity corresponds to value? What do the black dots mean? What is the direction of measurement for the "Distance from ALTZ boundary", i.e. are positive values inside and distance is taken normal to some boundary? If so, which and how determined?

7 – (and elsewhere) it would be useful to have the definitions of ALT, AMT, LTZ, XRD, XRF etc. recalled here. Capitalise XRF (L694)

10 – Please state (either here or in the main text) how Rcum is calculated. Please also use proper representations of units in your axes labels (e.g. "W/sqm")

11 – Please label the abscissa in D and give units. Please also use proper units for ordonate label ("kW/sqm").

Generally, there is no description of the figures in the text of the style "In Figure X we show…", rather the figures are referred to *en passant* (e.g. "pixels of Type 1 to 4 surface show a general increase of mean pixel temperatures from Type 4 to Type 1 surface by an average of 2 degrees (Figure 5).", L330-332). This may be a deliberate stylistic choice by the authors but it this makes it harder still to understand the figures, and leads to the impression that they are not very important, particularly given the paucity of the descriptions in the captions.