Response to Reviewer 3 comments

Interactive comments on “Volcanic cloud detection using Sentinel-3 satellite data by means of neural networks: the Raikoke 2019 eruption test case” by Petracca et al.

We would like to thank the Reviewer for her/his constructive comments and suggestions, which have improved the manuscript.
Please find our replies to each comment below. Referee comments are reported in black. Our replies are given in red.

The Authors present a neural network technique to detect volcanic ash clouds by combining visible and thermal infrared channels of moderate resolution spectroradiometers. A neural network trained on MODIS imagery acquired during the Eyjafjallajökull eruption in 2010 is applied to two pairs (nadir and oblique view) of SLSTR images of the Raikoke eruption that occurred in 2019.

The neural network method is compared to the classic brightness temperature difference (BTD) method and the accuracy of the two methods is evaluated against manually classified pixels. The results show a reasonable performance of the NN method in detecting ash clouds in nadir view, whereas I have some perplexities about its performance in the oblique view, as it seems to me that the NN underperforms for a fairly thick ash plume which I would expect to be easily detectable.

All in all, I think the paper can be published with minor revisions, although I recommend careful proofreading by a native English speaker, as the quality of the written English does not look impeccable to me.

A native English speaker proofread the revised version of the paper.

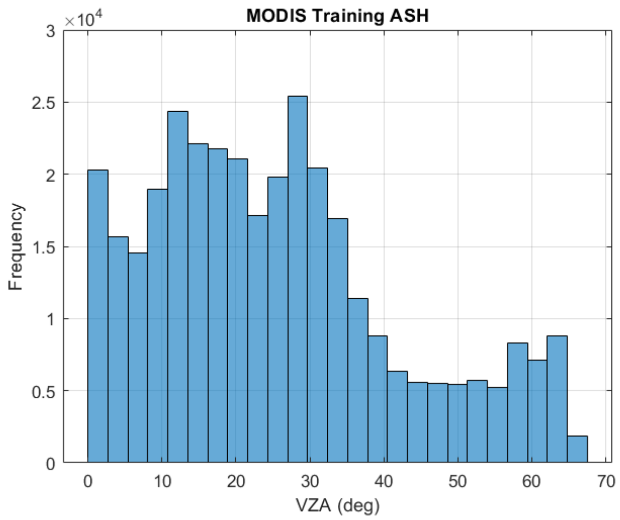
MAIN COMMENT

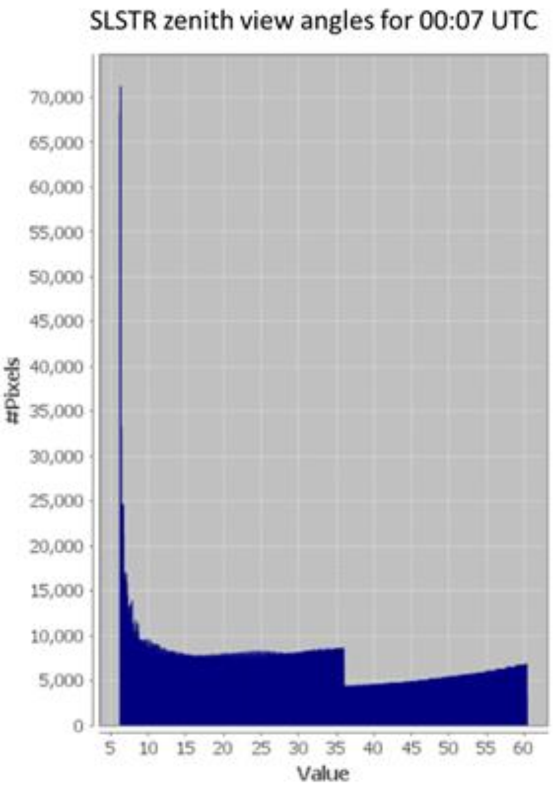
I think that the extension of the NN approach to oblique view needs further investigation.
What are the typical values of the viewing angles sampled in the oblique view and how do they compare to those of the nadir view? If the air mass sampled in the oblique view is much bigger than that sampled in nadir view, the difference in the slant optical depth may translate to a noticeable difference in top-of-atmosphere signal levels. Furthermore, if there is a large difference in the observed scattering angles you may be also sampling different ranges of (weather and ash) cloud phase functions, which also may lead to significant differences in the signal levels in VIS/NIR channels.
In this case, it looks far from obvious to me that the NN can still be applied reliably to oblique view situations that are probably not covered in the training set.
Therefore, I would recommend studying the sensitivity of the NN detection to the observation angle by generating synthetic top-of-atmosphere spectra of VIS/NIR radiance and thermal brightness temperature for a typical liquid water, ice and ash cloud. In my opinion, the results presented in the paper do not allow to draw reliable conclusions on the robustness of the NN method to off-nadir observations.

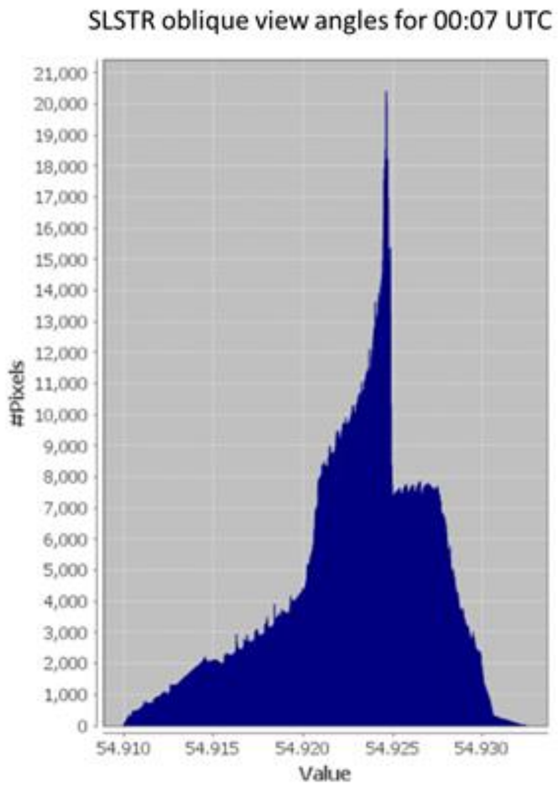
We thank the reviewer for all her/his interesting suggestions regarding the oblique view application.

Indeed, in case of the proposed work our intention was to preliminarily show an additional point with the idea to go in deep in future developments. For this reason, we moved the application of the NN model to the oblique view data in the conclusions section. As an anticipation we think it is interesting to show how the main features of the classification map (represented in Figure 7) obtained using a NN model trained only on near nadir view acquired products and used for classifying oblique view data are mostly conserved.

The complexity of the problem also involves the training dataset generation. In fact, below we report the histogram of the View Zenith Angles (VZA) used for MODIS Training (9 images) related to the pixels considered as ash. The VZAs greater than 40 degrees are undersampled with respect to the others and this could probably have an impact on the results of the off-nadir SLSTR view (SLSTR zenith angle in the oblique view is about 55°, as reported in figure below).







Finally, we modified the last sentence of the abstract to: "Finally, the results show that the NN developed for the SLSTR nadir view can produce reasonable image classification also for the SLSTR oblique view."

DETAILED COMMENTS

- P1, L15, spaceborne sensors acquired data -> satellite data

Done

- P1, L19, The classification of the clouds and of the other surfaces -> A classifications of clouds and other surfaces

Done

- P1, L22, foster the robustness of the approach, which allows overcoming -> allow to extend the approach to SLSTR, thereby overcoming

We changed to "...the robustness of the approach, which improves on the use of SLSTR products..."

- P2, L42, to detect the volcanic cloud -> to detect volcanic clouds

Done

- P2, L43, you can remove "problem" after detection

Done

- P2, L43, lies on -> relies on

Done

- P2, L44. There is no such thing as "water vapour clouds". I guess you mean "liquid water clouds"

Yes, we do mean that, but there are also “clouds of water vapour”. However, we changed to “liquid water clouds”.

- P2, L49, region -> regions

Done

- P2, L59, procedures described -> described procedures. Plus, is "among" really what you mean, or do you mean "in addition to"? Does "described" refer to Prata et al. (2001b) and Corradini et al. (2008,2009)?

Yes, thank you. We mean “in addition to” and “described” refers to the procedures mentioned in the introduction, including those in Prata et al. (2001b) and Corradini et al. (2008,2009). Now the text has been improved.

- P3, L70, statistical -> statistically

Done

- P3, L71. Is this real time capability really an advantage of the NN approach? Isn't the BTD method also in near real time, given that it involves taking a difference? Furthermore, in an emergency scenario is there really such a big advantage in correctly detecting a few more ash pixels than the BTD method?

The problem is not strictly related to the computation time of BTD which is actually very fast, but to the reliability and the time consumption associated to the choice of the threshold to be used, which is based on a subjective interpretation. Indeed, using simply BTD < 0 °C (as in standard procedure) not always gives good results. The choice of the BTD threshold needs more time (Radiative Transfer Model simulation) and the presence of an operator. We can say that the NN approach, keeping the operation fast, can be more reliable and objective compared with the BTD method in general.

- P3, L85-86, either "a vertically ascending cloud" or "vertically ascending clouds"

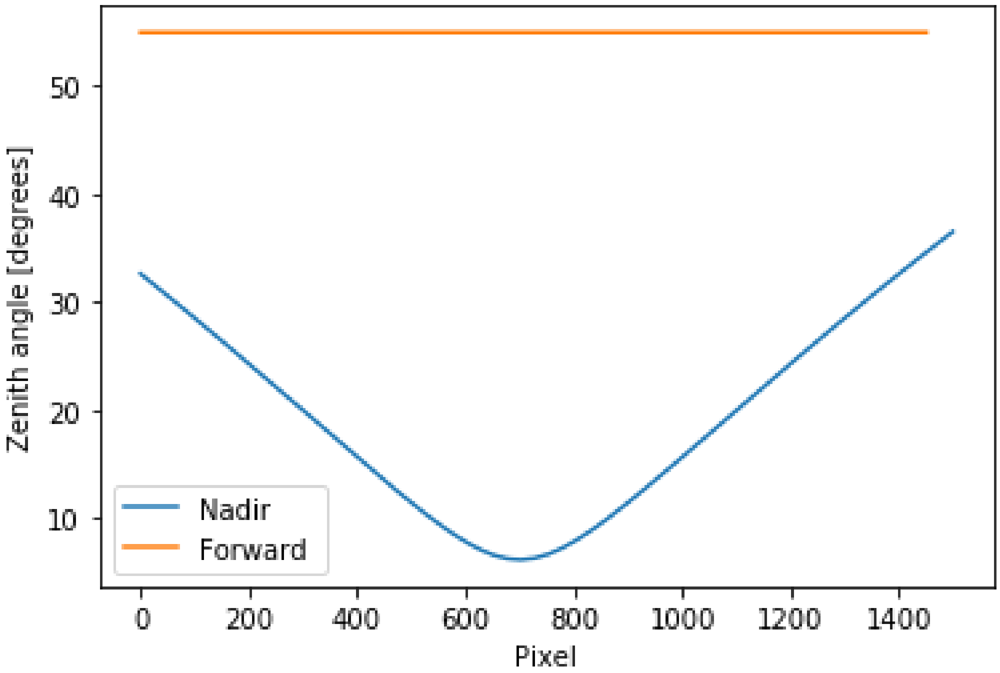
Done

- P4, L103, "water vapour" -> "liquid water"

Done

- P4, L107. At what angles does SLTR dual view observe?

SLSTR dual view observes at many angles. In general, the zenith angle in the oblique view is about 55° and the nadir view angles range as shown in the following plot (see also the plots in the reply to the main comment).



- P4, L109-110. I don't understand the use of "since" here. What do you mean when you say that the feasibility of the method was confirmed for high latitudes "since" your study area is at medium-high latitudes.
We now improved the understandability of the sentence removing "since our study area is located in medium-high latitudes" as it was unnecessary.

-P8, L159-160. How is each percentage in the confusion matrix computed? Furthermore, overall accuracy is not a particularly informative parameter. Given that the main focus is on ash, it may be useful to provide statistics on the task of ash detection (probability of detection, false alarm ratio, critical success index).

The accuracy percentages in the confusion matrix (Figure 4) are computed according to (Fawcett, 2006) and those values are related to the training phase with MODIS data. To evaluate the performances of the trained NN model for classifying the SLSTR products we inserted more informative indexes in Table 4 (and Table 5).

Ref:

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

-P8, L166. What do you mean by "commission" and "omission" errors? I guess one is "false detection" and the other is "missed detection", but it is not clear which is which.

“Commission” and “omission” were changed with “False positives (false detection) and false negatives (missed detection)..”.

-P9, Figure 5. If I look at panel a, it seems to me that on the edges of the plume there are quite a few pixels that the NN classifies as "cloud ice". Do you have an idea why this happens?

This could be due to pixel heterogeneity (some parts of the pixel are ash and some parts may be ice), or maybe there could be some cloud ice at the edges.

-P11, L186, emphasizes -> shows

Done

-P11, L198-199, "even if some pixels are misclassified as ash on land". For such a thick ash cloud I would indeed expect that there is hardly any information in the signal to distinguish ash over land from ash over sea or cloud. Does it really make sense to introduce such a fine distinction between ash classes? What do you gain from that?

We suppose that since we know where the land is, it makes logical sense to have two classes. In case of very thick ash cloud could be not relevant to know if it’s “ash on land” or “ash on sea” or “ash on cloud”. On the other hand, in other cases (semi-transparent or thin ash cloud), the differentiation in these 3 classes can be useful for many retrieval procedures.

-P11, L199, less false positives -> fewer false positives. On top of that, are they really false positives? Doesn't the BTD detect fewer ash pixels compared to the NN?

We agree on this point where actually we are comparing two methods, neither of which is considered as “truth”. We modified the sentence to:

“.. the NN algorithm is able to detect a wide volcanic cloud area and more ash, especially in the opaque regions, compared to the BTD approach”.

-P11, L206, water vapour cloud -> liquid water cloud

Done

-P11, L213, aerial trails -> aircraft condensation trails

We changed to “aircraft contrails”.

-P11, L214. What causes the BTD method to give false positives over contrails?

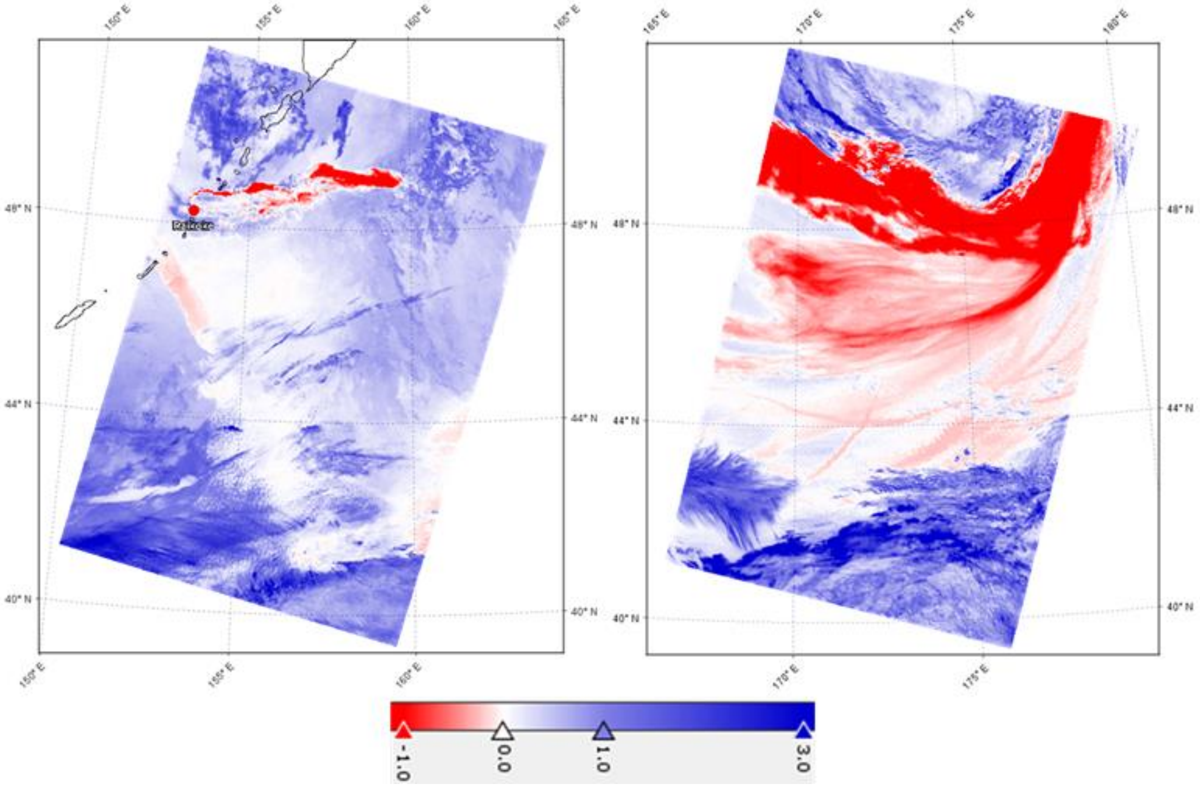
It shouldn’t. We would expect contrails to have a positive BTD (that is the 11-12 µm difference should be positive). It could happen if the ash is below the contrail and then the contrail might look like ash, or could be related to thermal contrast, and perhaps noise, pixel heterogeneity and viewing angle effects. In general, the broader question of false positives needs a deeper discussion.

Here a reference related to pitfalls with the BTD approach:
Prata, F. Bluth, G., Rose, W. I., Schneider, D. and A. Tupper (2001). Comments on “Failures in detecting volcanic ash from a satellite-based technique”., 78(3), 341–346. doi:10.1016/s0034-4257(01)00231-0

-P11, L218, "produces good results". I would say "reasonable". The ash cloud looks so thin here that I doubt you have a very good reference to compare your results against. How does the BTD approach perform for this image?

The role of this part has been completely transformed. Now is it reported in the Conclusions as a very preliminary result encouraging full dedicated studies addressing this topic.

However, we report below a figure showing the BTD map for the S3/SLSTR oblique view products for 00:07 UTC (left panel) and 23:01 UTC (right panel).



-P12, Fig. 7(a). Here the NN seems to detect a much smaller portion of the plume compared to what happens in the nadir image. Interestingly, a large fraction of the ash cloud is again classified as cloud ice (see my previous comment about Fig. 5).
Do you have any explanation for this apparently systematic tendency to confuse ash with cloud ice? Again, how does the BTD approach perform for this case?

As in the reply to the previous comment, the role of this part has been completely transformed. Now is it reported in the Conclusions as a very preliminary result encouraging full dedicated studies addressing this topic.

As in the reply to the comment P9 about Fig. 5, this could be due to pixel heterogeneity (some parts of the pixel are ash and some parts may be ice), or maybe there could be some cloud ice at the edges.

Please see the figure attached to the previous comment for the BTD approach for the image.

-P12, L220, "this is due to the opacity of the volcanic cloud". Why is the opacity of the volcanic cloud a problem for the NN detection? I would expect a more opaque cloud to provide a better contrast against weather clouds.
Even visually, the plume in image (a) looks easier to detect than the faint plume in image (c).

An opaquer volcanic cloud could be easier to detect in the RGB bands, but this does not mean that it is easier to be discriminated from other species, in particular meteorological clouds, in the other spectral channels, which are used as input to the neural network. While, usually, the information in these other bands is crucial to resolve ambiguities, in our opinion this is a rather anomalous case (SLSTR image with volcanic cloud particularly thick and wide distribution of meteorological clouds) where the information coming from the infrared may generate some confusion in the NN output. In fact, as in the figure of the reply to comment P11 (L218), the BTD approach has a similar issue. The NN training dataset relies also on BTD and this also has an impact on the resulting classification.

-P13, L249, matching -> agreement

Done

- P14, Table 4. How would such a table compare to a similar one for BTD vs MPM?

We added in Table 4 the classification metrics derived from the comparison BTD<0°C vs MPM in addition to those obtained from the comparison NN vs MPM.

- P16, L286, includes also -> also includes

Done