RC1: Anonymous Referee #1, 04 Sep 2023
General comments

This paper describes a new measurement technique for identifying pollen grains that combines holography and fluorescence. Overall, the paper reports methods and laboratory results that will be useful to the pollen observation community, although with the caveat that this has only been tested and trained in the lab and has not yet been used on ambient samples. It would be helpful to include something to this effect (e.g, laboratory only, needs to be tested in ambient air) in the abstract, as it only is raised briefly in the discussion. I would have preferred to see some ambient samples presented in the paper but recognize that this increases the scope of work substantially. Otherwise, the technique holds promise for distinguishing different types of pollen for real-time sampling, which is an exciting result. I have several minor presentation comments below that would make this manuscript acceptable for publication.

We thank the referee for their positive feedback on the manuscript. We agree on the importance of clarifying the scope of the study. The measurements were performed outdoor and not in the laboratory as suggested by the Referee. To this effect, material collected on neighboring plants was used. We have added the following comments to the revised version of the manuscript to clarify this point:

L18: Added ", using manually generated data," to underline that we do not use operational data.

L105-106: "After collection, pollen was brought to the outdoor measurement site and aerosolised.", to precise the outdoor measurements.

L232-233: "The results based on manually generated data show that combining FL with holography leads to a substantial identification performance gain.", to underline the manually generated data.

L270-272: "The present study demonstrates the potential of using FL measurements as a complementary input to holographic images for single-grain pollen identification using the SwisensPoleno and ML algorithms for the most important allergy causing pollen taxa in Central Europe.", changed benefit to potential to nuance the message.

Below we address the presentation comments by the Referee in the order as they appear.

Minor comments

1. Title: "measurement" -> "measurements"

   Done

2. Line 61: "… for the main allergenic species…." – is this the main species in Switzerland, or more broadly in Europe? More clarity for the selection of these seven pollen types would be beneficial.

   It is true for Switzerland but also representative for central Europe. Grass pollen (here *Dactylis glomerata*) is the most impacting pollen type in central Europe. Ragweed (*Ambrosia artemisiifolia*) is the most allergenic and the other five (*Corylus avellana, Fagus sylvatica*, *Fraxinus excelsior*, *Pinus sylvestris*, *Quercus robur* and *Urtica dioica*) are amongst the most common pollen types in this region. This list should also be completed by Birch (*Betula*

pendula) but no sample was available at the time of the study. We have added the following comment to clarify this point:

L61: precision "… for eight of the main allergenic pollen species in central Europe …"

3. Line 75: What size of particles trigger the detector? While this study is specifying the types of pollen evaluated, what would happen if it were an ambient air sample?

The detector triggers particles of a minimal size of 0.5 µm, however below 2 µm there is very limited information for holography due to the resolution of 0.595 µm/pixel. If the model was applied to ambient air samples (operational data), we would filter the data before giving it to the model. It is usually done this way because Neural Networks are a supervised classification method, meaning they can classify correctly only what they learned to distinguish. Usually for pollen, a simple filter on the particle aera and solidity is sufficient to remove debris and smaller particles. The general concept of prefiltering before classifying is described in Sauvageat et al. 2020.

L77: added "…, in the size range from 0.5 to 300 µm, …"

L176-178: "This means that a degradation of scores is possible when applying the model to operational data as all sorts of pollen taxa can be encountered considering that other particles are filtered out before the classification.", to precise the process in case of operational data

4. There is some terminology in the paper that is rather confusing:
    1. "event" – this is defined on line 89, but that makes it sound like you are sampling ambient air versus a controlled emission. Also confusing in Table 1, where it really seems to be the number of particles counted and evaluated. I would suggest something like "number of particles counted", "number of images", or even just "pollen count".

The term "event" represents the input data to a model. We defined the term "event" for simplicity as it represents the data recorded for each particle measured by the instrument independently of the nature of the data. This way, we do not have to specify that we have "holographic images and fluorescence spectra" each time we speak about a measured particle. An event should be seen as an occurrence of the measurement of a particle and does correspond to the number of pollen grains that were measured by the SwisensPoleno. We used the term "event" as it also corresponds to the terminology of the SwisensPoleno. For consistency with previous work, we prefer to keep that term. However, we checked that its use is clear along the paper and made a few corrections in that regard:

L225: Caption of Table 2: "… ratio of correct prediction over total count for each class… "

L407: Caption of Figure 4: "Uniform Manifold Approximation and Projection (UMAP) of particle features (morphology or/and FL features) …"

    2. "class" – defined on line 95 as the same as plant taxa, although it was unclear why this specific term was used – why not just keep "taxa"? Also, later throughout the paper (e.g., line 218, 241, y-axis label on Figure 5) these are used interchangeably and makes it rather confusing.

The term “class” comes from the Machine Learning field. Using the word class instead of taxa makes the paper less specific to pollen as other biological or non-biological aerosols can be measured using the same instrument and method.

5. Table 1: If using the term “class” in the paper, I would suggest changing the header on the first column from “Common name” to “Class (common name)” to clarify the terms more clearly.

Change done as suggested.

6. Line 179: eccentricity seems to be a useful metric, but is there also one about symmetry that might be helpful for more complex grains?

We thank the Referee for the useful suggestion. We have added the following comment to the revised version of the manuscript.

“While the eccentricity is used to give a hint on the symmetry of the pollen grain, further metrics could be introduced to further quantify symmetry. This was not implemented in the present study as feature extraction is done automatically by the convolutional neural network.”

7. In Figure 3c, the standard deviation on the relative fluorescence is extremely large. Can the authors comment on why they think that this metric improves the ML model?

In Figure 3c, we displayed the relative fluorescence measured when exciting the particle with a 280nm laser and receiving on emission channel at 357nm. This is one metric among the complete spectrum of the particle. Even though the variation on this metric is large, it still gives useful information when considering the complete set of intensities (13 in total). When considering the spectrum, we have a combination of values, and this is the full pattern that is given to the model.

8. Line 183: “superior” -> “larger”

L183: replaced “superior to” by “larger than”

9. Line 247-248: Please rephrase

L247-248: rephrased as “The coherence between our results and those from Pöhlker et al. 2013 brings confidence into our measurements and the stability of the Poleno.”

We thank the referee for their time and implication in our work. The comments were constructive and helped us improve the quality of the paper.