Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-432-AC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Automatic pollen recognition with the Rapid-E particle counter: the first-level procedure, experience and next steps" *by* Ingrida Šaulienė et al.

Ingrida Šaulienė et al.

ishauliene@gmail.com

Received and published: 1 April 2019

First of all, we would like to thank the reviewers for their detailed comments to the paper. We used their comments and criticism to improve the manuscript. Below, we provide point-by-point response to the issues raised by the reviewers.

Reviewer 2

1. With three different research centers, three different training/test dataset, three different ANNs and three different experimental set-ups, it becomes unrealistic to make any meaningful comparisons. As a result, I'm afraid that the paper reads as if it were

Printer-friendly version

Discussion paper

three separate studies intertwined, without any benefit of joining them into one.

We respectfully disagree. The studies were indeed following their own paths, primarily reflecting the differences in the three independent projects supporting the work of the teams. It is also important that there is no previous experience with Rapid-E and no "good practice" could therefore be followed – development of such practice is one of the purposes of this paper. The main added value was that the different approaches resulted in similar (but not completely identical) skills of the pollen recognition. The differences in the skills were, where possible, traced to the different approaches, ANN construction and data pre-processing. As a result, all groups participating in the paper were able to assess efficiency of their approaches and look into alternative ways of analysis. Now, we would like to share the findings with others, who are just considering to enter the field of online pollen monitoring. We hope that the revised paper is more clear in this regard.

2. The paper is quite lengthy and can be condensed considerably, improving the readability of the paper and preventing reader fatigue.

We have reviewed the paper structure and shortened it

3. How do the results compare to the Poleno method that integrates both image recognition and laser fluorescence?

There was no comparison with POLENO monitor, it was not the purpose of the paper to compare different devices. We were building the procedures for Rapid-E and analyse the "absolute" skills of recognition that can be obtained with the Rapid-E technology.

4. The timing of the fluorescence imaging is apparently such that it captures only a very limited amount of particles detected by the 400 nm laser. As presented in Table 1, the recall rate is not only low in general; the percentage of analysed particles was varying considerably among plant groups, ranging from 51% (for Quercus) to 87% (for Salix alba). Therefore, this raises the concern whether the first step already introduces

a bias towards certain pollen types.

The recall rate for fluorescence is indeed one of the significant concerns. For the confusion tables, however, the recall rate was not important: the initial number of pollen noticed by the scattering laser is not used in the analysis, i.e. the algorithms only dealt with fluorescent particles. The point is explicitly commented in the revised paper as a note to Table 1.

5. Experiments were carried out with fresh samples of collected pollen grains of separate plants and dried. Could this have skewed the results, as this procedure may decrease the amount of damaged pollen and aggregates as compared to daily practice?

This seems to be improbable, in fact, shaking pollens off the inflorescences and storing them in a package may rather increase the number of aggregates and small-size debris. This is why we put attention to microscopic analysis confirming that the fraction of these additions is very small.

6. What was the accuracy of the particle morphology step for preclassifying pollen vs non-pollen? Is there any selection bias just based on particle size and therefore pollen type?

In fact, the pollen-nonpollen pre-classification (so-called “pollen mode” of Rapid-E) is solely based on a requirement of the particle being larger than 5 um of optical diameter. Therefore, introduction of bias at this step is very improbable.

7. The reason for choosing three different designs of neural networks is not clear to me. Especially the choices in designing the first network from Siauliai seem to be a bit ‘ad hoc’. In the network for analyzing the scattering images, it is surprising that three fully-connected layers were needed and that batch normalization was combined with dropout. In the network for fluorescence, it is not clear why no 1D convolution layers were used.

As stated above, there is no guidance of optimal practices of ANN construction for Rapid-E. Therefore, all participating groups were building their own approaches and all have tried several options before arriving at the ones shown in the paper. Harmonization effort is certainly in need but was left for the future when the problem of compatibility of the devices themselves is solved. In the Siauliai case, various options have been tested for the neural network design. The three layers gave the best results. To avoid overtraining, we used regularization layers that are currently often combined: dropout and batch normalization. The 1D convolutional layers are suitable when the entire signal is homogenous. Pollen fluorescence peaks at different wavelengths are different, so it seemed expedient to use a fully-connected layer.

Introduction Generally, the introduction starts off quite clear, but at the end it needs more structure to clarify the goals of the paper.

The goals of the paper have been refined.

8. line 52: “Hirst-type pollen traps” needs a bit more introduction on its methodology, before discussing its limitations.

An outline of the Hirst device is added.

9. line 72: Is there any literature on the evaluation of the Poleno device, that needs referencing?

Not really. This is a very new device and no consistent evaluation has been published.

10. line 77: the goal to evaluate the Rapid-E is described in too general terms. What aspect will be evaluated: sensitivity & specifity (and compared to which gold standard?), reproducibility of the system, processing speed or general applicability (Is this why three different centers participated?)?

Clarification is added, goals have been refined, as well as the roles of each centre. As stated above, the centres are operating within completely independent projects, which initiated certain diversity in the approaches.

11. line 78: what are “the Rapid-E data”?

Clarified.

12. line 81-85: Apparently, the different centers had different tasks in this project, but it is not clear which. As I understand it from the text, the system was assessed only in Siauliai, but then it is not clear how Novi Sad and Payerne determined their best classification.

Clarification is added, goals have been refined, as well as the roles of each centre. As stated above, the centres are operating within completely independent projects, which initiated certain diversity in the approaches.

13. line 82: Why did you compare the best classification from the three centers, because they have different training or different procedures or pollen population? And what is the definition of ‘the best classification’?

Clarification added.

14. line 84: What was the hypothesis that led you to compare the Swiss classifies results to the Hirst data? And what is ‘the Hirst’ data?

The selection of the group was opportunistic: the Swiss team had the Rapid-E and Hirst devices working next to each other during full seasons in the previous years, so it was possible to put the time series next to each other. However, the paper does not have a goal of the Rapid-E – Hirst comparison. The provided time series were added to discuss the “false positive” identifications. We made it clearer in the revised version.

15. line 96: which range do you refer to?

Clarified: changed to “this saturation level will not be reached in realistic ambient conditions”.

16. line 103-104: Do I understand correctly, that this follow-up analysis is then comparable to the Hirst-type of analysis?

Well, to some extent. The geometry of the slide and flow is strongly different but the principle is the same: impaction on a sticky slide followed by the microscopic analysis.

17. line 109-110: What is the conclusion of this sentence? Is the 0.5-100 um range not being used because of hardware life expectancy, or just because it does not apply to pollen? In the latter case, it is only logical to use the Pollen-mode.

In this section, we did not derive conclusions, only described the device. The pollen mode as the default one is brought up later when we describe the setup of the experiment.

18. line 111: Before entering the experiments section I would first expect a section on the ANN methods that do the classification, because these are evaluated in the experiments.

The data analysis sections and the experiment description sections have been switched.

19. line 111: Chapter 2.2. The scheme of the experiment. Here I would expect first an overview of the different experiments that have been conducted, and especially the rationale behind them. As it is now, we dive into the details of the different experiments performed at the three sites, without a clue of the bigger picture.

The experimental section has been rearranged to provide the overall picture first and then specificity of each site.

20. line 113: What was the goal of this experiment (To test the accuracy)and what was the gold standard (separate purified samples of a particular plants)? Why was training needed in the first place (apparently the system does not come with a trained classifier for a specific area)?

Indeed, the system does not come with any classifier to any of the machines, plus the machines proved to be substantially different.

21. line 117: “the pollen recognition algorithm” is not introduced yet. This is due to comment 14.

The sentence is reformulated.

22. line 125: “repeated twice”, so in total 3 times?

No, just twice. The word “repeated” is replaced with “performed”.

23. line 144: What does “practically identical” mean? So, the same classifier was trained (again) and the experiment repeated twice?

No, just the practical side, i.e. the pollen exposure. Clarified.

24. lines 150-156: It is not clear what the rationale is behind adjusting the test (and training?) set; is it to adapt to the local pollen population in Serbia or was it to test other hypotheses?

This is indeed just reflecting the pollen availability in Serbia.

25. line 158: “For practical reasons” is a bit vague.

The meaning clarified: “Low ambient concentration of coarse particles allowed a less laborious approach: pollen calibrations in Payerne . . .”

26. line 160: What were the threshold-based criteria? To what parameter was the threshold applied?

Clarified.

27. line 234: Why was a threshold of 10um used instead of 5?

Partly, for the historical reasons, following the work of Crouzy et al. For the tested pollen it is the same: they all are much larger. But the tighter threshold was effective in eliminating the ambient particles, which, albeit in low numbers, were still reaching the device.

28. lines 258: “Results obtained in Siauliai”. Shouldn’t there then be a chapter on the results of the other two sites?

The results from all three sites are collected together.

29. lines 259-265: These research question should be presented in the introduction, not in the results section.

Moved to introduction.

30. line 270: Why have the spectra been normalized, isn’t the amplitude a characteristic in itself?

Each spectrum was normalised separately prior to any further analysis to eliminate the effect of imperfect hit of the pollen by the UV laser impulse. Normalization allows to reduce variation of fluorescence amplitude for one pollen taxon at fixed wavelength.

31. line 270-274: It is not clear what is tested statistically: difference between genera or within genera? The Student test is not appropriate for testing data that is highly inter-correlated, like a spectrum.

In the shortened paper, we removed the Student test remark as a weakly-connected part of the analysis.

32. line 272: “The uncertainties [. . .] were a fraction of a %..” This sentence is not clear to me: you mean less than 1%, but that is not true.

Corrected to “percentage”.

33. line 273: “statistically significant difference”, what was the differences and what was the p-value?

Updated to $p < 0.001$.

34. line 323: Which of the 8 test data tests have been used here, what is n? (idem for section 3.2.2.)

In Siauliai experiment with each genus consisted of up to 8 sample tests (approximately 5 mg of pollen/test). n of registered data by Rapid-E and analysed in research is indicated in Table 1 in column “Fluorescent particles”. They were all subjected to the recognition analysis, together.

35. line 372: “very similar way”. I don’t think you can say that since both the training and test data were different, and the more classes are included the more difficult the classification by an ANN will be.

The sentence is changed to “showed similar recognition skills”.

Technical Corrections:

36. line 372: typo, Swissens ->Swisens

37. line 89: typo, “analyses” -> “it analyses”

38. line 126: typo, “with” ->”with a”

39. line 128: typo, “fit” -> “fitted”

40. line 136: typo’s, “compare” -> “compared” and “used new”-> “used a new”

41. line 202: typo: Rapid-e ->Rapid-E

42. line 205: typo, “founding” -> “finding”

43. line 150: style, “same or similar pollen morphotypes” -> “an adjusted set of pollen morphotypes”

44. line 217: NLL (negative log-likelihood) in full

45. line 810: Figure 1, the font size in this figure are quite small and difficult to read in a printed version.

Thanks you! All corrected.

46. line 820: The results of the two ANNs in Figure 3 and 4 were simply summed,

whereas the other centers used concatenation. Then it is more consistent to show Figures 3 and 4 in one figure, and connect the two networks with a summation component. Please use the same conventions as in figures 5 and 6; the network in Figure 3 can be shortened by using Convolution Blocks, like in Figure 5.

Corrected.

47. line 825: Figure 6, for consistency with Figure 5, the lifetime and spectrometer sub-network need to be swapped.

48. line 835: Figure 8, font size on the x- and y-axis is too small to read in printed version.

49. line 830: y-axis in Figure 7, normed ->normalized.

50. line 845: y-axis in Figure 9, normed ->normalized.

Thank you! All corrected.

On behalf of authors Ingrida Šauliene

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-432, 2019.

Printer-friendly version

Discussion paper