

Interactive comment on “Characterisation of biofluorescent aerosol emissions over winter and summer periods in the United Kingdom” by Elizabeth Forde et al.

Anonymous Referee #1

Received and published: 12 September 2018

General comments

In this paper, the authors report on the deployment of WIBS-3 and WIBS-4 sensors in four ground sites in the U.K. The collected dataset is very extensive, covering different locations and seasons. Records like these of fluorescent particle concentrations are of current interest in the community.

However, I found some of the work on the HAC clustering and particle identification to be speculative at times. I think it could be made more convincing with the use of laboratory data that the authors reference, but don't quite show. From what I understand, the cluster types were initially assigned to HAC-derived clusters following broad

C1

observations of similarity to laboratory types. I think the analysis would be significantly strengthened if a direct statistical comparison of lab and field clusters were presented. Since HAC method validation is a major part of this paper, I also feel that presenting clustering quality metrics would help.

I recommend the publication of this paper in ACP, after the following major comments are addressed:

(1) Can you discuss further how the clusters are initially assigned to types following laboratory work? Has the proximity of lab-derived clusters to field clusters been assessed or calculated? Can you use a distance metric to directly compare them? Basically, how sure are you of the assignments of the field-derived clusters to cluster types shown in Table 2?

(2) Similarly, an inter-cluster distance metric could be used to support the segregation of initial clusters into distinct groups in Table 3. Maybe using a cluster dendrogram plot to show similarity would be a good idea?

(3) Can you present any metrics on how successful the HAC algorithm was in segregating particles into different clusters? How distinct are the clusters? Consider presenting criteria such as Calinski-Harabasz index or Davies-Bouldin index to demonstrate the cluster separation quality.

(4) There is a discussion in section 3.6 that shows that assigning a different fluorescence threshold caused a completely different clustering solution for the Chilbolton dataset. Is the extra cluster found with the 3SD threshold comprised solely of interferences? I would be interested to see more discussion of how this affects the other datasets.

(5) More discussion of possible interferences would help. What do you think they are? How do they compare to previous laboratory studies?

Specific comments

C2

Abstract, pg. 1 line 18 (last sentence): consider rephrasing, not sure what this means.

Introduction, pg. 3 line 28: principal should be principle

Introduction, pg. 5 line 5: can you discuss thresholding here briefly? Why was a different threshold used in this work? What are the advantages?

Methods, pg. 5, line 16: this sentence (starting with "Whilst. . .") seems unfinished.

Section 3.4.1: Temperature and Relative Humidity: consider providing more figures for this analysis. It would be useful to see if all temperature and RH trends for clusters identified as fungi vs. bacteria match each other. It is much harder to see from just a text description.

Section 3.4.2: Wind Speed and Wind Direction: similarly, would it be possible to show wind roses of each cluster (or cluster group) to make the similarities discussed in text more obvious?

Section 3.5: Statistical relationship between fluorescent particles per site and meteorological data: what is the purpose of this analysis? In particular, why were third-order polynomial fits used? Given that the statistical model is not fully clear and that all of the r-squared values are low, consider either significantly expanding this section or eliminating it.

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2018-513>, 2018.