RESPONSES TO ANONYMOUS REFEREE #1

(1) Reviewer comments are in black text.
(2) Author responses are in blue text.
(3) Additions/modifications made to the manuscript.

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 under-stand, the cluster types were initially assigned to HAC-derived clusters following broad 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:

We thank the reviewer for this comment and, in the following responses, clarify why this will be the subject of future work. This work utilised data and analysis from published studies (e.g. Savage et al 2017, Hernandez et al 2016) and whilst some of the other laboratory data used in this study needs to be subject to the peer review process, we felt alluding to recent results, which we will publish imminently, would help the classification process.

(1) Can you discuss further how the clusters are initially assigned to types following laboratory work?

The fluorescent signals of ambient derived clusters have been compared to laboratory data using trends in fluorescent channels in order to initially group the 18 clusters for further analysis. Specifically, the clusters from each field site were compared to existing ('Dstl experiment 2014') and new ('Dstl experiment 2017') laboratory data, depending upon the instrument used, in addition to published data e.g. Savage et al 2017. Laboratory data was available from a WIBS-3 for the 2014 Dstl dataset, and from a WIBS-4 for the 2017 Dstl dataset (the results from such are to be published in the new year). Prior to comparing the ambient clusters to the laboratory data according to broad fluorescent signature, the process for deriving these clusters was the same as used in all previous ambient studies. Additionally, data from published laboratory experiments (e.g. Savage et al 2017, Hernandez et al 2016) were used to provide some further support, and aid the initial grouping of these clusters into suspected particle type groups.

Has the proximity of lab-derived clusters to field clusters been assessed or calculated?

These have not been calculated in this paper directly; rather a qualitative comparison made between trends across fluorescent channels, size and shape. To perform this in a quantitate manner requires consideration of a number of issues which require further laboratory data to be published and subject to the peer review process. Firstly, rather than using unsupervised methods [in this case hierarchical clustering, HAC], supervised techniques would be able to assign any sampled ambient particle to a class that has been studied in the laboratory depending on choice of parameters use for any given technique. Part of this procedure includes choice of appropriate distance metric between each fluorescence signal [which the reviewer refers to as a proximity metric]. These methods demonstrate exciting potential for improved and more detailed bio-aerosol classification. However, as noted in Ruske et al (2018), before recommendations can be given to choice of method and distance metric, more laboratory data is needed to reduce the chance of misclassification. Indeed, even for HAC, Ruske et al 2018 studied a range of model permutations, demonstrating the variability in laboratory signatures according to how samples were prepared, for example. We are planning on a conducting a much more thorough evaluation of statistical methods once we have published and had this data appropriately peer reviewed. In this paper we use the current recommended configuration HAC as used in all current bio-aerosol publications.

Can you use a distance metric to directly compare them?

Please see our response to the previous questions. It is entirely possible to employ a distance metric to directly compare ambient clusters and laboratory data; this information would be explicitly used within supervised learning techniques to perform direct classification. However we feel it would serve no real benefit to detail those distances without then using the supervised techniques. Indeed, this goes beyond the scope of this particular piece of work and is inline the current state of the literature. The idea behind the use of both laboratory data and published data in this study was to qualitatively compare the fluorescent profiles of known biological types to ambient data to group these clusters into suspected particle types for further analysis. This is similar to the approach taken by Kasparian et al. (2017).

Basically, how sure are you of the assignments of the field-derived clusters to cluster types shown in Table 2?

There is undoubtedly a level of error with this method, not least by qualitatively using laboratory data, which may not be fully representative of ‘real-world’ conditions i.e. not accounting for the effects of atmospheric transport, aggregation and fragmentation of particles. However, the use of such a method has been employed previously to determine potential cluster particle types (e.g. Crawford et al 2017, Kasparian et al. 2017) and is still a valuable method for sanity checking profiles.

Here, we did not rely only on the expected fluorescent signals from laboratory data to determine the type of particles these clusters comprise. Instead, we built upon this by considering abundance, the size and shape of the particles within the cluster, the diurnal variation of the cluster, the response to the meteorological variables (temperature and relative humidity), and the land cover category for the site location in question. The assignments following this process (Table 4 in the manuscript) either reinforces the initial assumptions made using only the laboratory data, or disproves it at each stage. A similar approach was used by Crawford et al., (2014) where the identified PBAP clusters were found to correlate well with other bioaerosol detection techniques (Gosselin et al., 2016).

(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?

The use of the Calinski-Harabasz index to segregate the clusters has been employed in this paper, similar to previous studies (e.g. Crawford et al 2017). The segregation of clusters into the six groups as seen in Table 3 are based on the fluorescent profile analysis and comparison to laboratory data between the four sites as conducted in section 3.2.

An example cluster dendrogram can be seen for Weybourne (Fig.1), which is accompanied by the Weybourne centroid figure. Using the cluster dendrogram plot it can be seen that the merging of Cluster 3 and Cluster 2 occurs first, which is interesting considering the differences in fluorescent signal which can be seen in the centroid figure. As Cluster 3 was the dominant cluster at this site, the merge with Cluster 1 is not unexpected, owing to the similar fluorescent profile between the two clusters. The clusters then merge with Cluster 4 last which has a slightly similar signal to Cluster 1 and Cluster 3 in terms of the higher FL3 signal, but differs with some signal in FL2 and FL1.

Though this illustrates the process of merging the clusters, this does not have the advantage of showing the characteristics of the clusters, such as the particle size and shape, compared to the centroid figures used in this study. A more detailed discussion of the application, interpretation and limitation of the Calinski–Harabasz index applied to these instruments may be found in the related publication by Ruske et al. 2018. Additionally, a more detailed description of HAC clustering using dendrograms is described in Ruske et al. (2018). The hierarchy using the original strategy suggested in Crawford et al. (2015) compared with the modification using a 9 sigma threshold as suggested by Savage et al. (2017) is also discussed.

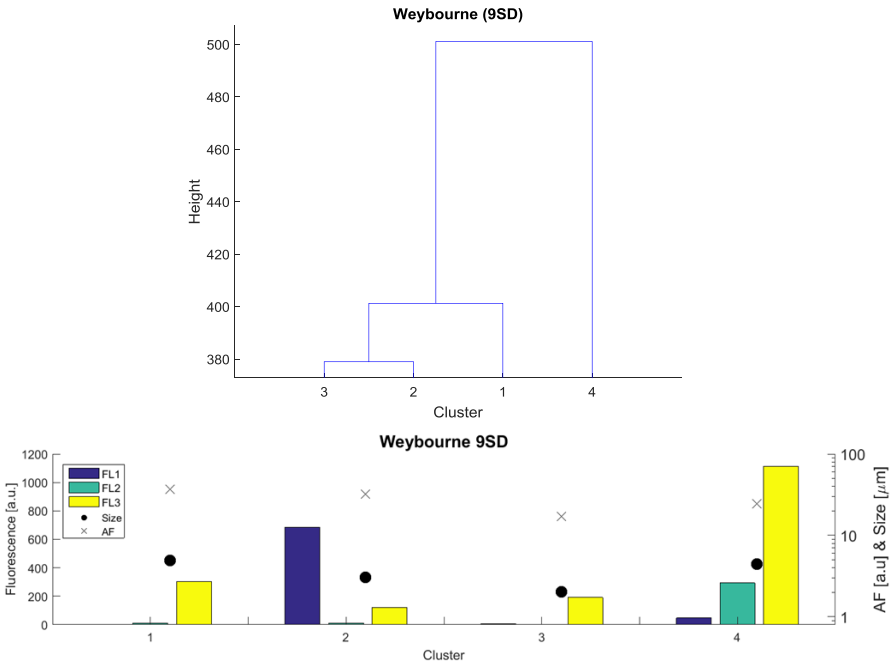

Figure 1: Example cluster dendrogram for Weybourne in addition to the centroid figure used in this study.

(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.

As shown in Figure 2, when using the Calinski-Harabasz (CH) criterion for segregating the clusters for Weybourne, the optimum value suggests a four-cluster solution for the data from this site. As highlighted in the paper, it is common that similar clusters are often subsets, segregated by particle size and shape.

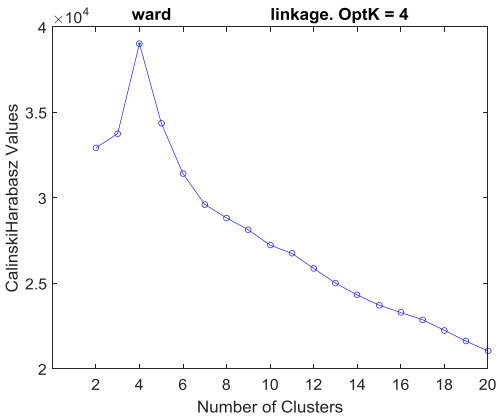

Figure 2: Example Calinski-Harabasz cluster solution following clustering of Weybourne.

(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 interferents? I would be interested to see more discussion of how this affects the other datasets.

When using the 3SD threshold for the Chilbolton site, we note that the change in threshold does not result in a completely different clustering solution (Figure 3). Rather, it can be seen when using 9SD that Cluster 3 and 4 are representative of Cluster 4 and 5 when using 3SD, but at lower concentrations. Considering the presence of some signal in channel FL1, but a dominant signal in channel FL3, it can be assumed that Cluster 2 (3SD) is representative of Cluster 1 (9SD). This leaves Clusters 1 and 3 (3SD), which may have been merged as they appear to be represented by Cluster 2 (9SD). This extra cluster was determined to be a wet-discharged fungal spore following the complete analysis as opposed to interferent particles.

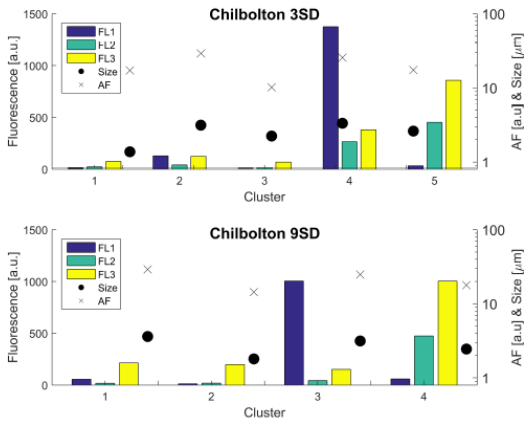

Figure 3: Comparison figure showing the difference between 3SD and 9SD for Chilbolton (Figure 7 in manuscript).

At the other sites there is not always a loss of a cluster as in Chilbolton. In comparison, some sites retained the same number of clusters (e.g. Davidstow and Weybourne), whilst the change from 3SD to 9SD for Capel Dewi resulted in the gain of a cluster, from four clusters when using 3SD to five clusters when using 9SD.

As a result of challenges in interpreting cluster solutions when using FT + 3SD and 9SD at each site, these plots been added to the supplementary materials (on pages 16-17).

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

The four sites in this paper are similar in that they are not closely located to any major cities or towns and are similarly situated in agricultural/grassland locations. This reduces the potential impacts of vehicle emissions from city traffic and fuel burning and other sources, but does not rule out some episodic emissions from roads or access points located close to a few of the sites.

In the Hernandez et al 2016 laboratory study the dominance of the FL2 channel (referred to as Type B in their study following ABC analysis) was determined to be representative of potential interferents . By using a higher 9SD threshold for analysis and comparing this to the 3SD the fluorescent signal intensity decreases, even when, for some sites, the number of clusters does not change. For Capel Dewi, the amount of clusters increases, while Weybourne and Davidstow stay the same.

Though, given that there are no loss of clusters for the other sites, and even the production of an additional cluster for Capel Dewi, the use of 9SD here produces only a reduction in the fluorescent fraction of each cluster. The use of 9SD for Capel Dewi appears to split Cluster 4 (3SD) into two different clusters (Cluster 1 and Cluster 2) when using the 9SD threshold.

SPECIFIC COMMENTS:

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

This sentence is a comment on the lack of available published information of different particle species and the influence of meteorological variables (such as RH and Temp) on their abundance.

This has been changed on page 1 and page 25 from:
"More knowledge of the reaction of speciated biological particles to differences in meteorology, such as relative humidity and temperature would aid characterisation studies such as this."

To:
"More published data and information on the reaction of different speciated biological particles types to fluctuations in meteorological conditions, such as relative humidity and temperature, would aid particle type characterisation in studies such as this."

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

We have changed principal to principle on page 3.

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

A brief description has been added to page 5 to describe the use of the 9SD threshold.
"Contrary to previous work, this is additionally the first use of a differing fluorescent threshold of 9 standard deviations (SD) compared to traditionally 3SD, in an ambient setting, ***to reduce the impact of interferents from potential anthropogenic sources, following Savage et al 2017."***

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

The sentence in this paragraph on page 5 has been re-worded from:

"Whilst Skjøth et al. (2012) utilised the Corine Land Cover 2000 dataset to identify agricultural areas under rotation and in harvest in relation to Alternaria spore concentrations in Denmark."

To:

"***In addition, Skjøth et al. (2012) utilised the Corine Land Cover 2000 dataset to identify agricultural areas under rotation and in harvest in relation to Alternaria spore concentrations in Denmark"***

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.

The authors acknowledge that inclusion of these plots would aid interpretation of the paper, as opposed to a text-based description. We were keen to include these plots in the manuscript, however owing to the quantity of plots, these would take up a considerable amount of space. **We have added the suggested figures, showing the differences observed in relation to temperature and relative humidity, to the supplementary materials (pp. 2 – 9).**

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?

Due to the amount of clusters including these figures in the manuscript would consume a considerable amount of the paper, which is the reason why only the total fluorescent polar plots from each site were included. **As a result these plots have been added to the supplementary materials (pp. 12 – 15).**

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.

The purpose of this analysis was to produce a value to be used to calculate an emission factor, similar to Crawford et al (2014), due to the sparsity of data relating to bioaerosol emission and various meteorological drivers. Due to the variance in the total fluorescent data from each site and the cluster variability third-order polynomial fits were chosen. **As a result of the already lengthy analysis in the manuscript, this section has been removed.**

References

Crawford, I., Robinson, N. H., Flynn, M. J., Foot, V. E., Gallagher, M. W., Huffman, J. A., Stanley, W. R., and Kaye, P. H.: Characterisation of bioaerosol emissions from a Colorado pine forest: results from the BEACHON-RoMBAS experiment, Atmos. Chem. Phys., 14, 8559-8578, https://doi.org/10.5194/acp-14-8559-2014, 2014.

Crawford, I., Gallagher, M. W., Bower, K. N., Choularton, T. W., Flynn, M. J., Ruske, S., Listowski, C., Brough, N., Lachlan-Cope, T., Flemming, Z. L., Foot, V. E., and Stanley, W. R.: Real Time Detection of Airborne Bioparticles in Antarctica, Atmospheric Chemistry and Physics Discussions, pp. 1–21, https://doi.org/10.5194/acp-2017-421, http://www.atmos-chem-phys-discuss.net/acp-2017-421/, 2017.

Gosselin, M. et al., 2016. Fluorescent bioaerosol particle, molecular tracer, and fungal spore concentrations during dry and rainy periods in a semi-arid forest. *Atmospheric Chemistry and Physics*, 16(23), pp.15165–15184.

Hernandez, M., Perring, A. E., McCabe, K., Kok, G., Granger, G., and Baumgardner, D.: Chamber catalogues of optical and fluorescent signatures distinguish bioaerosol classes, Atmospheric Measurement Techniques, 9, 3283–3292, https://doi.org/10.5194/amt-9-3283-2016, 2016.

Kasparian, J. *et al*. Assessing the Dynamics of Organic Aerosols over the North Atlantic Ocean. *Sci. Rep.* **7**, 45476; doi: 10.1038/srep45476 (2017).

Ruske, S., Topping, D. O., Foot, V. E., Morse, A. P., and Gallagher, M. W.: Machine learning for improved data analysis of biological aerosol using the WIBS, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2018-126, in review, 2018.

Savage, N., Krentz, C., Könemann, T., Han, T. T., Mainelis, G., Pöhlker, C., and Huffman, J. A.: Systematic Characterization and Fluo- rescence Threshold Strategies for the Wideband Integrated Bioaerosol Sensor (WIBS) Using Size-Resolved Biological and Interfering Particles, Atmospheric Measurement Techniques, pp. 1–41, 2017.