We thank the referees for their thoughtful and constructive comments. Please find the detailed response to each comment, including text changes to the manuscript. Our response is in blue font while changes to the manuscript are underlined. The line numbers used here refer to those in the change-tracked manuscript file.

Reviewer #1

This paper applies a machine learning approach to try to interpret the POA and SOA data that is observed at the Mace Head atmospheric observatory. These have been reported in previous publications, but this uses a different approach to try to analyse and interpret a long-term dataset.

There are advantages to this type of approach over other methods like PMF, in particular computational cost, interpreting dependencies on independent variables and the ability to gap-fill where necessary. However by using a supervised approach, this can be said to be less objective. This is a particular concern because quite a lot of a priori information is brought into the analysis, in particular how POA is defined, where a number of assumptions are made based on previous experience analysing data. So that being the case, it is perhaps not surprising that the model performs so well. Although it should be noted a certain amount of objectivity is brought in through the use of clustering to generate the representative POA. Furthermore, given that this specific site has proved fairly unique in producing observations such as these (owing to its location), it would remain to be seen whether this technique was applicable to other sites.

All that said however, I still found this an interesting and informative paper. The relationships between the POA and SOA and the other variables in particular, but it was also interesting comparing these data products to the outputs of the HTDMA data, which could have implications for marine CCN populations. I do have some concerns, perhaps the biggest being that I didn't find enough technical information accompanied how the models had been set up (see below) but besides that my comments are pretty minor and I recommend publication after these are addressed.

Response: We thank the reviewer for the insightful and positive comments. Indeed, we acknowledge the model itself might be suitable for this specific site, but we believe the similar methodology is can to be applied to other locations.

Major comments:

The article, as it stands, is severely lacking in the technical details of how the FCM and SVR methods were applied to the data. While it is not necessary to post the code that was used, certain technical details were missing from the article and the supplement. Specific things relating to the FCM I found missing were the distance metric (while Euclidian distance is the default it cannot be inferred), whether there

was any pre-treatment to weight or linearise variables, and how the optimum number of factors were determined. Likewise with the SVR, more explanation should be given regarding the specifics of the input data and parameters. More details such as these should be included, in the supplement if necessary.

Response: We thank the reviewer for these important technical points. We have now included the requested details in the revised manuscript.

We have provided a detailed description of the settings for the FCM model in the revised section 2.3.

Line 154: Fuzzy C-Means (FCM) is a clustering algorithm that enables the grouping of data points into multiple clusters with varying degrees of membership. In this study, the input variables for the FCM model included chemical components (SeaSalt, Org, NO_3, SO_4, NH_4, MSA, and BC), as well as meteorological parameters — temperature (temp), relative humidity (rhum), wind speed (wdsp), and wind direction (wddir). We randomly selected 10,000 samples and retained only those with positive concentrations for all chemical components. The data were then log10-transformed and Z-score standardized. The FCM model was configured with 5 clusters, a fuzziness exponent of 1.2, and Euclidean distance as the distance metric.

The SVR model settings are detailed in the revised Section 2.2.

Line 118: SVR was chosen for its generalizability in handling small datasets and its resistance to overfitting (Ghimire et al., 2022; Juang and Hsieh, 2009). Unlike tree-based models like random forest (Breiman, 2001), SVR model can predict continuous values (Ma et al., 2003; Tang et al., 2024). The hyperparameters, including the penalty coefficient (C) and gamma (γ) of the Radial Basis Functions (RBF) kernel were tuned via grid search. The model targeted OA concentration, using nss-SO_4, MSA, NH_4, and meteorological parameters (temperature, relative humidity, boundary layer height, wind direction, and pressure) as predictors. We also included hours of the day to capture diurnal variations. Predictors not directly linked to secondary production, e.g., sea salt, NO_3, and BC, are excluded to avoid over-fitting and ensure generalizability, even though including these might have enhanced the model performance in the training dataset. Wind speed was used to select the SOA production period, therefore, it was not suitable as a predictor. A summary of the variables employed as predictors is shown in Table 1.

Line 128: The SVR were trained using 'tidymodel 1.3.0' framework using R programming software (version 4.4.3)

Reference

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/a:1010933404324, 2001.

Ghimire, S., Bhandari, B., Casillas-Pérez, D., Deo, R. C., and Salcedo-Sanz, S.: Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia, Engineering Applications of Artificial Intelligence, 112, 104860, https://doi.org/10.1016/j.engappai.2022.104860, 2022.

Juang, C.-F. and Hsieh, C.-D.: TS-fuzzy system-based support vector regression, Fuzzy Sets and Systems, 160, 2486–2504, https://doi.org/10.1016/j.fss.2008.11.022, 2009.

Ma, J., Theiler, J., and Perkins, S.: Accurate On-line Support Vector Regression, Neural Computation, 15, 2683–2703, https://doi.org/10.1162/089976603322385117, 2003.

Tang, D., Zhan, Y., and Yang, F.: A review of machine learning for modeling air quality: Overlooked but important issues, Atmospheric Research, 300, 107261, https://doi.org/10.1016/j.atmosres.2024.107261, 2024.

Minor comments:

L185: While the Monte Carlo bootstrapping is a powerful method of assessing random uncertainties in the data, it does not address the issue of systematic uncertainties, which given that these AMS OA types have been reported from few locations, is potentially substantial. This should be noted.

Response: we thank the reviewer for pointing this out, we have now included the following statement in the revised manuscript to highlight the potential uncertainties associated with Monte Carlo bootstrapping.

Line 225: “While it should be noted that the Monte Carlo bootstrapping is used to assess the random uncertainties, potential uncertainties associated with possible systematic uncertainties require further investigation.”

L224: A note of caution should be added regarding treating nss-SO4 as a secondary marker because while it is indeed formed through secondary timescales, the precursors and formation timescales are different to SOA and the relationship between the two is inconsistent when looking at terrestrial environments. There is a case to be made for them being correlated in the marine boundary layer if they are both assumed to originate from biological activity in the sea surface, but explicit correlation should not necessarily be expected.

Response: We thank the reviewer for the important comment. Indeed, the nss-SO4 formation timescale and dynamics are expected to be different to marine SOA. However, nss-SO4 were selected as a secondary marker as it is almost ubiquitous in the marine boundary layer and represents the most well-known secondary formation pathway of marine secondary aerosols. Actually we have not constrained the marine SOA to be correlated with nss-SO4, but the decoupled correlation between marine POA and SOA with nss-SO4 have led further confidence on the OA apportionment.

MSA, another typical marine secondary marker, seems to be equally important as nss-SO4. Although MSA and nss-SO4 have different formation pathways when DMS underwent atmospheric reactions, their high correlation suggests the marine SOA is also expected to be highly relevant in terms of correlation. We have now included additional statement in the revised manuscript to highlight possible bias by using the selected set of variables, including nss-SO4.

Line 194: "It should be noted that the nss-SO4, which albeit being marine secondary species, exhibit different formation dynamics and timescales with marine SOA, which might induce some extent of uncertainty. However, both marine SOA and nss-SO4 might originate from marine biological activities. Furthermore, marine air masses arriving in MHD are expected to advected over Northeast Atlantic for several days. The use of nss-SO4 can also be supported by the high correlation between nss-SO4 and MSA, which exhibit different atmospheric formation dynamics."

Figure S3: I actually found this one of the more interesting aspects of this work, and I would consider moving it to the main article.

Response: We thank the reviewer for the suggestions, we have now removed the Figure S3 to the main text.

Figure S7: The figure caption doesn't currently make sense. Reword.

Response: We have now reworded the caption of Figure S7:

Figure S7. Observed OA versus predicted (SOA +POA_{bg}) for (a) training, (b) validation and (c) test datasets. Data density is illustrated using a color gradient with darker colour indicating lower data density. Black lines denote the 1:1 correspondence lines, blue lines represent regression lines. The ML model does not use MSA as predictors.

Reviewer #2

Overview and recommendations

Chen et al., 2025 provides a technique for estimating marine POA by using a machine learning technique to estimate marine SOA. I found the layout and logic in this study confusing and hard to follow at times. My interpretation of the work is that ML was used to predict marine SOA, and then POA = OA - SOA. But the authors focused on filtering the data down to periods in which marine SOA were most likely and little marine POA was predicted during these times. It was not clear to me whether this was used to train the model and that then the model was applied to the whole time period, or if the model was only used in these highly filtered times. If the latter case (seems unlikely!), then the ML work seems unnecessary. I've tried to point out where I found the methods and application confusing in my comments below. I do think that if the authors can make their intent and applicability of methods more clear that this study could be published in EGUsphere.

Response: We thank the reviewer for the comments and tried our best to clarify the concerns, point-by-point response is below.

Major comments

Why did the authors choose to predict SOA with the ML model, rather than POA? Neither seems inherently correct or incorrect but the logic should be explained.

Response: We thank the reviewer for pointing this critical issue. We choose to predict the SOA because there are some variables is more related to secondary production processes. As for the POA production, less is known about its emission strengthen, source regions etc.

We have now included the following statement in the revised manuscript.

Line 106: "marine SOA was chosen as the ML model predictor, because SOA is expected to be impacted by environmental factors and less is known about POA, emission strengthen and source regions."

Line 85: Downscaling hourly meteorological measurements to 10 minutes could introduce significant bias to the study, depending on how sensitive the ML predictions were to the various met parameters. Were any sensitivity studies performed on this assumption? Can the authors comment on any potential limitations from this assumption?

Response: We thank the reviewer for this critical technical point. We have conducted additional sensitivity test for the meteorological downscaling. We test different downscaling techniques, including using hourly average values, linear downscaling.

As shown in Figure below, using different interpolation techniques introduces minor bias to the POA results.

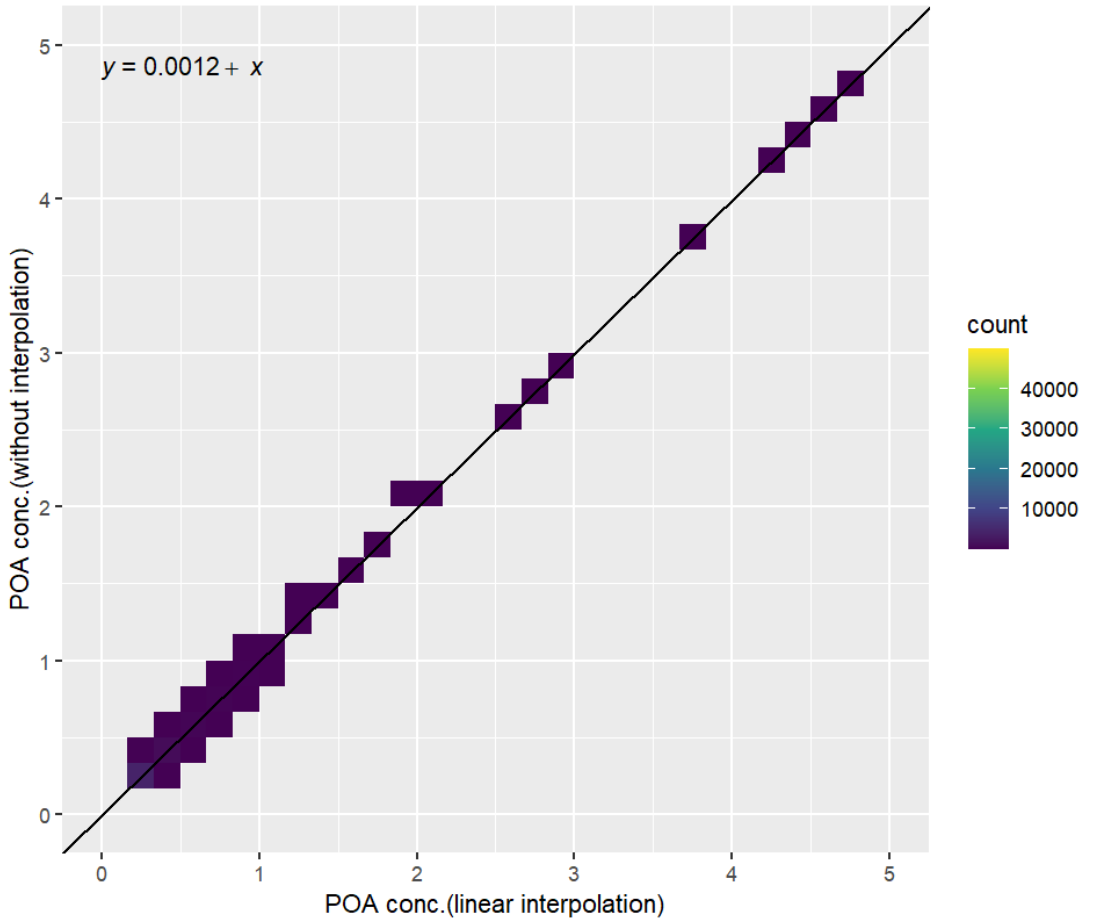

Figure R1. POA concentration using linear interpolated meteorology data versus POA concentration using hourly average meteorology data.

Sect 2.2 The filtering processes don't make sense how they're presented - for example the paragraph that begins "We then applied additional filtering processes to reduce the impact of POA production..." If this study is intended to use ML to predict SOA and then obtain POA via POA = OA - SOA....then what is done for the periods that are filtered out to exclude POA production times? Are those times assumed to be POA only or just removed from the dataset? Basically overall, this section reads as "We find the periods of time in which SOA is most likely and then predict POA from that ." Which then begs the question - wouldn't POA be very low and why is a ML model needed? Presumably this is not the authors' intent so clarity is needed throughout this section. Perhaps once SOA is quantified in the "SOA-only" times, then the authors deploy their methods to the whole data set?

Response: We thank the reviewer for rising this important point. The ML-model was trained using SOA-only periods and then applied to whole dataset, the anomaly

between predicted SOA and measured OA is then attributed to marine POA. We have now included additional statement in the revised manuscript for clarity:

Line 141: "Finally, the ML method was applied on the clean marine air masses data to predict marine SOA concentrations…."

Assuming that the authors trained the model on SOA production periods only and then applied the model to the whole dataset - have similar aerosol/ML studies been performed? Does it seem valid that the model can easily sort out anthropogenic influences to correctly predict marine SOA (mSOA) and therefore mPOA? Another way of asking this question is: what happened to the time periods with anthropogenic influence? How is the model supposed to know that everything that is not mSOA could be a mixture of mPOA and aSOA/aPOA? Again, what portions of the dataset were used needs to be clarified throughout.

Response: Indeed, similar aerosol/ML studies have been performed. For example, Lei et al. (ACS ES&T Air 2025), trained a ML model strictly on local-emission dominated events and applied the ML on whole dataset to distinguish between local and regional transport.

It has to be noted that the model does not sort out anthropogenic influences. We have utilized MHD clean sector criteria to the data to ensure minimum influence from anthropogenic influences. The time periods with anthropogenic influences were removed from further analysis. However, we expected that it is possible to train an additional anthropogenic OA model to further expand the approach.

We have included the following statement in the revised manuscript:

Line 141: "Finally, the ML method was applied on the clean marine air masses data to predict marine SOA concentrations…."

Line 155 - the SVR model is used to predict OA. I thought that OA was available via AMS data, and that the model is being used to predict SOA to find POA.

Response: The reviewer is correct; we have revised the statement for clarity.

Line 179: demonstrating the model's accuracy in predicting total OA (SOA + POA-bg) concentration using the selected predictors.

Line 181

Section 3.3 - I'm confused where the k_HTDMA and spread values come from. Is this from the observations? The ML model? Other? This section seems somewhat out of place at the moment.

Response: As stated in the k_HTDMA were obtained using a humidified tandem differential mobility analyzer. The detailed description has been included in the revised manuscript.

The humidified tandem differential mobility analyzer (HTDMA) (Swietlicki et al., 2000) was used to measure aerosol hygroscopic growth at a fixed relative humidity of 90% for aerosol with selected dried sizes of 35, 50, 75, 110, and 165 nm. The growth factors measured by HTDMA were inverted using a piecewise linear function (Gysel et al., 2009).

$$GF_{mean}^{a,b} = \frac{1}{nf^{a,b}} \int_a^b GF\ c(GF, D) dGF$$

The arithmetic mean GF (GF_{mean}) was calculated as:

$$GF_{mean} = \int GF\ c(GF, D) dGF$$

To quantify the mixing state, the GF spread factor (SF), defined as the standard deviation of the GF-PDF divided by the GF_{mean}, was calculated as:

$$SF = \frac{\left(\int_0^\infty (GF - GF_{mean})^2\ c(GF, D) dGF\right)^{1/2}}{GF_{mean}}$$

The GF was measured at 90% RH, however, the RH of the second DMA fluctuated slightly with the ambient temperature. The data between 88-92% RH were corrected using the κ-Köhler theory according to formula derived from Petters and Kreidenweis (2007):

$$\kappa = \frac{(GF^3 - 1)(1 - a_w)}{a_w} \quad \rightarrow \quad GF(a_w, \kappa) = \left(1 + \kappa \frac{a_w}{1 - a_w}\right)^{1/3}$$

where a_w is the water activity, and obtained by Köhler theory:

$$a_w = \frac{RH}{\exp\left(\frac{4\sigma_s v_w}{RTD}\right)}$$

Where σ_s is the surface tension of the droplet, v_w is the partial molar volume of water, R is the universal gas constant, T is the temperature, and D is the diameter of the droplet. The surface tension is assumed to be 0.072 mN m^{-1}.

We believe the inclusion of HTDMA data analysis is useful as it is related to marine aerosol-cloud interaction, which is associated with the climate impact of both marine POA and SOA.

Minor comments

The introduction refers to numerous marine studies but does not provide any geographical context to any of them. It would be helpful to understand to what regions each studies' conclusions are most likely appropriate to.

Response: We thank the reviewer for pointing this out. Marine aerosols indeed have shown a large geographical heterogeneity. We have now included more details on regional specific studies.

Line 36: Ovadnevaite et al. (2011a) documented a marine POA plume with a peak POA concentration of up to 3.8 μg m^{-3} in Northeast Atlantic,

Line 46: O'Dowd et al. (2015) observed significant changes in the CCN activities of sea spray aerosol during a phytoplankton plume over Northeast Atlantic, whereas Quinn et al. (2014) and Bates et al. (2020) reported no substantial alterations in CCN activity over Northwest Atlantic.

Lines 78-79: Please note what was used to provide meteorological conditions.

Response: The source of meteorological conditions is now implemented.

Line 101-110: Please provide citations for the assertions about SVR's skills and merits.

Response: Thanks for your comment. The references have been added, and the relevant details are provided below.

SVR was chosen for its generalizability in handling small datasets and its resistance to overfitting (Ghimire et al., 2022; Juang and Hsieh, 2009). Unlike tree-based models like random forest (Breiman, 2001), SVR model can predict continuous values (Ma et al., 2003; Tang et al., 2024). The hyperparameters, including the penalty coefficient (C) and gamma (γ) of the Radial Basis Functions (RBF) kernel were tuned via grid search.

The model targeted OA concentration, using nss-SO_4, MSA, NH_4, and meteorological parameters (temperature, relative humidity, boundary layer height, wind direction, and pressure) as predictors. We also included hours of the day to capture diurnal variations. Predictors not directly linked to secondary production, e.g., sea salt, NO_3, and BC, are excluded to avoid over-fitting and ensure generalizability, even though including these might have enhanced the model performance in the training dataset. Wind speed was used to select the SOA production period, therefore, it was not suitable as a predictor. A summary of the variables employed as predictors is shown in Table 1. The SVR were trained using 'tidymodel 1.3.0' framework using R programming software (version 4.4.3).

Reference

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/a:1010933404324, 2001.

Ghimire, S., Bhandari, B., Casillas-Pérez, D., Deo, R. C., and Salcedo-Sanz, S.: Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia, Engineering Applications of Artificial Intelligence, 112, 104860, https://doi.org/10.1016/j.engappai.2022.104860, 2022.

Juang, C.-F. and Hsieh, C.-D.: TS-fuzzy system-based support vector regression, Fuzzy Sets and Systems, 160, 2486–2504, https://doi.org/10.1016/j.fss.2008.11.022, 2009.

Ma, J., Theiler, J., and Perkins, S.: Accurate On-line Support Vector Regression, Neural Computation, 15, 2683–2703, https://doi.org/10.1162/089976603322385117, 2003.

Tang, D., Zhan, Y., and Yang, F.: A review of machine learning for modeling air quality: Overlooked but important issues, Atmospheric Research, 300, 107261, https://doi.org/10.1016/j.atmosres.2024.107261, 2024.

Line 113: please explain why 2015 was used; are there any indicators that POA and SOA could or could not have shifted in any way over the entire time period (2009-2018)? For example, given that 27.8% of SOA periods happened in the 4-year challenge period (2015-2018) but this period represents ~44% of the total time, that seems like it could indicate that a change in aerosol balance between POA and SOA may have occurred over the time period. Did the authors attempt using data from each year in the training and challenge datasets as a sensitivity study?

Response: The year 2015 was randomly selected. We have not noticed any shift for the entire time periods.

We now included leave-one-year-out cross validation to identify any shift over the entire data periods. As shown in the figure R2 below.

The model performances on each year are similar, indicating minimum shift over the

period.

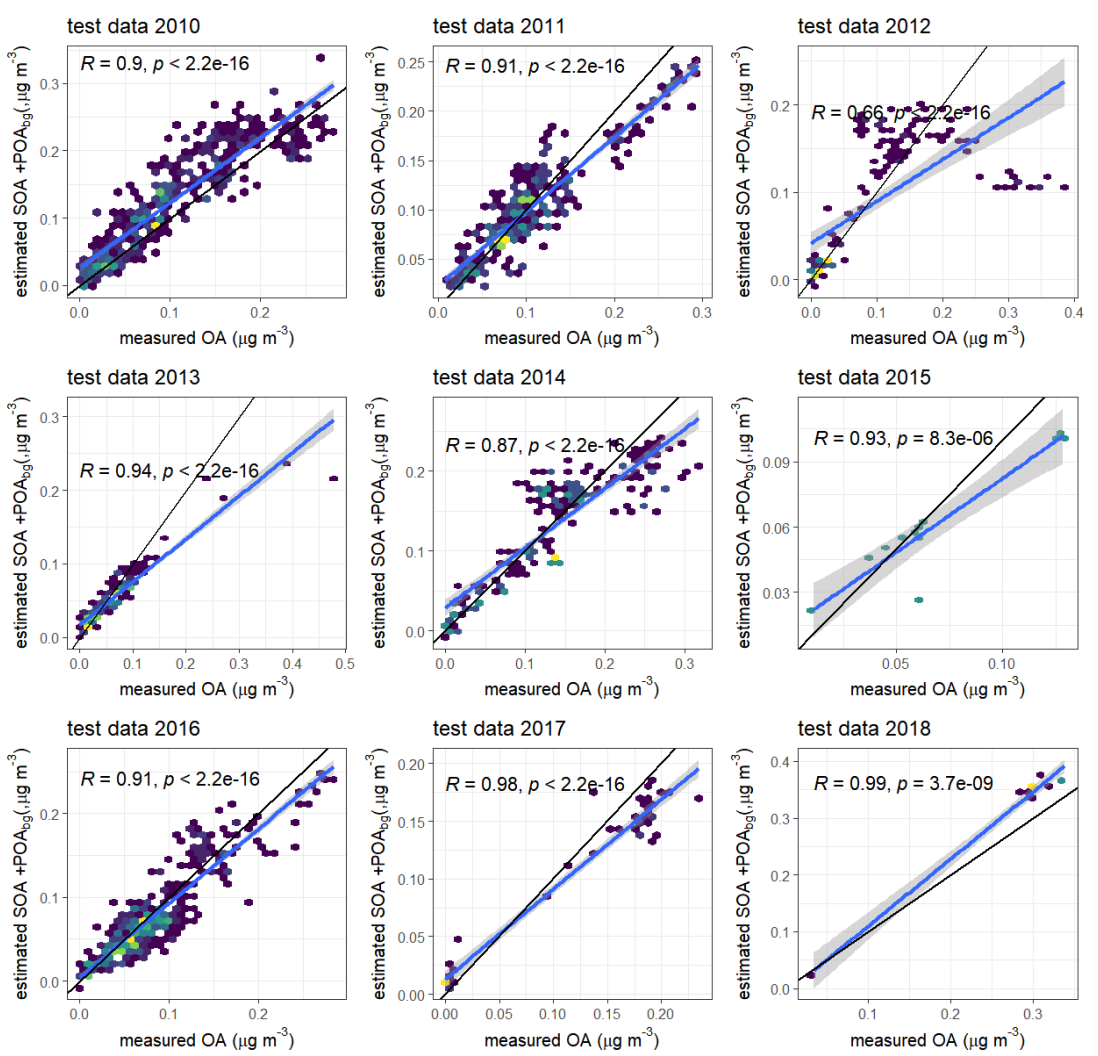

Figure R2. Measured OA versus estimated OA (SOA +POA_{bg}) for each year using other years as training dataset.

Figure 2 - is the y-axis mislabeled? Isn't it predicted SOA + POA_bg?

Response: We thank the reviewer for pointing this out, the y-axis label has been corrected.

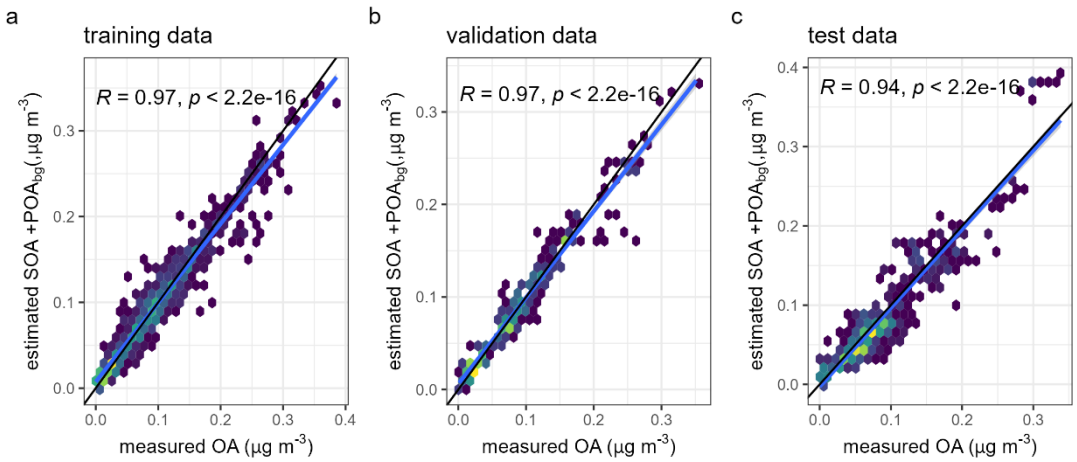

Figure R3. Observed OA versus predicted (SOA +POAbg) for (a) training, (b) validation and (c) test datasets. Data density is illustrated using a color gradient with darker colour indicating lower data density. Black lines denote the 1:1 correspondence lines, blue lines represent regression lines.

Line 172 - 183 and Figure S5: Figure S5 would be more meaningful with an indication of % of data points the y axis represents. Having 25,000 negative values as the optimal POA_bg assumption still seems quite large. Why are there so many negative values - is this to be expected with the ML methods deployed? How were the negative values handled?

Response: We thank the reviewer for pointing this out. As shown in the Figure below, the percentage number with negative values are lower than 6%. The negative values from originated from the uncertainty of the machine learning model and measurement. Nevertheless, the negative values were mainly found in winter time, in which the POA and SOA concentrations were too low to have any significant climate impacts.

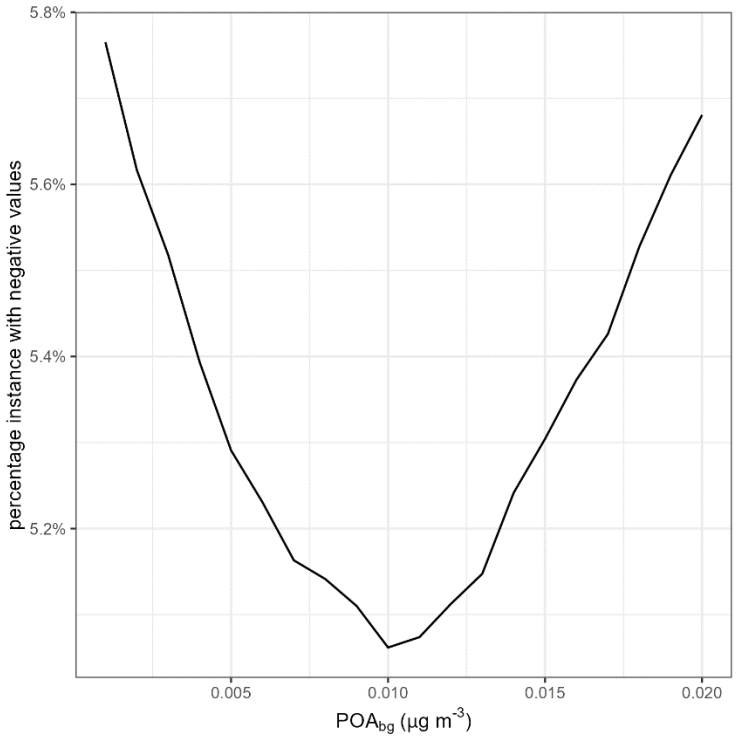

Figure R4. The different POA_{bg} value used in ML model and the percentage of non-physical prediction instance including POA and SOA (lower than 0). This indicates the POA_{bg} of 0.01 µg m^{-3} is the optimum value.

Technical comments

Line 16-17: " ...introduce a machine learning approach to differentiate and quantify the contribution of marine POA from marine secondary organic aerosol (SOA)." - the phrasing here makes it sound like the team is quantifying the contribution of POA that is formed from SOA. Suggest rephrasing (something like quantify the contribution of marine POA to total OA).

Response: The sentence has been rephrased: " ...introduce a machine learning approach to ~~differentiate and quantify the contribution of~~ distinguish between marine POA and marine secondary organic aerosol (SOA)."

Line 89: MHD - I assume the Mace Head site? Please be sure to define.

Response: Mace Head site defined in the revised manuscript.

Line 120 - please provide reference for FCM

Response: Reference included:

Bezdek, J. C., Ehrlich, R., and Full, W.: FCM: The fuzzy c-means clustering algorithm,

Computers & Geosciences, 10, 191–203, https://doi.org/10.1016/0098-3004(84)90020-7, 1984.

Line 139 - “periods” not “period”

Response: Corrected.

Throughout the manuscript (including figures) the authors refer to “train” or “training” data - I

believe training is more typically used. Please use one or the other throughout.

Response: We have now revised the terminology to ‘training’.

Figures and Tables

Please provide in the SI a key for the abbreviations used on Figures S1, S2 and elsewhere (e.g. bih, rhum...)

Response: The abbreviation table has been included in Table 1.