Review of "Towards understanding the mechanisms of new particle formation in the Eastern Mediterranean" by Anonymous Referee #1

The manuscript titled "Towards understanding the mechanisms of new particle formation in the Eastern Mediterranean" presents yearlong observations of NPF events at a rural background location in Cyprus. Observations are based on various instrumentation providing information about NPF events since the early cluster sizes. These are very important observations in the poorly presented in the literature region of East Mediterranean and Middle East and it is worth being published after some minor revisions.
I think however that the title is rather misleading since the manuscript is focused on the description of NPF events in Cyprus and their general characteristics and it does not contribute to actually understanding the underlying processes governing the formation of atmospheric particles and therefore I recommend a more modest title.

We thank the reviewer for the constructive comments. We believe that this manuscript not only unveils the characteristics of NPF but also provides insight into the underlying processes governing the formation of atmospheric particles. We analyzed the seasonality of sulfuric acid, NPF sinks, and formation rates, which are all important factors for understanding the processes of NPF. We also show that the formation rates measured at this sit cannot be explained by ammonia-sulfuric acid nucleation alone. That said, we agree with the reviewer that a more modest title is more suitable. As such, we have changed the title of the manuscript to: "*Towards understanding the characteristics of new particle formation in the Eastern Mediterranean".*

We provide our point-to-point replies to the general and specific comments below. The reviewer comments are in black, our replies are in green, the text before adjustment is in orange, and the adjustments made to the manuscript are in blue. The corresponding changes are noted in the manuscript by track changes. The referred line and figure numbers in these replies denote the new ones in the revised manuscript and Supplementary Data. All references are provided at the end of the replies.

General comments

Comment 1: The authors are only briefly describing observations of NPF events during periods that desert dust was present in the atmosphere. Although it has been pointed out that mixed conditions of dust and pollution may result to the formation of new particles even under conditions with high preexisting aerosol loadings, the observations reported in the literature are scarce and only in few locations around the world. During the study period, 37 out of the 50 dust days were categorized as NPF days. This is an extraordinary figure and these events should have been prioritized in their analysis, given that under dust conditions it is more possible to have an NPF event (74%) compared to the average situation (57%). On the contrary, the authors choose not to present a single event. Even if it is chosen to present these events in a separate research article, the intention of the present work to introduce the scientific community to a novel location under the EMME atmospheric conditions which are greatly affected by the presence of desert dust makes the presentation of such NPF events in more detail necessary.

We agree with the reviewer on the importance of presenting dust events in more depth. As such, we have added the subsequent section to the manuscript.

Addition to manuscript:

Figure 10 shows the temporal variation of $PM_{10-2.5}$, $PM_{2.5}$, and particle number size distribution measured during three of the dust episodes with ±5 days window before and after the dust episode. NPF took place at high dust loadings, and there is no obvious threshold for the dust loading above which NPF does not occur. In addition, the formation rates (Figure S9) and growth rates (Figure S10) between NPF event days not affected by high dust loading, and NPF event days affected by high dust loadings seem to be comparable. J_7 was slightly higher on days affected by high dust loading, but this could be related to the lower number of dust cases compared with the non-dust cases. High dust loadings can affect NPF in opposing ways. On the one hand, it can suppress photochemical processes by scavenging reactive gases and condensable vapors (De Reus et al., 2000;Ndour et al., 2009). On the other hand, it can provide particles that can act as a site for heterogeneous photochemistry promoting the formation of gaseous OH radicals, which initiate the conversion of SO_2 to

H_2SO_4 (Dupart et al., 2012;Nie et al., 2014). However, a clear association between high dust loading and NPF was not found from the data set presented here.

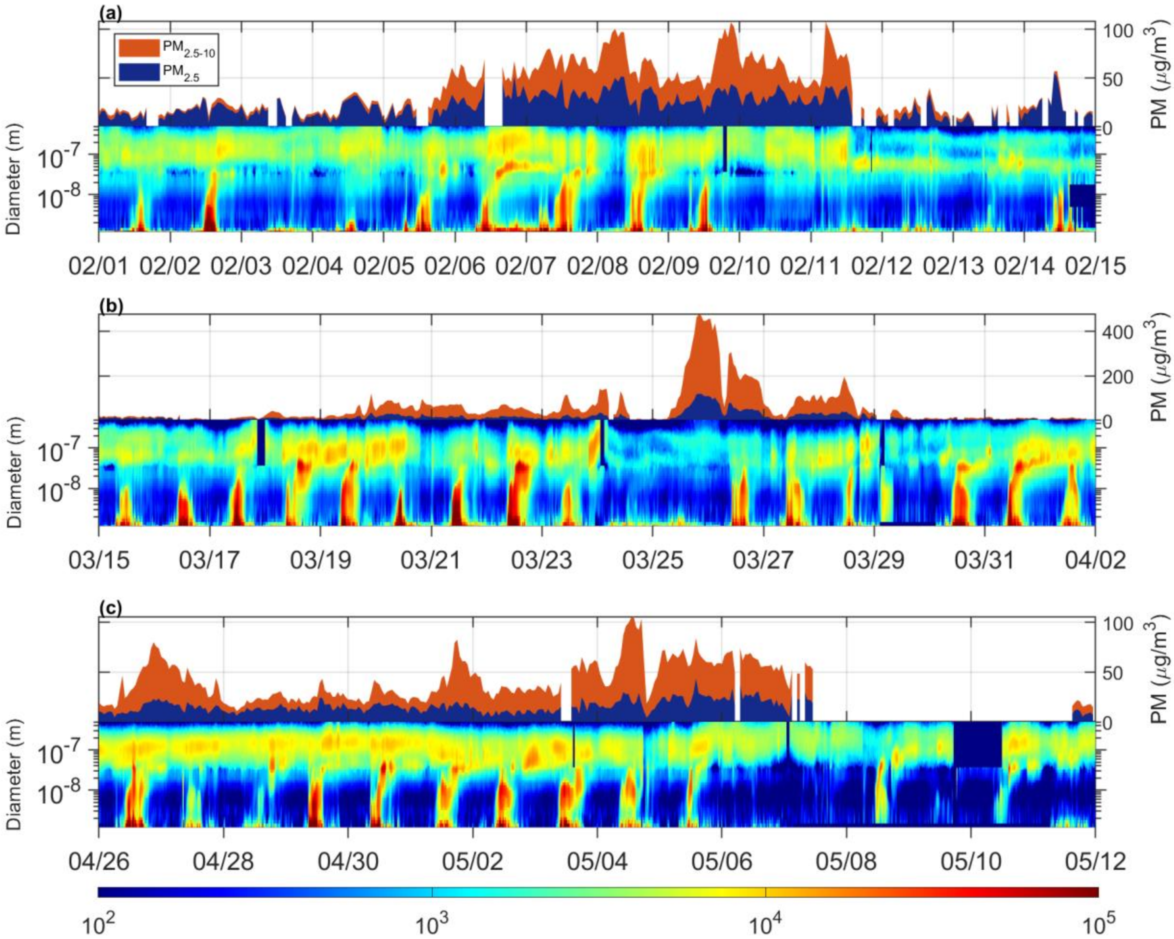

Figure 10. Temporal variations of aerosols during dust episodes with 5 days before and 5 days after the dust episode. (a) Time series of particle size distribution, PM_{10}, and $PM_{2.5}$ between Feb 1, 2018 and Feb 15, 2018 (dust episode: Feb 6 to Feb 10). (b) Time series of particle size distribution, $PM_{10\text{-}2.5}$ (coarse PM), and $PM_{2.5}$ between Mar 15, 2018 and Apr 2, 2018 (dust episode: Mar 20 to Mar 28). (c) Time series of particle size distribution, PM_{10}, and $PM_{2.5}$ between Apr 26, 2018 and May 15, 2018 (dust episode: Apr 26 to Apr 27 and May 1 to May 7).

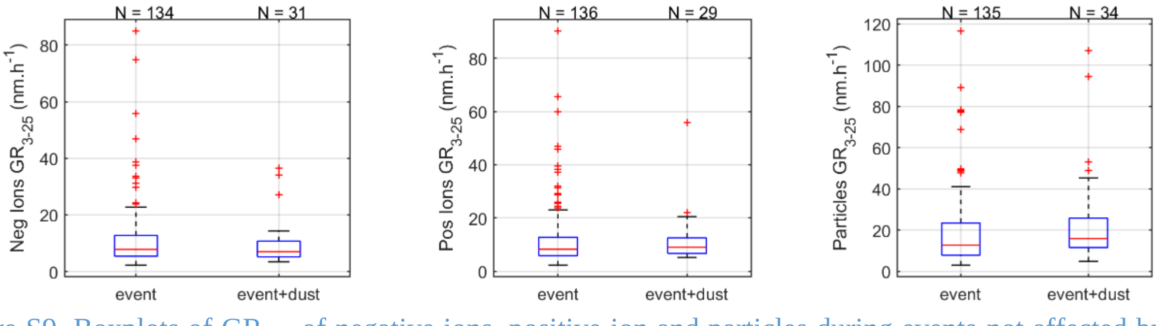

Figure S9. Boxplots of $GR_{3\text{-}25}$ of negative ions, positive ion and particles during events not affected by high dust loading and events affected by high dust loading

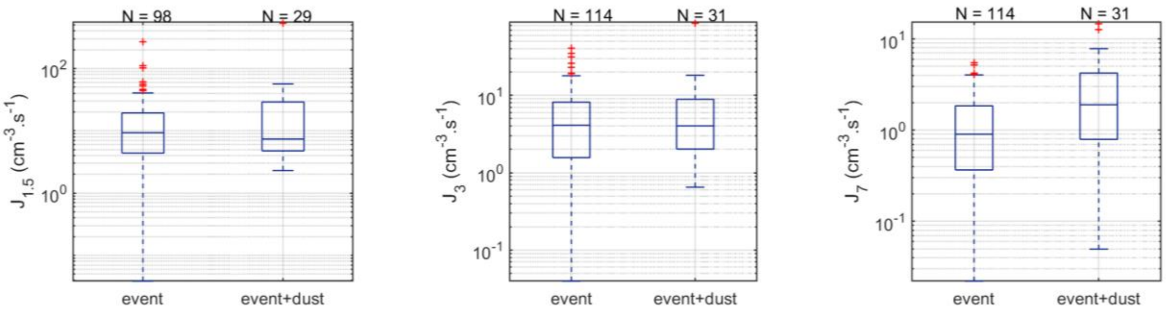

Figure S10. Boxplots of formation rates ($J_{1.5}$, J_3, J_7) during events not affected by high dust loading and events affected by high dust loading.

Comment 2: Another general comment has to do with the presentation of the driving parameters of NPF in the atmosphere of EMME. The authors have available a great set of complementary measurements to examine which atmospheric conditions favor or suppress NPF. The authors choose to present annual variability of each parameter rather than utilizing simple statistical tests to explore possible correlations. Visual inspection of event vs non-event conditions is not enough to contribute to the understanding the mechanisms of NPF and I would like to see some more in depth analysis such as PMF, PCA or simply regression analysis, for instance of cluster mode number concentration vs the various atmospheric components.

We agree with the reviewer that the visual inspection is not sufficient to understand the mechanisms of NPF. The visual inspection was crucial to analyze the seasonality of the different parameters and we have complemented it now with an in depth analysis as suggested by the reviewer. We have explored the applicability of several statistical methods that would help us understand which atmospheric variables are important for the occurrence of NPF. We used both a regression method that can model continuous values and a classification model that can predict or classify discrete (categorical) values. We chose the formation rate of 1.5 nm particles ($J_{1.5}$) as the modeled variable because it had a more evident difference between event and non-event days as opposed to the cluster number concentration. Alternatively, the event classification (is this day an NPF event day or a non-event day) was used for the classification model.

We did not use PCA or PMF analysis per se as suggested by the reviewer because both are statistical dimension reduction techniques. They reduce complex data into a small number of factors, each of which is a linear combination of the original variable but they cannot answer the question of which parameters are important for NPF. To answer this question, we could use the principle component scores, which are an output of PCA as independent variables in a regression analysis thus called principal component regression (PCR). However, after evaluating this method for our data, we found that a regression with PCA provides very little enhancement over the model without PCA.

<u>Addition to main text:</u>

To further understand the occurrence of NPF events, we present in the section 12 of the SI material, two types of analysis: the first is a linear regression analysis of formation rate of 1.5 nm particles ($J_{1.5}$) and the second is a decision tree classification model to indicate whether each day is an NPF event day or a non-event day. Both analyses have shown that NO_2, H_2SO_4 and wind direction (mainly from N to E direction which is the direction where the main agglomerations and livestock farming lands are situated) are the most important parameters that are associated with NPF occurrence (Figures S20-S23). While the role of H_2SO_4 in NPF is well known in literature, the role of NO_2 is not that clear. Most studies have focused on NO_x (contrary to NO_2 alone) role in NPF. NO_x has been shown to play contrasting roles in NPF depending on the associated pool of gas molecules. On the one hand, when oxidized to nitric acid, it can enhance NPF in the presence of ammonia vapors (Wang et al., 2020). On the other hand, it can suppress NPF by reducing autoxidation and low-volatility HOM dimer formation (Wildt et al., 2014;Zhao et al., 2018). Nevertheless, Yan et al. (2020) have shown that this effect is weak when NH_3 and H_2SO_4 are additionally present and that NO is more effective than NO_2 in changing the

HOM composition and volatility. Xie et al. (2015) have revealed that NO_2 can play an important role, not only in surface catalytic reactions of SO_2 but also in dust-induced photochemical heterogeneous reactions of NO_2, which produces additional sources of OH radicals and promote new particle formation and growth. However, while NPF seems to occur more frequently at higher NO_2 concentrations in our study, we cannot conclude if it plays a role in NPF or if it is a proxy of some other pollutant, especially that NO_2 concentrations were mostly lower than 4 ppb. What is evident however is that H_2SO_4 does not nucleate on its own at the concentrations reported in this study, thus an unknown stabilizer and possibly other compounds participating in NPF are missing in this analysis (as explained in the regression analysis). We hypothesize that these unknown compounds (e.g. NH_3 / amine /HOM) are associated with the North to east wind directions and higher NO_2 concentrations.

Addition to supplementary:

1. **Stepwise linear regression analysis:**

Data pretreatment:

We used hourly data for the regression analysis. Before performing the analysis, we applied a logarithmic transformation to the predictor variables having a skewed distribution. Then, we used Belsley collinearity diagnostics for assessing the strength and sources of collinearity among the predictor variables (Belsley et al., 1980). The remaining predictor variables after removing the variables that exhibited collinearity were NO, NO_2, CO, RH, temperature, solar radiation, wind direction, wind speed, $PM_{2.5}$ and sulfuric acid. We transformed the wind direction data into a categorical variable with four levels (N to E; E to S; S to W; W to N) to avoid data circularity. We removed the data corresponding to nighttime hours (solar radiation < 50 W/m^2), and undefined days from the analysis for a better separation between events and non-events. We further excluded any observation with any missing variable. Finally, we normalized all variables to make sure that all variables are of equal weight. Figure S19 shows the available data presented as correlation matrices.

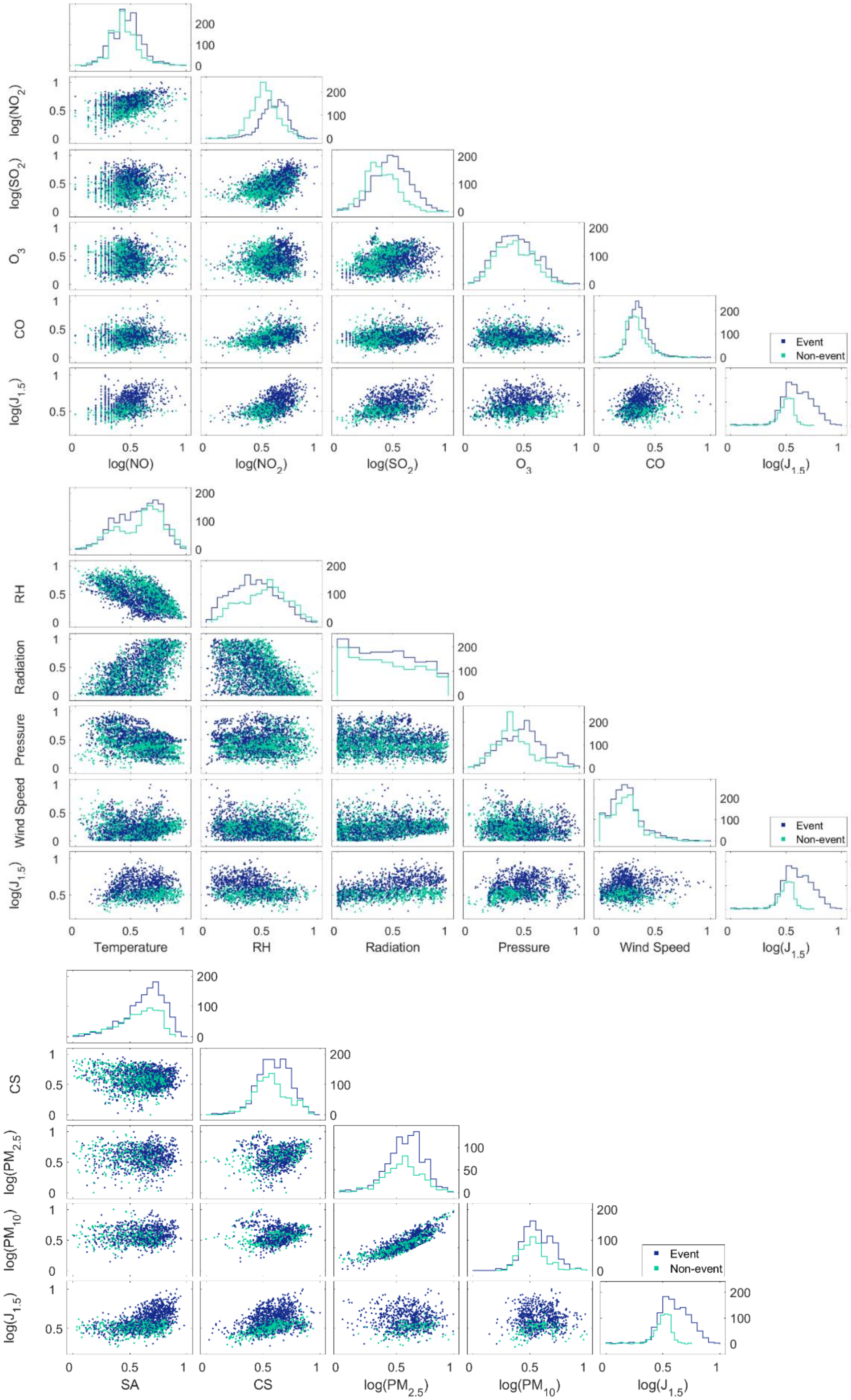

Figure S19. Correlation matrix of hourly atmospheric variables during event and non-event days.

Analysis:

We performed stepwise regression on the hourly data set to fit a linear model that would best describe $J_{1.5}$, which is the response variable. The steps were bidirectional starting from a model having no predictor terms and at each step, searching for terms to add to the model or remove from the model based on a pre-specified optimization criterion. Here we used the Akaike information criterion (AIC) as the optimization criterion, which is an estimator of prediction error and thereby the relative quality of the statistical model (Yamashita et al., 2007). First, we ran the stepwise linear regression by setting the upper bounds of the model to have an intercept and a linear term for each predictor (basic linear model; Model 1). Based on AIC criteria, NO_2, RH, solar radiation, H_2SO_4 and wind direction are the most important terms for the model. Here, we only show the change of AIC in the first step and last step of the stepwise regression (Table S4). A negative change implies that the addition or removal of a certain term decreases the model AIC and thus enhances the model. The importance of the aforementioned variables is also reflected by the coefficients for these terms and their p-values; Figure S20.a). CO and NO had a positive influence on $J_{1.5}$ but they were not as important as the aforementioned variables based on AIC criteria. Temperature and wind direction from the east to south sector had very little effect and were excluded from the final model. Wind speed, $PM_{2.5}$, RH and wind from the south to west sector had a negative influence on $J_{1.5}$, but the coefficients for the last two terms did not pass the significance level.

Table S4. The change in AIC during the first and last step of the stepwise linear regression.

	First step Change in AIC for adding	Last step Change in AIC for removing
NO	-139.2208	21.8408
NO_2	-259.412	101.3157
CO	-50.4108	4.3911
Wind speed (WS)	-1.5997	7.4704
RH	-104.7738	101.3157
Temperature	-40.7115	1.5184 (adding)
Solar radiation (SR)	-272.4999	34.4826
$PM_{2.5}$	1.22	13.4375
H_2SO_4	-410.1395	29.6485
Wind direction (WD)	-239.3033	99.1564

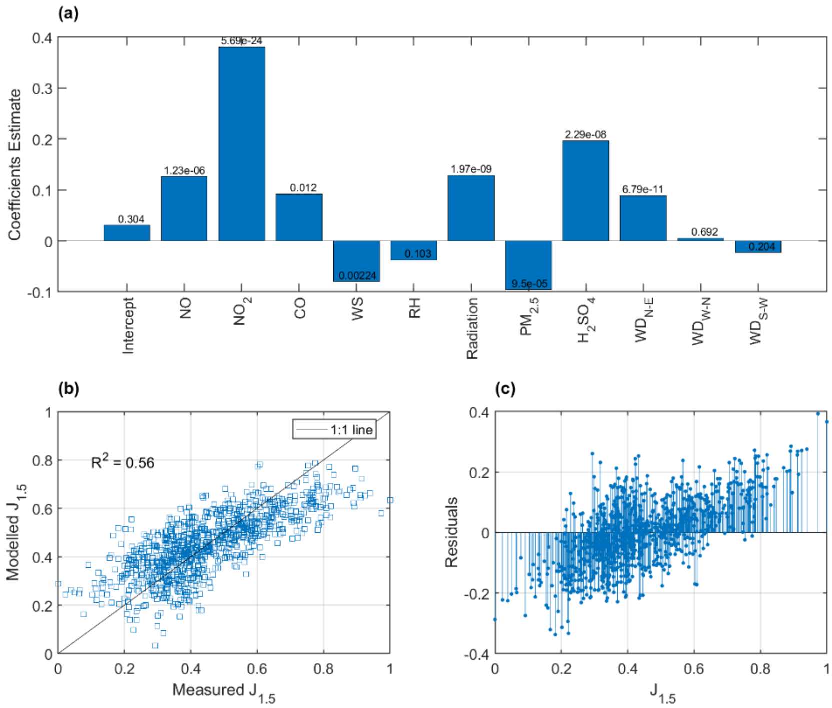

Figure S20. Model 1: (a) coefficients of model terms with p-values presented above the bars. (b) modelled versus measured $J_{1.5}$. (c) distribution of model residuals with respect to measured $J_{1.5}$.

The basic model had a coefficient of determination (R^2) of 0.56, a root mean square error (RMSE) of 0.1135 and it was not able to fit the data at high $J_{1.5}$ properly (Figure S20.b). The residuals of the model (Figure S20.c) indicate that there might be a missing predictor variable. We tried optimizing the regression model by allowing interaction terms (Model 2) and quadratic terms (Model 3). These models show enhanced RMSE and R^2 but they were also not able to fit the lowest and highest $J_{1.5}$ values properly. Figure S21 shows the optimized interactions model (Model 2; RMSE=0.0966). In general, this model is more complex to analyze because of overlapping terms. However, it shows that H_2SO_4 (represented as SA in the figure) has a positive influence when coupled with NO_2 and solar radiation and a negative influence when coupled with RH and temperature. It also shows that H_2SO_4, NO_2, and wind direction are the most important variables to explain $J_{1.5}$. In fact, we get a reasonable response when including these terms only in the regression (Model 4; Figure S22). However, none of the models gave a good response for the lowest and highest $J_{1.5}$ values, which could be yet another indication that there is a missing variable not included yet in the regression.

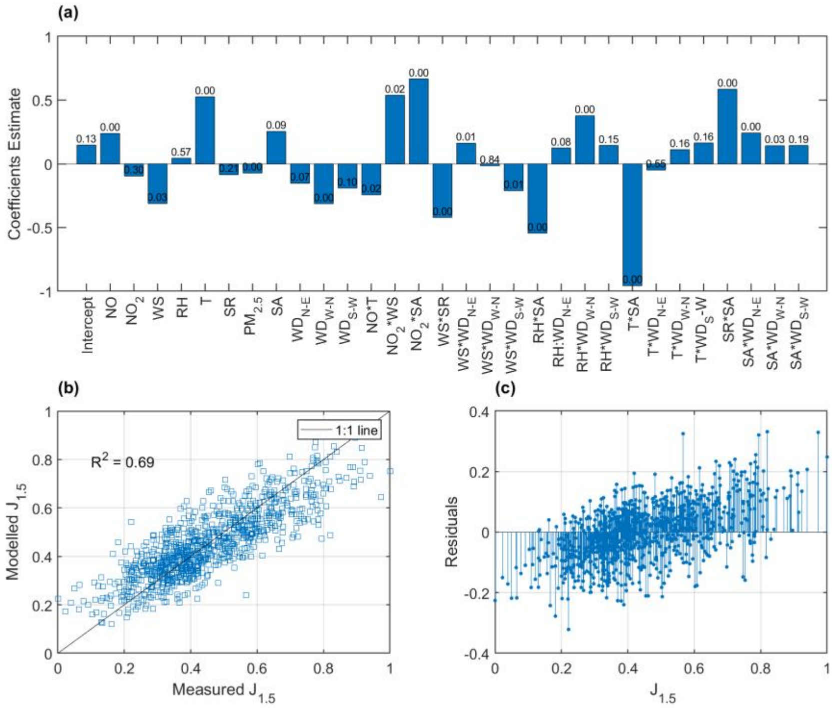

Figure S21. Model 2: (a) coefficients of model terms with p-values presented above the bars. (b) modelled versus measured $J_{1.5}$. (c) distribution of model residuals with respect to measured $J_{1.5}$.

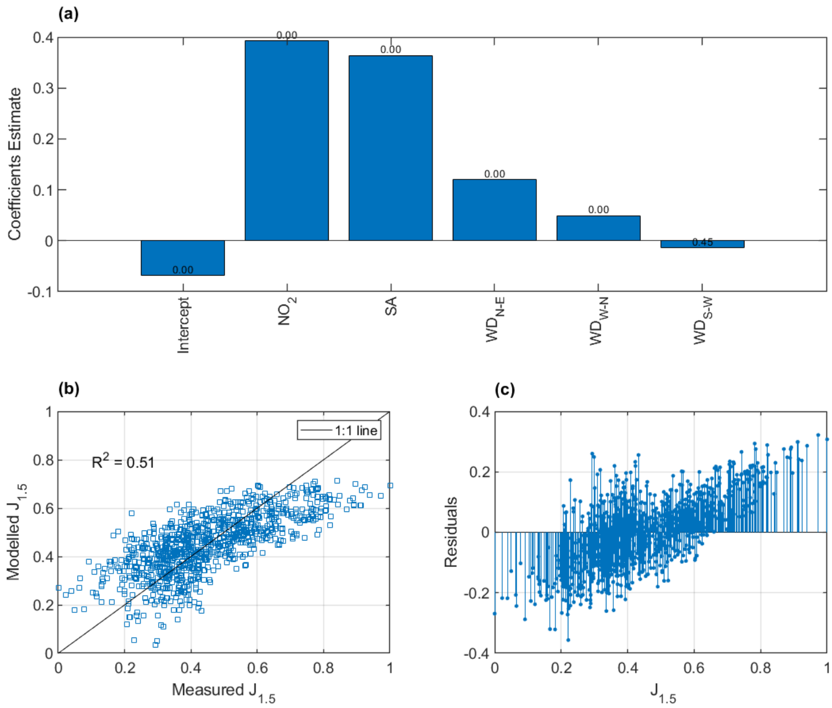

Figure S22. Model 4: (a) coefficients of model terms with p-values presented above the bars. (b) modelled versus measured $J_{1.5}$. (c) distribution of model residuals with respect to measured $J_{1.5}$.

2. **Classification decision trees:**

Classification trees comprise one of the most commonly used non-parametric classification approaches in machine learning and data mining. They recursively partition the feature space into a set of leaves with the most homogeneous collection of outcome possible (Breiman et al., 1984).

Data pretreatment:

We used daily data for the classification analysis. The daily data was computed as the mean of daytime (solar radiation < 50 W/m^2) hourly observations and was only calculated if there are more than 75% of hourly observations within the appropriate time window. Similar to the regression analysis, we removed the variables that exhibited multicollinearity. We also removed both PM_{10} and $PM_{2.5}$ data because they exhibited many missing values. The remaining predictor variables were NO, NO_2, CO, O_3, RH, temperature, solar radiation, wind direction, wind speed and H_2SO_4. We further excluded the bump events from the analysis because the decision trees usually misclassified them. The number of days after removing undefined events and bump events was 279. Finally, we excluded any observation with any missing variable. The total observation days were thus reduced to 184 days (115 event days and 79 non-events days).We did not normalize or log-transform the data because these procedure are not necessary for decision trees.

Analysis:

The classification tree hyperparameters (maximum depth, minimum number of samples required to split an internal node, minimum number of samples required to be at a leaf node and the split criterion) were tuned until the best performance was reached. The performance of the trees was evaluated using performance metrics (accuracy, sensitivity, specificity and precision), 10-fold cross-validation error, and re-substitution error. The outcome decision tree is shown in Figure S23, while the confusion matrix and predictor importance is shown in Figure S24, and the statistics for each node are shown in table S5. The decision tree model had an accuracy (ratio of the correctly labeled days to the whole pool of days) of 89%, a sensitivity (percentage of the labeled events to true events) of 89.6%, a specificity (percentage of labeled non-events to true non-events) of 88.4%, a resubstitution error of 0.1087, and a 10-fold cross-validation error of 0.2337. The decision tree shows that when NO_2 > 0.88 ppb, H_2SO_4 concentration is the most important criteria to determine NPF occurrence, such that events coincided with H_2SO_4 >1.4e^6 molecules.cm^{-3}. When NO_2 < 0.88 ppb, the event occurrence seems to be explained by a combination of wind direction, solar radiation, RH and H_2SO_4 concentration. When the wind is from the N-E or S-W direction, events coincided with RH<54.4%. While at E-S and W-N wind direction, events coincided with solar radiation between 584 and 620 W.m^{-2} and H_2SO_4 concentration>3.1e^6 molecules.cm^{-3}. The analysis also showed that events did not occur when the solar radiation was > 620 W.m^{-2}. Since these last observations cannot be fully supported to explain NPF occurrence, we believe that the analysis is missing further components.

Final remarks:

While the analysis shown in this section provides an important insight into the parameters governing NPF events, it is important to present them with caution because causality cannot be inferred by this analysis. Additionally, the sample size and limitations of available predictors heavily affected the output of both models. A yearlong dataset with missing values cannot provide adequate counting statistics for every NPF case. Therefore, this analysis is useful to discern the importance of certain chemical atmospheric components or physical properties on NPF but cannot be used to predict future nucleation events characteristics

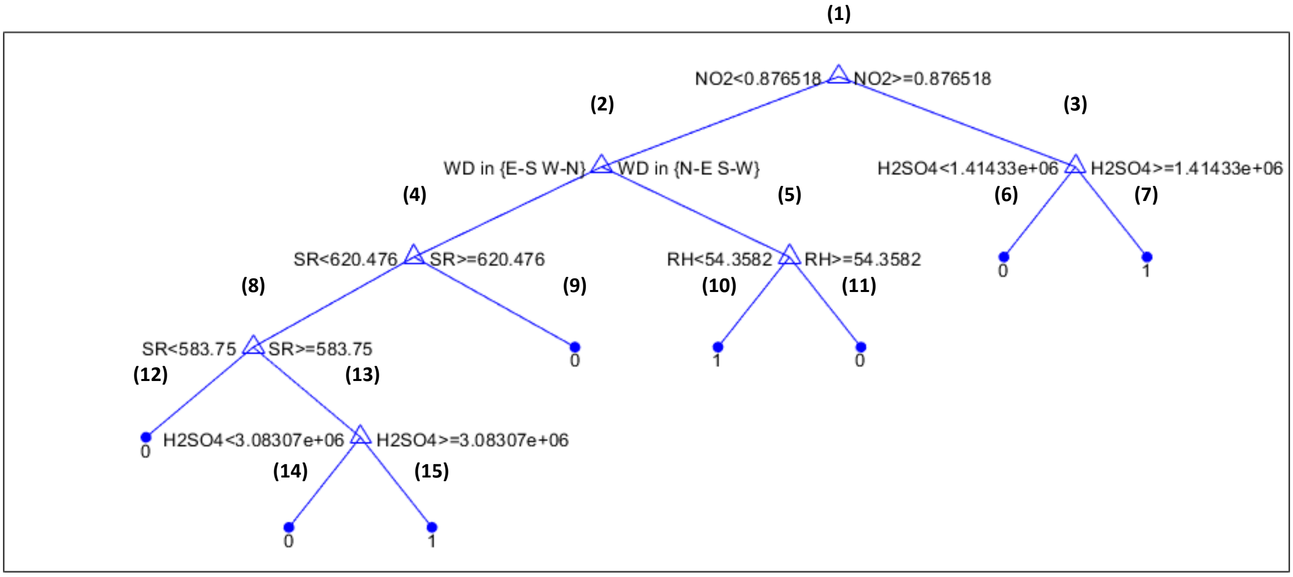

Figure S23. Daily NPF occurrence decision tree. The number of each node is displayed between parentheses above the node. Branch nodes are represented with triangles while leaf nodes are presented with circles. Non-events are represented by 0 and events are represented with 1. WD is wind direction; SR is solar radiation in (W.m^{-2}); NO_2 is in ppb; RH is in % and H_2SO_4 is in molecules.cm^{-3}.

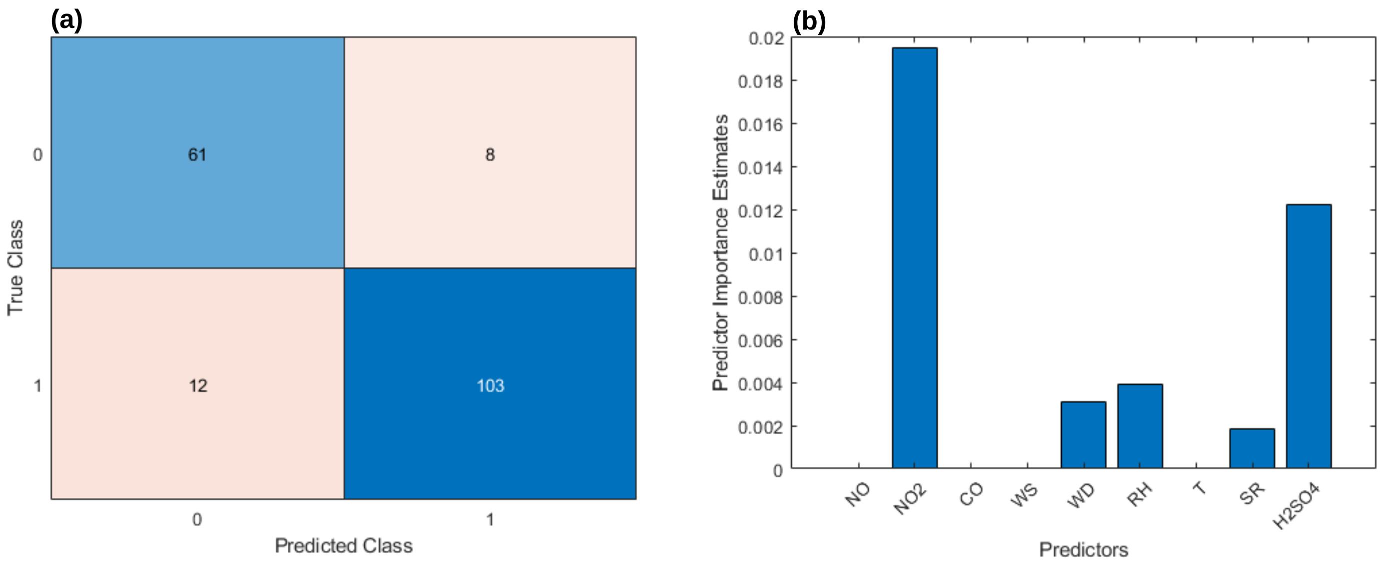

Figure S24. (a) The classification tree confusion matrix where class 0 represents non-event days and class 1 represent event days. (b) The importance estimate of predictors.

Table S5. Description and content of the decision tree nodes

Node no	Node type	Node class	Node size	No of non-events	No of events	Misclassified	Node error
1	branch	event	184	69	115	-	0.38
2	branch	non-event	78	53	25	-	0.32
3	branch	event	106	16	90	-	0.15
4	branch	non-event	60	46	14	-	0.23
5	branch	event	18	7	11	-	0.39
6	leaf	non-event	14	11	3	3	0.21

7	leaf	event	92	5	87	5	0.05
8	branch	non-event	48	35	13	-	0.27
9	leaf	non-event	12	11	1	1	0.08
10	leaf	event	11	1	10	1	0.09
11	leaf	non-event	7	6	1	1	0.14
12	leaf	non-event	33	27	6	6	0.18
13	branch	non-event	15	8	7	-	0.47
14	leaf	non-event	7	6	1	1	0.14
15	leaf	event	8	2	6	2	0.25

Specific comments.

Comment 3: L. 101: The most populated island in the Mediterranean is Sicily, Cyprus is the third most populous.

Adjusted.

Comment 4: L. 103: Also Isreal to the southeast.

Adjusted.

Comment 5: L. 139: How were the data prior to June 2018 treated with regard to activation efficiencies distortion?

We noticed that the original text did not reflect properly the actual problem faced with the CPC. The problem that arose during the summer was that high water vapor content in the sampled air was mixing with the butanol of the CPC resulting in sub-saturated conditions. Thus, the CPC was not able to activate particles at all and was reporting zero concentrations. This problem did not occur before the summer. We updated the text accordingly. We also provide an explanation about the effect of the addition of the diluter in the answers to reviewer 2 (comment 4).

Before correction:

From June 2018 onwards, the PSM was additionally equipped with a diluter to reduce the humidity of the sampled air. This procedure was necessary because the water content of the air at the measurement site was too high, and it affected the activation efficiency inside the CPC and therefore distorted the size distribution measurements for the smallest sizes. Further information about the diluter design and operation can be found in the supplementary information (SI) Sect. 2.2.

After correction:

From June 2018 onwards, the nCNC was additionally equipped with a diluter to reduce the humidity of the sampled air. This procedure was necessary because the water content of the air at the measurement site was too high. The water present in the sample air was mixed with butanol inside the CPC of the nCNC and rendered it measuring zeros. Further information about the diluter design, its operation and effect on the data can be found in the supplementary information (SI) Sect. 2.2.

Comment 6: L. 208: A reference is needed here to support this statement.

We added the suitable references and elaborated on the text as per below

Before correction:

In addition, spectrums of total particles (both neutral and charged) 206 are usually less ambiguous to classify than charged particle spectra (ion mode of NAIS), and the classification 207 of event days may be different if one only looks at these charged spectrums.

After correction:

In addition, spectra of total particles (both neutral and charged) are often easier to visually classify than the charged particle spectra (ion mode of NAIS) because atmospheric nucleation is dominated by neutral processes (Kontkanen et al., 2013;Kulmala et al., 2013;Wagner et al., 2017). In addition, the concentration of the growing mode in the charged spectra is lower for the smaller particle sizes, and increases with diameter as the probability of cluster ions attaching to the growing neutral particles increases (Gonser et al., 2014). Thus, it could be visually difficult to determine if particle nucleation starts from the smallest sizes when looking at charged spectra only. In contrast, one should not neglect looking at charged spectra because it might show sign preference or ion induced nucleation events (Rose et al., 2018).

Comment 7: L. 231: The start and end time are not fully described here, more details should be given.

A description of the start and end time has been added.

Addition to the manuscript:

An event start is determined by an increase in the 2-4 nm particle concentration above the nighttime level, which last for at least an hour. An event end time is determined when the 2-4 nm particle concentration decrease to background levels.

Comment 8: L. 350: The calendar does not contribute to the discussion of the results, it rather occupies a great extent of the given page. I would prefer to move the diurnal patterns from Supplementary material next to annual variations and remove the calendar.

The calendar was moved to the supplement and the diurnal patterns of particle modes were moved from the supplement to the main text as suggested by the reviewer.

Comment 9: L. 387: Since there are only few references of dust relevant NPF event in the literature, these 37 events should be described in more detail and compared to dust free days. At least an example of such possible events should be given.

This comment was addressed in the first comment.

Comment 10: L. 405: How do you support your hypothesis? This is highly speculative.

We deleted the speculative hypothesis.

Comment 11: L.407: I would like to see all these information about Js in a Figure like 2 or 9.

We replaced the tables for formation rates and growth rates with figures and the data will be available on Zenodo.

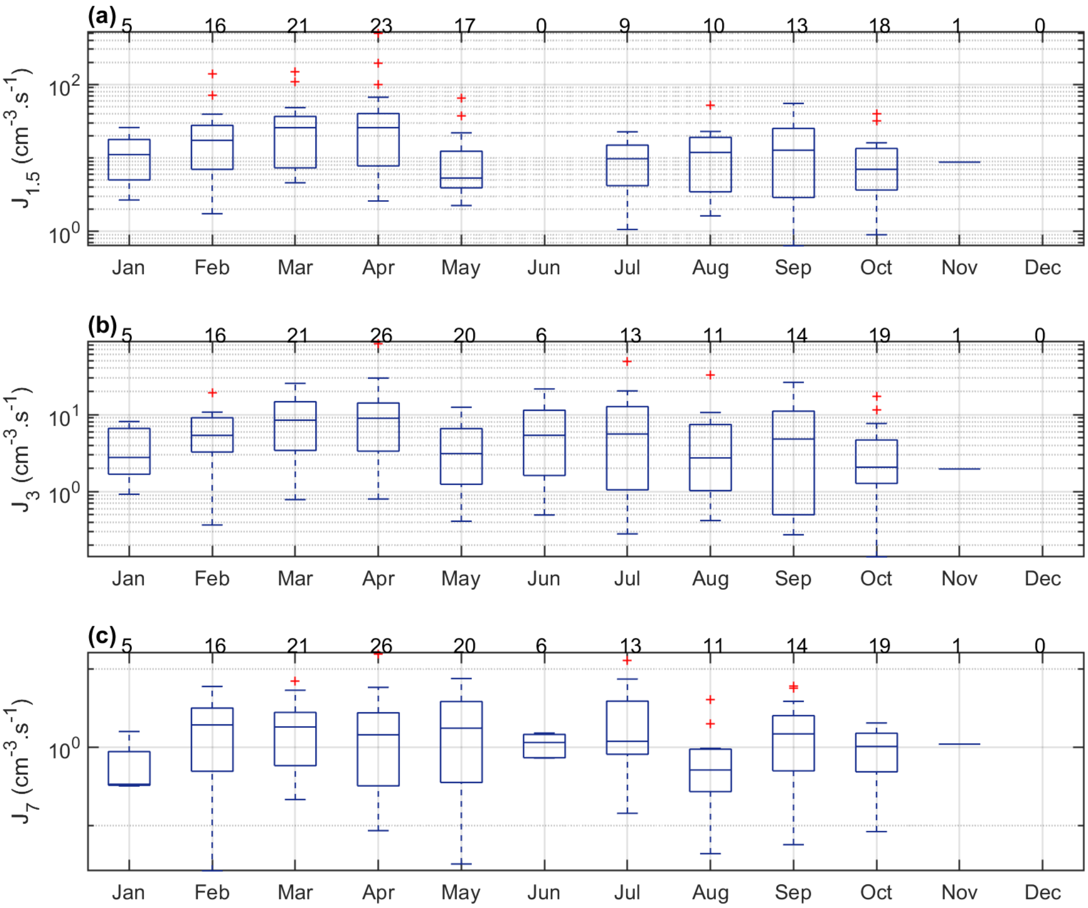

Figure 11. Monthly variation of particle formation rates during NPF events: (a) $J_{1.5}$, (b) J_3 and (c) J_7. The central marks indicate the median, the blue small boxes indicate the mean, the bottom and top edges of the big box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol .The numbers above the box plot represent the number of data points within each boxplot. Data presented have daily time resolution. Daily J values were calculated by taking the mean of hourly values within event duration times.

Comment 12: L. 408: How have the J values reported in Table1 been calculated, ie from average daily J values, maximum daily values, average values during event duration or something else?

We first calculated the daily J values by taking the mean of hourly values within event duration times. Then we calculated the monthly values as medians of daily values within a month. This calculation is now better explained in the caption of designated figure as shown in the responses to comment 11.

Comment 13: L. 433: Once again a figure for GR would be nice here.

The growth rates are now presented in a figure as suggested by the reviewer. Based on reviewer 2 comments, we additionally calculated the growth rates from the positive charged particles and total particles (charged + neutral). The discussion of the growth rates section was updated accordingly (refer to responses to reviewer 2 comment 7).

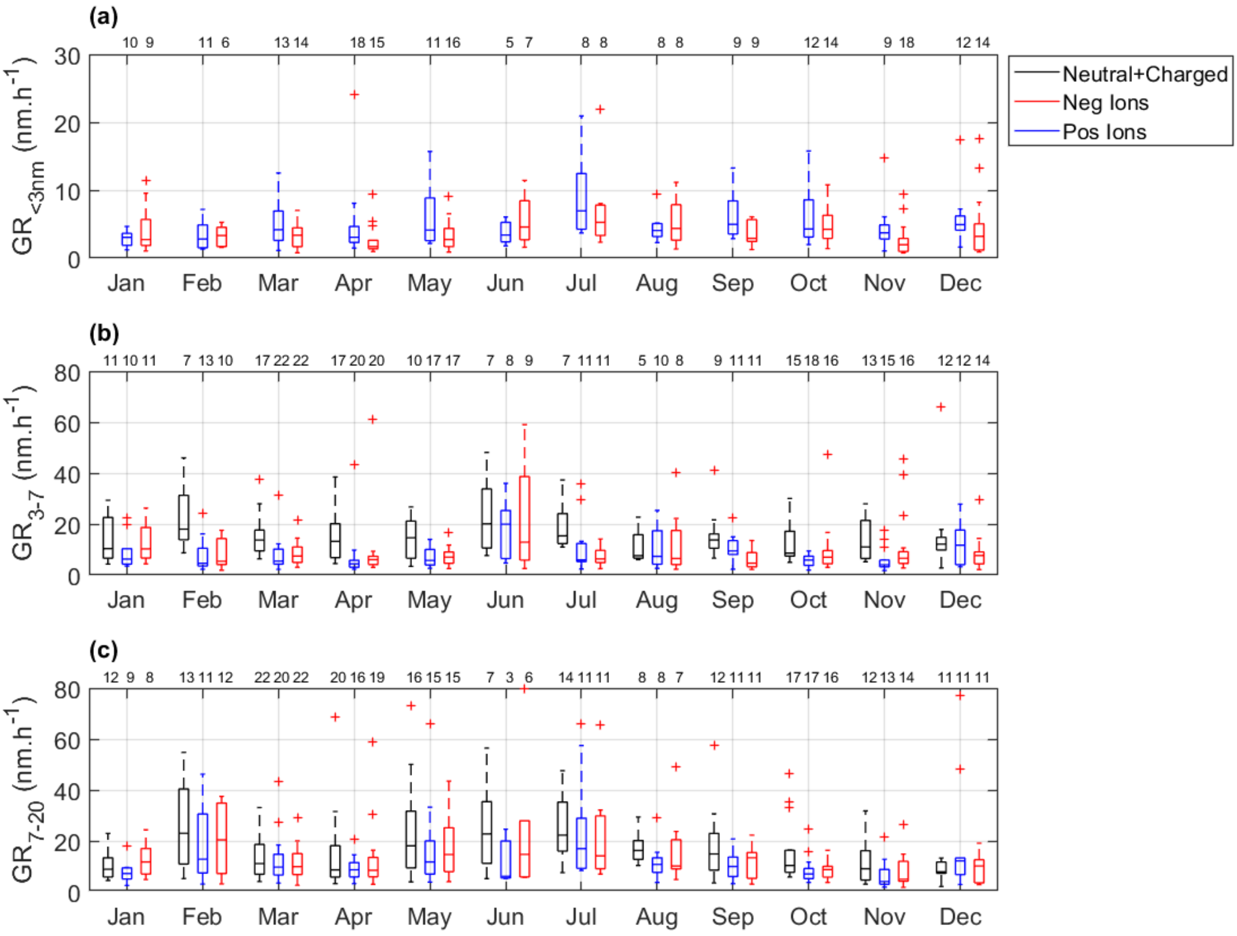

Figure 13. Monthly variation of growth rates during NPF events in three size ranges: (a) <3nm (b) 3-7nm and (c) 7-20nm. The central marks indicate the median, the bottom and top edges of the big box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol .The numbers above the box plot represent the number of data points within each boxplot. Black boxes represent the total particles (neutral+charges), blue boxes represent negative ions and red boxes represent positive ions.

Comment 14: L. 493: However, during the same period, SO_2 concentrations are much higher during events than during non-events, it seems that the SO_2 abundance does make a difference.

Indeed SO_2 abundance makes a difference on an intra-monthly nucleation scale, i.e. when comparing event to non-event days within one particular month. However, SO_2 concentrations could not explain the seasonality in NPF because the month that exhibited lower NPF frequency did not particularly have lower SO_2 concentrations than other month.

Comment 15: L. 515: What compounds could that be? Such an assumption may be investigated looking for instance at SO_2 charts for the region.

The missing compound is most likely a base as later explained in the last paragraph of section 3.4. H_2SO_4 binary nucleation with water requires high H_2SO_4 vapor concentrations that are not atmospherically relevant within the lower parts of the troposphere. Additional species are required to stabilize H_2SO_4 clusters, such as ammonia, amines or ions. From figure 18, we concluded that the NPF is proceeding at formation rates that cannot be explained by H_2SO_4-NH_3 alone, thus additional compounds like organics could be missing (Lehtipalo et al., 2018).

Comment 16: Table 1: Remove the period punctuation mark from the units of J.

The table was replaced by a figure as shown in the response to comment 11 and the period punctuation was removed.

References

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and regression trees, CRC press, 1984.

De Reus, M., Dentener, F., Thomas, A., Borrmann, S., Ström, J., and Lelieveld, J.: Airborne observations of dust aerosol over the North Atlantic Ocean during ACE 2: Indications for heterogeneous ozone destruction, Journal of geophysical Research, 105, 15263-15275, https://doi.org/10.1029/2000JD900164, 2000.

Dupart, Y., King, S. M., Nekat, B., Nowak, A., Wiedensohler, A., Herrmann, H., David, G., Thomas, B., Miffre, A., Rairoux, P., D'Anna, B., and George, C.: Mineral dust photochemistry induces nucleation events in the presence of SO2, Proceedings of the National Academy of Sciences, 109, 20842-20847, https://doi.org/10.1073/pnas.1212297109 2012.

Gonser, S. G., Klein, F., Birmili, W., Größ, J., Kulmala, M., Manninen, H. E., Wiedensohler, A., and Held, A.: Ion – particle interactions during particle formation and growth at a coniferous forest site in central Europe, Atmospheric Chemistry and Physics, 14, 10547-10563, https://doi.org/10.5194/acp-14-10547-2014, 2014.

Kontkanen, J., Lehtinen, K. E. J., Nieminen, T., Manninen, H. E., Lehtipalo, K., Kerminen, V. M., and Kulmala, M.: Estimating the contribution of ion–ion recombination to sub-2 nm cluster concentrations from atmospheric measurements, Atmospheric Chemistry and Physics, 13, 11391-11401, https://doi.org/10.5194/acp-13-11391-2013, 2013.

Kulmala, M., Kontkanen, J., Junninen, H., Lehtipalo, K., Manninen, H. E., Nieminen, T., Petäjä, T., Sipilä, M., Schobesberger, S., Rantala, P., Franchin, A., Jokinen, T., Järvinen, E., Äijälä, M., Kangasluoma, J., Hakala, J., Aalto, P. P., Paasonen, P., Mikkilä, J., Vanhanen, J., Aalto, J., Hakola, H., Makkonen, U., Ruuskanen, T., Mauldin, R. L., Duplissy, J., Vehkamäki, H., Bäck, J., Kortelainen, A., Riipinen, I., Kurtén, T., Johnston, M. V., Smith, J. N., Ehn, M., Mentel, T. F., Lehtinen, K. E. J., Laaksonen, A., Kerminen, V.-M., and Worsnop, D. R.: Direct Observations of Atmospheric Aerosol Nucleation, Science, 339, 943-946, https://doi.org/10.1126/science.1227385 2013.

Lehtipalo, K., Yan, C., Dada, L., Bianchi, F., Xiao, M., Wagner, R., Stolzenburg, D., Ahonen, L. R., Amorim, A., Baccarini, A., Bauer, P. S., Baumgartner, B., Bergen, A., Bernhammer, A.-K., Breitenlechner, M., Brilke, S., Buchholz, A., Mazon, S. B., Chen, D., Chen, X., Dias, A., Dommen, J., Draper, D. C., Duplissy, J., Ehn, M., Finkenzeller, H., Fischer, L., Frege, C., Fuchs, C., Garmash, O., Gordon, H., Hakala, J., He, X., Heikkinen, L., Heinritzi, M., Helm, J. C., Hofbauer, V., Hoyle, C. R., Jokinen, T., Kangasluoma, J., Kerminen, V.-M., Kim, C., Kirkby, J., Kontkanen, J., Kürten, A., Lawler, M. J., Mai, H., Mathot, S., Mauldin, R. L., Molteni, U., Nichman, L., Nie, W., Nieminen, T., Ojdanic, A., Onnela, A., Passananti, M., Petäjä, T., Piel, F., Pospisilova, V., Quéléver, L. L. J., Rissanen, M. P., Rose, C., Sarnela, N., Schallhart, S., Schuchmann, S., Sengupta, K., Simon, M., Sipilä, M., Tauber, C., Tomé, A., Tröstl, J., Väisänen, O., Vogel, A. L., Volkamer, R., Wagner, A. C., Wang, M., Weitz, L., Wimmer, D., Ye, P., Ylisirniö, A., Zha, Q., Carslaw, K. S., Curtius, J., Donahue, N. M., Flagan, R. C., Hansel, A., Riipinen, I., Virtanen, A., Winkler, P. M., Baltensperger, U., Kulmala, M., and Worsnop, D. R.: Multicomponent new particle formation from sulfuric acid, ammonia, and biogenic vapors, Science Advances, 4, eaau5363, https://doi.org/10.1126/sciadv.aau5363 2018.

Ndour, M., Conchon, P., D'Anna, B., Ka, O., and George, C.: Photochemistry of mineral dust surface as a potential atmospheric renoxification process, Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL036662, 2009.

Nie, W., Ding, A., Wang, T., Kerminen, V.-M., George, C., Xue, L., Wang, W., Zhang, Q., Petäjä, T., Qi, X., Gao, X., Wang, X., Yang, X., Fu, C., and Kulmala, M.: Polluted dust promotes new particle formation and growth, Scientific Reports, 4, 6634, https://doi.org/10.1038/srep06634, 2014.

Rose, C., Zha, Q., Dada, L., Yan, C., Lehtipalo, K., Junninen, H., Mazon, S. B., Jokinen, T., Sarnela, N., Sipilä, M., Petäjä, T., Kerminen, V.-M., Bianchi, F., and Kulmala, M.: Observations of biogenic ion-induced cluster formation in the atmosphere, Science Advances, 4, eaar5218, https://doi.org/10.1126/sciadv.aar5218, 2018.

Wagner, R., Yan, C., Lehtipalo, K., Duplissy, J., Nieminen, T., Kangasluoma, J., Ahonen, L. R., Dada, L., Kontkanen, J., Manninen, H. E., Dias, A., Amorim, A., Bauer, P. S., Bergen, A., Bernhammer, A. K., Bianchi, F., Brilke, S., Mazon, S. B., Chen, X., Draper, D. C., Fischer, L., Frege, C., Fuchs, C., Garmash, O., Gordon, H., Hakala, J., Heikkinen, L., Heinritzi, M., Hofbauer, V., Hoyle, C. R., Kirkby, J., Kürten, A., Kvashnin, A. N., Laurila, T., Lawler, M. J., Mai, H., Makhmutov, V., Mauldin Iii, R. L., Molteni, U., Nichman, L., Nie, W., Ojdanic, A., Onnela, A., Piel, F., Quéléver, L. L. J., Rissanen, M. P., Sarnela, N., Schallhart, S., Sengupta, K., Simon, M., Stolzenburg, D., Stozhkov, Y., Tröstl, J., Viisanen, Y., Vogel, A. L., Wagner, A. C., Xiao, M., Ye, P., Baltensperger, U., Curtius, J., Donahue, N. M., Flagan, R. C., Gallagher, M., Hansel, A., Smith, J. N., Tomé, A., Winkler, P. M., Worsnop, D., Ehn, M., Sipilä, M., Kerminen, V. M., Petäjä, T., and Kulmala, M.: The role of ions in new particle formation in the CLOUD chamber, Atmospheric Chemistry and Physics, 17, 15181-15197, https://doi.org/10.5194/acp-17-15181-2017, 2017.

Wang, M., Kong, W., Marten, R., He, X.-C., Chen, D., Pfeifer, J., Heitto, A., Kontkanen, J., Dada, L., Kürten, A., Yli-Juuti, T., Manninen, H. E., Amanatidis, S., Amorim, A., Baalbaki, R., Baccarini, A., Bell, D. M., Bertozzi, B., Bräkling, S., Brilke, S., Murillo, L. C., Chiu, R., Chu, B., De Menezes, L.-P., Duplissy, J., Finkenzeller, H., Carracedo, L. G., Granzin, M., Guida, R., Hansel, A., Hofbauer, V., Krechmer, J., Lehtipalo, K., Lamkaddam, H., Lampimäki, M., Lee, C. P., Makhmutov, V., Marie, G., Mathot, S., Mauldin, R. L., Mentler, B., Müller, T., Onnela, A., Partoll, E., Petäjä, T., Philippov, M., Pospisilova, V., Ranjithkumar, A., Rissanen, M., Rörup, B., Scholz, W., Shen, J., Simon, M., Sipilä, M., Steiner, G., Stolzenburg, D., Tham, Y. J., Tomé, A., Wagner, A. C., Wang, D. S., Wang, Y., Weber, S. K., Winkler, P. M., Wlasits, P. J., Wu, Y., Xiao, M., Ye, Q., Zauner-Wieczorek, M., Zhou, X., Volkamer, R., Riipinen, I., Dommen, J., Curtius, J., Baltensperger, U., Kulmala, M., Worsnop, D. R., Kirkby, J., Seinfeld, J. H., El-Haddad, I., Flagan, R. C., and Donahue, N. M.: Rapid growth of new atmospheric particles by nitric acid and ammonia condensation, Nature, 581, 184-189, https://doi.org/10.1038/s41586-020-2270-4, 2020.

Wildt, J., Mentel, T. F., Kiendler-Scharr, A., Hoffmann, T., Andres, S., Ehn, M., Kleist, E., Müsgen, P., Rohrer, F., Rudich, Y., Springer, M., Tillmann, R., and Wahner, A.: Suppression of new particle formation from monoterpene oxidation by NOx, Atmospheric Chemistry and Physics, 14, 2789-2804, https://doi.org/10.5194/acp-14-2789-2014, 2014.

Xie, Y., Ding, A., Nie, W., Mao, H., Qi, X., Huang, X., Xu, Z., Kerminen, V.-M., Petäjä, T., Chi, X., Virkkula, A., Boy, M., Xue, L., Guo, J., Sun, J., Yang, X., Kulmala, M., and Fu, C.: Enhanced sulfate formation by nitrogen dioxide: Implications from in situ observations at the SORPES station, Journal of Geophysical Research: Atmospheres, 120, 12679-12694, https://doi.org/10.1002/2015JD023607, 2015.

Yamashita, T., Yamashita, K., and Kamimura, R.: A Stepwise AIC Method for Variable Selection in Linear Regression, Communications in Statistics - Theory and Methods, 36, 2395-2403, https://doi.org/10.1080/03610920701215639, 2007.

Yan, C., Nie, W., Vogel, A. L., Dada, L., Lehtipalo, K., Stolzenburg, D., Wagner, R., Rissanen, M. P., Xiao, M., Ahonen, L., Fischer, L., Rose, C., Bianchi, F., Gordon, H., Simon, M., Heinritzi, M., Garmash, O., Roldin, P., Dias, A., Ye, P., Hofbauer, V., Amorim, A., Bauer, P. S., Bergen, A., Bernhammer, A.-K., Breitenlechner, M., Brilke, S., Buchholz, A., Mazon, S. B., Canagaratna, M. R., Chen, X., Ding, A., Dommen, J., Draper, D. C., Duplissy, J., Frege, C., Heyn, C., Guida, R., Hakala, J., Heikkinen, L., Hoyle, C. R., Jokinen, T., Kangasluoma, J., Kirkby, J., Kontkanen, J., Kürten, A., Lawler, M. J., Mai, H., Mathot, S., Mauldin, R. L., Molteni, U., Nichman, L., Nieminen, T., Nowak, J., Ojdanic, A., Onnela, A., Pajunoja, A., Petäjä, T., Piel, F., Quéléver, L. L. J., Sarnela, N., Schallhart, S., Sengupta, K., Sipilä, M., Tomé, A., Tröstl, J., Väisänen, O., Wagner, A. C., Ylisirniö, A., Zha, Q., Baltensperger, U., Carslaw, K. S., Curtius, J., Flagan, R. C., Hansel, A., Riipinen, I., Smith, J. N., Virtanen, A., Winkler, P. M., Donahue, N. M., Kerminen, V.-M., Kulmala, M., Ehn, M., and Worsnop, D. R.: Size-dependent influence of

NOx on the growth rates of organic aerosol particles, Science Advances, 6, eaay4945, https://doi.org/10.1126/sciadv.aay4945 2020.

Zhao, D., Schmitt, S. H., Wang, M., Acir, I. H., Tillmann, R., Tan, Z., Novelli, A., Fuchs, H., Pullinen, I., Wegener, R., Rohrer, F., Wildt, J., Kiendler-Scharr, A., Wahner, A., and Mentel, T. F.: Effects of NOx and SO2 on the secondary organic aerosol formation from photooxidation of α-pinene and limonene, Atmospheric Chemistry and Physics, 18, 1611-1628, https://doi.org/10.5194/acp-18-1611-2018, 2018.