Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-859-RC1, 2016

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Long-term analysis of clear-sky new particle formation events and non-events in Hyytiälä" *by* Lubna Dada et al.

Anonymous Referee #1

Received and published: 4 November 2016

Summary review

This manuscript uses a large dataset of nucleation events from the Hyytiala site to determine a criterion for new particle formation on days with low cloud cover, when most nucleation events at Hyytiala occur. The conclusion of the data analysis might be best paraphrased by "the probability of a new particle formation event on a clear-sky day in Hyytiala springtime is determined almost exclusively by the condensation sink on that day". It should be emphasised that very similar conclusions were reached before by Hyvonen et al in 2005, although with a less sophisticated analysis of a smaller dataset, and the text of the present manuscript could be made more coherent and complete. However, the analysis of the authors does nicely complement previous studies, the data are valuable and appropriately summarised by the plots that are presented, and, overall, the manuscript does represent useful new work.

The paper is within the scope of the journal, the title and abstract are reasonable, and due credit is generally given to previous studies. The quality of the written English is also generally fine. After the authors have improved the interpretation of the data in accordance with my comments, better justified the criterion for new particle formation, and added the full dataset as supplementary materials, I believe the paper would be appropriate for publication in ACP.

General comments

The research follows other studies considering the effect of cloud cover on NPF events (Baranizadeh et al in Boreal Env. Research) and several criteria for predicting new particle formation (papers by Kuang (ACP 10 8469), and Hyvonen, Nieminen as cited in the manuscript), some of which have also considered solar radiation levels. This research shows that cloudiness, condensation sink and temperature, when used together, can effectively predict the probability of a nucleation event at this important field station in spring, but rather less effectively in summer, autumn and winter.

The authors should be more explicit in motivating their work. Yes, aerosols are important for climate and nucleation is an important source of aerosols, but it should be explained more clearly why predicting the probability of a nucleation event in springtime at Hyytiala will help people improve their understanding of the atmosphere. For example, the authors could point out that Hyytiala is reasonably representative of semi-clean forested environments in the Northern Hemisphere, and that it is a suitable site without too many highly localised sources of aerosol that are difficult to model. And, since the authors present a criterion that is only effective in spring, explain that this is the most important season for NPF. And that 20 years of detailed observation data are not readily available at other sites.

The authors investigate several other variables that should be correlated to NPF event probability, but they do not explain why CS and temperature, and not the other variables, feature in their final criterion. For example, the box-and-whisker plots (where the

cloudiness parameter is than 0.7, indicating clear skies) do not clearly suggest that on clear sky days, T offers a better separation between NPF events and non-events than RH. The correlation matrix shows RH and CS are less correlated than T and CS, so RH might have more discriminating power. Moreover, RH on non-event days is almost always higher than on event days, while temperature on non-event days is higher in winter and spring but lower in autumn and winter. Therefore, in principle one might expect RH to be a better second variable than T when all seasons are considered, in line with Hyvonen et al, even after one has separated clear sky days and cloudy days as suggested by Baranizadeh et al already. The authors should quantitatively demonstrate why their criterion offers better discriminating power than a few other obvious possibilities, such as RH/CS.

In addition to the criterion for new particle formation events, the paper also aims to quantify the effect of cloudiness on NPF event frequency. However, this was done already by Baranizadeh et al and it should be made clearer what this manuscript adds to this relatively comprehensive previous work. The authors should either remove all but a very brief summary of this from the paper, or state clearly how their analysis relates to that of Baranizadeh et al with a sentence like “Our work confirms the conclusions of Baranizadeh et al with a complementary dataset”.

The paper also aims to “find out the connection between nucleating precursor vapours and new particle formation rates”. The authors should re-think this part of the manuscript. The analysis has the potential to provide interesting conclusions, but currently it is not well connected to the rest of the paper and the approach of the authors does not match the stated aim. The sentence I quote here is misleading because the connection is assumed by the manuscript, not “found”: the new particle formation rates presented in the paper are not calculated from the rate of change of particle concentration, but from a parameterisation of the nucleating precursor vapour proxy concentrations. This is not a bad approach but just needs to be described more carefully: the comparison of the probability of a new particle formation event to the parameterised

nucleation rate is still a useful exercise.

To connect this to the rest of the paper, the authors could consider presenting this study as a comparison of the effectiveness of their condensation sink based criterion and their nucleation rate parameterisation at determining whether or not a nucleation event will occur on a given day. Then the conclusion might indicate explicitly that the parameterisation is a poor criterion for NPF compared to the condensation sink-temperature criterion (except perhaps in winter, from Figure 10?). Since parametrisations of this form might be considered reasonable starting points to determine whether or not NPF should occur (naively, a high parameterised rate ought to imply a nucleation event is likely), this would seem to be an interesting message. While Figure 10 is helpful in providing this message, further evidence could be obtained by re-plotting some of the data so that the criteria can be compared more directly: Figures 5a and 9a can be compared, but it would be better if the combined criterion including temperature were plotted on the y axis of a new version of Fig. 5a, since this should further improve the separation.

The careful statistical summaries presented in this manuscript do convince the reader that the underlying dataset is valuable. The large size allows statistically significant results to be extracted. A csv table (or similar) containing the full dataset, or at a minimum a list of dates studied over the 20 years with, on each day, the condensation sink, temperature, RH, cloudiness parameter, and whether or not the day contained a nucleation event, should be included in supplementary materials. While much of this information is already available, via the smartSMEAR website for example, a carefully compiled dataset specific to this paper would still be very useful. It would allow, for example, modellers comparing event frequency in models and observations to split up the dataset into individual years and compare model to measurements selected from the overall dataset to match their model simulations. A brief explanation of where subsets of the data have been published before should accompany the data file.

Specific comments

Summary: the sentence “This study serves as basis for scientists aiming at improving their understanding towards new particle formation” should be rephrased to improve the written English, for example “This study serves as a basis for scientists aiming to improve their understanding of new particle formation.”

Abstract: “utilizing”->”building on” “In this comparison we considered, for example, the effects of calculated particle formation rates, condensation sink, trace gas concentrations and various meteorological quantities.” -> considered the effect on what? “The formation rate of 1.5 nm particles was calculated by using proxies for gaseous sulfuric acid and oxidized products of low volatile organic compounds”-> add “and a nucleation rate parameterization” after “compounds” “As expected, our results indicate an increase in the frequency of NPF events under clear-sky conditions.“-> “increase under clear-sky conditions compared to cloudy conditions”

“The calculated formation rate of 3 nm particles showed a notable difference between the NPF event and non-event days during clear-sky conditions, especially in winter and spring”-> so in cloudy conditions do you get high NPF rates but no events? Please be more explicit here.

Line 59: “That study” -> which of the three cited?

Line 88: The title of section 2.1.1 should be amended to make it clear that it is this section which explains how events are categorised, and this section should be extended with a very brief summary of how Dal Maso et al decide whether a day is an event, non-event or undefined day.

Line 127: At least four possible MT proxies are presented in Kontkanen et al. Which one did you use? Is it the recommended proxy MTproxy1,doy? Please specify.

Line 140-147: this section needs more detail on the data analysed and the characteristics of NPF in Hyytiala. Table 1 caption implies all of the data analysed are from the months of March to May, but this seems not to be true. However, it is clear that the

instruments would not be running every single day between 1 January 1996 and 31 December 2015. While Figure 2 is helpful here, it also needs some additional explanation and referencing early in the text. The brief statements about the seasonal cycle at lines 227-232 are confusing without this additional context.

Line 160: Somewhere here it would be good to state why you calculate J3, why not just use J1.5?

Line 174: Please summarise very briefly the improvement made by Kontkanen (2016).

Line 197: It would be helpful to state the number of undefined days here, so the reader does not wonder how it can be that 877 days are events, 229 are non-events, and 55% of days are event days.

Line 205 “days having less” -> “days with fewer”

Lines 204-209: Please rewrite or combine with the previous paragraph to ensure the message of this paragraph does not repeat the message of the previous paragraph

Line 213: “In order to find out clear results and conclusions, we will focus on comparison between NPF events and non-events in following sections.“-> this is long-winded, could shorten to “Undefined events are not considered further in the analysis”

Line 230: If the annual trend is important to note, state explicitly what is the annual trend.

Line 235: What is the median and percentile of a trajectory? The median compass direction at the point on the trajectory where it arrives at Hyytiala, or the median compass direction of some kind of average over the length of the trajectory? Does “at every half hour” mean for the arrival of the air masses at Hyytiala every half hour or for one trajectory, moving back along it by half an hour at a time?

Line 252: “However, the monthly cycle of CS on non-event days had two maxima, one in spring and another one in autumn”- what is the reader meant to conclude from this

sentence?

Line 257: “The temperature at which clear-sky NPF events occurred was different for each month“ - > The following sentences are not really ‘examples’ of this sentence. I would delete this.

Line 266: “even though it might also be attributed to the presumable increase” -> but it might also be attributed to the increase”

Line 279-281: this sentence needs a verb outside the “while” clause

Line 281: Increased RH leads to increased production of H2SO4. Additionally, even with constant H2SO4 concentrations, nucleation rates increase with RH in flow tube or chamber studies (e.g. Duplissy et al, JGR 2015) and are expected to from theory (Merikanto et al, JGR 2015, Vehkamaki et al 2002). However, it is indeed clear from Fig 5c that RH is negatively correlated with nucleation. This could be due to any number of reasons, but it seems odd to point out the Boy & Kulmala study on RH limiting VOC ozonolysis without discussing the far more robust and well-established evidence from atmospheric chemistry that RH should promote nucleation of sulphuric acid.

Line 303: Specify that the OxOrg proxy concentration depends on temperature via the MT proxy in Kontkanen et al. Also see previous comment concerning this proxy.

Line 323: The comparison between J1.5 and J3 is interesting but should be made more explicit – how much later is the peaking time of J3 than J1.5? Are Figures 8b and 9b the same, or are there differences? Would you expect differences, based on how long it takes particles to grow from 1.5 to 3nm in general?

Line 331 It is stated that figure 10 represents “median diurnal cycles”. The description of what this actually means is currently a bit hard to follow, and it needs to be repeated in the figure caption. If I understand correctly, the median CS is calculated for each half hour, and plotted against the corresponding median J3 value. Perhaps it would be clearer to describe the plot by saying “the J3 and CS data were divided into 30 groups

according to the time of day at which the data were recorded, and the median J3 and CS values for each group were calculated. The first group of data were recorded between 5am and 5.30am, the next from 5.30am to 6am, and so on until 8pm local time” (with adjustments for the precise times/numbers of groups). The figure would also be clearer if the scales of the axes were better optimised so that the data extend closer to the extremes of the axis ranges.

Line 342. From Figure 10, it is interesting that in summer, autumn and winter the highest J3 on non-event days is almost as high as on event days, and one would therefore expect the J1.5 to be very similar to the J1.5 on event days. This is in sharp contrast to the large differences between event and non-event days shown in Figure 9b for spring. For autumn, winter and summer, figure 10 would imply that the J rate by itself is a poor predictor for whether or not an event will occur. This is surely a useful message for your paper: it could be used to emphasise the importance of your new discriminating variable, based on condensation sink, which from Figure 10 clearly should perform better than the nucleation rate, which naively sounds like a more obvious variable to determine whether or not a nucleation event is occurring.

Figure 2: In addition to the helpful rows of numbers presented below the box plot, it would be helpful to state the number of event and non-event days with P > 0.7 in two additional rows.

Figure 12: why does the criterion for NPF you have developed perform badly in summer, autumn and winter?

It would be possible to determine quantitatively the benefit of the clear-sky classification by applying the NPF criterion in the clear-sky case and also without first separating clear sky events from non-events. Please state the effectiveness of the criterion in the case where you do not distinguish clear-sky and cloudy events, in order to prove the usefulness of the clear-sky distinction by showing that the NPF criterion is less effective without it.

References: should cite Kuang et al, ACP 2010, “An improved criterion for new particle formation in diverse atmospheric environments” somewhere

Table 1: +/-0.45 does not really indicate “high correlation”: for this description to be justified I think you need +/-0.7 at least! Also, since the tables are symmetric about the diagonal, please remove the lower triangle (or replace with “-“) so the reader does not have to check the upper and lower triangles are the same.

Figures 1/2: what is the “relevant statistical limit”?

Figure 3 caption: “5.4%, (add comma) making the classification biased.” Please state more explicitly what you mean here : do you have only global radiation data for 5.4% of the days in 1998?

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-859, 2016.