

We thank both reviewers for catching some important oversights in the paper and suggesting other improvements. We feel that the paper is much improved because of their efforts. Our responses to their comments are below.

*Reviewer 2*

*In their manuscript Pierce et al. present the statistics on new particle formation (NPF) and the growth of the formed particles to sizes where they can act as cloud condensation nuclei, i.e. to diameters > 50-100 nm. The authors calculate the survival probabilities of the particles to these sizes and estimate the share of NPF originated particles from total > 50-100 nm particle concentrations. They also examine how the main variables related to NPF and particle growth (except for organic vapors) affect the existence and type of the NPF event.*

*I have some detailed comments on the manuscript, as listed below. In general, the manuscript doesn't have very novel ideas or perspectives. However, it brings forth valuable data over a full year for modeling and data analysis purposes and thus I find it would be worth publishing in ACP after my comments have been satisfactorily responded to and the required changes have been made.*

*General comments*

*I think the way the authors have applied the NPF event classification based on Dal Maso et al (2005) is not as close to the original Dal Maso classification as the authors claim, and it is not as meaningful as it could be if applied a bit more strictly. Dal Maso et al. state that both classes I and II are regional NPF events, out of which class I events are those for which growth and formation rates can be determined with good confidence level (and class II those for which not). The division to Ia and Ib is secondary process, in which the events are chosen to class Ia if the event can be easily distinguished from the pre-existing particles and/or particles from other sources. In Dal Maso et al. (2005) the undefined class consists of those days during which the occurrence of regional NPF event cannot be clearly determined. Here, Pierce et al. have, as they repetitively state, in class Ib days for which they are not certain whether regional NPF event occurred or not. Additionally, they have the class II days, named as events, during many of which it seems quite certain that regional NPF event did not occur. If the new nucleation mode does not show signs of growth it most likely is not NPF event (as in this context the word event points to a regional scale phenomenon). Why is this important? Firstly, because the authors show figure in which there is e.g. ~95 % event frequency on February, which is definitely not the case. I do not find it enough, or even reasonable, to state in the text that also undefined days are counted in the event class II. Secondly, they claim to use Dal Maso et al methodology, except for class II definition, but the division between Ia, Ib and II is not based on the same definitions. These problems the authors can of course get rid of with few modifications in the text (e.g. replacing class II events with class II days, and changing few words on page 714). However, following the Dal Maso classification more strictly would make the comparison to different sites easier, and could also make clearer differences for different categories in figures 4-10.*

We thank the reviewer for this clarification. We had initially classified the events based on what we thought the definitions in Dal Maso were (lesson learned that it pays to re-look through a paper before you go through a time-consuming task that is based on that paper), and we realized that we were not correct after we had done all of the classifications. However, we should have realized, as the reviewer points out, that our class 1b days are really Dal Maso's class II days (likely regional events, but little confidence in the formation and growth rates) and our class 2 days are really closer to Dal Maso's undefined events (low confidence that they are regional events). The reviewer suggests “replacing class II events with class II days”, we assume this is a typo and that they mean that we should move our class 1b days to class II and move our class 2 days to undefined. We did this change and called our 1a

days just class I days (we left out the distinction of 1a vs 1b as the growing nucleation mode was generally easy to distinguish from the pre-existing background).

*The second general comment is that the quality of the language fluctuates too much. Especially Sect. 2 is partly difficult to read. E.g. the tense changes from present to past and back within several paragraphs and there are some sentences up to over 6 lines (page 714 lines 19-25). Some needs for corrections are listed in Technical comments, but the authors should read the manuscript through more times in detail.*

Regarding tense. The co-authors and I (Jeff) went back and forth on what was best for tense. I ultimately ended up trying to use past tense when referring to measurements that are complete and present tense for discussion about our methods and general phenomena that always occur (e.g. Egbert air is influenced by Toronto when winds are from the south). Thus, this means that for Section 2, the tense may change between sentences if we are discussing the measurements in one sentence and our analysis technique in the next. However, after going back through the text, I definitely was not strictly consistent with this.

We have gone through the text and have changed most sentences to present tense (unless it was inappropriate to do so).

*Specific comments:*

*The authors should not discuss nucleation rates but new particle formation rates, as the measurements are from 10 nm up. Now these terms are continuously mixed, first time in abstract (page 708, lines 8 and 20). Additionally, it should be kept in mind that new particle formation means nucleation and initial growth, thus saying “nucleation and new particle formation” (page 710, line 7) is not reasonable.*

Changed to new-particle formation throughout.

*Page 714, L15-18: The example in Fig 1a does not show homogenous air mass in terms of particle growth, as there are clear jumps and plateaus in nucleation mode diameters.*

We have updated the text to read as, “Class 1a days are most likely widespread, regional new-particle formation events with a relatively homogeneous air mass advecting over the Egbert measurement site. The example in Figure 1a shows an air mass that is not completely homogeneous as the growth in the nucleation mode is not smooth. However, we are still able to retrieve formation and growth rates on these days.”

*Page 715, L 26: How is peak diameter defined? Maximum concentration or mode fitting?*

Maximum concentration. We have clarified in the text.

*It would be useful to determine a symbol for 24-h average J10. Now e.g. page 719 lines 11 and 12 are confusing, first “24-mean” seems to refer to single day mean, and on the next line 24 h mean is the average of these daily mean values.*

Yes, this was written confusingly. We have updated the text to be more clear, “For 1a days, the annual mean and median values of the 24-h mean formation rates were 0.13 and 0.12  $\text{cm}^{-3} \text{s}^{-1}$ , respectively. Westervelt et al. (2013) presented 24-hour-mean new-particle formation rate statistics at 3 nm (J3) for 5 locations (Pittsburgh, Hyytiälä, Atlanta, St. Louis and the Po Valley). The observed annual means for

the 24-hour J3s at these locations ranged from 0.58 to 8.7 cm<sup>-3</sup> s<sup>-1</sup>, and the annual medians ranged from 0.09 to 0.55 cm<sup>-3</sup> s<sup>-1</sup>.”

*The authors note that the comparison between this work and Westervelt et al. (2013) is not straight forward because of different cut of diameters in measured size distributions. The authors should be able to back calculate J3 with their model PUG to have better comparison between Egbert and sites studied by Westervelt et al., since the growth rates should be roughly comparable (Westervelt et al. used growth rate from 3 to 25 nm and here authors use from 10 to 25 nm). Even though there were some additional uncertainties due to this, it would be useful for comparing the SPs.*

Good idea, we have done this and added the sentence, “We estimate the mean survival probability between 3 nm and 10 nm to be ~0.25 for the Egbert study, which gives us estimated mean/median J3s of about 0.5 cm<sup>-3</sup> s<sup>-1</sup>, in line with the estimates of Westervelt et al. (2013).”

*Page 722, line 19-27: Comparison between variables during the event time and 24 h average on non-event days is not relevant. Even though the authors state on the next page that this difference in time window causes probably a part of the differences e.g. in Fig. 4, I find it would be more appropriate to use the most typical event time window for the non-event days as well.*

This is a good idea. We have updated Figure 4 and the discussion using data from non-event days taken between the mean start and end times of class I days (~11:00-16:00). This changed the solar radiation and temperature distributions on non-event days, as would be expected.

*Page 726, L 26-27. Please, describe better what the statistically significant means here. Is it that occurrence of e.g. 1a class days correlates positively with surface pressure with the referred significance?*

Modified text to say, “Regions with a statistically significant (95% confidence relative to randomly chosen sets of days) high surface pressure anomaly are shaded in pink, and regions where there is a statistically significant (95% confidence) low surface pressure anomaly are shaded in blue.”

The full explanation is in Appendix A, and we refer to it in the sentence following the one above, “Statistical significance is computed following the bootstrap method (Efron, 1979; see Appendix A for details).”

*Figures 5 and 6 should have panels also for non-event days for comparison, or if those figures do not show anything worth showing, it should be mentioned.*

We have added, “(non-event days show only small deviations from the mean, so we have not plotted non-event days here)” to the text.

*Trajectory analysis. The authors show four figures on trajectories describing from which directions different types of air masses come to Egbert. I find this too many, e.g. figures 8 and 10 could be just explained in text. More importantly, the authors do not discuss the frequencies and shares of different event classes and non-events related to air masses from different directions. From figures 7-10 it looks like most often the air mass comes from west or northwest (when non-events are not considered), but especially the class 1a events are as often related to southern air masses than to north(west)ern. For class 1a western trajectories are very rare. It would be important also to show, for comparison, the trajectories for non-event days e.g. for the hour most frequently applied for class 1a and 1b event trajectories.*

We have removed the figure for SO<sub>2</sub> and the figure for solar radiation, and we just discussed them in the text as you have suggested. Additionally, we added more discussion about the directionality of event days.

*Would it be more informative to calculate sum of diurnal J in units #/cm<sup>3</sup>/d instead of average diurnal J in units #/cm<sup>3</sup>/s?*

We see advantages of both units. We choose to leave as the latter to be consistent with Westervelt et al. (2013)

*Technical comments:  
Page 709, line 19: effect, not affect*

Fixed.

*Page 711, line 21: CAN enhance*

Fixed

*Page 712, line 22-24: this sentence could be removed, it is explained in the next subsection.*

Ok

*Page 714, line 2-3: e.g. here I find past tense should be used*

See the main comment on tense

*Page 715, line 22: in all examples in Fig. 1 new particle formation period is longer than 2-4 h.*

Removed in text

*Page 719, line 22: This sentence does not sound right (“This” and “range as”)*

Changed. “These mean and median values are at the low end of the range in Westervelt et al. (2013) at the 5 locations.”

*Page 724, line 24: Should probably be “higher than class 1b and 2 days”*

Yes, fixed.

*E.g. Page 722 Many paragraphs end with “ in the next subsection. . .”. The authors should think whether it is necessary to state it that repetitively.*

We looked through this. In each case we feel that leaving it out may make the reader feel that this discussion is ending short of where it should, but we want to re-assure them that we will pick up the discussion shortly.

*Page 725, line 20: should this be “both SR and SO<sub>2</sub>”?*

Yes, fixed.

*Page 725, line 21-22: In figure 4 the values indicate statistical difference also at 95 % level*

Correct, fixed.

*Page 726, line 9-12: Note that the logarithmic dependence of J on H<sub>2</sub>SO<sub>4</sub> is also in agreement with nucleation theorem, and as H<sub>2</sub>SO<sub>4</sub> is linearly dependent on SO<sub>2</sub> and CS (and SR), also the dependence between J and these variable should be logarithmic.*

Added discussion.

*Page 727, line 8: Should be "(Fig 4d)"*

Fixed

*Page 727, line 14-20: Does this (Fig 6) bring any additional information to Fig 5?*

We feel it strengthens the case about what the synoptic meteorology is on Class I days.

*Page 727, line 26-27 and earlier. Class 1a could be called clear and strong regional NPF events. Most probably at least many of 1b are also regional events.*

Discuss once I've updated.

*Page 727, line 28: something strange in words here.*

Yes, the word "subsidence" was missing from the sentence.