We thank both referees for their thorough review on our manuscript. The comments were extremely valuable and we have redone most of the analysis based on them. The main weakness of the previous submitted version was the poor performance in estimating the growth rate with the mode fitting method, which further meant poorly predicted time lag and poor performance in the time resolved formation rate comparison. We have now reanalyzed the growth rates with the so-called maximum concentration method and the results are overall muck better. All figures and table 1 are modified accordingly, and we also removed the old Figure 3 comparing time-lags, as we believe that it is unnecessary in the new version. We also removed the standard error color coding, related to the uncertainties when determining GR, from Figure 1.

Below we give our detailed responses to the referees' comments.

The revised version of the manuscript, showing all the changes made in the text, is included after our replies.

Anonymous Referee #1

General Comments:
A thorough, data-based evaluation of whether particle formation rates can be extrapolated from measurements at larger sizes, as attempted by this paper, is vital for the aerosol community as so many data exist with only larger size information available. While the method used to tackle this problem is valid and useful, the evaluation requires development and more nuanced analysis before the substantial conclusions stated in the paper can fairly be reached (see below for specific comments).

Specific Comments:
A major assumption, that the two measurement sites are directly comparable with the method used for extrapolating nucleation rates, is made in the paper. Kurten et al showed that the method used, while valid in many circumstances, may not be valid for situations where pre-existing populations of aerosols do not dominate the coagulation sink and newly formed particles play a larger role in this sink. The differences in background aerosols at the two sites should be discussed in relation to this. Differences between the two site may also influence the magnitude of growth rates and coagulation sinks, which may affect the accuracy of the J extrapolation, which should be addressed.
The referee has a valid point that, generally, differences in aerosol dynamics (i.e which processes are dominant for the growth and loss of the newly formed particles) between two sites could potentially lead to erroneous conclusions when comparing the scaled J values. However, here this is not the case. The background distributions in our two sites are quite similar, both in total concentration and mode location. The mean values of CoagS of 7 nm particles are 5.3077e-5 s^{-1} and 5.3272e-5 s^{-1} in Hyytiälä (mean value of all analyzed NPF event days during 2002-2012) and Puijo (mean value of all analyzed NPF event days during 2007-2015), respectively. For both sites the nucleation mode concentrations are so small that both the contribution of self-coagulation on growth as well as the contribution of newly formed particles on the sink are negligible. The contribution of particles of different sizes on the sink has been investigated by Lehtinen et al. (Boreal Environment Research 8, p. 405-411, 2003 – see fig. 3) for Hyytiälä size distributions. Particles below ca. 50 nm in diameter

have typically negligible effect on condensation/coagulation sinks. As Puijo size distributions are very similar, this conclusion holds also there.
We added after equation 1 on line 110 “Lehtinen et al. (2003) studied the contribution of particles of different sizes to the condensation sink at Hyytiälä and found that particles below 50 nm in diameter have typically negligible contribution. This is a reasonable assumption at Puijo also as the concentrations and size distributions are similar to those at Hyytiälä. The mean values of CoagS of 7 nm particles are 5.31e-5 s^{-1} and 5.33e-5 s^{-1} in Hyytiälä (event days during 2002-2012) and Puijo (event days during 2007-2015), respectively.”

Line 59: The assumptions that the coagulation sink is time independent and the growth rate size independent should be more fully investigates. Kurten et al. highlights the possibility and affect of time dependent coagulation sinks. If this is not a problem for these two sites it should be explained more explicitly.
This is true and in our analysis we do not take time dependence of CoagS and GR into account. This is, however, intentional from our part since we wish to follow the procedure of Kulmala et al. (Nature Protocols) in order to analyze formation rates consistently with most other studies previously analyzed.
Below in fig. R1 we show the median diurnal variation for CoagS (3nm). We also added in the revised manuscript Fig. 2 the CoagS time evolution for each of the three example events. It is clear that there may be significant time evolution in the CoagS/GR term of Equation 1, which is of course one of the key reasons why the simple approximation equation is not perfect.
The mentioned assumptions are mentioned in the text after Equation 1, but to clarify this, we added to the conclusions (lines 253-256 of the revised manuscript): “when considering detailed daily time evolution, the agreement is not as good. This is caused by three main things: 1. there are significant fluctuations in experimental size distribution data, 2. the extrapolation method assumes a constant value for CoagS/GR, and 3 there is a time lag between J_3 and J_7 and a poor estimation of the growth rate *GR* results in comparing values at different times.”

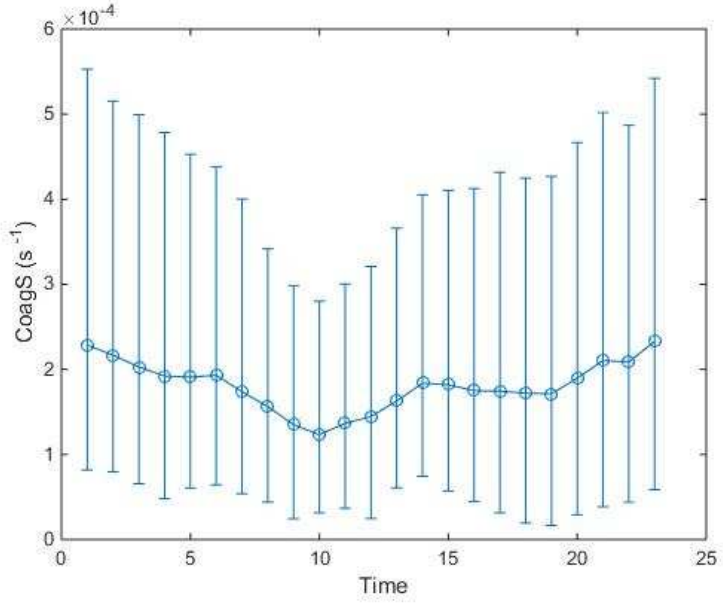


Figure R1. The median diurnal variation of coagulation sink of 3nm particles for all the NPF

events analyzed in this study in Hyytiälä. The error bars indicate the 25th and 75th percentiles of the CoagS data.

Line 75: some discussion of how the different environments of Hyytiala and Kuopio affect the average size distributions and patterns of nucleation would be helpful here. This can affect how accurately equation 1 can be applied. Equation 1 assumes that the coagulation sink is dominated by larger pre-existing populations, which is less applicable in cleaner environments. If the Hyytiala environment is much cleaner, for example, than Kuopio, then the two situations are less comparable for this method of calculating formation rates. This assumption is mentioned on line 114, but it's validity for both situations requires further discussion.
See our reply to the first comment above: nucleation mode has negligible contribution to CoagS both in Hyytiälä and in Puijo. The coagulation sink levels in Hyytiälä and Puijo are very similar.

Line 145: averaging of m and CoagS(d1) between t and t' may be inaccurate, especially for high Js and low GRs – is the any indication of this in the data? How much do m and CoagS(d1) differ between t and t'?
The median variation of the CoagS over all the analyzed NPF event days is shown in Figure R1 above (not included in the revised manuscript). To illustrate the temporal variation we added into Figure 2 also the time evolution of CoagS for the selected three NPF events. This variation naturally limits the validity of the constant CoagS assumption when applying Eq. 1. We also added at the discussion of figure 1 (lines 196-199 of the revised manuscript): "This is caused by three main things: 1. there are significant fluctuations in experimental size distribution data, 2. the extrapolation method assumes a constant value for CoagS/GR, and 3 there is a time lag between J_3 and J_7 and a poor estimation of the growth rate *GR* results in comparing values at different times."
The effect of the CoagS variation on m, and further on gamma is, however, minor. We now mention this when discussing the result in Figure 1 (lines 199-200 of the revised manuscript): "The variation of CoagS with time also affects *m* and γ in equation 1. This is, however, negligible as CoagS(7 nm)/CoagS(3 nm) is a very weak function of time."

Line 172: This discussion of how well J3,est and J3,obs agree needs further development. Suggest removing qualitative judgement of 'reasonably well', and leaving only quantitative measurements of this. While the 0.78 correlation coefficient is helpful, the (linear?) fit result that this relates to would give a better measurement of the systematic difference between the two, this needs to be given here and on figure 1. This would then also quantify the following assertion that equation 4 overestimates the formation rate.
We have now removed 'reasonably well' and added the linear regression line to Figure 1 as the referee suggested. We also now show the results for both growth rate ranges studied: 3-10 nm and 7-20 nm. Note that, as mentioned at the beginning, we have redone all calculations – now using the maximum concentration method to determine the growth rate. Now, especially the time resolved comparison shows a much better result than previously. The slopes and correlation coefficients for the regression lines are 0.90 and 0.90 for the mean J_3 values and 0.87 and 0.83 for the time resolved ones, respectively. There is a slight overestimation bias for small and underestimation for large J_3 values. We have added this to the discussion of

Figure 1 in the manuscript (beginning of Section 3.1): "Figure 1 shows the comparison of estimated formation rates $J_{3,est}$ (Eq. (4)) with the observed ones $J_{3,obs}$, as calculated directly from the measured size distribution evolution according to Eq. (2) in Hyytiälä. In the top figures, the range 3-10 nm is used to evaluate the growth rate, in the bottom ones 7-20 nm. We analyzed 65 NPF event days for which the formation and growth rates could be quantified. Each data point in Figures 1-b and 1-d represents the arithmetic mean of the 3-nm particle formation rates ($J_{3,est}$ and $J_{3,obs}$) for a single NPF day during the time window from 07:00 to 19:00 local time. The mean is also a measure of the total particle production strength of each event. The results show that, when using GR in the range 3-10 nm, the estimated mean $J_{3,est}$values correlate with $J_{3,obs}$ with a correlation coefficient of 0.90 and a slope of 0.90 using bilinear fitting. Furthermore, 91 % of estimated $J_{3,est}$ are within a factor of two of the observed $J_{3,obs}$. The corresponding numbers when using GR in the range 7-20 nm are 0.92, 0.87 and 93%. Equation (4) seems to have a tendency of slightly overestimating the formation rate of 3-nm particles. There is not much difference in the results with different GR size ranges. The total means of $J_{3,obs}$ and $J_{3,est}$ (not shown in the figure) calculated using $GR_{3\text{-}10}$ are 0.57 and 0.61 $cm^{-3}\,s^{-1}$, confirming the tendency of Eq. (4) in slightly overestimating the 3-nm particle formation rates."

Line 175: Standard deviation should be given along with the daily means. These means are taken over a long period of time, during which I suspect J varies quite a bit. If J does vary a lot of this time period, then taking a daily mean is not very meaningful.
We chose not to add the standard deviations to the plot with means as the time resolved all-data-plot reveals the variation in J-values. The daily mean is meaningful in the sense that it is a measure of the overall strength of a nucleation event. Another variable choice would be the total number of particles produced at some size, but as most of the existing literature reports rates, we chose this approach.

Line 184: The quoted daily median values of J are very small. Where in the day do they actually occur? Is it actually before a significant nucleation event occurs? If so these values are not very meaningful – suggest either cutting data to only encompass the nucleation event or finding a more meaningful statistic here.
The median values were calculated over the same time window (07-19) as the mean values given in the text. We, however, agree with the referee that median is not necessarily the best statistic to use here, and decided to remove the median values from the text altogether.

Line 188: I would argue that reduction of percentage of points within a factor 2 of Jobs from 85% to 78% still reasonably significant and could indeed indicate that GR3-10 different and more accurate than GR7-20, which has strong implications for the conclusion that it's ok to use this GR in extrapolating results from Puijo. It would be more meaningful here to look at the fit equation again rather than simply correlation and percentage within factor 2 to understand this difference better.
We have now fitted the scatter plots of $J_{3,est}$ vs. $J_{3,obs}$ as suggested by the referee, and explained before, and show the fits in Figure 1 of the revised manuscript. Using the maximum-concentration GR for calculating $J_{3,est}$, the effect of using $GR_{7\text{-}20}$ instead of $GR_{3\text{-}10}$ is much smaller and overall the results are better (91% and 93% of the daily mean $J_{3,est}$ are within factor of 2 from $J_{3,obs}$).

Lines 191-195: Would prefer to see a full comparison of difference between observed and calculated Js here using GR3-10, GR3-7 and GR7-20, as well as a developed discussion of the degree of agreement and implications of this for using this method for J extrapolation. "Did not affect the results . . .by much" is too qualitative and glosses over what could be an important result here.

This is an excellent suggestion, and based on it we expanded Figure 1 in the revised manuscript to include results using both size ranges for the growth rate GR calculation, 3-10 nm and 7-20 nm. Below is the new Figure 1, and we also made the discussion related to it more quantitative.

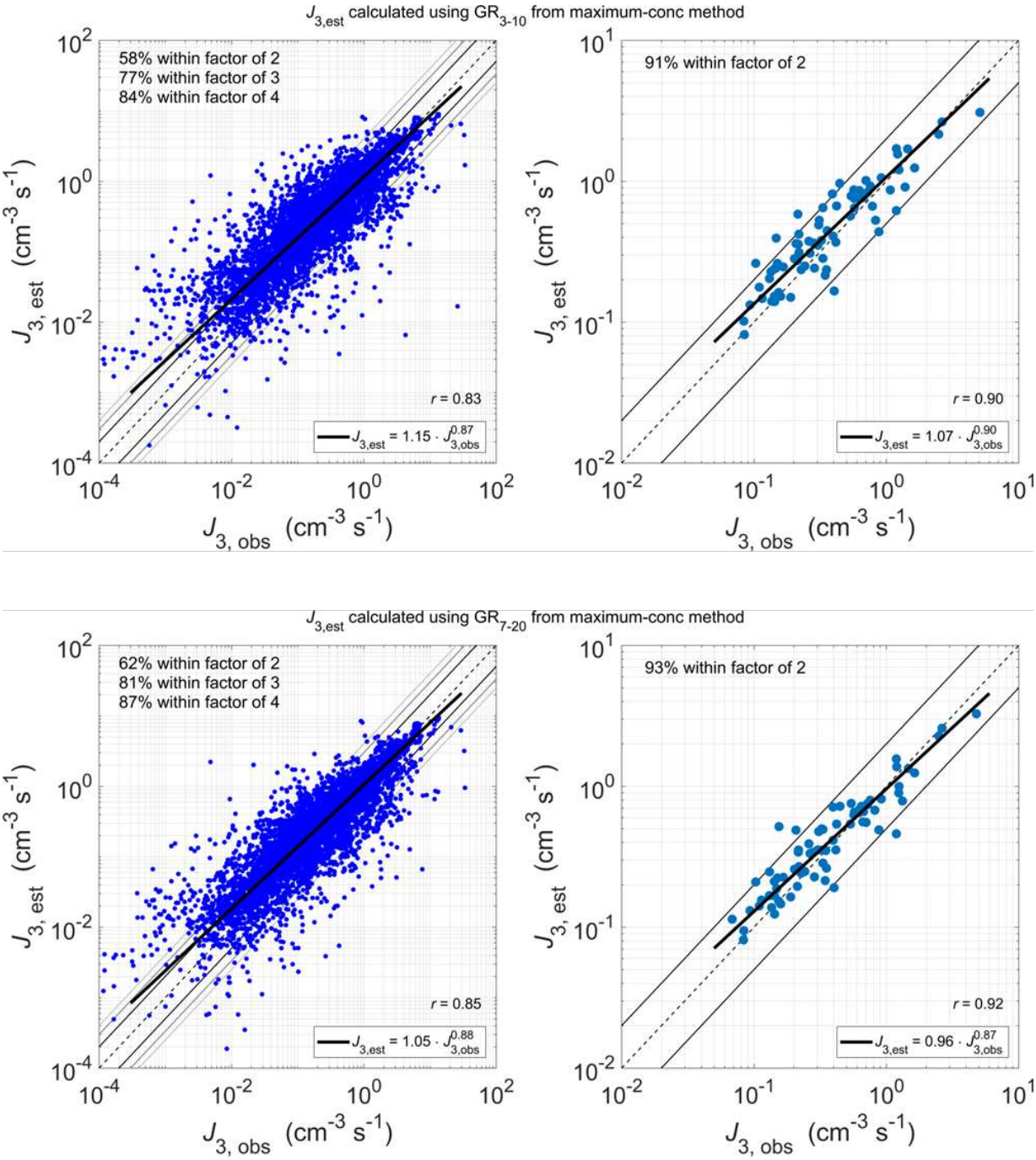


Figure R2. Comparison of the estimated (J3,est) and observed (J3,obs) formation rates of 3 nm particles in Hyytiälä, calculated using GR determined by maximum-concentration method. The top panels show the results for J3,est calculated using GR from 3-10 nm size

range, and the lower panels using GR from 7-20 nm size range. This figure is included in the revised manuscript as Figure 1.

Lines196-203: The lack of correlation on temporarily resolved data here may indicate that the growth rates are wrong – this should be discussed here. It could also be because, by taking averages over long events where J varies significantly, the correlation seen early was simply an artifact of such heavy 'smoothing'. Is there another, meaningful measure of J (e.g. peak J of an event) that could be compared to asses this?
The reanalysis of growth rates with the maximum concentration method has improved the time resolved results significantly as explained before (see figure above). We also choose to stick with looking at mean formation rates as they are a good measure of overall event strength.

Lines 205-208: Given the lack of correlation for time resolved Js, testing the affect of different GRs here does not have much meaning – suggest leaving this out completely.
Regarding the original version of the manuscript, we agree with the referee. However, now with much improved performance with respect to time resolved formation rates, this comparison is meaningful, we think.

Lines 211-212: "For some NPF days, the estimated time dependence and values of Jest are in fairly good agreement with those of observed Jobs." This statement needs better quantification to be of value. What proportion of days (since we're looking at a relatively small number of event, suggest quoting both total number of events examined and number of those with time dependence and value agreement here instead of just a percentage). How 'fairly good agreement' was judged needs explanation Also, are there distinguishing features of this sub-group where agreement is good? E.g. slow growth, classic 'banana' nucleation pattern?
With the new analysis for GR these figures (Fig. 2 of the revised manuscript) have now also changed – and the results are generally much better. Still, the motivation behind figure 2 is the same: we show why for some events (and estimated GR) the analysis works better and for some worse. Thus we also chose not to give quantified information on the comparisons of Figure 2. We explain this now clearly with the discussion of Figure 2.

Line 213: quantify 'most of those days'
As the performance related to time resolved data is now much better we have modified this part of the text: "However, the time-dependency of $J_{3,est}$ is not consistent with $J_{3,obs}$ for some most of the days and, instead, typically the $J_{3,est}$ peak occurs earlier than the $J_{3,obs}$ peak (see e.g. Figure 2-e), indicating that our method of estimating GR is not always satisfactory perfect and typically underestimates the GR values."

Line 221: why does this burst of particles of 3-7nm occur and not then grow? Is this indicated by the calculated GR and coagulation sink? Or is it perhaps a transport artifact? If it is the later it should be removed from the analysis as it is not nucleation. If it's the former then the equation used to calculate J3 should be able to handle it. Therefore this needs full investigation and explanation.

We investigated the event in more detail and found that it is a transport artifact. This is, of course, one of the general problems when analyzing events measured at one fixed location. In the figures we do not see the same aerosol growing, but particles formed at various location appearing at the measurement site at various stages of their growth. If we have a large enough homogeneous region of similar formation and growth, there is no problem. However, if there are inhomogeneities and the air mass transport direction changes during an event, we see dynamics as in fig. 2c and f. As this day was still classified as an event according to the protocol by Kulmala et al., we chose to include it – also to show what kind of challenges there can be.
We added at the end of section 3.1: "This is one of the general problems when analyzing events measured at one fixed location. We do not observe the same aerosol growing, but particles formed at various location appear at the measurement site at various stages of their growth. If we have a large enough homogeneous region of similar formation and growth, there is no problem. However, if there are inhomogenities and the air mass transport direction changes during an event, we see dynamics as in fig. 2c and f."

Lines 225-226: Estimated time-lag longer than observed time-lag indicates that the GR used is too low, which has implications for the calculated J and the time-dependence of the nucleation event. Can this explain the poor ability of this method to reproduce the time-evolution of nucleation events? This should be discussed.
Yes, true. As mentioned before, now the growth rate analysis has been redone and gives much better results.

Line 227: 15 days out of how many in total?
We removed the old figure 3 as we think it is now not necessary and the related text.

Line 227: 1.5 hours difference: what percentage of the total time lag is this?
With our new results, having improved growth rates as well as a much better match between the observed and estimated formation rates, we decided to remove (the old) figure 3 as well as related text.

Line 229: quantify 'reasonably good accuracy'
For the $J_{3,est}$ calculated using $GR_{3\text{-}10}$ from maximum-concentration method, the fraction of data points $J_{3,est}$ vs. $J_{3,obs}$ which are within factor of 2 is now 91% (67 out of 74 events) for the daily mean values. For all the 10-min data points it is 58% within factor of 2 (77% within factor of 3, and 84% within factor of 4). We now focus on the numbers (when discussing our results) and leave out these more vague statements.

Line 244: This monotonic increase in number of event days per year with time is indeed worth noting. Is this because of improvements in instrumentation/data quality? Change of activity or climate in the local area? Some discussion warranted. Do other things, such as total number of nucleation mode particles, size of coagulation sink, or anything else also monotonically change over this time period that might indicate why this is happening?
We took a look at this once again and now feel that this trend is far too short to be considered a significant trend. In Hyytiälä this time period shows a decrease both in SO2 and CS, which

have opposing effects on nucleation event probability. As we cannot quantify/justify such a trend with our supporting measurements we decided to remove the sentence.

Line 245: Given the lack of correlation shown early between median J3 est and obs, using J3 est here for analysis does not seem justified. Surely mean J, where some correlation between estimated and observed values was calculated is the value to use in figure 6?
This was a typo. The presented values are means, which makes much more sense.

Line 249: How does lower average GRs in Puijo affect the analysis? Lower GR gives larger time difference between J7 and J3, mean that inaccuracies in coagulation sinks and neglecting of time dependence of some quantities plays a larger role. Discuss.
True. We added to the revised manuscript (lines 235-237): "As the growth rates in Puijo are on average higher, there is less time needed for the particles to grow from 3 to 7 nm. This means that our assumption of time independent growth rate and coagulation sink during growth should hold in Puijo as good as in Hyytiälä."

Technical Corrections:
Line 35: commas needed around 'at several locations' Line 134: 'used a parabolic differentiation method ON the measured number concentration' instead of TO
Done.

Line 276: new paragraph needed for "the ultimate aim of this work"
Done.

Anonymous Referee #2

GENERAL

The manuscript is basically suitable for ACP but a few important points need to be addressed first. I'll rate this "major revision" for now but remain skeptical that all concerns can be addressed to my satisfaction.

1. The authors present a method to estimate formation rates for smaller particles based on measurements of the formation rate for larger particles. This is no doubt necessary to compare measurements with different instruments and to gain information on the actual nucleation rate which happens at sizes which are often outside the measurement range. However, a few years back, Kulmala et al. have published a sort of how-to guide on nucleation measurements in Nature Protocols. And that looks very, VERY similar to what is presented in this manuscript while the text reads as if something new is shown. So my question would be: What is actually new in the approach that the authors present? Or is this just another application of the same formula that has been used in lots and lots of papers for quite a number of years? If this were the case, the manuscript's content would be very slim indeed (and all the description of methods used obsolete) and one would have to ask if just running an old formula on a new set of data would justify a scientific publication.

The method used in our manuscript for scaling the formation rates is indeed the same as described by Kulmala et al. in their Nature Protocols paper. While the method itself is not new, it has not been tested with atmospheric particle number size distribution data before (to our knowledge), although widely used in e.g. global modelling of aerosol dynamics. In our work, we present comparisons with the scaled formation rates to those calculated directly from the measurement data, and thus evaluate the applicability of this method with real atmospheric data.

2. The method to determine GR obviously doesn't work as the authors themselves point out. That is not a surprise since mode-fitting at the edge of a size distribution is always a bad idea. However, there are other methods. How can you justify using a method that clearly produces bad results while there are alternatives available? Sure, other methods typically are much more labour-intensive but given the current state of affairs it seems quite clear to me that other approaches MUST be employed. At least a preliminary test on a smaller subset of the data is absolutely and totally necessary.

As the referee notes, the main weakness of the previous submitted version was the poor performance in estimating the growth rate with the mode fitting method, which further meant poorly predicted time lag and poor performance in the time resolved formation rate comparison. We have now reanalyzed the growth rates with the so-called maximum concentration method and the results are overall much better (the updated Figure 1 included in the revised manuscript is shown above as Fig. R2 in our replies to referee #1). All the figures and Table 1 are modified accordingly, and we also removed the old Figure 3 comparing time-lags, as we believe that it is unnecessary in the new version.

3. I do wonder how there can be only 65 days with good enough data from 12 or 13 years of Hyytiälä observations. Assuming 12 full years à 365 days and an NPF frequency of 23% (Nieminen et al., 2014), that's a tiny 6.5% of all nucleation events observed during that time (about 1000). How can that be? Are we supposed to believe that one of the longest and

probably the most published data set of aerosol size distributions is actually total crap? I mean, I have worked with DMPS and SMPS data quite a bit but never ever has the data been so terrible that a proper analysis was possible for less than 10% of events. And even if I was to accept this low percentage (which I can't and won't) then the question arises if this kind of cherry-picking doesn't introduce a bias into the analysis that would make all and any results highly questionable.
The number of days included in our analysis is not limited by data availability, but rather the criteria of the NPF event analysis: the growing nucleation mode needs to be clearly observable for several hours (i.e. no changes in air masses). Typically in Hyytiälä the number of these "well-behaved" NPF events is around 10% of all days (Dal Maso et al., 2005). In the Dal Maso classification the NPF events in Hyytiälä are classified as Ia, Ib, and II. For the analysis of this manuscript, we only chose class Ia NPF events, producing the number of events analyzed here. This strict selection of NPF events was done because we wanted to eliminate the possible effect of e.g. changes in air masses in our results.

CONTENTS
line 31f "(e.g. Almeida et al., 2013; Berndt et al., 2014; Kirkby et al., 2016)" –> Bianchi et al., 2016 should probably be added there.
We add this reference.

line 51 "we aim to estimate 3 nm particle formation rates" –> why? the 3 nm limit has no physical meaning, it's just a tradition born out of instrumental limitations from two decades back. i understand that you do that for the hyytiälä data since the point is testing the approach. but for puijo?
The referee is correct that nowadays the aerosol instruments are able to measure particles down to 1.5 nm. However, since the vast majority of particle formation rates reported in the literature is at 3 nm, we chose to scale also the Puijo data to this size. That way the Puijo results can be more directly compared to observations at other sites.

line 88ff –> this whole section is useless if this is the same approach as in the Nature Protocol
The method is the same as described in the Nature Protocol paper, however we feel that presenting the method in our manuscript makes it easier for the user to read our paper. One other reason for this choice is the need to use growth rates in the equations. For Hyytiälä data, growth rates are available down to 3 nm while for Puijo only above 7 nm. Thus we wanted to make clear in the equations what size ranges we are using.

line 164 "the size dependence of the growth rate in the range 3-20 nm is typically weak" –> really? or is this just an artefact of the GR approach not working (which we know is true). certainly you could cite some previous studies that have found this; hyytiälä isn't exactly under-studied after all.
The referee has a valid point here. As we recalculated the growth rates using maximum-concentration method, there is indeed a size-dependency in the growth rates, as is shown in Figure R3 below. However, this does not affect greatly the correlation between $J_{3,obs}$ and $J_{3,est}$ (calculated either using $GR_{3\text{-}10}$ or $GR_{7\text{-}20}$), as can be seen from Fig 1 in the revised manuscript.

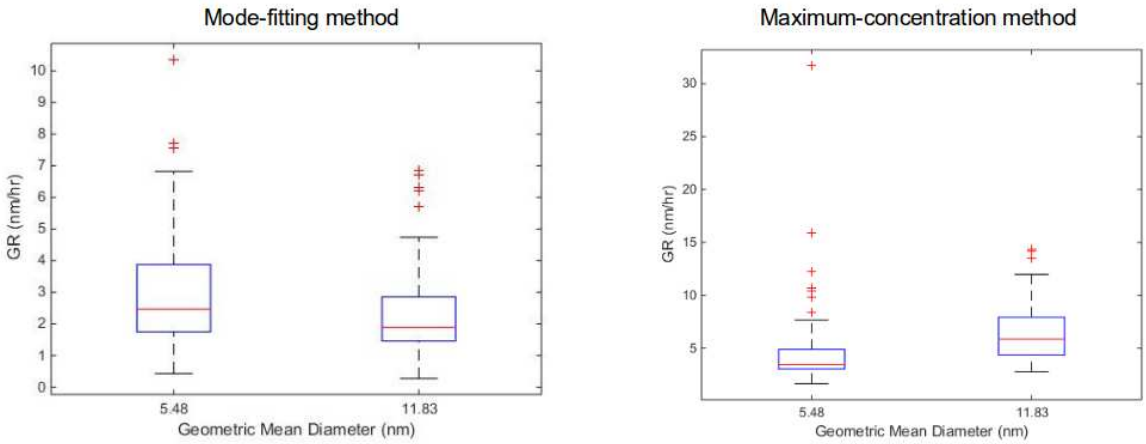


Figure R3. The central mark (red) shows the median, the edges of the box the 25th and 75th percentiles, and the error bars the 10th and 90th percentiles of the GR values of NPF events in Hyytiälä (the red data points are all GR values larger than the 90th percentile). The values on the x-axes are the geometric mean diameters of the two size ranges 3-10 and 7-20 nm, respectively. Left panel shows the GR calculated from mode-fitting method, and right panel shows the GR from maximum-concentration method.

line 173 "85 %" –> wasn't it 84 in the abstract?
All the results related to comparison of the estimated and observed formation rates are updated in the revised manuscript according to the new $J_{3,est}$ values, which are calculated using the GR from maximum-concentration method. The correlation coefficient between $J_{3,est}$ and $J_{3,obs}$ is now 0.90 and 91% of the $J_{3,est}$ daily mean values are within factor of 2 from $J_{3,obs}$.

line 181ff –> the whole median thing seems a bit silly. i mean, you take the median over 12 hours during most of which there is no formation of 3 nm particles. Of course the result will be close to 0 (as it is). a pointless exercise which tells us nothing.
We agree, and have left the medians out from the revised manuscript.

line 201ff, line 210ff, and lots of other places –> i won't comment on the GR stuff here, see general comments above.
As mentioned previously, we chose a better method for the GR analysis, and the results improved a lot.

line 216f "the days during which a clear peak in each of the [different] time evolution curves could be observed (39 days out of 65 days)" –> 39 out of 65 sounds quite good. but really it is 39 out of 1000, and that is not acceptable.
As mentioned before we were very selective with the events analyzes and chose only the so-called 1A events based on the Dal Maso et al. classification.

line 224 "It can be also concluded that visual inspection of the data is still valuable" –> that's sound advise that you might want to follow with regard to the GR business.
So true. As we redid our analysis with the improved growth rate analysis, we also checked each analyzed event visually.

line 227 "There are 15 NPF days for which the estimated time-lag is within 1.5 hours of the observed time-lag." –> please take a moment and think what you have written there. with an average GR of roughly 4 nm/h, the average time-lag should be around 1 hour, right? that 15 cases lie within 1.5 hours is no proof that the method works sometimes but rather that the method does not work AT ALL.
With our new results, having improved growth rates as well as a much better match between the observed and estimated formation rates, we decided to remove (the old) figure 3 as well as related text.

line 227ff "Overall these results from analyzing Hyytiälä data show that Eq. (4) can be used to estimate the mean formation rates of 3-nm particles with reasonably good accuracy." –> but maybe things could be much better with an improved determination of GR?
Yes, true. Now they are.

Estimation of atmospheric particle formation rates through an analytical formula: Validation and application in Hyytiälä and Puijo, Finland

Elham Baranizadeh[1], Tuomo Nieminen[1], Taina Yli-Juuti[1], Markku Kulmala[2], Tuukka Petäjä[2], Ari Leskinen[1,3], Mika Komppula[3], Ari Laaksonen[1,4], Kari E. J. Lehtinen[1,3]

[1]Department of Applied Physics, University of Eastern Finland, Kuopio, Finland

[2]Department of Physics, University of Helsinki, Finland

[3]Finnish Meteorological Institute, Kuopio, Finland

[4]Climate research Unit, Finnish Meteorological Institute, Helsinki, Finland

Correspondence to: Elham Baranizadeh (elham.baranizadeh@~~uef.fi~~gmail.com)

Abstract. The formation rates of 3-nm particles were estimated at SMEAR IV, Puijo (Finland) where the continuous measurements extend only down to 7 nm in diameter. We extrapolated the formation rates at 7 nm (J_7) down to 3 nm (J_3) based on an approximate solution to the aerosol general dynamic equation, assuming a constant condensational growth rate, a power-law size dependent scavenging rate and negligible self-coagulation rate for the nucleation mode particles. To evaluate our method, we first applied it to new-particle formation (NPF) events in Hyytiälä (Finland), which extend down to 3 nm, and, therefore, J_3 and J_7 can be determined directly from the measured size distribution evolution. The Hyytiälä results show that the estimated daily mean J_3 slightly overestimate the observed mean J_3, but a promising 91~~84~~% of the estimated J_3 are within a factor of 2 from the measured ones. However, when considering detailed daily time evolution, the agreement is ~~typically poor~~ not as good ~~presumably~~ due to fluctuations in data as well as uncertainties in estimated growth rates which are required in order to calculate the time-lag between formation of 3-nm and 7-nm particles. At Puijo, the mean ~~and median~~ J_7 for clear NPF days during April 2007-December 2015 were ~~0.23~~0.44 ~~and 0.07~~ $cm^{-3}s^{-1}$~~, respectively~~, while the extrapolated mean ~~and median~~ J_3 were ~~0.47~~0.61 ~~and 0.13~~ $cm^{-3}s^{-1}$~~, respectively~~.

1 Introduction

Atmospheric new particle formation (NPF) events, i.e. nucleation and subsequent growth of newly formed particles have received increasing attention due to their impact on climate and human health (Kulmala et al., 2004; Merikanto et al., 2009; Nie et al., 2014, Kerminen et al., 2012; Fuzzi et al., 2015, Minguillón et al., 2015 and references therein). Many studies have been conducted to find out which variables cause and which possibly inhibit NPF events. Sulfuric acid, water and ammonia

have already long been considered important molecules for atmospheric new particle formation (Weber et al., 1995; Weber et al., 1996; Korhonen et al., 1999; Kulmala et al., 2000; Laaksonen et al., 2008; Xiao et al., 2015). More recently, studies show that amines, ions and volatile organic vapors can play an important role in NPF events either by participating in the nucleation itself or by stabilizing the nucleated clusters (e.g. Almeida et al., 2013; Berndt et al., 2014; Bianchi et al., 2016; Kirkby et al., 2016). However, several features at the nucleation level including the actual mechanism and other possible vapors involved (Kulmala et al., 2006; Lehtinen et al., 2007) remain unknown.

The lack of exact knowledge of NPF mechanisms is partly because, at several locations, particle size distribution measurements do not extend to nucleation size range but instead start at ca. 3 nm or even at larger sizes (e.g. 7 or 10 nm). This limits the use of the particle data in NPF studies and poses a challenge in understanding NPF globally. In addition, the actual nucleation rates of critical clusters sizes (sub-2-nm in diameter) remain unknown. Even with data obtained by the new condensation particle counters (CPC), that have cut-off mobility diameters of sub-2 nm (Sgro and Fernández de la Mora, 2004; Iida et al., 2009; Vanhanen et al., 2011; Kuang et al., 2012; Wimmer et al., 2013), the determination of nucleation rates still involves approximation.

Measuring sub-3-nm particles is a challenging task because of their diffusion loss during transporting the sample, difficulties in collecting representative samples for electrical detection, difficulties in charging them for electrical size-selection (classification), their insufficient amount to be chemically analyzed, and the need for a very high supersaturation condition to grow them to large enough sizes that they can be optically detected (Kulmala et al., 2012). Because of these challenges in measuring small particles, methods to extrapolate size distributions and formation rates below the measurement range have been suggested by McMurry and Friedlander (1979), McMurry (1982; 1983), Weber et al. (1996); Kerminen and Kulmala (2002); Kerminen et al. (2003); Lehtinen et al. (2007) and most recently by Kürten et al. (2015). We are, however, not aware of another study in which these methods have been tested with atmospheric measurement data.

Our study has two main goals. Firstly, we aim to estimate 3 nm particle formation rates J_3 for Puijo, where continuous size distribution measurements have been going on since 2006. We estimate the J_3 by a scaling method based on aerosol dynamics theory for the range 3 - 7 nm, because the measured size range at Puijo has been only down to 7 nm in diameter. Therefore, our second main goal is to validate our method to estimate J_3. For this, we use size distributions measured at Hyytiälä, where detailed particle size distribution measurements down to 3 nm have been performed since 1996. From the Hyytiälä data we can thus evaluate formation rates both at 3 nm and 7 nm. The fraction of particles that survives the scavenging by larger aerosols is determined by the ratio of their growth and scavenging rates (Kerminen et al., 2004b). In this study, we use the method of Lehtinen et al. (2007) in which time and size independent particle growth rate and, time independent but size dependent coagulation sink are assumed.

2 Methods

2.1 Data sets and site descriptions

In this study we use the aerosol size distribution measurements at two different SMEAR (Station for Measuring Ecosystem-Atmosphere Relations) stations in Finland: SMEAR II located in Hyytiälä and SMEAR IV in Kuopio. SMEAR II (Hyytiälä, southern Finland; 61°51´ N, 24°17´ E, 181 m a.s.l.) is characterized by boreal coniferous forest. The main pollution sources are the city of Tampere (60 km away) and the buildings at the station. These sources are most effective when the wind is from the southwest direction (Kulmala et al., 2001). For this study we analyzed aerosol size distributions measured at SMEAR II with a Differential Mobility Particle Sizer (DMPS; Aalto et al., 2001), with a cut-off size at 3 nm, between years 2000-2012.

At SMEAR IV the instruments are set up at the top of the Puijo observation tower (62°54′34″ N, 27°39′19″ E), 306 m and 224 m above the sea level and the surrounding lake level, respectively). Puijo tower is located in the city of Kuopio (Eastern Finland), a semi-urban environment with surroundings characterized by forest with conifer and deciduous (mostly birch) trees, and many lakes. The main local sources surrounding the tower are a paper mill (direction 35°, distance >1.4 km), the city center (direction 120-155°, distance 1.6-3.2 km), a heating plant (direction 160°, distance 3.5 km), a highway and residential areas (see Leskinen et al. (2009) and Portin et al. (2014) for more details). The aerosol size distribution is measured with a twin-DMPS (Winklmayr et al., 1991; Jokinen and Mäkelä, 1997) covering the size range 7-800 nm (Leskinen et al., 2009). The twin-DMPS consists of two differential mobility analyzer (DMA) tubes, one shorter with 11-cm length and another one longer with 28-cm length, and a condensation particle counter (TSI Model 3010 CPC) after each DMA tube. In both DMPS systems, the sample is neutralized (before it enters to the DMA) into charge equilibrium by a beta radiation source (Ni-63 10 mCi=370 MBq). The size range measured by the longer tube is 27-800 nm with 29 discrete bins and 7-49 nm with 17 discrete bins for the shorter tube. The full particle size distribution (7-800 nm) is measured every 12 minutes (Leskinen et al., 2009). At Puijo there is a twin-inlet system for aerosol-cloud interaction studies: one inlet removes cloud droplets (when the station is in a cloud) and collects only the interstitial particles and the other inlet collects the total aerosol, i.e. cloud droplets and interstitial particles. When the station is not in a cloud, the size distribution measured from both inlets are the same. In this study, we used the data from the total aerosol inlet and analyzed aerosol size distributions measured between April 2007 and December 2015.

2.2 Data analysis method

Kerminen and Kulmala (2002) derived an analytical formula which links the “real” particle formation rate and the “apparent” formation rates of particles of larger sizes for which measurements are available (typically above 3 nm). The formula was later improved by Lehtinen et al. (2007) by (1) correcting the slightly inaccurate size dependence of the coagulation sink, and (2) removing the unnecessary assumption of the identity of the condensing vapor. According to the formula (equation (7) in

Lehtinen et al., 2007) one can estimate the formation rate of smaller particles (J_{d1}) with diameter d_1, for which no measurements are available, from the formation rate of measured larger particles (J_{d2}) with diameter d_2, as follows:

$$J_{d1} = J_{d2} \cdot \exp\left(\gamma \cdot d_1 \cdot \frac{CoagS(d_1)}{GR}\right), \quad (1)$$

with $\gamma = \frac{1}{m+1}\left(\left(\frac{d_2}{d_1}\right)^{m+1} - 1\right)$ and $m = \frac{\log[CoagS(d_2)/CoagS(d_1)]}{log[d_2/d_1]}$,

where *CoagS* is the coagulation sink of smaller particles (diameter d_1) onto the background particles, and *GR* is the particle growth rate (which is assumed to be constant from diameter d_1 to diameter d_2).

In this study, we apply the Eq. (1) to estimate the apparent formation rates of particles of 3 nm in diameter at Puijo where the size distribution of particles below 7 nm is not measured. To derive Eq. (1) (i.e. equation (7) in Lehtinen et al., 2007), it was assumed that the growth rate between d_1 and d_2 is constant. This assumption, however can fail especially for sizes below 3 nm, where some recent studies have indicated strong size dependence of *GR* (Kuang et al., 2012; Kulmala et al., 2013).

Korhonen et al. (2014) modified Eq. (1) to also include either linear or power-law type size dependent growth rate and tested the method by using modelled NPF events. In their studies especially the method assuming power-law type growth rate gave promising results with various types of size dependent growth profiles. However, in this study, we assume a constant *GR* because as mentioned earlier a strong size-dependency of *GR* has been reported for very small particles typically below 3 nm (e.g. Kuang et al., 2012) rather than for larger sizes. The other assumption when deriving Eq. (1) is that the nucleating particles are lost only by coagulation onto larger pre-existing particles. Lehtinen et al. (2003) studied the contribution of particles of different sizes to the condensation sink at Hyytiälä and found that particles below 50 nm in diameter have typically negligible contribution. This is a reasonable assumption at Puijo also as the concentrations and size distributions are similar to those at Hyytiälä. The mean values of CoagS of 7 nm particles are ~~5.3e-5~~5.41×10^{-5} s^{-1} and ~~5.3e-5~~5.29×10^{-5} s^{-1} in Hyytiälä (event~~s~~ days during 2002-2012) and Puijo (event~~s~~ days during 2007-2015), respectively.

To evaluate Eq. (1) against measurements, we use the particle size distribution evolution data during nucleation event days from SMEAR II. There the measurements have extended down to 3 nm in diameter, and therefore, one is able to get apparent formation rates at 7 nm (J_7) and at 3 nm (J_3) directly from measurements. We then set d_1= 3 nm and d_2 = 7 nm in Eq. (1) and calculate $J_{3,obs}$ and $J_{7,obs}$ as outlined in Kulmala et al. (2012) and slightly improved in Vuollekoski et al. (2012). Here we use the subscript *obs* to indicate *observed* apparent formation rates *J*. The formation rate of particles of 3 nm ($J_{3,obs}$) and 7 nm ($J_{7,obs}$) in diameter from measured aerosol size distribution were calculated as follows:

$$J_{3,obs} = \frac{dN_{3-7}}{dt} + n_7 . GR_{7-20}, + N_{3-7} . CoagS(d_{GMD}), \quad (2)$$

where $n_7 = \frac{N_{5-9}}{9-5}$ and $d_{GMD} = \sqrt{3 \times 7}$ nm.

$$J_{7,obs} = \frac{dN_{7-10}}{dt} + n_{10} . GR_{7-20} + N_{7-10} . CoagS(d_{GMD}), \quad (3)$$

where $n_{10} = \frac{N_{8-12}}{12-8}$ and $d_{GMD} = \sqrt{7 \times 10}$ nm.

Here $N_{3-7,}$ N_{5-9}, N_{7-10} and N_{8-12} are the number concentration of particles within size ranges 3-7 nm, 5-9 nm, 7-10 nm and 8-12 nm, respectively, and n_7 and n_{10} are the size distribution function at 3 nm and 7 nm, respectively. The coagulation sink (*CoagS*) terms were calculated directly from the measured particle size distributions, taking into account the hygroscopicity effects using the parametrization of Laakso et al. (2004) who used the hygroscopic growth factor parametrization by Zhou (2001). We used a parabolic differentiation method onto the measured number concentration to obtain its time-derivative (the first term in Eq. (2) and Eq. (3)). The method fits a second order polynomial to seven data points centered at the data point where derivative is calculated while at the edges a parabola is fit through the first or last six data points, from which the derivative is calculated directly. Also, to avoid spurious fluctuations in the second and third terms in equations 2 and 3, the N_{3-7}, N_{5-9}, N_{7-10} and N_{8-12} were smoothed using a moving average (over five data points) filter.

The estimated formation rate J_3 was then calculated based on Eq. (1):

$$J_{3,est}(t) = J_{7,obs}(t') . \exp\left(\gamma(t).3nm.\frac{CoagS(d_1=3nm)}{GR_{3-10}}\right), \quad (4)$$

Note $J_{3,est}$ at time t is calculated based on $J_{7,obs}$ at time t', where $t = t' - \frac{4nm}{GR_{3-10}}$, thus accounting for the growth time of the 3 nm particles to 7 nm particles. To average over this time interval needed for growth, the m and $CoagS(d_1)$ values are calculated as medians of the corresponding values during time t to t'.

To determine the growth rates required in this study, we first used the automated algorithm developed by Hussein et al. (2005) for fitting log-normal modes to the measured size distributions. The algorithm assumes that the size distribution is a superposition of 1-3 log-normal modes and at each measurement time optimizes three unknown parameters for each mode to fit the measurements. The parameters for each individual log-normal mode are the mode number concentration N_i, geometric variance σ_g^2, and geometric mean diameter D_{pg}. We then estimate GR by fitting the geometric mean diameter D_{pg} of the

~~growing nucleation mode as a function of time; the slope of the fitted line determines the *GR* in the desired particle size ranges. We also determined the standard error (*SE*) of the *GR* estimates when fitting D_{pg} values respect to time to obtain *GR*.~~ we used the maximum-concentration method (Lehtinen et al., 2003; Yli-Juuti et al., 2011). In this method, the particle growth rates are determined from the times of the concentration maxima in each of the size-bins of the measured particle number size distributions. A linear function is fitted to the datapoints of the geometric mean diameters of the size-bins as function of the determined times of the concentration maxima in the size-bins, and the growth rate GR is the slope of this linear function. We also tested another GR determination method, which uses log-normal mode-fitting of the measured size distributions to follow the growth of the particles (Yli-Juuti et al., 2011). However, when comparing in Hyytiälä the observed 3 nm particle formation rates to those estimated using GR from both maximum-concentration and mode-fitting methods, it become apparent that the maximum-concentration method yielded better results. Therefore, we chose to use the GR from maximum-concentration method in Equations 2, 3 and 4. We left out the days where the growth rates required in the aforementioned equations (i.e. GR_{3-10} and/or $GR_{7\text{-}20}$) were not quantifiable. We chose the size range 3-10 nm rather than 3-7 nm to determine the GR in the exponential term of equation 4 (denoted as GR_{3-10}). This was done to increase the number of data points in the *GR* fitting and thereby to improve the reliability of the fitted *GR*.

After evaluating the analysis method with SMEAR II data, we applied the method for Puijo where the DMPS detection range extended only down to 7 nm. To estimate the formation rate of 3-nm particles at Puijo we adapted Eq. (4) by replacing GR_{3-10} with GR_{7-20} due to lack of DMPS measurements below 7 nm. However, as it will be shown in section 3.1, using GR_{7-20} instead of GR_{3-10} does not affect the accuracy of estimated J_3 for NPF events in Hyytiälä, which is an indication that the size dependence of the growth rate in the range 3-20 nm is typically weak. The $J_{7,obs}$ was calculated with the same method as was used for Hyytiälä (i.e. using equation 3).

3 Results and discussion

3.1 Analysis of estimated J_3 in Hyytiälä (Finland)

Figure 1 shows the comparison of estimated formation rates $J_{3,est}$ (Eq. (4)) with the observed ones $J_{3,obs}$, as calculated directly from the measured size distribution evolution according to Eq. (2) in Hyytiälä. In the top figures, the range 3-10 nm is used to evaluate the growth rate, in the bottom ones 7-20 nm. We analyzed 65 NPF event days for which the formation and growth rates could be quantified. Each data point in Figures 1-b and 1-d~~a~~ represents the arithmetic mean of the 3-nm particle formation rates ($J_{3,est}$ and $J_{3,obs}$) for a single NPF day during the time window from 07:00 to 19:00 local time. The mean is also a measure of the total particle production strength of each event. The results show that, when using GR in the range 3-10 nm, the estimated mean $J_{3,est}$ values ~~agree reasonably well~~ correlate with $J_{3,obs}$ with a correlation coefficient of 0.90 and a slope of 0.90 using

bilinear fitting. ~~78 and~~ Furthermore, ~~85~~ 91 % of estimated $J_{3,est}$ are within ~~a~~the factor of two of the observed $J_{3,obs}$. The corresponding numbers when using GR in the range 7-20 nm are 0.92, 0.87 and 93%. Equation (4) seems to have a tendency of slightly overestimating the formation rate of 3-nm particles. There is not much difference in the results with different GR size ranges. The total means of $J_{3,obs}$~~,~~ and $J_{3,est}$ ~~and $J_{7,obs}$~~ (not shown in the figure) calculated using $GR_{3\text{-}10}$ are ~~0.21~~0.57~~,~~ and ~~0.27~~0.61 ~~and 0.14~~ # cm^{-3} s^{-1}, respectively, confirming the tendency of Eq. (4) in slightly overestimating the 3-nm particle formation rates.

One interesting and important result is that there is not much difference in the estimated formation rates with different GR size ranges. This is both an indication of the weak size independence of GR as well as an encouragement for using GR for the size interval 7-20 nm for Puijo to extrapolate *J* below 7 nm. The correlation coefficient and the fraction of points within a factor of two for the mean formation rates even increase (from 0.90 to 0.92 and from 91% to 93%, respectively) – however, the regression slope decreases from 0.90 to 0.87. ~~We calculated the arithmetic mean of all data points (the total mean) presented in Figure 1 a. The total means of $J_{3,obs}$, $J_{3,est}$ and $J_{7,obs}$ (not shown in the figure) are 0.21, 0.27 and 0.14 # cm^{-3} s^{-1}, respectively, confirming the tendency of Eq. (4) in overestimating the 3-nm particle formation rates.~~ ~~The color code of Figure 1 a indicates the ratio of the relative standard error (SE) of GR_{3-10} and GR_{3-10} (i.e. SE/GR_{3-10}). According to Figure 1-a no relationship between uncertainty in GR estimates and formation rate estimates is seen. For example the data points close to the 1:1 line consist of both days with high and days with low values for SE/GR_{3-10}.~~

~~Moreover, we also compared the estimated and observed J_3 values using daytime median values (not shown here), resulting in a correlation coefficient of 0.73 between $J_{3,obs}$ and $J_{3,est}$. The median $J_{3,est}$ values are, however, even more overestimated than the corresponding mean values and only 38% of estimated $J_{3,est}$ are within a factor of two of the observed $J_{3,obs}$. Total medians (median of daily median values) of $J_{3,obs}$, $J_{3,est}$ and $J_{7,obs}$ are 0.05, 0.106 and 0.047 # cm^{-3} s^{-1}.~~

~~In addition, we replaced GR_{3-10} with GR_{7-20} in Eq. (4) as will be needed to estimate $J_{3,est}$ in Puijo. Results show that although the correlation coefficient improves to 0.93, a smaller fraction (78%) of $J_{3,est}$ data points are within the factor of two of $J_{3,obs}$ values, and subject to more bias (overestimation). However, these changes are minor and do not significantly affect the results. The total mean of $J_{3,est}$ changed from 0.27 to 0.31 # cm^{-3} s^{-1} after replacing GR_{3-10} with GR_{7-20} in Eq. (4).~~

~~Furthermore, we also tested how replacing $GR_{3\text{-}10}$ with $GR_{3\text{-}7}$ in Eq. (4) affected the estimated $J_{3,est}$ values. Replacing $GR_{3\text{-}10}$ with $GR_{3\text{-}7}$ resulted in similar bias (i.e. towards overestimation) and agreement between $J_{3,est}$ and $J_{3,obs}$ mean values, to what was obtained from replacing $GR_{3\text{-}10}$ with $GR_{7\text{-}20}$: the correlation coefficient slightly improved (0.82) with slightly less $J_{3,est}$ data points (80%) within the factor of two of $J_{3,obs}$. In general replacing $GR_{3\text{-}10}$ with $GR_{3\text{-}7}$ did not affect the results (i.e. agreement level between of $J_{3,est}$ and $J_{3,obs}$) by much.~~

Figure 1-~~ba~~ and 1-~~dc~~ show~~s~~ $J_{3,obs}$ versus $J_{3,est}$ values with the same 10-minute temporal resolution as for the measured size distribution. The points are within the time window from 07:00 to 19:00 local time. With this higher temporal resolution $J_{3,obs}$ and $J_{3,est}$ are clearly correlated (with correlation coefficients of 0.83 and 0.85 for the $GR_{3\text{-}10}$ and $GR_{7\text{-}20}$ cases, respectively) ~~not correlated (correlation coefficient = 0.17) despite that~~ but the match is not as good as for ~~their~~ daily mean values presented in Figure 1-~~ab~~. ~~correlate clearly; only~~ For the time resolved data ~~32~~58% (60% for the $GR_{7\text{-}20}$ case) of the estimated $J_{3,est}$ are within a factor of two of the observed $J_{3,obs}$. There are three key reasons for this: 1. there are significant fluctuations in time resolved experimental size distribution data, 2. the extrapolation method assumes a constant value for CoagS/GR, and 3 there is a time lag between J_3 and J_7 and a poor estimation of the growth rate GR results in comparing values at different times.~~The main reason for this is the lack of success in estimating the time lag between the formation of 3-nm and 7-nm particles (see for example the Figure 3-b presented later in this section), which then results in an incorrect time shift for the time evolution of J_3, even though the daily average values agree reasonably well.~~ The variation of CoagS with time also affects m and γ in equation 1. This is, however, negligible as CoagS(7 nm)/CoagS(3 nm) is a very weak function of time.

~~After replacing GR_{3-10} with GR_{7-20} in Eq. (4), still 31% of estimated 10-minute $J_{3,est}$ are within factor of two of the observed $J_{3,obs}$, correlation coefficient slightly worsens to 0.13 and, the $J_{3,est}$ data are subject to more bias (positive bias thus overestimation). We, therefore, conclude that replacing GR_{3-10} with GR_{7-20} both for mean $J_{3,est}$ values and 10-minute values, has only a minor effect on the results thus using GR_{7-20} to estimate J_3 values in Puijo is reasonable.~~

Figure 2 shows examples of the time evolution of the particle size distribution, ~~as well as~~ the different formation rates J and CoagS (3nm) on three NPF days in Hyytiälä. For most ~~some~~ NPF days, the estimated time-dependence (or time-lag between 3-nm and 7-nm particle formation rates) and values of $J_{3,est}$ are in fairly-good agreement with those of observed $J_{3,obs}$ (see e.g. Figure 2-d). However, the time-dependency of $J_{3,est}$ is not consistent with $J_{3,obs}$ for some ~~most~~ of the days and, instead, typically the $J_{3,est}$ peak occurs earlier than the $J_{3,obs}$ peak (see e.g. Figure 2-e), indicating that our method of estimating GR is not always ~~satisfactory~~ perfect and ~~typically~~ ~~understimates~~ underestimates the GR values. ~~In order to investigate how well Eq. (4) estimates the time evolution of the 3-nm particles we visually chose the days during which a clear peak in each of the $J_{7,obs}$, $J_{3,obs}$ and $J_{3,est}$ time evolution curves could be observed (39 days out of 65 days). For these events, we extracted the time difference between 3-nm and 7-nm particle formation from 1) the observed time between peaks in $J_{3,obs}$ and $J_{7,obs}$ (named here observed time lag), and 2) from growth time $t' - t = 4\ \mathrm{nm}/GR_{3-10}$ (named estimated time lag) which is also equal to the time difference between $J_{3,est}$ and $J_{7,obs}$.~~ Figure 2-f shows an example of a NPF day for which the $J_{3,est}$ and $J_{3,obs}$ are dramatically different. This is due to the burst in the number concentration which appeared mostly within the size range 3-7 nm (chosen to calculate $J_{3,obs}$) and is thus not included in the size range 7-10 nm from which $J_{7,obs}$ is calculated and then scaled to $J_{3,est}$. Therefore, Eq. 4 can give quite inaccurate results for NPF days associated with e.g. this type of inhomogeneity

in the particle number concentrations in different size ranges. This is one of the general problems when analyzing events measured at one fixed location. We do not observe the same aerosol growing, but particles formed at various location appear at the measurement site at various stages of their growth. If we have a large enough homogeneous region of similar formation and growth, there is no problem. However, if there are inhomogenities and the air mass transport direction changes during an event, we see dynamics as in fig. 2c and f. It can be also concluded that visual inspection of the data is still valuable - cases like this are very challenging for automatic data analysis routines. ~~Figure 3 shows the estimated time lag versus the observed time lag. As can be seen from the figure, the estimated time lag is mostly longer (toward earlier times of $J_{3,obs}$) than the observed time lag. There are 15 NPF days for which the estimated time lag is within 1.5 hours of the observed time lag. Overall these results from analyzing Hyytiälä data show that Eq. (4) can be used to estimate the mean formation rates of 3 nm particles with reasonably good accuracy. However, the performance in predicting detailed time evolution of the 3 nm particle formation rate is poor in most NPF days with the methods that we use for *GR* estimation.~~

3.2 Estimation of J_3 in Puijo (Finland)

For the aerosol size distribution data in Puijo, the NPF event days were first recognized visually and classified as "quantifiable" and "non-quantifiable" based on whether or not the event is homogeneous enough to allow quantification of the basic characteristics such as formation and growth rates (Dal Maso et al., 2005). Therefore, our data pool consists of event (E), non-event (NE) and undefined days, the last being days during which the evolution of the size distribution is too unclear for definitive determination of whether or not NPF has been occurring. We noticed that there are two types of undefined days in Puijo. One is characterized with a burst in the number concentration of particles of the smallest detectable sizes but doesn't seem to show the characteristics of a NPF event day (i.e. growth to larger sizes, see e.g. Figure ~~4~~3-a) and most likely originate from local emissions. In the other type, some particles appear in larger sizes (with minor growth), which may or may not be originated from NPF processes. (e.g Figure ~~4~~3-b) like the first type. Note that 48 and 44% of the days are missing during years 2010 and 2012, respectively. The monthly number and yearly fraction of NPF event days recorded in Puijo from year 2007 to 2015 are shown in Figure ~~5~~4. The figure shows that a maximum number of event days occurred during spring time similar to NPF events reported in Hyytiälä (Dal Maso et al., 2005~~). It is also worth noting that the fraction of event days is monotonically increasing from 2012 to 2015.~~ There are ~~75~~ 105 quantifiable NPF event days for which we calculated the $J_{3,est}$ at Puijo. Figure ~~6~~ 5 shows the seasonal ~~medians~~ mean values of $J_{3,est}$ and $J_{7,obs}$, GR_{7-20} and coagulation sink for 7-nm particles (*CoagS(d=7 nm)*) for the quantifiable NPF event days in Puijo. The total mean ~~and median~~ of $J_{3,est}$ ~~are~~ is ~~0.47~~0.61 ~~and 0.13, respectively~~, while the corresponding values for $J_{7,obs}$ ~~are~~ is ~~0.23~~0.44 ~~and 0.07~~ #cm^{-3}s^{-1}, respectively. Total means of GR_{7-20} and *CoagS* of 7-nm particles for NPF days are ~~2.34~~5.77 nm/h and ~~1.5~~1.84×10^{-4} 1/s, respectively. Thus, the mean *GR* at Puijo is somewhat ~~lower~~ higher compared to Hyytiälä where ~~mean~~ median value of *GR* = 4.3 nm/h is reported for the period April 2003-December 2009 (Yli-Juuti et al., 2011). As the growth rates in Puijo are on average ~~smaller~~ higher, there is ~~more~~less time

needed for the particles to grow from 3 to 7 nm. This means that our assumption of time independent growth rate and coagulation sink during growth should hold in Puijo as good as in ~~-~~Hyytiälä.

Table 1 summarizes the seasonal means of parameters presented in Figure ~~6~~5. The seasonal mean 3-nm particle formation rates seem to have the highest values during ~~spring~~ summer (~~0.52~~1.12 cm^{-3} s^{-1} for ~~50~~ 17 NPF days) and ~~summer~~ spring (~~0.53~~0.70 cm^{-3} s^{-1} for ~~12~~ 68 NPF days) and drops significantly in fall and winter. The seasonal ~~median~~ mean of the growth rate has its maximum in summer (~~4.14~~7.96 nm/h) and minimum in ~~spring~~ winter (2.30 nm/h). The seasonal ~~median~~ mean of *CoagS* values~~, however, seem to be rather constant in Puijo in contrast to Hyytiälä.~~ for 7 nm during NPF event days are highest in summer and lowest in winter in Puijo.

4 Conclusions

In this study, the formation rates of 3-nm particles in SMEAR IV, Puijo (Finland) were estimated. The measurements at Puijo extend only down to 7 nm in diameter, which means that we had to extrapolate to 3 nm using aerosol dynamics theory. The approach used here is based on the competing processes of condensational growth and scavenging onto background aerosols, assuming time and size independent growth rate and time independent coagulation sink in the range 3 to 7 nm.

To first evaluate our extrapolation method, we applied it to particle formation events at Hyytiälä, where DMPS measurements extend down to 3 nm and formation rates at 3 nm ($J_{3,obs}$) and 7 nm ($J_{7,obs}$) can thus be determined directly from the measured size distribution evolution. The results show that the estimated daily mean values of J_3 are in reasonably good agreement with observed mean J_3, with ~~84~~91% of the estimated J_3 within a factor of two from the measured ones and, mostly overestimated. However, when considering detailed daily time evolution, the agreement is ~~typically poor~~not as good. This is caused by three main things: 1. there are significant fluctuations in experimental size distribution data, 2. the extrapolation method assumes a constant value for CoagS/GR, and 3~~the fact that~~ there is a time lag between J_3 and J_7 and ~~to take this into account in the comparison an~~ a poor estimation of the growth rate *GR* ~~is needed~~results in comparing values at different times. ~~.~~ Estimating $GR_{3\text{-}10}$, as was shown from Hyytiälä data, does not seem to give in all cases satisfactory results for this purpose. It should be noted that we have to estimate *GR* from the data above 7 nm for Puijo site due to the lack of the measured data below 7 nm.

At Puijo, the mean ~~and median~~ of J_7 for quantifiable particle formation days ~~were~~ was ~~0.23~~0.44 ~~and 0.07~~ $cm^{-3}s^{-1}$~~, respectively~~, while the extrapolated mean ~~and median~~ J_3 ~~were~~ was ~~0.47~~0.61 ~~and 0.13~~ $cm^{-3}s^{-1}$~~, respectively~~. These are about two times greater than the corresponding values in Hyytiälä. Asmi et al. (2011) reported monthly mean 7-nm particle formation rate between 0.1 and 0.2 $\#cm^{-3}s^{-1}$ for the NPF events in the sub-Arctic Pallas station, Finland. The ultimate aim of this work is to predict nucleation rates from size distribution measurements that do not extend to sizes lower than 7nm. The results obtained in this study suggest this is very challenging, in large part due to the difficulty in reliably predicting the growth rate down to around 1.5nm. It is noted that the possible size dependence of this growth rate further complicates the matter.

Acknowledgments:

We gratefully acknowledge the financial support by the Academy of Finland Center of Excellence program (project numbers 272041, 1118615), the Nordic Centre of Excellence CRAICC, University of Eastern Finland (UEF) strategic funding, and ACTRIS (under the European Union Seventh Framework Programme (FP7/2007–2013) grant agreement no. 262254 and Horizon 2020 research and innovation programme under grant agreement no. 654109).

References

Aalto, P., Hämeri, K., Becker, E., Weber, R., Salm, J., Mäkelä, J., Hoell, C., O'dowd, C., Hansson, H.-C., Väkevä, M., Koponen, I., Buzorius, G., and Kulmala, M.: Physical characterization of aerosol particles during nucleation events, Tellus B [Online], Volume 53, 2001.

Asmi, E., Kivekäs, N., Kerminen, V.-M., Komppula, M., Hyvärinen, A.-P., Hatakka, J., Viisanen, Y., and Lihavainen, H.: Secondary new particle formation in Northern Finland Pallas site between the years 2000 and 2010, Atmos. Chem. Phys., 11, 12959-12972, doi:10.5194/acp-11-12959-2011, 2011.

Almeida, J., Schobesberger, S., Kürten, A., Ortega, I. K., Kupiainen-Määttä, O., Praplan, A. P., Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, M., David, A., Dommen, J., Donahue, N. M., Downard, A., Dunne, E., Duplissy, J., Ehrhart, S., Flagan, R. C., Franchin, A., Guida, R., Hakala, J., Hansel, A., Heinritzi, M., Henschel, H., Jokinen, T., Junninen, H., Kajos, M., Kangasluoma, J., Keskinen, H., Kupc, A., Kurtén, T., Kvashin, A. N., Laaksonen, A., Lehtipalo, K., Leiminger, M., Leppä, J., Loukonen, V., Makhmutov, V., Mathot, S., McGrath, M. J., Nieminen, T., Olenius, T., Onnela, A., Petäjä, T., Riccobono, F., Riipinen, I., Rissanen, M., Rondo, L., Ruuskanen, T., Santos, F. D., Sarnela, N., Schallhart, S., Schnitzhofer, R., Seinfeld, J. H., Simon, M., Sipilä, M., Stozhkov, Y., Stratmann, F., Tomé, A., Tröstl, J., Tsagkogeorgas, G., Vaattovaara, P., Viisanen, Y., Virtanen, A., Vrtala, A., Wagner, P. E., Weingartner, E., Wex, H., Williamson, C., Wimmer, D., Ye, P., Yli-Juuti, T., Carslaw, K. S., Kulmala, M., Curtius, J., Baltensperger, U., Worsnop, D. R., Vehkamäki, H., and Kirkby, J: Molecular understanding of sulphuric acid-amine particle nucleation in the atmosphere. Nature, 502: 359-363, doi:10.1038/nature12663, 2013.

Berndt, T., Sipilä, M., Stratmann, F., Petäjä, T., Vanhanen, J., Mikkilä, J., Patokoski, J., Taipale, R., Mauldin III, R. L., and Kulmala, M.: Enhancement of atmospheric H_2SO_4 / H_2O nucleation: organic oxidation products versus amines, Atmos. Chem. Phys., 14, 751-764, doi:10.5194/acp-14-751-2014, 2014.

Bianchi, F., Tröstl, J., Junninen, H., Frege, C., Henne, S., Hoyle, C.,R., Molteni, U.,Herrmann, E., Adamov, A., Bukowiecki, N., Chen, X., Duplissy, J., Gysel, M., Hutterli, M., Kangasluoma, J., Kontkanen, J., Kürten, A., Manninen, H. E., Münch, S., Peräkylä, O., Petäjä, T., Rondo, L., Williamson, C., Weingartner, E., Curtius, J., Worsnop, D. R., Kulmala, M., Dommen, J., and Baltensperger, U.: New particle formation in the free troposphere: A question of chemistry and timing. Science 352, 1109-1112, 2016. DOI: 10.1126/science.aad5456

Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P. & Lehtinen K. E. J.: Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiälä, Finland, Boreal Environment Research. 10, 5, p. 323-336. 14 p, 2005.

Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., Denier van der Gon, H., Facchini, M. C., Fowler, D., Koren, I., Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M., and Gilardoni, S.: Particulate matter, air quality and climate: lessons learned and future needs, Atmos. Chem. Phys., 15, 8217-8299, doi:10.5194/acp-15-8217-2015, 2015.

Hussein, T., Dal Maso, M., Petäjä, T., Koponen, I. K., Paatero, P., Aalto, P. P., Hämeri, K. & Kulmala, M.: Evaluation of an automatic algorithm for fitting the particle number size distributions. Boreal Env. Res. 10: 337–355, 2005.

Iida, K., Stolzenburg, M. R., McMurry, P. H.: Effect of Working Fluid on Sub-2 nm Particle Detection with a Laminar Flow Ultrafine Condensation Particle Counter, Aerosol Science and Technology, 43:1, 81-96, doi: 10.1080/02786820802488194, 2009.

Jokinen, V. and Mäkelä, J. M.: Closed-loop arrangement with critical orifice for DMA sheath/excess flow system, J. Aerosol Sci., 28, 643–648, 1997.

Kerminen, V.-M., & Kulmala, M.: Analytical formulae connecting the "real" and the "apparent" 25 nucleation rate and the nuclei number concentration for atmospheric nucleation events. Journal of Aerosol Science, 33, 609–622, doi: http://dx.doi.org/10.1016/S0021-8502(01)00194-X, 2002.

Kerminen, V.-M., Lehtinen, K. E. J., Anttila, T. and Kulmala, M.: Dynamics of atmospheric nucleation mode particles: a timescale analysis. Tellus B, 56, 135-146, doi: 10.1111/j.1600-0889.2004.00095.x, 2003.

Kerminen, V.-M., Anttila, T., Lehtinen, K. E. J. and Kulmala, M.: Parameterization for atmospheric new-particle formation: Application to a system involving sulfuric acid and condensable water-soluble organic vapors, Aerosol Sci. Tech. 38, 1001-1008, doi: 10.1080/027868290519085, 2004.

Kerminen, V.-M., Paramonov, M., Anttila, T., Riipinen, I., Fountoukis, C., Korhonen, H., Asmi, E., Laakso, L., Lihavainen, H., Swietlicki, E., Svenningsson, B., Asmi, A., Pandis, S. N., Kulmala, M., and Petäjä, T.: Cloud condensation nuclei production associated with atmospheric nucleation: a synthesis based on existing literature and new results, Atmos. Chem. Phys., 12, 12037-12059, doi:10.5194/acp-12-12037-2012, 2012.

Kirkby, J., Duplissy, J., Sengupta, K., Frege, C., Gordon, H., Williamson, C., Heinritzi, M., Simon, M., Yan, C., Almeida, J., Tröstl, J., Nieminen, T., Ortega, I. K., Wagner, R., Adamov, A., Amorim, A., Bernhammer, A.-K., Bianchi, F., Breitenlechner, M., Brilke, S., Chen, X., Craven, J., Dias, A., Ehrhart, S., Flagan, R. C., Franchin, A., Fuchs, C., Guida, R., Hakala, J., Hoyle, C. R. , Jokinen, T., Junninen, H., Kangasluoma, J., Kim, J., Krapf, M., Kürten, A., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Mathot, S., Molteni, U., Onnela, A., Peräkylä, O., Piel, F., Petäjä, T., Praplan, A. P. , Pringle, K., Rap, A., Richards, N. A. D., Riipinen, I., Rissanen, M. P. , Rondo, L., Sarnela, N., Schobesberger, S., Scott, C. E. , Seinfeld, J. H., Sipilä, M., Steiner, G., Stozhkov, Y., Stratmann, F., Tomé, A., Virtanen, A., Vogel, A. L., Wagner, A., Wagner, P. E., Weingartner, E., Wimmer, D., Winkler, P. M., Ye, P., Zhang, X., Hansel, A., Dommen, J., Donahue, N. M., Worsnop, D. R., Baltensperger, U., Kulmala, M., Carslaw, K. S., Curtius, J.: Ion-induced nucleation of pure biogenic particles, Nature, doi:10.1038/nature12663, 2016.

Korhonen, P., Kulmala, M., Laaksonen, A., Viisanen, Y., McGraw, R. and Seinfeld, J. H.: Ternary nucleation of H2SO4, NH3, and H2O in the atmosphere,Journal of Geophysical Research: Atmospheres,104, 26349—26353, doi = 10.1029/1999JD900784,1999.

Korhonen, H., Kerminen,V. M., Kokkola, H., and Lehtinen, K. E. J.: Estimating atmospheric nucleation rates from size distribution measurements: Analytical equations for the case of size dependent growth rates, Journal of Aerosol Science, 69, 13-20, doi:http://dx.doi.org/10.1016/j.jaerosci.2013.11.006, 2014.

Kuang, C., Chen, M., McMurry, P. H., and Wang, J.: Modification of laminar flow ultrafine condensation particle counters for the enhanced detection of 1 nm condensation nuclei, Aerosol Sci. Technol., 46, 309–315, 2012.

Kürten, A., Williamson, C., Almeida, J., Kirkby, J., and Curtius, J.: On the derivation of particle nucleation rates from experimental formation rates, Atmos. Chem. Phys., 15, 4063-4075, doi:10.5194/acp-15-4063-2015, 2015.

Kulmala, M., Liisa, P, and Mäkelä, J. M.: Stable sulphate clusters as a source of new atmospheric particles, Nature 404, no. 6773, 66-69, doi: 10.1038/35003550, 2000.

Kulmala M., Hämeri K., Aalto P. P., Mäkelä J. M., Pirjola L., Nilsson E. D., Buzorius G., Rannik Ü., Maso M. D., Seidl W., Hoffman T., Janson R., Hansson H.-C., Viisanen Y., Laaksonen A. & O'Dowd C. D.: Overview of the international project on biogenic aerosol formation in the boreal forest (BIOFOR). Tellus B, 53: 324–343. doi: 10.1034/j.1600-0889.2001.530402, 2001.

Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, and W., McMurry, P.H.: Formation and growth rates of ultrafine atmospheric particles: a review of observations, J Aerosol Sci., 35:143-76, doi:10.1016/j.jaerosci.2003.10.003, 2004.

Kulmala, M., Lehtinen, K. E. J., and Laaksonen, A.: Cluster activation theory as an explanation of the linear dependence between formation rate of 3 nm particles and sulphuric acid concentration, Atmos. Chem. Phys., 6, 787–793, doi:10.5194/acp-6-787-2006, 2006.

Kulmala, M., Petäjä, T., Nieminen, T., Sipilä, M., Manninen, H. E, Lehtipalo, K., Dal Maso, M., Aalto, P. P, Junninen, H., Paasonen, P., Riipinen, I., Lehtinen, K. E J, Laaksonen, A., Kerminen, V.-M.: Measurement of the nucleation of atmospheric aerosol particles, 7, 1651–1667, doi: http://dx.doi.org/10.1038/nprot.2012.091, 2012.

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., Kurten, T., Johnston, M. V., Smith, J. N., Ehn, M., Mentel, T. F., Lehtinen, K. E. J., Laaksonen, A., Kerminen, V.-M., Worsnop, D. R.: Direct Observations of Atmospheric Aerosol Nucleation, 339, 943—946, doi: 10.1126/science.1227385, 2013.

Laakso, L., Petäjä, T., Lehtinen, K. E. J., Kulmala, M., Paatero, J., Hõrrak, U., Tammet, H., and Joutsensaari, J.: Ion production rate in a boreal forest based on ion, particle and radiation measurements, Atmos. Chem. Phys., 4, 1933-1943, doi:10.5194/acp-4-1933-2004, 2004.

Laaksonen, A., Kulmala, M., Berndt, T., Stratmann, F., Mikkonen, S., Ruuskanen, A., Lehtinen, K. E. J., Dal Maso, M., Aalto, P., Petäjä, T., Riipinen, I., Sihto, S.-L., Janson, R., Arnold, F., Hanke, M., Ücker, J., Umann, B., Sellegri, K., O'Dowd, C. D., and Viisanen, Y.: SO_2oxidation products other than H_2SO_4 as a trigger of new particle formation. Part 2: Comparison of ambient and laboratory measurements, and atmospheric implications, Atmos. Chem. Phys., 8, 7255-7264, doi:10.5194/acp-8-7255-2008, 2008.

Lehtinen, K. E. J. and Kulmala, M.: A model for particle formation and growth in the atmosphere with molecular resolution in size. Atmos. Chem. Phys. 3, 251-257, 2003.

Lehtinen, K. E. J., Korhonen, H., Dal Maso, M. and Kulmala, M.: On the concept of condensation sink diameter. Boreal Env. Res. 8, 405-411, 2003.

Lehtinen, K. E. J., dal Maso, M., Kulmala, M., and Kerminen, V.- M.: Estimating nucleation rates from apparent particle formation rates and vice versa: Revised formulation of the Kerminen–Kulmala equation, J. Aerosol Sci., 38, 988–994, 2007.

Leskinen, A., Portin, H., Komppula, M., Miettinen, P., Arola, A., Lihavainen, H., Hatakka, J., Laaksonen, A., Lehtinen, K. E. J.: Overview of the research activities and results at Puijo semiurban measurement station, Boreal Env. Res., 14, 576–590, 2009.

McMurry, P. H., and S. K. Friedlander: New particle formation in the presence of an aerosol, Atmos. Environ., 13, 1635 – 1651, 1979.

McMurry, P. H., & Wilson, J. C.: Growth laws for the formation of secondary ambient aerosols: Implications for chemical conversion mechanisms. Atmospheric Environment (1967), 16(1), 121-134. doi:10.1016/0004-6981(82)90319-5, 1982.

McMurry, P. H.: New particle formation in the presence of an aerosol: Rates, time scales and sub-0.01 mm size distributions, J. Colloid Interface Sci., 95(1), 72 – 80, 1983.

Merikanto, J., Spracklen, D. V., Mann, G. W., Pickering, S. J., and Carslaw, K. S.: Impact of nucleation on global CCN, Atmos. Chem. Phys., 9, 8601-8616, doi:10.5194/acp-9-8601-2009, 2009.

Minguillón, M. C., Brines, M., Pérez, N., Reche, C., Pandolfi, M., Fonseca, A. S., Amato, F., Alastuey, A., Lyasota, A., Codina, B., Lee, H.-K., Eun, H.-R. , Ahn, K.-H. , Querol, X.: New particle formation at ground level and in the vertical column over the Barcelona area, Atmospheric Research, Volumes 164–165, Pages 118-130, doi:10.1016/j.atmosres.2015.05.003, 2015.

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., Kulmala, M.: Polluted dust promotes new particle formation and growth, 4, 6634, doi:10.1038/srep06634, 2014.

Portin, H., Leskinen, A., Hao, L., Kortelainen, A., Miettinen, P., Jaatinen, A., Laaksonen, A., Lehtinen, K. E. J., Romakkaniemi, S., and Komppula, M.: The effect of local sources on particle size and chemical composition and their role in aerosol–cloud interactions at Puijo measurement station, Atmos. Chem. Phys., 14, 6021-6034, doi:10.5194/acp-14-6021-2014, 2014.

Sgro , L. A., Fernández de la Mora, J.: A Simple Turbulent Mixing CNC for Charged Particle Detection Down to 1.2 nm, 38, 1-11, 10.1080/02786820490247560, 2004.

Vanhanen, J., Mikkilä, J., Lehtipalo, K., Sipilä, M., Manninen, H. E., Siivola, E., Petäjä, T., Kulmala., M.: Particle Size Magnifier for Nano-CN Detection, Aerosol Science and Technology, 45:4, 533-542, 2011.

Vuollekoski, H., Sihto, S.-L., Kerminen, V.-M., Kulmala, M., and Lehtinen, K. E. J.: A numerical comparison of different methods for determining the particle formation rate, Atmos. Chem. Phys., 12, 2289-2295, doi:10.5194/acp-12-2289-2012, 2012.

Weber, R. J., McMurry, P. H., Eisele, F. L. , Tanner, D. J.: Measurement of Expected Nucleation Precursor Species and 3–500-nm Diameter Particles at Mauna Loa Observatory, Hawaii, Journal of the Atmospheric Sciences 1995 52:12, 2242-2257, 1995.

Weber, R. J., Marti, J. J., and McMurry, P. H., Eisele, F. L., Tanner, D. J., and Jefferson, A.: Measured atmospheric new particle formation rates: implications for nucleation mechanisms, Chemical Engineering Communications, 151, 53-64, doi: 10.1080/00986449608936541, 1996.

Wimmer, D., Lehtipalo, K., Franchin, A., Kangasluoma, J., Kreissl, F., Kürten, A., Kupc, A., Metzger, A., Mikkilä, J., Petäjä, T., Riccobono, F., Vanhanen, J., Kulmala, M., and Curtius, J.: Performance of diethylene glycol-based particle counters in the sub-3 nm size range, Atmos. Meas. Tech., 6, 1793–1804, doi:10.5194/amt-6-1793-2013, 2013.

Winklmayr, W., Reischl, G. P., Linder, A. O., and Berner, A.: A new electromobility spectrometer for the measurement of aerosol size distribution in the size range 1 to 1000 nm, J. Aerosol Sci., 22, 289–296, 1991.

Yli-Juuti, T., Nieminen, T., Hirsikko, A., Aalto, P. P., Asmi, E., Hõrrak, U., Manninen, H. E., Patokoski, J., Dal Maso, M., Petäjä, T., Rinne, J., Kulmala, M., and Riipinen, I.: Growth rates of nucleation mode particles in Hyytiälä during 2003−2009: variation with particle size, season, data analysis method and ambient conditions, Atmos. Chem. Phys., 11, 12865-12886, doi:10.5194/acp-11-12865-2011, 2011.

Xiao, S., Wang, M. Y., Yao, L., Kulmala, M., Zhou, B., Yang, X., Chen, J. M., Wang, D. F., Fu, Q. Y., Worsnop, D. R., and Wang, L.: Strong atmospheric new particle formation in winter in urban Shanghai, China, Atmos. Chem. Phys., 15, 1769-1781, doi:10.5194/acp-15-1769-2015, 2015.

Zhou, J.: Hygroscopic Properties of Atmospheric Aerosol Particles in Various Environments, PhD thesis, University of Lund, Division of Nuclear Physics, Sweden, http://lup.lub.lu.se/record/41435, 2001.

Table 1. Overall and seasonal mean values of the observed formation rates of 7-nm particles ($J_{7,obs}$), the estimated formation rates of 3-nm particles ($J_{3,est}$), the growth rates of particles in size range 7-20 nm (GR_{7-20}) and the coagulation sink of 7 nm particles onto larger particles ($CoagS_7$) for 105 NPF days which occurred at Puijo during Apr 2007 - Dec 2015. The $J_{7,obs}$, $J_{3,est}$ and $CoagS_7$ include data during 07:00–19:00 on each NPF day.

	$J_{7,obs}$ ($cm^{-3}\ s^{-1}$)	$J_{3,est}$ ($cm^{-3}\ s^{-1}$)	GR_{7-20} (nm/h)	$CoagS_7$ (s^{-1})
Winter (Dec-Feb)	0.16	0.22	4.33	1.36×10^{-4}
Spring (Mar-May)	0.49	0.70	5.04	1.94×10^{-4}
Summer (Jun-Aug)	0.85	1.12	7.96	2.34×10^{-4}
Fall (Sep-Nov)	0.27	0.40	5.74	1.70×10^{-4}
Overall	0.44	0.61	5.77	1.84×10^{-4}

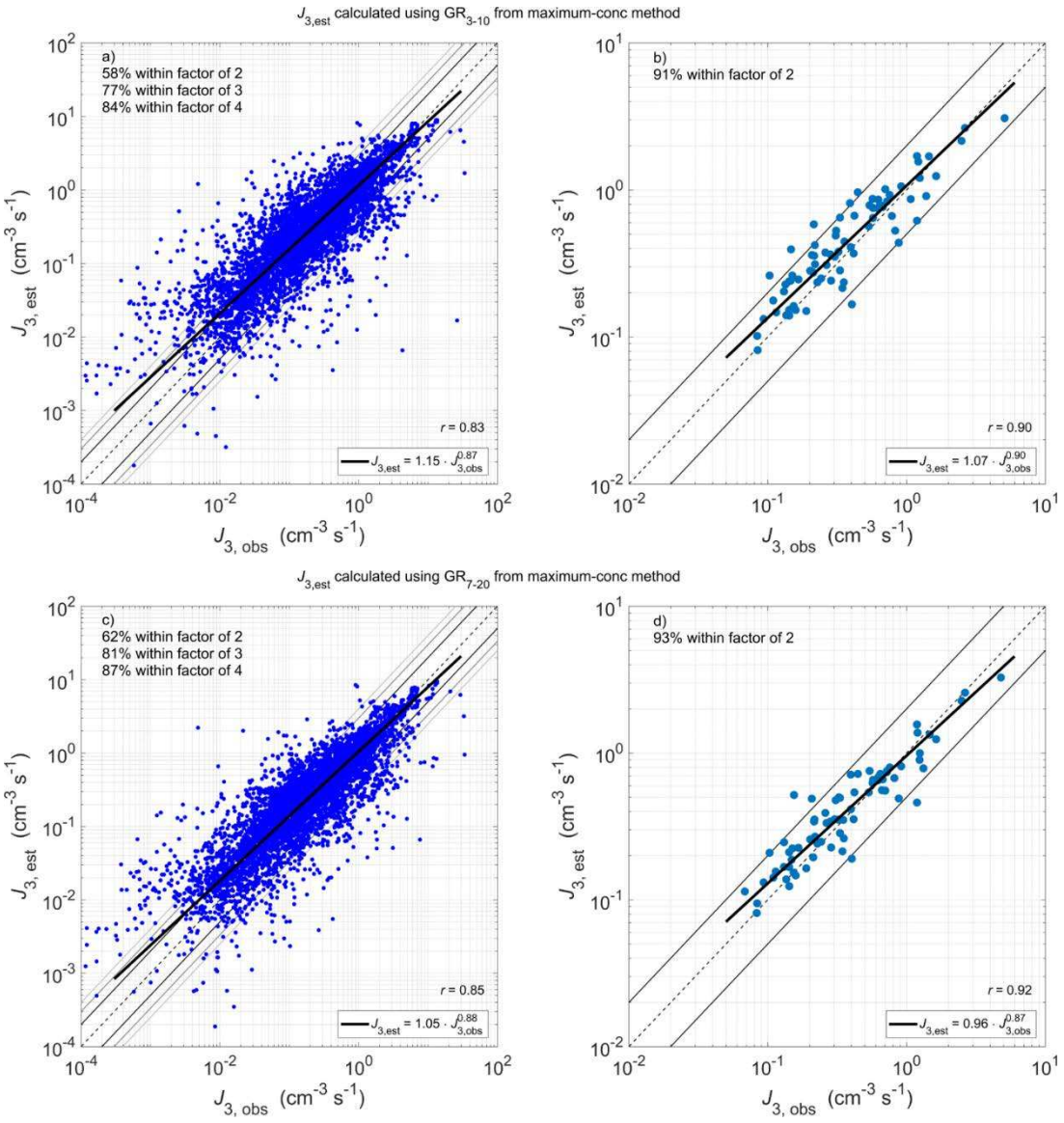


Figure 1. Comparison of the estimated ($J_{3,est}$) against observed ($J_{3,obs}$) formation rates of 3 nm particles during the selected new-particle formation (NPF) event days in Hyytiälä. Panels a), c) show formation rates $J_{3,est}$ and $J_{3,obs}$ calculated at 10-minute time resolution between 7:00–19:00. Note that the time-lag during which particles grow from 3 nm to 7 nm is taken into account in the $J_{3,est}$. Panels b), d) show arithmetic mean of formation rates between 07:00 to 19:00 local time for each NPF day. Panels in the top row refer to $J_{3,est}$ results calculated using GR_{3-10}, and the bottom row to those calculated using GR_{7-20}. The black lines shows the bivariate linear fits to the logarithmic data values; the corresponding fit equations are and the correlation coefficients r are given in each panel.

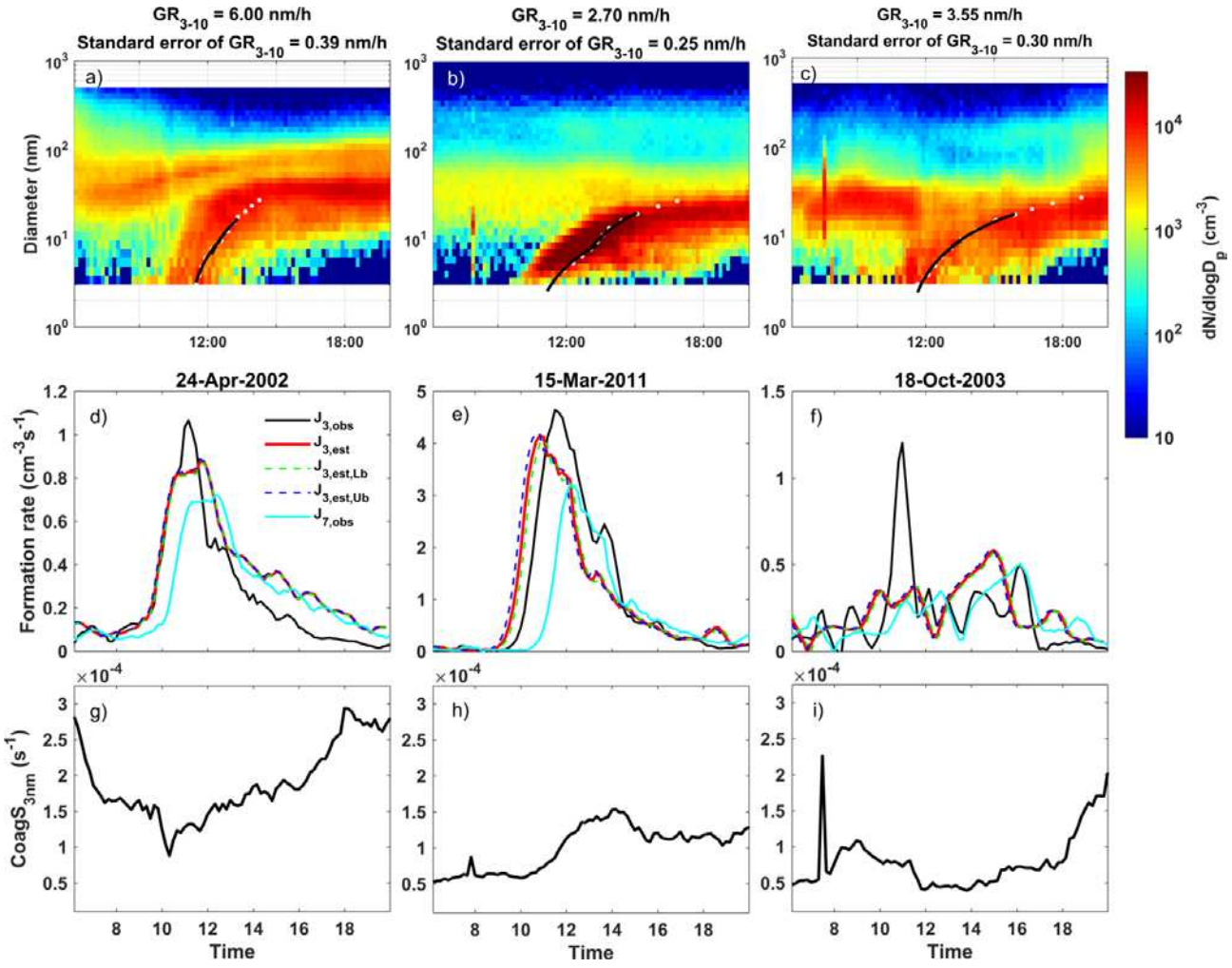


Figure 2. Examples of Hyytiälä NPF events. a, b, and c) the evolution of the particle number size distribution. White dots represent the times of the maximum concentration in each size-bin of the measured size distribution, and the solid black line shows the first-order polynomial fit to those data points. Figures d), e) and f) are the corresponding time-evolution of 3 nm particle formation rates $J_{3,obs}$ obtained from Eq. (2) (black line), observed formation rates of 7 nm particles $J_{7,obs}$ obtained from Eq. (3) (cyan line), and the estimated formation rate of 3 nm particles $J_{3,est}$ calculated by Eq. (4) (red line). The dashed lines show the upper bound ($J_{3,est,Ub}$) and lower bound ($J_{3,est,Lb}$) calculated by Eq. (4) using the lower ($GR_{3–10}$ - SE) and upper ($GR_{3–10}$ + SE) bound of $GR_{3–10}$, respectively. Figures g), h) and i) show the time evolution of the coagulation sink of 3 nm particles.

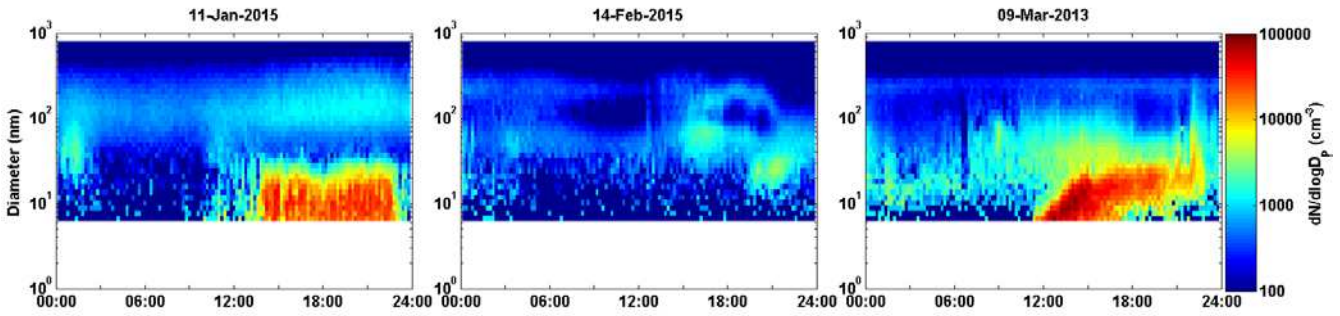


Figure 3. Examples of the time evolution of the aerosol size distribution in Puijo for (a) an undefined day characterized by a burst in the number concentrations of the small particles which doesn't have the characteristics of a typical NPF event day (b) a typical undefined day, and (c) a clear NPF event day.

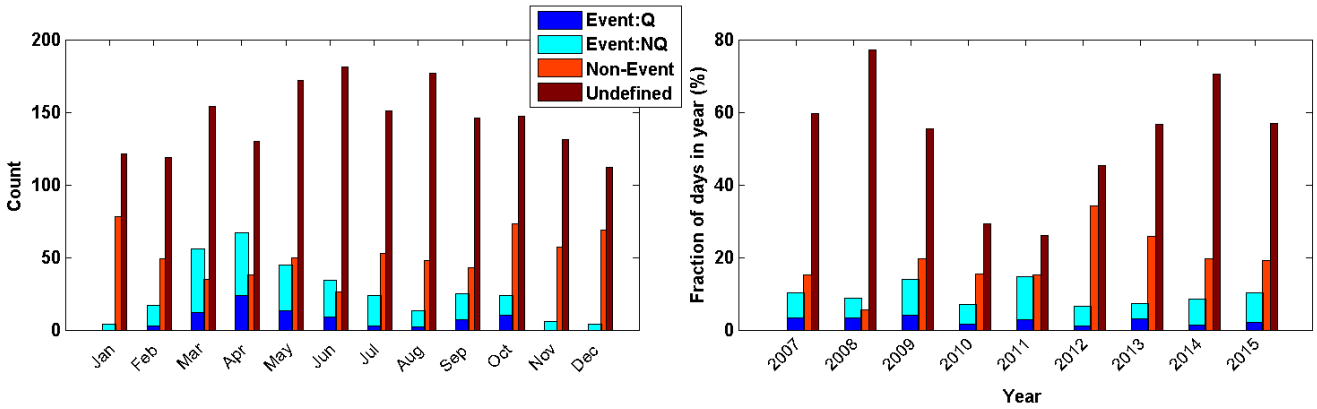
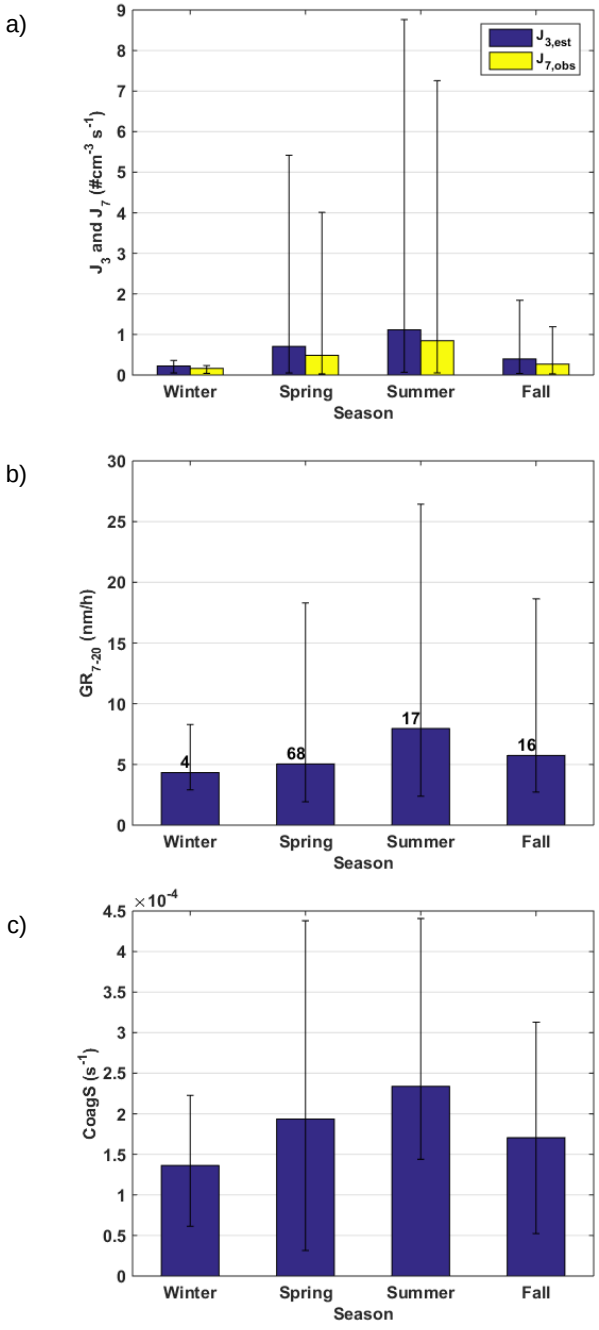


Figure 4. Monthly number (left panel) and yearly fraction (right panel) of NPF event days (divided into Quantifiable Events (Q) and Non-Quantifiable events (NQ)), Non-Events (NE) and undefined days recorded in Puijo during period 2007-2015. Fraction of (e.g. NE) days in year is the ratio of number of NEs and number of days within the year. Note that the days for which bad or no data were recorded are not shown here. Note that 48 and 44% of the days are missing during years 2010 and 2012, respectively.



525

Figure 5. Seasonal mean values of different parameters for NPF days at Puijo: a) estimated formation rates of 3-nm particles ($J_{3,est}$) and observed formation rates of 7-nm particles ($J_{7,obs}$). b) growth rate of the particles within size range 7-20 nm c) coagulation sink (*CoagS*) of 7-nm particles. The height of the bars shows the mean values of data points (i.e. mean values

during 7:00 to 19:00 of the *J* and *CoagS* values for 105 NPF event days) within each season, and the error bars indicate the values between minimum and maximum of the data points. The numbers on top of each bar in middle panel indicate the number of the NPF events in corresponding season. The same applies to the Figures 5-a and 5-c.