Atmospheric
Chemistry
and Physics
Discussions

Interactive comment on "Estimation of atmospheric particle formation rates through an analytical formula: Validation and application in Hyytiälä and Puijo, Finland" *by* Elham Baranizadeh et al.

Anonymous Referee #1

Received and published: 27 January 2017

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.

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.

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.

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'?

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.

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.

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.

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.

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.

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?

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.

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?

Line 213: quantify 'most of those days'

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.

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.

Line 227: 15 days out of how many in total? Line 227: 1.5 hours difference: what percentage of the total time lag is this?

Line 229: quantify 'reasonably good accuracy'

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?

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?

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.

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