

Kupiainen-Määttä presents a Markov chain Monte Carlo (MCMC) study to derive sets of evaporation rates from observed cluster distributions of negatively charged sulfuric acid ammonia clusters. The simulations are expanded by also treating the fragmentation rates of the clusters in the mass spectrometer as unknown parameters that are varied with MCMC as well.

The paper is generally well written. It presents a useful modelling exercise to gain insight into cluster evaporation rates that are difficult to access. The MCMC is especially useful to realize that several different sets of fitting parameters are well suited to describe a set of experimental cluster measurements, and finding one well-fitting solution does not necessarily mean that this is the correct set of parameters. Exploring MCMC for this type of data is valuable. For larger data sets, covering larger ranges of conditions, hopefully in the future more and more firm conclusions can be drawn from this type of analysis.

The paper is publishable in ACP after addressing the following comments:

- 1) p1 l12: The Sipilä et al. 2010 paper is not a good reference for this statement because it claimed that the H₂SO₄/H₂O system alone would be sufficient to explain the nucleation rates as observed in the BL.
- 2) p1 l17: The high res ToF mass spectrometers certainly allowed a lot of advances for characterizing the clusters during nucleation, but also earlier MS studies such as described by Hanson and Eisele, JGR, 2000 and 2002, already allowed to study the first steps of cluster formation for the sulfuric acid/water and sulfuric acid/ammonia systems.
- 3) p2 l26 and line 30/31: Besides Olenius et al., 2013b, also other references for the CLOUD data should be included: At least Kirkby et al., Nature, 2011, Schobesberger et al., ACP, 2015, and Duplissy et al., JGR, 2016, should be cited here as well. These papers are from the experimental groups and describe the experimental set-up and the experimental data in much more detail. Referencing only Olenius et al. does not give credit to the many other groups that contributed in order to set up and perform the CLOUD experiments and to obtain the experimental data that are used here (note, for example, that only authors from U Helsinki are part of Olenius et al. but many more groups were involved running the experiments and obtaining the H₂SO₄ and NH₃ concentrations that are used here).
- 4) p5 l110: The assumption of a size-independent wall loss coefficient is problematic. The diffusion coefficient is strongly size dependent, and a cluster consisting of 5 sulfuric acid molecules will diffuse much slower to the walls than the monomer or dimer. This needs to be mentioned, and it should be discussed in how far it may influence the results.
- 5) Section 3.2.1. and Section 4.3: Besides fragmentation also the transmission efficiency of the mass spectrometer should be discussed (see, e.g. Heinritzi et al., AMT, 2016). The mass dependent transmission efficiency also influences the observed cluster distributions. While fragmentation can only lead to an overestimation of the measured small clusters and underestimation of the large clusters, changes in the transmission efficiency can also have the opposite effect. Transmission efficiency is very dependent on the tuning of the individual mass spectrometer. Influences on the observed distributions due to uncertainties of the transmission efficiency or mass discrimination should be discussed.
- 6) Section 3.3. At some point the limits of the MCMC should be discussed in more detail. Currently this discussion is distributed over the paper and limitations become evident from the results but it would be helpful to state the limitations already in the beginning of Section 3.3. When just 22 experimental distributions can be used to derive a large set of parameters, and additionally the input parameters are correlated, then the solutions will not be unambiguous. More discussion of this is needed.

- 7) Table 1 and section 4.2: The “alternative solutions” and cases (A)-(E) are listed but not explained at all. The differences need to be briefly described so that the reader has some idea about what is different in these cases without reading the Supplementary Material (see also comment #11).
- 8) Figure 6 shows the total fragmentation probabilities, e.g. the upper left panel, displaying $A_2A^- \rightarrow A_{0-1}A^-$, should be formed from #18 and #19 from Fig 5. Why does the peak at about 0.2, where #19 has its maximum, not show up in the upper left panel of Fig 6? Adding a scale to the y-axis could be helpful.
- 9) P17 l11-21: a) An unexpected result is the high stability of the pentamer while the tetramer is less stable. It is mentioned that the stability could be due to hydration of the pentamer but hydration should also stabilize the tetramer. Please discuss.
b) Could it be that the pentamer forms in a “closed shell” cluster configuration that is more stable than the tetramer?
c) The stabilities can also be compared with the lifetimes of clusters discussed in Hanson and Eisele, JGR, 2002, Section 2.3.2 and 3.1.
- 10) Acknowledgment: p18 l14-16. The CLOUD team and CERN resources should be acknowledged for provision of the experimental data.
- 11) Supplementary Material. I am lost in section S2.6. It is not clear how the separation was made and why it was made in the way it was made. At the end of p7 the separation of several cases is briefly explained. I do not understand why parameters 3 and 5 are selected for the separation of the synthetic data and why is parameter 6 selected for the posterior distributions with 1 ppt ammonia and parameter 5 for the 5 ppt ammonia simulations, respectively.
It is stated that “First, it can be seen in Fig. S4 that the posterior distribution of coefficient number 3 has two peaks.” I think, Figure 4 is meant here. But even then, only the purple line (5ppt) has two peaks (are we supposed to look only at the purple line? Why not blue and green?). The selection process seems to be arbitrary.
Furthermore, the five lines of description of S3.3 on page 9 are much too short. It is still unclear what makes the difference for cases (A) and (B), and (C)-(E).
- 12) Section S2.6. Second line: “consider a case were” \rightarrow “consider a case where”