Aerosol formation and growth rates from chamber experiments using Kalman smoothing

Matthew Ozon[1], Dominik Stolzenburg[2], Lubna Dada[2,3,4], Aku Seppänen[1] and Kari E.J. Lehtinen[1,~~3~~5]

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

[2] Institute for Atmospheric and Earth System Research/ Physics, University of Helsinki, 00014 Helsinki, Finland

[3] EPFL, School of Architecture, Civil and Environmental Engineering, Sion, 1951, Switzerland

[4] Paul Scherrer Institute, Laboratory of Atmospheric Chemistry, 5232 Villigen PSI, Switzerland

[5] Atmospheric Research Centre of Eastern Finland, Finnish Meteorological Institute, Kuopio, Finland

Correspondence to: Kari E.J. Lehtinen (kari.lehtinen@uef.fi)

Keywords. Particle number size distribution, nucleation, growth, chamber experiments, Bayesian state estimation, Kalman smoother, DMA train

Abstract.

Bayesian state estimation in the form of Kalman smoothing was applied to Differential Mobility Analyser Train (DMA-train) measurements of aerosol size distribution dynamics. Four experiments were analysed in order to estimate the aerosol size distribution, formation rate and size-dependent growth rate, as functions of time. The first analysed case was a synthetic one, generated by a detailed aerosol dynamics model, and the other three chamber experiments performed at the CERN CLOUD facility. The estimated formation and growth rates were compared with other methods used earlier for the CLOUD data and with the true values for the computer-generated synthetic experiment. The agreement in the growth rates was ~~remarkably~~ very good for all studied cases: estimations with an earlier method fell within the uncertainty limits of the Kalman smoother results. The formation rates matched also well, within roughly a factor of 2.5 in all cases, which can be considered very good ~~especially~~ considering the fact that they were estimated from data given by two different instruments, the other being the Particle Size magnifier (PSM), which is known to have large uncertainties close to its detection limit. The presented Fixed Interval Kalman Smoother (FIKS) method has clear advantages compared with earlier methods that have been applied to this kind of data. First, FIKS can reconstruct the size distribution between possible size gaps in the measurement in such a way that it is consistent with aerosol size distribution dynamics theory, and second, the method gives rise to direct and reliable estimation of size distribution and process rate uncertainties if the uncertainties in the kernel functions and numerical models are known.

1 Introduction

Atmospheric new particle formation and growth are important phenomena when considering global aerosol concentrations. Aerosol number ~~Their~~ concentration together with their size distribution and chemical composition determines how aerosols affect visibility, health and climate (Albrecht, 1989; Appel et al., 1985; Daellenbach et al., 2020; Pope and Dockery, 2006; Twomey, 1974). These are determined by atmospheric dynamics and aerosol dynamics such as new particle formation and growth as well as removal rates. Nieminen et al. (2018) reviewed the existing literature on the formation and growth rates ranging from polar sites, with very small aerosol concentrations to polluted urban sites with

extremely high concentrations. The rates have been typically estimated using the methodology reviewed in Kulmala et al. (2012) and adjusted for chamber experiments by Dada et al. (2020). Both are ,based on rather simple regression or balance equation approaches, suffering from potentially crude approximations and permitting no proper estimation of the uncertainties. At the same time, however, instrument development, especially advances in particle detection efficiency and mass spectrometry, has developed rapidly (Kangasluoma et al., 2020). Potentially superior advanced data analysis methods have not been used, and, it is likely that there are significant inaccuracies in the estimated particle formation and growth rates estimated previously (Kürten et al., 2018).

There have been some attempts to estimate aerosol formation and growth rates with different inverse methods (Henze et al., 2004; Kuang et al., 2012; Lehtinen et al., 2004; Sandu et al., 2005; Verheggen and Mozurkewich, 2006; Viskari et al., 2012). We are, however, not aware of any of the above-mentioned methodology being used widely. The most promising ones in our view, that include also estimations of uncertainties, have been the ones by Kupiainen-Määttä (2016) and Shcherbacheva et al. (2020), who used Markov Chain Monte Carlo methodology to estimate evaporation rates as well as their uncertainties from synthetic cluster dynamics data. In addition, the INSIDE-method by Pichelstorfer et al. (2018), which is based on numerical solution of the aerosol general dynamic equation and matching the solution optimally to integrated measured concentrations of selected size intervals, has been successfully applied to determining growth rates at the CLOUD (Cosmics Leaving OUtdoor Droplets) experiments at CERN (European Organization for Nuclear Research) (Stolzenburg et al., 2020). Furthermore, the results in very recent manuscript by McGuffin et al. (2020), in which nucleation, growth and emission rates of are estimated using techniques from the field of nonlinear process control, seem promising.

In a recent paper, Ozon et al. (2020) presented BAYROSOL, a Julia software package that combines a finite difference solution to the general dynamic equation for aerosols (GDE; Seinfeld and Pandis, 2016) to Bayesian state estimation in order to estimate unknown size dependent process rates (nucleation, condensation, losses) from known time evolution of the aerosol size distribution. Bayesian state estimation is a general framework for estimating time-dependent variables (state variables) based on (direct or indirect) noisy observations that are collected sequentially during the temporal evolution of the state variables (Gelb, 1979). The state estimation is based on the so-called state-space representation, which consists of the state evolution model and observation model. In this work, the state variables consist of the particle size distribution – the temporal evolution of which is modelled with GDE – and the nucleation, growth and deposition rates which are parameters of the GDE. The observation model is the mapping from the size distribution to DMA-train measurements. In Bayesian formulation, both the state variable and the observations are modelled as stochastic processes; their randomness reflects their uncertainty, which decreases, when measurement data is accounted for in the state estimation – formally speaking, this is done by conditioning the state variables with respect to measurement data (realized observations) sequentially. The result of Bayesian state estimation is the posterior probability density which reflects the uncertainty of the state variables after accounting for the measurement data.

A large variety of state estimation schemes exits, and the choice between them depends on 1) the type of the state-space model (linearity, gaussianity, etc); 2) the type of data available when computing an estimate at time t (if data is available up to time $k < t$, the problem is of prediction type, while cases where $k = t$ and $k > t$ are referred to as filtering and smoothing, respectively), and 3) the approximations which are sometimes needed to lower the computational demand of state estimation. In the case of linear and gaussian state-space models, a Bayesian filtering problem can be solved recursively by the well-known Kalman filter algorithm. In non-linear and non-gaussian cases, the rigorous choice is to use so-called particle filters and smoothers (Särkkä, 2013). However, because these MCMC-based estimators are highly time

consuming in large-dimensional cases, approximative methods are often used – such as the extended Kalman filter and smoother adopted in this paper. These recursive algorithms use sequential linearization to approximate the non-linear models and non-gaussian probability distributions.

In previous work ~~Using synthetic SMPS measurements~~ Ozon et al., (2020) showed that the Fixed Interval Kalman Smoother (FIKS; Kaipio and Somersalo, 2005) performed very well in estimating the process rates of the GDE in two distinct cases. First in a case in which continuous nucleation, growth and losses lead to a nearly steady state size distribution and second also in a case in which there is a growing nucleation mode after a nucleation burst. In the method, the unknowns (such as the discretized particle size distribution) are modeled as random variables, and their prior probability distributions are incorporated in the solution of the inverse problem. One important key feature of the Kalman smoother method is that it estimates also the error covariance matrices of the process rates that is their uncertainties, if the uncertainties of the measurement device are known.

In this manuscript, we show results of applying BAYROSOL, for the first time to real experimental data. We use experiments performed with the differential mobility analyser-train (DMA-train; Stolzenburg et al., 2017) measuring new particle formation and growth at the CERN CLOUD chamber (Duplissy et al., 2016; Kirkby et al., 2011). In addition to testing the method with synthetic DMA-train data (in which the 'correct' results are known), we estimate formation and growth rates from three different formation and growth experiments: sulphuric-ammonia (Stolzenburg et al., 2020), alpha-pinene ozonolysis (Heinritzi et al., 2020) and iodic acid (He et al., 2021; Sipilä et al., 2016). We compare the formation rates with results obtained by using the methodology by Dada et al. (2020) based on Particle Size Magnifier (PSM) measurements and the growth rates with the results obtained by the INSIDE method (Pichelstorfer et al., 2018). We chose the DMA-train measurements for three main reasons: first, the high time resolution makes it an ideal instrument for nucleation studies due to a more accurate estimate of dN/dt. Second, the collection efficiencies of the channels have been carefully characterized (Stolzenburg et al., 2017; Wlasits et al., 2020) and yield higher sensitivities to low particle number concentrations (Kangasluoma et al., 2020), which are often faced in atmospherically relevant nucleation studies, and third, the DMA train is, at the same time, interesting and challenging instrument for detailed data analysis because of the gaps in the measured size spectrum.

2   Bayesian framework for parameter estimation

2.1   Aerosol measurement and evolution models

~~The~~ Typical quantities of interest in chamber experiments studying new particle formation are the particle number size-distribution, the formation rate at the critical cluster size and the growth rate of the freshly formed particles. The available measurements to infer these quantities usually size classify the aerosol and measure the size classified number concentrations. While the retrieval of the particle number size distribution from such measurements is a classical inverse problem (Chambolle and Pock, 2011; Fiebig et al., 2005; Kandlikar and Ramachandran, 1999; Wolfenbarger and Seinfeld, 1990), the estimation of the process parameters (formation and growth rate) is often done by analysis of the time evolution of the retrieved particle size distribution (Dada et al., 2020). Here, we focus on the formulation of the problem within a statistical Bayesian framework, where the state parameters described by a measurement model and an evolution model are treated as multivariate random processes and are estimated from measurements using a FIKS (Ozon et al., 2020).

113 A measurement can be described by a vector y^k representing m indirect observations (channels of the instrument) of the

114 particle size distribution $n(d_p, t_k)$. The observations are linked to the size distribution by the so-called instrument transfer

115 (or kernel) functions $\mathcal{H}$ such that:

116 $$y^k = \int \mathcal{H}(d_p) n(d_p, t_k) dd_p \qquad (1)$$

117 The transfer functions $\mathcal{H}$ can be inferred from calibration experiments and instrument design considerations. Direct

118 inversion of Eq. (1) for every time instant t_k is typically an underdetermined and ill-posed inverse problem and requires

119 some additional assumptions in order to avoid oscillatory and unstable solutions (Kandlikar and Ramachandran, 1999).

120 At the same time, the time evolution of the particle size-distribution $n(d_p, t)$ can be described by the aerosol general

121 dynamic equation (GDE):

122 $$\frac{\partial n}{\partial t}(d_p, t) = -\frac{\partial g(d_p, t) n(d_p, t)}{\partial d_p} - \lambda(d_p, t) n(d_p, t) - \text{CoagSink}(\beta, d_p, t) + \text{CoagSrc}(\beta, d_p, t) \qquad (2)$$

123 Here, $g(d_p, t)$ is the condensation growth/evaporation rate, $\lambda(d_p, t)$ is the particle loss by deposition or dilution and

124 CoagSink and CoagSrc are the sink and source rates due to particle coagulation within the size distribution with the

125 coagulation coefficients β. An exact expression of all terms can be found in e.g. Ozon et al. (2020) and Seinfeld and

126 Pandis (2016). The boundary conditions of Eq. (2) are given by the apparent formation rate $J_{d_{min}}(t) =$

127 $g(d_{min}, t) n(d_{min}, t)$ of newly formed particles at the minimum detectable size d_{min} and a zero numerical flux condition

128 $g(d_\infty, t) n(d_\infty, t) = 0$ at very large sizes. Altogether, the process parameters $g(d_p, t), J(t), \lambda(d_p, t)$ and $\beta(d_i, d_j)$ as well

129 as the initial and boundary conditions determine completely the evolution of the size distribution, but especially $g(d_p, t)$

130 and $J(t)$ are usually not known. The coagulation coefficients $\beta(d_i, d_j)$ can often be obtained from theory (and coagulation

131 above d_{min} can be even neglected in many applications with low particle concentrations) and the loss parameters $\lambda(d_p, t)$

132 are well quantified for controlled aerosol chamber experiments.

133 A single measurement of the size distribution y^k does not depend explicitly on the process parameters, but as $g(d_p, t)$

134 and $J(t)$ determine the temporal evolution of $n(d_p, t_k)$ the estimation of the process parameters is feasible from a

135 sequence of l measurements y^k at several time instances.

136 **2.2 State estimation with Kalman smoothing**

137 Following Ozon et al. (2020) we formulate the problem as a Bayesian state estimation problem. After discretization of

138 the problem in size space, i.e. particle diameter ($i = 1, .., q$) and time ($k = 1, ... l$), we can define the state variable $X^k =$

139 $[N_i^k \quad g_i^k \quad \lambda_i^k \quad J^k]$ $X^k = [N^k \quad g^k \quad \lambda^k \quad J^k]^T$ for each time step k with the particle concentrations N_i^k per size

140 discretization bin i, the condensation and loss terms g_i^k and λ_i^k, respectively, for each size discretization bin i and the

141 nucleation rate J^k. Here, we have denoted $N^k = [N_1^k, ..., N_q^k], N^k = [g_1^k, ..., g_q^k]$ and $\lambda^k = [\lambda_1^k, ..., \lambda_q^k]$. The problem can

142 then be formulated as:

143 $$X^{k+1} = F(X^k) + w^k \qquad (3)$$

144 $$y^k = HX^k + v^k \qquad (4)$$

145 Eq. (3) represents the discretized non-linear evolution model, which is based on the general dynamic equation for N^k, on

146 second order processes for g^k and J^k, and a random walk evolution for λ^k (see Section 2.3, Eq (5)-(7)). Eq. (4) represents

147 the discretized linear observation model. The terms w^k and v^k are the error terms, which are approximated as normally

148 distributed $\mathcal{N}(0, \Gamma_w^k)$ and $\mathcal{N}(0, \Gamma_e^k)$ with the covariance matrices Γ_w^k, Γ_v^k, which not only include stochastic noise, but also

149 errors due discretization, model and parameter uncertainties.

150 We note that the above description of the state-space model (3)-(4) is slightly simplified for the sake of notational

151 convenience. Namely, two additional features – both described in detail by Ozon et al. (2020) – are included in the model:

152 First, we assume that the process rates are positive quantities and incorporate this positivity constraint into the evolution

153 model by reparametrizing these quantities in the model. For example, for the nucleation rate J^k, we write $J^k =$

154 $\frac{1}{\alpha}\ln(1 + e^{\alpha\xi_J^k})$, where ξ_J^k is an unconstrained random variable and α is a scaling constant. Respective parametrizations

155 are written for g_i^k and λ_i^k. Secondly, as noted above, second order models are written for rates J^k and g^k. More

156 specifically, we consider the respective state parameters ξ_J^k, ξ_g^k as second order Markov processes; for example

157 $\xi_J^k = \psi_1\xi_J^{k-1} \pm \psi_2\xi_J^{k-2} + \eta$, where ψ_1 and ψ_2 are model parameters and η is Gaussian state noise. The second order models

158 are written, because they promote temporal smoothness of the processes. When the positivity constraint and the second

159 order models are included in the model, the state variable X^k in the state-space model (3)-(4) is rewritten in the form

160 $X^k = \begin{bmatrix} N^k & \xi_g^k & \xi_g^{k-1} & \lambda^k & \xi_J^k & \xi_J^{k-1} \end{bmatrix}^T$, and at each time step the above logarithmic functions are used for mapping the

161 unconstrained variables ξ_J^k, ξ_g^k and ξ_λ^k to respective quantities J^k, g_i^k and λ_i^k. For the details on the above modifications

162 as well as discretization of the GDE model, we refer to Ozon et al. (2020).

163

164 The GDE, i.e. the non-linear evolution model for N^k (Eq. (2)), is similar to an advection equation. Therefore, its

165 numerical discretization schemes are often unstable and must be treated carefully to avoid oscillation and divergence or

166 to minimize numerical diffusion (Shen et al., 2007; Smolarkiewicz, 1984). Thus, we show detailed considerations on the

167 magnitude of the different error terms in the Supplement.

168

Algorithm 1 Extended Kalman Filter (EKF)	**Algorithm 2** Fixed Interval Kalman Smoother (FIKS)
Initial state: Expectation $X^{0\|0}$ and covariance $\Gamma^{0\|0}$	**Initialization:** Run Algorithm 1, store all variables
for $k = 1, \dots, l$	**for** $k = l - 1, \dots, 1$
Prediction: expectation and covariance	Backward gain matrix
$X^{k\|k-1} = F(X^{k-1\|k-1})$	$A^k = \Gamma^{k\|k}(\partial F)^T(\Gamma^{k+1\|k})^{-1}$
$\Gamma^{k\|k-1} = \partial F^{k-1}\Gamma^{k-1\|k-1}(\partial F^{k-1})^T + \Gamma_w^{k-1}$	Smoother expectation and covariance
Kalman gain matrix:	$X^{k\|K} = X^{k\|k} + A^k(X^{k+1\|K} - X^{k+1\|k})$
$K^k = \Gamma^{k\|k-1}(H^k)^T(H^k\Gamma^{k\|k-1}(H^k)^T + \Gamma_v^k)^{-1}$	$\Gamma^{k\|K} = \Gamma^{k\|k} + A^k(\Gamma^{k+1\|K} - \Gamma^{k+1\|k})(A^k)^T$
Measurement update: filter expectation and covariance	**end**
$X^{k\|k} = X^{k\|k-1} + K^k(Y^k - H^kX^{k\|k-1})$	
$\Gamma^{k\|k} = (I - K^kH^k)\Gamma^{k\|k-1}$	
end	

169 **Table 1:** Extended Kalman Filter and Fixed Interval Kalman Smoother algorithms for estimation of the state variables and their
170 variances X^k and Γ^k.

171 Considering this structure of the problem, a non-linear extension to the Kalman Filter (Extended Kalman Filter; EKF) is

172 a well suited algorithm for solving the unknown size-distribution and process parameters (Gelb, 1974; Kaipio and

Somersalo, 2005). It is a two stage recursive procedure, where in the first stage the future state and propagation of error is predicted based on the state evolution model (Eq. (3)). In the second stage, the state variable and its covariance are estimated by updating the predicted state variable and covariance. This so-called measurement update accounts for the discrepancy between the realized measurements at time t_k and modelled measurements corresponding to the predicted state variable. This procedure is repeated until the final measurement $k = l$. After finishing the EKF recursions, we utilize a Fixed Interval Kalman Smoother (FIKS), which consists of a backward recursion from a backward gain matrix and smooths the results by backwards recursion from l to 1. The workflow of the EKF and FIKS are illustrated in Table 1 and more details on this algorithm can also be found in Ozon et al. (2020).

2.3 Adaption to chamber experiments

The state space model has been adjusted to represent best the evolution of an aerosol system during new particle formation experiments in an atmospheric simulation chamber like CLOUD. For the time evolution of the process parameters, we assume a rather smooth evolution for the nucleation and growth rates, approximated by a second order process (Ozon et al., 2020):

$$J^{k+1} = (r_1 + r_2)J^k - r_1 r_2 J^{k-1} + w_J^k \tag{5}$$

$$g^{k+1} = (r_1 + r_2)g^k - r_1 r_2 g^{k-1} + w_g^k \tag{6}$$

The constants r_1, r_2 depend on the characteristic time of change, discretization time and a dampening factor and their definition can be taken from (Ozon et al., 2020) and the corresponding values for our experiments are listed in Table S1 in the Supplement.

In contrast to the growth and formation rates, the loss rates in a chamber experiment do not depend on time Different to the growth and formation rates, Tthe loss rates in a chamber experiment do not depend on time, but can be decribed by time independent wall and dilution losses $\lambda = \lambda_{dil} + \lambda_{wall}(d_p)$. These loss rates are well characterized by dedicated wall loss experiments (Stolzenburg et al., 2020) and the dilution rate of the chamber $\lambda_{dil} = Q_{tot}/V_{chamber}$, where Q_{tot} is the total flow rate to the chamber to maintain constant pressure, and $V_{chamber}$ is the chamber volume. The time evolution is described by a random walk with a small stochastic noise term w_λ^k, and the expectation of the initial state (see Table S1 in the Supplement) is set to the experimentally determined value with a standard deviation of $\pm 10\%$:

$$\lambda^{k+1} = \lambda^k + w_\lambda^k \tag{7}$$

For fully defining the problem (Eq. (3)-(4)), an estimate of the covariance matrices corresponding to the error terms v^k and w^k is needed. The definition of the covariance matrices corresponding to the state noise w^k on the size-distribution evolution w_N^k, the growth rate w_g^k, the formation rate w_J^k and the wall loss rate w_λ^k follow the consideration of Ozon et al. (2020). The covariances of the wall losses, the growth rate and the size-distribution are dominantly diagonal with some additional off-diagonal terms in order to account for a correlation in size. The formulation given by Ozon et al. (2020) was slightly altered to give a stronger correlation between the closest size bins due to the sparser size-resolution of the DMA-train compared to the simulated SMPS system (values for the different experiments are given in Table S1 in the Supplement):

$$\Gamma_{N/g/\lambda}^k(i,j) = \sigma_{i,N/\lambda/g} \sigma_{j,N/\lambda/g} \, \exp\left(-\left(\frac{i-j}{\delta_{N/\lambda/g}}\right)^{a_{N/\lambda/g}}\right) \tag{8}$$

For the size-distribution evolution, we find that $\sigma_{i,N}^2 = (\delta^k)^2 \text{Var}(W_i^k)$, with δ^k the discretization time-step and W_i^k the error of the discretization of the size-distribution evolution. A detailed derivation of W_i^k is given in the Supplement.

210 The modelling error of the observation model and the measurement noise both contributing to v^k are assumed to be

211 mutually independent. For this reason, the covariance of the error term Γ_v^k in the measurement model is written as a sum

212 of the covariances of these two random variables, i.e. $\Gamma_v^k = \Gamma_{mod}^k + \Gamma_y^k$. For a detailed derivation we refer to the

213 Supplement, where we also show that the discretization error is negligible compared to model and measurement error if

214 a fine enough size discretization is chosen. We approximate Γ_v^k with uncorrelated processes, and hence the covariance

215 matrices are of the diagonal form. For the measurement error, the variance is given by Poisson counting statistics in the

216 case of a single-particle counting detector such as a condensational particle counter (CPC):

217 $$\Gamma_y^k(i, i) = y^k(i) \tag{9}$$

218 For the model uncertainty, we assume the variance of the kernel $\mathrm{Var}(\Delta H_{i,j})$ is composed of an uncertainty proportional

219 to $H_{i,j}$ (for example due to an offset in the absolute calibration of the instrument) and a shifting size information error (for

220 example discrepancy between set and actual classified size in a mobility spectrometer). It can then be formulated as

221 (detailed values for the experiments under investigation can be found in Table S1 in the Supplement):

222 $$\Gamma_{mod}^k(i, i) = \sum_{j=1}^q \left(n(d_j) \Delta_j \right)^2 \mathrm{Var}(\Delta H_{i,j}) \tag{10}$$

3   Experimental methods

224 We use experimental data from the CERN CLOUD experiment (Duplissy et al., 2016; Kirkby et al., 2011) where we

225 measured particle size-distributions in the sub-10 nm range with a DMA-train (Stolzenburg et al., 2017). The raw data

226 obtained from the DMA-train is used as input for the analysis of three different sets of experiments performed in the

227 atmospheric simulation chamber. The DMA-train instrument kernels are also used for modelling an instrument response

228 to simulated size-distribution data in order to verify the general performance of the FIKS to DMA-train like data.

3.1   DMA-train

230 The DMA-train is an electrical mobility spectrometer, specifically designed to measure sub-10 nm size-distributions

231 (Stolzenburg et al., 2017). Six identical DMAs are applied in parallel i.e. they sample through the same inlet. They are

232 set to six distinct but fixed voltages and hence classified sizes. The charging state of the aerosol is pre-conditioned in two

233 TSI Inc. Advanced Aerosol Neutralizers (Model 3088), each supplying three DMAs at 5.5 litre per minute (lpm) total

234 flow. We use the Wiedensohler approximation (Wiedensohler, 1988) to describe the steady-state charge distribution at

235 the DMA inlets. Kallinger and Szymanski (2015) showed that for the used neutralizers the steady-state charge distribution

236 is still achieved for flow rates up to 5 lpm and we assume that this holds true for 5.5 lpm flow, too. After size classification,

237 the aerosol is detected in condensation particle counters. Four channels are equipped with TSI Inc. Model 3776 ultrafine

238 CPCs for detection of aerosols down to 2.5 nm. They were operated at reduced temperature settings in order to increase

239 the detection efficiency of the smallest particles, achieving a 50% detection efficiency for particles as small as 2 nm

240 (Wlasits et al., 2020). Two channels of the DMA-train were operated with particle counters specifically designed for sub-

241 2 nm particle detection using diethylene glycol (DEG), an Airmodus Ltd. particle size magnifier (Model A10, PSM) and

242 a TSI Inc. nano-Enhancer (Model 3777). ~~Both are~~ Each is used as a booster stage~~s~~ to activate the particles, which are

243 subsequently detected by either an Airmodus Ltd. CPC (Model A20, for the PSM) or a TSI Inc. CPC (Model 3772, for

244 the nano-Enhancer). Both channels have a higher aerosol flow rate of 2.5 lpm resulting in a broader transfer function and

245 higher transmission at the DMA compared to the 1.5 lpm sample flow in the other four channels. The sheath flow at the

246 DMAs is kept constant at 15 lpm for all six channels.

247 The constant sampling at fixed sizes allows for either a higher time-resolution at large aerosol number concentrations or
248 a higher sensitivity towards low number concentrations due to longer signal averaging times compared to a scanning or
249 stepping differential mobility spectrometer. To increase the number of measured particle sizes, one DMA is still operated
250 in an alternating mode, switching between 6.2 and 8 nm every ten seconds. The other DMAs are set to classify particles
251 of 4.3, 3.2, 2.55, 2.2 and 1.8 nm. The instrument kernels are obtained from calibration experiments, where we use the
252 DMA transfer function and sampling loss characterization from Stolzenburg et al. (2017), the CPC activation efficiencies
253 from Wlasits et al. (2020) and the charging efficiency was tested to follow the Wiedensohler approximation in Tauber et
254 al. (2020). The kernel functions for all seven classified sizes are shown in Fig. 1 for an instrument averaging time of 120
255 seconds, sulphuric acid-like test particles (using the Ammonium Sulfate detection efficiencies from Wlasits et al. (2020))
256 and including the detector flow rates of each condensation particle counter. Therefore, the kernels can be used to convert
257 raw particle counts at the detecting CPCs into a particle size-distribution (within an inverse problem) and vice-versa. Note
258 that for different chemical composition of the input particles, the CPC response might be different. Therefore, the kernels
259 used for analysing experiments where particles were formed from oxidized organics (Kirkby et al., 2016) or from iodic
260 acid (He et al., 2021) are different and approximated by the calibration curves for oxidized beta-caryophyllene and sodium
261 chloride from Wlasits et al. (2020).

262
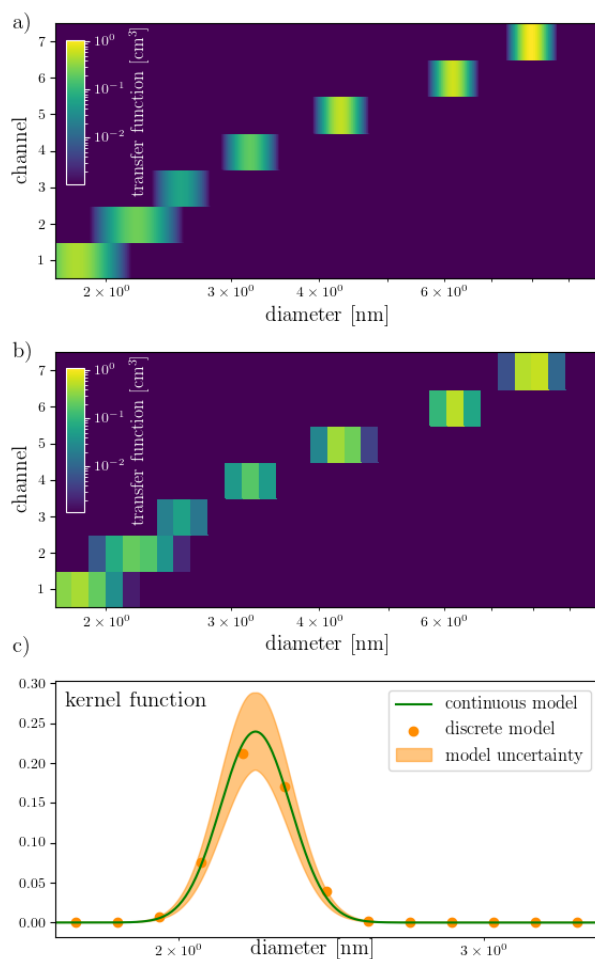
Figure 1: Kernel functions of the DMA-train when classifying sulphuric acid particles at a signal averaging time of 120 seconds. (a)
shows the continuous form of the transfer function (color code) of the seven DMA-train channels (y-axis), inferred from instrument
calibrations (Stolzenburg et al., 2017). (b) shows the discretization into 32 size bins from 1.7 to 10 nm used in the Kalman smoother
(c) shows the comparison between continuous form and used discretization for channel 2 with centroid diameter at 2.2 nm, together
with the model uncertainty, which is used for the error estimate (see Table S1 in the Supplement).

3.2 CLOUD

We use experimental data from the DMA-train measuring new particle formation in nucleation experiments at the CERN CLOUD chamber. The 26.1 m³ stainless-steel chamber provides a high-purity, temperature-controlled environment in order to perform experiments under atmospherically relevant conditions, where trace gases can be added precisely at pptv (parts per trillion by volume) level and sunlight can be simulated by UV-illumination of the chamber. We use three different sets of experiments, of varying chemical composition, in order to demonstrate the performance of the FIKS in reconstructing formation and growth rates. See also Kirkby et al. (2011) and Duplissy et al. (2016).

First, a nucleation experiment using sulphuric acid and ammonia was performed at 5℃ by adding SO_2 and O_3 to the chamber and through the photo-dissociation of ozone, the formation of OH radicals and sulphuric acid was induced which resulted in strong new particle formation (Stolzenburg et al., 2020). Second, we performed nucleation and growth experiments at 5℃ using oxidized organics from dark (i.e. no UV-illumination) ozonolysis of alpha-pinene (Kirkby et al., 2016; Stolzenburg et al., 2018). Both experiments resulted in moderate new particle formation rate and thus, in the Kalman smoother, we used DMA-train data that was averaged over 120 seconds time intervals. Third, we studied nucleation from iodine oxides at 10℃ (He et al., 2021), which resulted in high particle formation rates and fast growth. For the third experiment, we reduced the DMA-train averaging time down to 20 seconds, while keeping high counting statistics over the averaging interval.

3.3 PSM derived formation rates

Particle formation rates (J_{dp}) are calculated from the time derivative of the total particle number concentration larger than 1.7 nm following the method introduced in Dada et al. (2020). The particle number size distribution is measured with the particle size magnifier (PSM) coupled with a condensation particle counter (1.5 – 2.5 nm), a TSI nano-SMPS (3 – 65 nm) and home-built long-SMPS (10 – 800 nm). The formation rates are corrected for the size dependent wall and coagulation losses. Additionally, since the chamber is operated in continuous flow mode, the particle concentrations are corrected for dilution losses. For more information on the PSM derived formation rates, see Dada et al. (2020). The uncertainty on $J_{1.7}$ was assumed to be 30% for the CLOUD chamber derived from the repetition of the same experiment. A procedure as described in Dada et al. (2020) calculated using propagation of error in the concentration measurement, dilution, coagulation and wall losses as well as the error on the time-derivative of the total particle concentration within a Monte-Carlo simulation could be used if such repetition experiments were not available. It needs to be noted, that for comparison of the formation rate value at the arbitrary minimum detectable size $d_{\min} = 1.7$ nm with a system inherent nucleation rate at the critical cluster size, additional sophisticated approaches might be necessary (Kürten et al., 2015).

3.4 Growth rates using INSIDE

In order to compare the particle growth rates derived by Kalman smoothing, we use the size- and time-dependent growth rate analysis tool INSIDE (Pichelstorfer et al., 2018). It uses input particle size-distributions at time t_1 in order to simulate the known aerosol dynamics (coagulation, wall losses and dilution) until a time t_2 (typically separated by one measurement cycle of an instrument, i.e. the 120 seconds averaging time mentioned above). At t_2, the simulated aerosol size-distribution is compared to the measured size-distribution and by evaluating the general dynamics equation above a certain diameter d_{eval} the growth term $\left.\frac{\mathrm{d}d_p}{\mathrm{d}t}\right|_{d_{\mathrm{eval}}} = \frac{\left.\frac{\mathrm{d}N_\infty}{\mathrm{d}t}\right|^\infty_{d_{\mathrm{eval}}} - \left.\frac{\mathrm{d}N_\infty^{sim}}{\mathrm{d}t}\right|^\infty_{d_{\mathrm{eval}}}}{n(d_p,t)|_{d_{\mathrm{eval}}}}$ can be computed. Evaluating this equation for

several evaluation diameters d_{eval} and at all measurement times t_i, this results in time- and size-resolved growth rates from a size-distribution measurement., which This is distinct from most others, integrative growth rate analysis approaches, which can only derive one growth rate value for a specified size-interval in a single run (Dada et al., 2020; Kulmala et al., 2012; Lehtipalo et al., 2014; Paasonen et al., 2018). Last it should be noted that for INSIDE,. However, compared to the Kalman smoothing, each time step is analysed individually and the growth rate analysis framework relies on already inverted size distributions. For this inversion only, where a simple point-by-point inversion procedure is used for the DMA-train data of this work, assuming narrow DMA transfer functions and little variation of the size-distribution across it (Stolzenburg and McMurry, 2008). Moreover, the INSIDE method does not provide an uncertainty estimate on the growth rate calculation and hence a comparison with other growth rate calculations is beneficial to verify the results the Kalman smoothing will provide valuable insights on the uncertainty related to growth rate measurements.

4   Results: Simulation and Experimental

4.1   Numerical simulation test

First, we we modelled the DMA-train instrument numerically and simulated a data set representing a typical nucleation experiment performed in an atmospheric simulation chamber like CLOUD and modelled the DMA-train response according to the above Kernel functions and then applied the FIKS to this synthetic dataset. We used the same framework as in Ozon et al. (2020) to simulate a nucleation experiment with formation rate at 1 nm $J_{1.0} = 5.25$ cm^{-3} s^{-1}, and size-independent growth rate $GR = 2.5$ nm h^{-1} and the loss rates equal to the CERN CLOUD experiment (Stolzenburg et al., 2020). The evolution of the simulated size-distribution is shown in Fig. 2a. The measurement data y^k are then simulated using the kernel functions from Fig. 1a and altered with a Poisson-distributed random counting error. The FIKS is then applied to the measurement data with the input parameters given in Table S1 in the Supplement. We use a resolution of 32 bins from 1.7 to 10 nm for the FIKS to keep the computational effort low. We tested also 16 to 64 size discretization bins, but higher resolution required additional adjustments in the size-correlation of the covariance given in Eq. (8), which would result in significant differences compared to the original work of Ozon et al. (2020) without providing significantly more accuracy.

Figure 2 also shows the Kalman smoother estimates for the size-distribution, growth rate and formation rate at 1.7 nm. The reconstructed size-distribution (Fig. 2b) is very similar to the true size-distribution (Fig. 2a), especially taking into account the sparser discretization of the former. Also the estimated growth and nucleation rates agree well with true values of the respective process rates specified in the simulation. Moreover, the uncertainty estimates are feasible: For both quantities, the true values are within the uncertainty limits given by FIKS. The reconstructed size-distribution and especially, growth rate (Fig. 2 b and c, respectively) show some temporal oscillations, which are related to periods for the particle population to grow from the size range visible for one DMA-train channel to the next one. That is, the oscillation is a result of an insufficient coverage of the size-range by the DMA-train kernels. The problem could be approached by application of a regularization scheme in the measurement model (Voutilainen and Kaipio, 2001) or by adjusting the kernels improving the overlap, which will be and is discussed in more detail in Section 4.5. Nevertheless, the overall good retrieval of the simulated size distribution and process rates demonstrates that Kalman smoothing approach is well suited for analysing DMA-train data from chamber nucleation experiments.
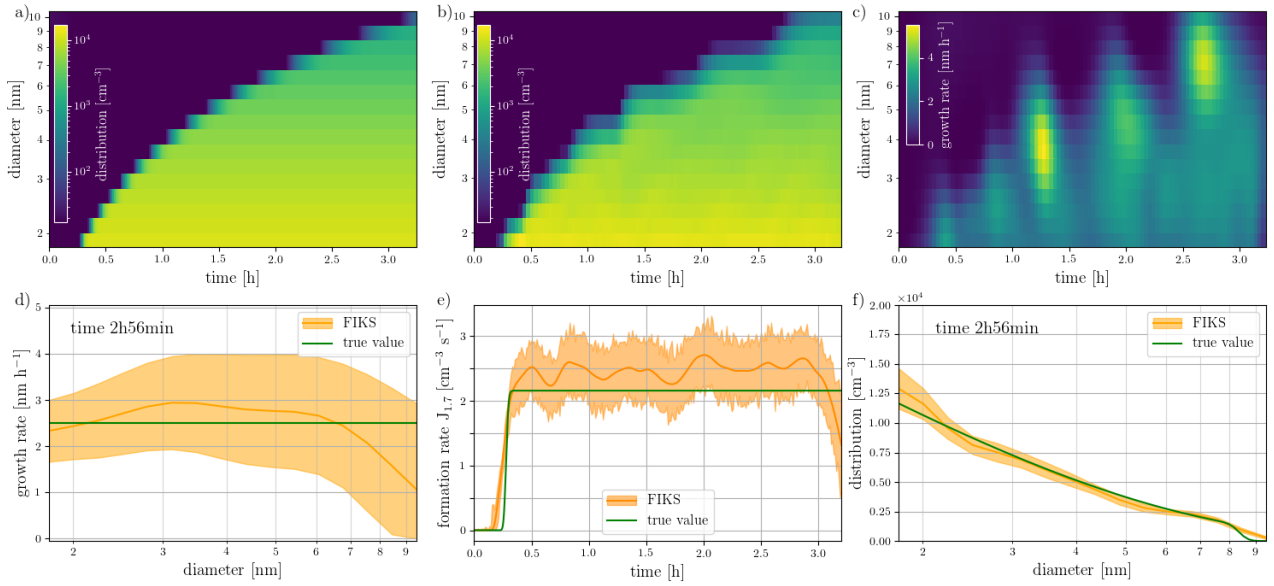
Figure 2: Results of the simulated chamber nucleation experiment: (a) The simulated, temporally evolving size distribution of aerosols. (b) The FIKS estimate for the size distribution. This estimate was computed based on the numerically simulated DMA-train data corresponding to the synthetic size distribution shown in (a). (c) The FIKS estimate for the growth rate of particles. (d) The growth rate corresponding to a single instant of time (2h 56min); here, the FIKS estimate and the associated uncertainty limits are plotted together with the true growth rate. (e) The FIKS estimate for the formation rate at 1.7 nm and its uncertainty, and the true formation rate. (f) FIKS estimate, its uncertainty and the true value of the particle size distribution corresponding to a single instant of time (2h 56min). In subfigures d-f, the orange lines represent the FIKS estimates (posterior expectations) and the orange shaded areas illustrate the uncertainties of the associated variables (more specifically, their approximate posterior standard deviation limits). The true values of the quantities are plotted in green.

4.2 Sulphuric Acid-Ammonia experiment

We applied the FIKS to experimental data from a sulphuric acid-ammonia nucleation and growth experiment performed at 5℃ at the CERN CLOUD chamber. The raw data measured with the DMA-train were averaged in 120-second time intervals in order to increase the counting statistics per channel and then used as input for BAYROSOL. The details of parameter choices in FIKS are given again in Table S1 in the Supplement. The results of applying the Kalman smoothing to this experimental data are shown in Fig. 3.
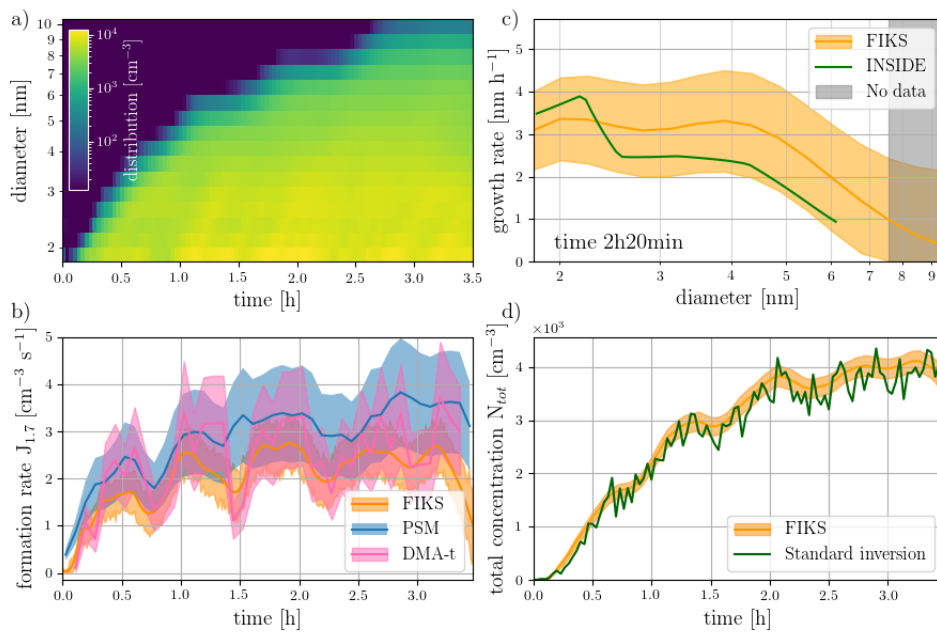
Figure 3: Results of an experiment performed at the CERN CLOUD chamber: The DMA-train data was acquired during sulphuric acid-ammonia nucleation and growth processes. Experimental conditions: 5°C, 60% RH, 5 ppb SO_2, 120 ppb O_3 and 40 ppt NH_3. (a) The FIKS estimate for the size distribution. (b) The evolution of the formation rate during the experiment. The orange line and shaded area represent the FIKS estimate and its uncertainty, respectively, while the solid blue line and blue area show the nucleation rate estimate and the associated 30% uncertainty based on a different set of instruments (particle size magnifier, PSM) and the standard approach for inferring nucleation rates at 1.7 nm (Dada et al., 2020). The pink line and shaded area show the standard approach using the data from the DMA-train only. (c) The growth rate corresponding to time 2h20min after the start of the experiment. The FIKS estimate and the associated uncertainty are marked with orange line and shaded area, respectively. The green line represents the growth rate estimate, which is obtained from DMA-train data by traditional inversion (INSIDE method). (d) The total number concentration reconstructed from the standard inversion approach (green line) and FIKS (orange line and shaded area for the uncertainty range).

The size- and time-dependence of the FIKS estimate for the true size-distribution (Fig. 3a) is very smooth, and it is also able to bridge the information gaps between the largest size-distribution channels. This is a significant improvement from the traditional point-by-point inversion, where the data of each DMA-train channel is inverted independently: Here the reconstructed total number concentration is significantly more noisy. When point by point inversion was applied to the same dataset, those information gaps caused severe reconstruction errors (discontinuities) in the size distribution, see (Stolzenburg et al., 2020). The smoother FIKS reconstructed size-distribution and total number concentration (Fig. 3d) still shows also some small temporal fluctuations. The fluctuations are even more clearly visible in the reconstructed evolution of the formation rate (Fig. 3b). In contrast to the oscillations found in the growth rate for the simulated case, the fluctuations do not only occur when the size-distribution reaches a new DMA-train size channel. FurthermoreRemarkably, the same fluctuations are also recovered when the formation rate is inferred from the PSM and nano-SMPS, i.e. an entirely different set of instrumentation and different type of data-analysis. This result strongly suggests that the fluctuation of the particle formation truly occurs physically in the experiment and is not a reconstruction error caused by instrument noise or a bias caused by the inversion method. The absolute values of the two independent formation rate estimates agree upon a factor of 1.5. Furthermore, large portions of the uncertainty intervals of these two estimates overlap with each other, which is another indicator of the feasibility of FIKS to analysing DMA-train data. In addition, also the inferred growth rates agree within the systematic uncertainties for both approaches (Fig. 3c for one time instant). It is worth noting, however, that both growth rate estimates rely on data from the same instrument and are hence more interdependent than the formation rate estimates. Nevertheless, the good agreement between them corroborates the feasibility of Kalman smoothing to reconstructing nanoparticle growth rates from experimental data.

4.3 Alpha-pinene ozonolysis experiment

Third, we applied the FIKS to experimental data obtained by the DMA-train from an alpha-pinene ozonolysis experiment. Besides the different chemical composition of the growing particles (resulting in different assumptions on the DMA-train transfer functions) in comparison to the experimental results used in section 4.2, this experiment is characterized by a different size-dependency of the growth rate, due to the increased condensation of low- and semi-volatile organics with increasing particle size (Simon et al., 2020; Stolzenburg et al., 2018; Tröstl et al., 2016). The formation rate, however, remains rather similar to the sulphuric acid-ammonia experiment under these specific experimental conditions, but the slower initial growth rates result in generally lower produced particle concentrations. The results of applying FIKS to the data from this experiment are shown in Fig. 4. Again, the size-distribution given by FIKS is much smoother than that obtained with a standard inversion procedure. (see Heinritzi et al. (2020) for the same dataset inverted by the standard approach and Fig. 4d for the total number concentration evolution retrieved from the standard inversion and the FIKS).

Because in alpha-pinene ozonolysis experiment the nucleation and growth rates were known to be similar to those in the sulphuric acid-ammonia experiment, the parameters of their evolution models were selected as in Section 4.2 (Table S1 in the Supplement). The formation rate is again slightly lower by a factor of 2.5 than the one obtained from the PSM.

Also the formation rate retrieved with the method from Dada et al. (2020) but using the DMA-train data is significantly higher, but in-between the two estimates. The possible deviation has hence two plausible reasons: The instrumental differences can be caused by different calibration procedures for the DMA-train and PSM (Dada et al. (2020); direct cross-calibration using NAIS versus Wlasits et al., (2020) using beta-caryophyllene ozonolysis surrogates). And the methodological differences could arise from the very low counting statistics in the DMA-train during this experiment compared to the other two, which will cause the inherent Gaussian assumption of the FIKS to fail. As the deviation in the reconstructed total number concentration of the DMA-train data using two inversion procedures is only a factor of ~1.3 (Fig. 4d), the formation rate discrepancies could be largely due to the more difficult and uncertain calibration procedures. Considering the fact that inter-instrument deviations in sub-10 nm size-distribution measurements can be as large as one order of magnitude (Kangasluoma et al., 2020), the achieved agreement is remarkable, especially as some fluctuations within the chamber can again be reconstructed in both approaches. The retrieved growth rates from the FIKS estimate and the INSIDE method agree remarkably well and both show the increasing growth rates with increasing particle size up to 4.3 nm, an indication for a strong Kelvin-effect in organic condensation (Stolzenburg et al., 2018; Tröstl et al., 2016). The decreasing Kalman smoother estimate above 4.3 nm is related to the fact that the FIKS searches a smooth estimate of the growth rate (Eq. (6)), but at that point (2h30min after experiment start), practically no information from the size-distribution above 4.3 nm is available, resulting in a slow decrease towards zero. This is due to the strong smoothness a priori used in the FIKS algorithm (Table S1 in the Supplement). The INSIDE method does not report growth rates from regions with no information (Pichelstorfer et al., 2018) from the size-distribution and hence the estimate stops at 4.3 nm.
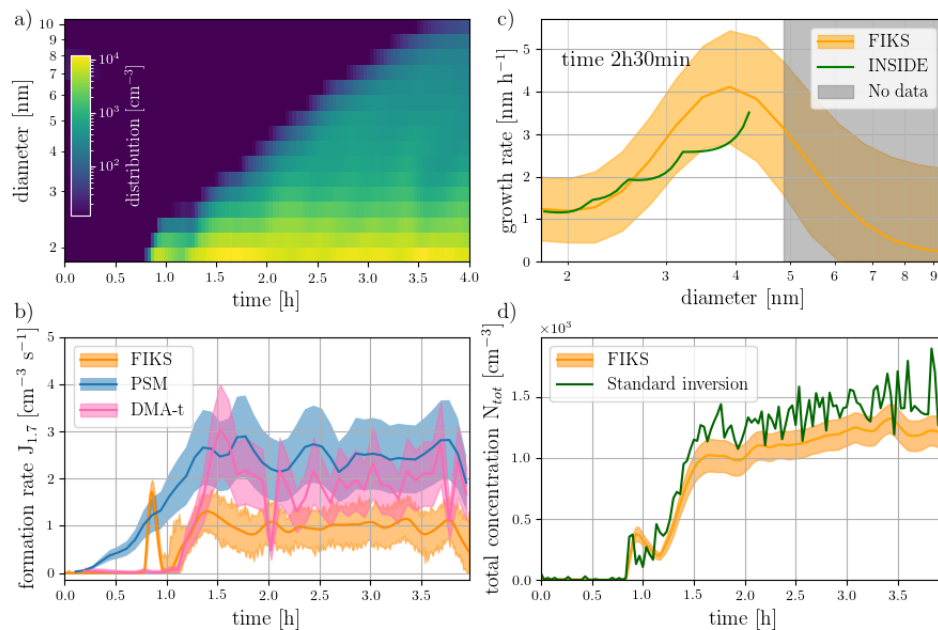
Figure 4: Results of an experiment performed at the CERN CLOUD chamber: The DMA-train data was acquired during an alpha-pinene ozonolysis. Experimental conditions: Experimental conditions: 5°C, 40% RH, 300 ppt alpha-pinene, 40 ppb O₃. (a) The FIKS estimate for the size distribution. (b) The evolution of the formation rate during the experiment. The orange line and shaded area represent the FIKS estimate and its uncertainty, respectively, while the solid blue line and blue area show the nucleation rate estimate and the associated 30% uncertainty based on a different set of instruments (particle size magnifier, PSM) and the standard approach for inferring nucleation rates at 1.7 nm (Dada et al., 2020). The pink line and shaded area show the standard approach using the data from the DMA-train only. (c) The growth rate corresponding to time 2h30 min after the start of the experiment. The FIKS estimate and the associated uncertainty are marked with orange line and shaded area, respectively. The green line represents the growth rate estimate, which is obtained from DMA-train data by traditional inversion (INSIDE method). (d) The total number concentration reconstructed from the standard inversion approach (green line) and FIKS (orange line and shaded area for the uncertainty range).

4.4 Iodic Acid experiment

Finally, we analysed a more dynamic experiment of nucleation and growth from iodic acid (He et al., 2020, 2021). The experiment is characterized by extremely high nucleation rates, which are two orders of magnitude higher than in the sulphuric acid-ammonia and alpha-pinene ozonolysis experiments. However, the growth rate is only half an order of magnitude higher compared to the other two example cases. Figure 5 shows that, in spite of the highly dynamic experiment, the formation rate recovered by Kalman smoothing agrees with the estimate obtained from the PSM (Fig. 5b). The usage of the DMA-train data with a time resolution of 20 seconds allows for the precise recovery of the spike in formation rate in the beginning of the experiment. This causes the build-up of a high condensation sink and vapour/cluster depletion almost shutting off any further nucleation during the continuation of the experiment. The four-minute time resolution of the data for the calculation of the nucleation rate from the PSM is limited in that respect. The reconstructed growth rates (Fig. 5c) show again agreement between the FIKS estimate and the INSIDE method, indicating a clear decreasing trend with size, which is expected for condensation at the kinetic limit if the vapour molecular size is taken into account (He et al., 2021; Lehtinen and Kulmala, 2003; Nieminen et al., 2010; Stolzenburg et al., 2020). The lower values towards 1.8 nm could be caused by a biased estimate of the PSM detection efficiency, because neither a calibration for iodic acid clusters nor for sodium chloride particles (which was used for the other detectors in the DMA-train) was available. The instabilities in the size-distribution at the smallest sizes and fluctuations of the formation rate are expected considering the highly dynamic process of this experiment. Overall, the good agreement for the inferred process parameters of the aerosol general dynamics equation, i.e. the formation and growth rates, is still remarkable. However, the reconstructed size-distribution from the FIKS estimate shows some discontinuities, especially during the growth above 3 nm. This is because the available instrument information from the DMA-train starts to get very sparse in that size range given the dynamic processes involved in the iodic acid nucleation and growth. This also causes the overshooting in the total number concentration using the standard inversion compared to the FIKS result (Fig. 5d), as the larger size channels cover a very broad range at too high resolution for the linear interpolation of the standard inversion used to obtain N_{tot}. More available size channels (hence more DMAs in the case of the DMA-train) would help to resolve such discontinuities. We will therefore provide an instrument design recommendation based on simulated data in the next section.
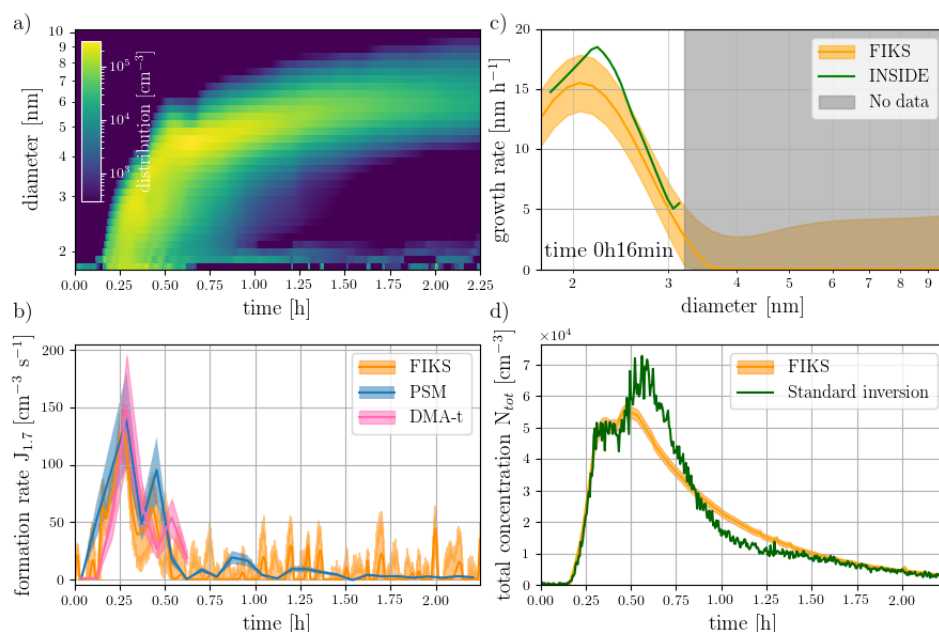
Figure 5: Results of an experiment performed at the CERN CLOUD chamber: The DMA-train data was acquired during iodic acid nucleation and growth processes. Experimental conditions: Experimental conditions: 10℃, 80% RH, 100 ppt I$_2$, 40 ppb O$_3$. (a) The FIKS estimate for the size distribution. (b) The evolution of the formation rate during the experiment. The orange line and shaded area represent the FIKS estimate and its uncertainty, respectively, while the solid blue line and blue area show the nucleation rate estimate and the associated 30% uncertainty based on a different set of instruments (particle size magnifier, PSM) and the standard approach for inferring nucleation rates at 1.7 nm (Dada et al., 2020). The pink line and shaded area show the standard approach using the data from the DMA-train only. (c) The growth rate corresponding to time 0h16min min after the start of the experiment. The FIKS estimate and the associated uncertainty are marked with orange line and shaded area, respectively. The green line represents the growth rate estimate, which is obtained from DMA-train data by traditional inversion (INSIDE method). (d) The total number concentration reconstructed from the standard inversion approach (green line) and FIKS (orange line and shaded area for the uncertainty range).

4.5 Instrument design recommendation from a signal processing point of view

The size range covered by the seven DMA-train channels was chosen semi-arbitrarily based on some external constraints: the lowest measured centroid diameter was supposed to be as close as possible to 1.7 nm where the formation rate is typically measured for experiments performed at the CLOUD chamber. In order to cover the sub-10 nm range the largest channels was set to 8 nm. The other channel diameters were chosen to yield sufficient coverage of the sub-3 nm range. The width of the transfer functions was fixed by the detector sample flow rates and the 15 lpm critical orifices provided by Grimm Aerosol for the DMA sheath air supply (also the standard flow rate used by the manufacturer for this type of DMA). However, the chosen centroid diameters (and hence fixed voltages at the DMAs) and selected sheath flow rates could easily be altered.

In order to study numerically the effect of choice for the DMA-train channels, we constructed a model corresponding to a channel choice different from that in the previous sections: The kernel for the DMA-train with seven channels having centroid diameters of 1.8, 2.2, 2.8, 3.6, 4.6, 6.1 and 8.0 nm is illustrated in Fig. 6a. Moreover, channels 3 to 7 have been altered to use a 2.5 lpm sample flow rate (could be achieved by a make-up flow at each CPC) and a reduced sheath flow rate of 10 lpm, only providing a ~~resolution of~~ theoretical non-diffusive resolution of ~4, which is significantly lower. However, this permits covering the entire size-range between 1.8 and 8 nm by overlapping channels. We revisited the numerical simulation described in Section 4.1 using DMA-train model corresponding to this configuration. The resulting size-distribution given by FIKS is shown in Fig. 6b. Comparison between the two reconstructions in Fig. 2b and Fig. 6b reveals that the new choice of DMA-channels stabilizes the size-distribution estimate, which is supported by the comparison of the estimated total number concentrations (Fig. 6d). While with the original design ~~While in Fig. 2b,~~ the size distribution evolves in step-wise manner when the growing particle mode reaches larger size-channels, the size distribution retrieved with the adjusted kernels ~~in Fig. 6b~~ is ~~very~~ smoother, also temporally, which is especially visible in the evolution of the total number concentration. ~~The~~ The improvement is also significant for the retrieval of the growth rate, where the estimate no longer overshoots as illustrated by the snapshot size dependence of the growth rate for a single instant in time, shown in Fig. 6c. ~~latter estimate also resembles the true (simulated) size distribution (Fig. 2a) more than the former one.~~ It should be noted, that also regularization schemes in the measurement model as proposed by Voutilainen and Kaipio (2001) could provide smoother estimates, but this would need a significant adjustment of the algorithm provided by Ozon et al. (2020) and is hence not implemented here.

The result of this additional numerical study thus demonstrates that in the chosen conditions of the simulation, the reconstruction quality improves when the resolution of individual channels is lowered. This seemingly counterintuitive effect stems from the fact that FIKS estimates do not rely only on the measurements, but are also advised by the GDE model, which makes the problem of optimizing the measurement design a somewhat cumbersome task. A rigorous investigation of the optimal experimental design is out of the scope of this paper, but the above observation is worth

noticing -- especially because a lot of recent experimental effort in the sub-10 nm range has been devoted to improving the instrument resolution (Kangasluoma et al., 2020). While on the individual channel level this might reduce systematic uncertainties, as discussed in Kangasluoma et al. (2020), signal processing rather requires a broad coverage of the size distribution than high resolution. However, the ideal instrument would combine both, full coverage of the size distribution but achieved with more, high-resolution channels. For the DMA-train principle, this would require new ideas in instrument design in order to incorporate more DMAs without the instrument becoming impractically bulky. ~~This would also be needed to E~~extending the measurement size-range above 10 nm would require even more DMAs, but this range is usually well-covered by commercially available instruments, which could easily be added to the FIKS measurement model, facilitating the measurement of experiments where particle grow well beyond 10 nm.
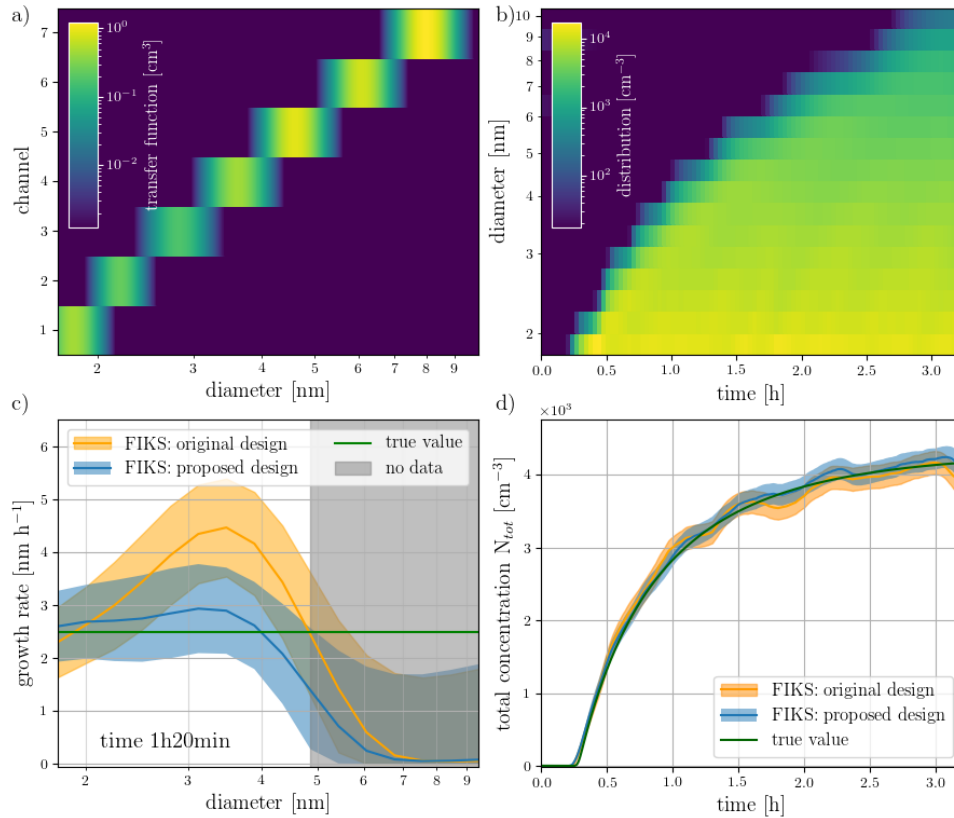
Figure 6: Adjusted DMA-train kernel for better signal processing (a), which could be achieved by choosing suitable centroid voltages and altering the resolution of the DMAs through sheath-flow and sample-flow adjustments. (b) shows the reconstructed simulated size-distribution of Fig. 2 using the adjusted kernel. (c) The growth rate corresponding to time 1h20min min after the start of simulation. The orange line and shaded area are the FIKS estimate using the original design of the DMA-train, while the blue line and shaded area indicate the FIKS result from the proposed new design. The green line shows the true value used to create the input size distribution. (d) shows the evolution of the total particle number concentration for the input size-distribution (green), the original kernel (orange line and shaded area for the uncertainty) and the proposed design (blue line and shaded area).

5   Conclusion

A recently developed methodology (BAYROSOL) applying the Fixed Interval Kalman Smoother (FIKS) to a finite difference solution to the aerosol GDE was used to analyze DMA-train measurements of aerosol dynamics at the CLOUD chamber facility at CERN. The overall aim of this methodology is to estimate unknown aerosol microphysical process rates as well as their uncertainties from size-distribution evolution measurements. In a previous paper, the methodology was shown to be able to predict new particle formation, growth and loss rates from synthetic computer-generated aerosol size distribution evolution data, while here the method has been applied to real experimental data for the first time. Four

524 experimental cases with particle formation and growth were tested: 1) a computer-generated synthetic case, 2) sulphuric

525 acid-ammonia, 3) alpha-pinene ozonolysis, and 4) iodic acid.

526 The DMA-train was selected for two main reasons: first, the instrument kernel functions are well characterized giving

527 rise to reliable estimation of the instrument uncertainties and second, new particle formation rates have not been estimated

528 directly from DMA-train measurements before. In addition, as the current version of the DMA-train is designed in such

529 a way that the individual DMAs have a rather narrow collection kernels for the channels, with significant gaps between

530 some of the channels, the FIKS can reconstruct the size distributions from the measured signals in such a way that the

531 distributions are rather smooth and consistent with the GDE.

532 We compared the growth rates, which with BAYROSOL can be estimated as functions of both size and time, with those

533 obtained by INSIDE, a method applied earlier to CLOUD data, and the agreement was remarkably good for all studied

534 cases. INSIDE is also based on matching the GDE solution to measured size distribution dynamics, however without the

535 capability of estimating uncertainties of the estimations. The FIKS-based estimates for the particle formation rates were

536 compared with those estimated from data obtained by a separate instrument, the Particle Size Magnifier (PSM), based on

537 the rate of change of the total number concentration measured by the instrument corrected by coagulation and wall-losses.

538 Again, the agreement was very good, especially considering the fact that instrument uncertainties are large at the very

539 lowest end of the measured size spectrum. For the iodic-acid case, the FIKS estimate of the formation rate was even able

540 to capture rapidly changing dynamics of the experiment. It was remarkable for all cases that some fluctuations in the

541 formation rates were recovered by both methods independently, indicating that these are physical variations during the

542 experiment

543 Finally, we utilized the FIKS from an instrument development point of view. Typically, an as-high-as-possible resolution

544 for the different measurement channels has been the aim when measuring nanometer-sized particles. This aim, however,

545 gives rise to gaps in the measured size range, as is the case in the DMA-train studied. Thus we studied whether a better

546 coverage of the size spectrum, but lower resolution of the individual channels would be advantageous for size distribution

547 estimation.

548 Summarizing, we believe that Bayesian state estimation methods such as FIKS can be very useful in the field of aerosol

549 science in many aspects. As mentioned above, they can be used to fill gaps in measurements in such a way that not only

550 the obtained size distributions but also unknown process rates are consistent with theory describing aerosol size

551 distribution dynamics. In addition, the methodology provides estimations of uncertainties both for size distributions as

552 well as process rates based on uncertainty estimations in the measurements and the used models, which is unfortunately

553 not common when reporting results of aerosol measurements. Finally, it is conceivable that the methodology presented

554 here will be superior to several previous approaches when combining measurement data obtained with several different

555 instruments that operate at different size ranges. This will be a topic of our forthcoming studies.

556 **Data availability**

557 The version of the implementation of the estimation method (BAYROSOL1.1)[https://doi.org/10.5281/zenodo.4450492]

558 is available under the MIT Expat License; it is the version used to generate the results discussed in this paper. The package

559 also contains the code used to obtain the results discussed in Ozon et al. 2020. It is possible to generate the simulated data

560 described in section 4.1. The experimental data (DMA-train kernels) and raw data, are also included in the repository.

Competing interest

The authors declare no competing interests.

Acknowledgements

This publication has been produced within the framework of the EMME-CARE project, which has received funding from the European Union's Horizon 2020 Research and Innovation Programme (under grant agreement no. 856612) and the Government of Cyprus. This research has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 895875 ("NPF-PANDA") as well as Academy of Finland Project #325647. The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

Author contributions

MO, DS, LD analysed the experimental data, MO developed the software, all authors were involved in the scientific discussion and in the writing of the manuscript.

References

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245(4923), 1227–1230, doi:10.1126/science.245.4923.1227, 1989.

Appel, B. R., Tokiwa, Y., Hsu, J., Kothny, E. L. and Hahn, E.: Visibility as related to atmospheric aerosol constituents, Atmos. Environ., 19(9), 1525–1534, doi:https://doi.org/10.1016/0004-6981(85)90290-2, 1985.

Chambolle, A. and Pock, T.: A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging, J. Math. Imaging Vis., 40(1), 120–145, doi:10.1007/s10851-010-0251-1, 2011.

Dada, L., Lehtipalo, K., Kontkanen, J., Nieminen, T., Baalbaki, R., Ahonen, L., Duplissy, J., Yan, C., Chu, B., Petäjä, T., Lehtinen, K., Kerminen, V.-M., Kulmala, M. and Kangasluoma, J.: Formation and growth of sub-3-nm aerosol particles in experimental chambers, Nat. Protoc., doi:10.1038/s41596-019-0274-z, 2020.

Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L.-E., Leni, Z., Vlachou, A., Stefenelli, G., Canonaco, F., Weber, S., Segers, A., Kuenen, J. J. P., Schaap, M., Favez, O., Albinet, A., Aksoyoglu, S., Dommen, J., Baltensperger, U., Geiser, M., El Haddad, I., Jaffrezo, J.-L. and Prévôt, A. S. H.: Sources of particulate-matter air pollution and its oxidative potential in Europe, Nature, 587(7834), 414–419, doi:10.1038/s41586-020-2902-8, 2020.

Duplissy, J., Merikanto, J., Franchin, A., Tsagkogeorgas, G., Kangasluoma, J., Wimmer, D., Vuollekoski, H., Schobesberger, S., Lehtipalo, K., Flagan, R. C., Brus, D., Donahue, N. M., Vehkamäki, H., Almeida, J., Amorim, A., Barmet, P., Bianchi, F., Breitenlechner, M., Dunne, E. M., Guida, R., Henschel, H., Junninen, H., Kirkby, J., Kürten, A., Kupc, A., Määttänen, A., Makhmutov, V., Mathot, S., Nieminen, T., Onnela, A., Praplan, A. P., Riccobono, F., Rondo, L., Steiner, G., Tome, A., Walther, H., Baltensperger, U., Carslaw, K. S., Dommen, J., Hansel, A., Petäjä, T., Sipilä, M., Stratmann, F., Vrtala, A., Wagner, P. E., Worsnop, D. R., Curtius, J. and Kulmala, M.: Effect of ions on sulfuric acid-water binary particle formation: 2. Experimental data and comparison with QC-normalized classical nucleation theory, J. Geophys. Res.-Atmos., 121(4), 1752–1775, doi:10.1002/2015JD023539, 2016.

Fiebig, M., Stein, C., Schröder, F., Feldpausch, P. and Petzold, A.: Inversion of data containing information on the aerosol

596 particle size distribution using multiple instruments, J. Aerosol Sci., 36(11), 1353–1372,
597 doi:https://doi.org/10.1016/j.jaerosci.2005.01.004, 2005.

598 Gelb, A.: Applied optimal estimation, MIT Press, Cambridge, Massachusetts., 1974.

599 He, X.-C., Iyer, S., Sipilä, M., Ylisirniö, A., Peltola, M., Kontkanen, J., Baalbaki, R., Simon, M., Kürten, A., Tham, Y.
600 J., Pesonen, J., Ahonen, L. R., Amanatidis, S., Amorim, A., Baccarini, A., Beck, L., Bianchi, F., Brilke, S., Chen, D.,
601 Chiu, R., Curtius, J., Dada, L., Dias, A., Dommen, J., Donahue, N. M., Duplissy, J., El Haddad, I., Finkenzeller, H.,
602 Fischer, L., Heinritzi, M., Hofbauer, V., Kangasluoma, J., Kim, C., Koenig, T. K., Kubečka, J., Kvashnin, A.,
603 Lamkaddam, H., Lee, C. P., Leiminger, M., Li, Z., Makhmutov, V., Xiao, M., Marten, R., Nie, W., Onnela, A., Partoll,
604 E., Petäjä, T., Salo, V.-T., Schuchmann, S., Steiner, G., Stolzenburg, D., Stozhkov, Y., Tauber, C., Tomé, A., Väisänen,
605 O., Vazquez-Pufleau, M., Volkamer, R., Wagner, A. C., Wang, M., Wang, Y., Wimmer, D., Winkler, P. M., Worsnop,
606 D. R., Wu, Y., Yan, C., Ye, Q., Lehtinen, K., Nieminen, T., Manninen, H. E., Rissanen, M., Schobesberger, S., Lehtipalo,
607 K., Baltensperger, U., Hansel, A., Kerminen, V.-M., Flagan, R. C., Kirkby, J., Kurtén, T. and Kulmala, M.: Determination
608 of the collision rate coefficient between charged iodic acid clusters and iodic acid using the appearance time method,
609 Aerosol Sci. Tech., 1–12, doi:10.1080/02786826.2020.1839013, 2020.

610 He, X.-C., Tham, Y. J., Dada, L., Wang, M., Finkenzeller, H., Stolzenburg, D., Iyer, S., Simon, M., Kürten, A., Shen, J.,
611 Rörup, B., Rissanen, M., Schobesberger, S., Baalbaki, R., Wang, D. S., Koenig, T. K., Jokinen, T., Sarnela, N., Beck, L.,
612 Almeida, J., Amanatidis, S., Amorim, A., Ataei, F., Baccarini, A., Bertozzi, B., Bianchi, F., Brilke, S., Caudillo, L., Chen,
613 D., Chiu, R., Chu, B., Dias, A., Ding, A., Dommen, J., Duplissy, J., El-Haddad, I., Gonzalez Carracedo, L., Granzin, M.,
614 Hansel, A., Heinritzi, M., Hofbauer, V., Junninen, H., Kangasluoma, J., Kemppainen, D., Kim, C., Kong, W., Krechmer,
615 J. E., Kvashnin, A., Laitinen, T., Lamkaddam, H., Lee, C. P., Lehtipalo, K., Leiminger, M., Li, Z., Makhmutov, V.,
616 Manninen, H. E., Marie, G., Marten, R., Mathot, S., Mauldin, R. L., Mentler, B., Möhler, O., Müller, T., Nie, W., Onnela,
617 A., Petäjä, T., Pfeifer, J., Philippov, M., Ranjithkumar, A., Saiz-Lopez, A., Salma, I., Scholz, W., Schuchmann, S.,
618 Schulze, B., Steiner, G., Stozhkov, Y., Tauber, C., Tomé, A., Thakur, R. C., Väisänen, O., Vazquez-Pufleau, M., Wagner,
619 A. C., Wang, Y., Weber, S. K., Winkler, P. M., Wu, Y., Xiao, M., Yan, C., Ye, Q., Ylisirniö, A., Zauner-Wieczorek, M.,
620 Zha, Q., Zhou, P., Flagan, R. C., Curtius, J., Baltensperger, U., Kulmala, M., Kerminen, V.-M., Kurten, T., et al.: Role of
621 iodine oxoacids in atmospheric aerosol nucleation, Science, 371(6529), 589–595, doi:10.1126/science.abe0298, 2021.

622 Heinritzi, M., Dada, L., Simon, M., Stolzenburg, D., Wagner, A. C., Fischer, L., Ahonen, L. R., Amanatidis, S., Baalbaki,
623 R., Baccarini, A., Bauer, P. S., Baumgartner, B., Bianchi, F., Brilke, S., Chen, D., Chiu, R., Dias, A., Dommen, J.,
624 Duplissy, J., Finkenzeller, H., Frege, C., Fuchs, C., Garmash, O., Gordon, H., Granzin, M., El Haddad, I., He, X., Helm,
625 J., Hofbauer, V., Hoyle, C. R., Kangasluoma, J., Keber, T., Kim, C., Kürten, A., Lamkaddam, H., Laurila, T. M.,
626 Lampilahti, J., Lee, C. P., Lehtipalo, K., Leiminger, M., Mai, H., Makhmutov, V., Manninen, H. E., Marten, R., Mathot,
627 S., Mauldin, R. L., Mentler, B., Molteni, U., Müller, T., Nie, W., Nieminen, T., Onnela, A., Partoll, E., Passananti, M.,
628 Petäjä, T., Pfeifer, J., Pospisilova, V., Quéléver, L. L. J., Rissanen, M. P., Rose, C., Schobesberger, S., Scholz, W.,
629 Scholze, K., Sipilä, M., Steiner, G., Stozhkov, Y., Tauber, C., Tham, Y. J., Vazquez-Pufleau, M., Virtanen, A., Vogel, A.
630 L., Volkamer, R., Wagner, R., Wang, M., Weitz, L., Wimmer, D., Xiao, M., Yan, C., Ye, P., Zha, Q., Zhou, X., Amorim,
631 A., Baltensperger, U., Hansel, A., Kulmala, M., Tomé, A., Winkler, P. M., Worsnop, D. R., Donahue, N. M., Kirkby, J.
632 and Curtius, J.: Molecular understanding of the suppression of new-particle formation by isoprene, Atmos. Chem. Phys.,
633 20(20), 11809–11821, doi:10.5194/acp-20-11809-2020, 2020.

634 Henze, D. K., Seinfeld, J. H., Liao, W., Sandu, A. and Carmichael, G. R.: Inverse modeling of aerosol dynamics:
635 Condensational growth, J. Geophys. Res.-Atmos., 109(D14), doi:https://doi.org/10.1029/2004JD004593, 2004.

Kaipio, J. and Somersalo, E.: Statistical and Computational Inverse Problems, Springer-Verlag, New York., 2005.

Kallinger, P. and Szymanski, W. W.: Experimental determination of the steady-state charging probabilities and particle size conservation in non-radioactive and radioactive bipolar aerosol chargers in the size range of 5-40 nm, J. Nanopart. Res., 17(4), 171, doi:10.1007/s11051-015-2981-x, 2015.

Kandlikar, M. and Ramachandran, G.: Inverse methods for analysing aerosol spectrometer measurements: a critical review, J. Aerosol Sci., 30(4), 413–437, doi:https://doi.org/10.1016/S0021-8502(98)00066-4, 1999.

Kangasluoma, J., Cai, R., Jiang, J., Deng, C., Stolzenburg, D., Ahonen, L. R., Chan, T., Fu, Y., Kim, C., Laurila, T. M., Zhou, Y., Dada, L., Sulo, J., Flagan, R. C., Kulmala, M., Petäjä, T. and Lehtipalo, K.: Overview of measurements and current instrumentation for 1–10 nm aerosol particle number size distributions, J. Aerosol Sci., 105584, doi:https://doi.org/10.1016/j.jaerosci.2020.105584, 2020.

Kirkby, J., Curtius, J., Almeida, J., Dunne, E., Duplissy, J., Ehrhart, S., Franchin, A., Gagné, S., Ickes, L., Kürten, A., Kupc, A., Metzger, A., Riccobono, F., Rondo, L., Schobesberger, S., Tsagkogeorgas, G., Wimmer, D., Amorim, A., Bianchi, F., Breitenlechner, M., David, A., Dommen, J., Downard, A., Ehn, M., Flagan, R. C., Haider, S., Hansel, A., Hauser, D., Jud, W., Junninen, H., Kreissl, F., Kvashin, A., Laaksonen, A., Lehtipalo, K., Lima, J., Lovejoy, E. R., Makhmutov, V., Mathot, S., Mikkilä, J., Minginette, P., Mogo, S., Nieminen, T., Onnela, A., Pereira, P., Petäjä, T., Schnitzhofer, R., Seinfeld, J. H., Sipilä, M., Stozhkov, Y., Stratmann, F., Tomé, A., Vanhanen, J., Viisanen, Y., Vrtala, A., Wagner, P. E., Walther, H., Weingartner, E., Wex, H., Winkler, P. M., Carslaw, K. S., Worsnop, D. R., Baltensperger, U. and Kulmala, M.: Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation, Nature, 476, 429–433, doi:10.1038/nature10343, 2011.

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. C., 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. and Curtius, J.: Ion-induced nucleation of pure biogenic particles, Nature, 533, 521–526, doi:10.1038/nature17953, 2016.

Kuang, C., Chen, M., Zhao, J., Smith, J., McMurry, P. H. and Wang, J.: Size and time-resolved growth rate measurements of 1 to 5 nm freshly formed atmospheric nuclei, Atmos. Chem. Phys., 12(7), 3573–3589, doi:10.5194/acp-12-3573-2012, 2012.

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. and Kerminen, V.-M.: Measurement of the nucleation of atmospheric aerosol particles, Nat. Protoc., 7(9), 1651–1667, doi:10.1038/nprot.2012.091, 2012.

Kupiainen-Määttä, O.: A Monte Carlo approach for determining cluster evaporation rates from concentration measurements, Atmos. Chem. Phys., 16(22), 14585–14598, doi:10.5194/acp-16-14585-2016, 2016.

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(8), 4063–4075, doi:10.5194/acp-15-4063-2015, 2015.

Kürten, A., Li, C., Bianchi, F., Curtius, J., Dias, A., Donahue, N. M., Duplissy, J., Flagan, R. C., Hakala, J., Jokinen, T.,

676   Kirkby, J., Kulmala, M., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Onnela, A., Rissanen, M. P., Simon, M., Sipilä,

677   M., Stozhkov, Y., Tröstl, J., Ye, P. and McMurry, P. H.: New particle formation in the sulfuric acid--dimethylamine--

678   water system: reevaluation of CLOUD chamber measurements and comparison to an aerosol nucleation and growth

679   model, Atmos. Chem. Phys., 18(2), 845–863, doi:10.5194/acp-18-845-2018, 2018.

680   Lehtinen, K. E. J. and Kulmala, M.: A model for particle formation and growth in the atmosphere with molecular

681   resolution in size, Atmos. Chem. Phys., 3(1), 251–257, doi:10.5194/acp-3-251-2003, 2003.

682   Lehtinen, K. E. J., Rannik, Ü., Petäjä, T., Kulmala, M. and Hari, P.: Nucleation rate and vapor concentration estimations

683   using a least squares aerosol dynamics method, J. Geophys. Res.-Atmos., 109(D21), doi:10.1029/2004JD004893, 2004.

684   Lehtipalo, K., Leppä, J., Kontkanen, J., Kangasluoma, J., Franchin, A., Wimmer, D., Schobesberger, S., Junninen, H.,

685   Petäjä, T., Sipilä, M., Mikkilä, J., Vanhanen, J., Worsnop, D. R. and Kulmala, M.: Methods for determining particle size

686   distribution and growth rates between 1 and 3 nm using the Particle Size Magnifier, Boreal Environ. Res., 19(suppl. B),

687   215–236, 2014.

688   McGuffin, D. L., Huang, Y., Flagan, R. C., Petäjä, T., Ydstie, B. E. and Adams, P. J.: Novel estimation of aerosol

689   processes with particle size distribution measurements: a case study with TOMAS algorithm, Geosci. Model Dev.

690   Discuss., 2020, 1–27, doi:10.5194/gmd-2020-281, 2020.

691   Nieminen, T., Lehtinen, K. E. J. and Kulmala, M.: Sub-10 nm particle growth by vapor condensation – effects of vapor

692   molecule size and particle thermal speed, Atmos. Chem. Phys., 10(20), 9773–9779, doi:10.5194/acp-10-9773-2010, 2010.

693   Nieminen, T., Kerminen, V.-M., Petäjä, T., Aalto, P. P., Arshinov, M., Asmi, E., Baltensperger, U., Beddows, D. C. S.,

694   Beukes, J. P., Collins, D., Ding, A., Harrison, R. M., Henzing, B., Hooda, R., Hu, M., Hõrrak, U., Kivekäs, N., Komsaare,

695   K., Krejci, R., Kristensson, A., Laakso, L., Laaksonen, A., Leaitch, W. R., Lihavainen, H., Mihalopoulos, N., Németh,

696   Z., Nie, W., O'Dowd, C., Salma, I., Sellegri, K., Svenningsson, B., Swietlicki, E., Tunved, P., Ulevicius, V., Vakkari, V.,

697   Vana, M., Wiedensohler, A., Wu, Z., Virtanen, A. and Kulmala, M.: Global analysis of continental boundary layer new

698   particle formation based on long-term measurements, Atmos. Chem. Phys., 18(19), 14737–14756, doi:10.5194/acp-18-

699   14737-2018, 2018.

700   Ozon, M., Seppänen, A., Kaipio, J. P. and Lehtinen, K. E. J.: Retrieval of process rate parameters in the general dynamic

701   equation for aerosols using Bayesian state estimation, Geosci. Model Dev. Discuss., 2020, 1–32, doi:10.5194/gmd-2020-

702   236, 2020.

703   Paasonen, P., Peltola, M., Kontkanen, J., Junninen, H., Kerminen, V.-M. and Kulmala, M.: Comprehensive analysis of

704   particle growth rates from nucleation mode to cloud condensation nuclei in boreal forest, Atmos. Chem. Phys., 18(16),

705   12085–12103, doi:10.5194/acp-18-12085-2018, 2018.

706   Pichelstorfer, L., Stolzenburg, D., Ortega, J., Karl, T., Kokkola, H., Laakso, A., Lehtinen, K. E. J., Smith, J. N., McMurry,

707   P. H. and Winkler, P. M.: Resolving nanoparticle growth mechanisms from size- and time-dependent growth rate analysis,

708   Atmos. Chem. Phys., 18(2), 1307–1323, doi:10.5194/acp-18-1307-2018, 2018.

709   Pope, C. A. and Dockery, D. W.: Health Effects of Fine Particulate Air Pollution: Lines that Connect, J. Air Waste

710   Manag., 56(6), 709–742, doi:10.1080/10473289.2006.10464485, 2006.

711   Sandu, A., Liao, W., Carmichael, G. R., Henze, D. K. and Seinfeld, J. H.: Inverse Modeling of Aerosol Dynamics Using

712   Adjoints: Theoretical and Numerical Considerations, Aerosol Sci. Tech., 39(8), 677–694,

713   doi:10.1080/02786820500182289, 2005.

714   Särkkä, S.: Bayesian Filtering and Smoothing, Cambridge University Press., 2013.

715   Seinfeld, J. and Pandis, S.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 3rd edition, 3rd

editio., Wiley., 2016.

Shcherbacheva, A., Balehowsky, T., Kubečka, J., Olenius, T., Helin, T., Haario, H., Laine, M., Kurtén, T. and Vehkamäki, H.: Identification of molecular cluster evaporation rates, cluster formation enthalpies and entropies by Monte Carlo method, Atmos. Chem. Phys. Discuss., 2020, 1–58, doi:10.5194/acp-2019-1036, 2020.

Shen, J., Shu, C. and Zhang, M.: High Resolution Schemes for a Hierarchical Size-Structured Model, SIAM J. Numer. Anal., 45(1), 352–370, doi:10.1137/050638126, 2007.

Simon, M., Dada, L., Heinritzi, M., Scholz, W., Stolzenburg, D., Fischer, L., Wagner, A. C., Kürten, A., Rörup, B., He, X.-C., Almeida, J., Baalbaki, R., Baccarini, A., Bauer, P. S., Beck, L., Bergen, A., Bianchi, F., Bräkling, S., Brilke, S., Caudillo, L., Chen, D., Chu, B., Dias, A., Draper, D. C., Duplissy, J., El-Haddad, I., Finkenzeller, H., Frege, C., Gonzalez-Carracedo, L., Gordon, H., Granzin, M., Hakala, J., Hofbauer, V., Hoyle, C. R., Kim, C., Kong, W., Lamkaddam, H., Lee, C. P., Lehtipalo, K., Leiminger, M., Mai, H., Manninen, H. E., Marie, G., Marten, R., Mentler, B., Molteni, U., Nichman, L., Nie, W., Ojdanic, A., Onnela, A., Partoll, E., Petäjä, T., Pfeifer, J., Philippov, M., Quéléver, L. L. J., Ranjithkumar, A., Rissanen, M. P., Schallhart, S., Schobesberger, S., Schuchmann, S., Shen, J., Sipilä, M., Steiner, G., Stozhkov, Y., Tauber, C., Tham, Y. J., Tomé, A. R., Vazquez-Pufleau, M., Vogel, A. L., Wagner, R., Wang, M., Wang, D. S., Wang, Y., Weber, S. K., Wu, Y., Xiao, M., Yan, C., Ye, P., Ye, Q., Zauner-Wieczorek, M., Zhou, X., Baltensperger, U., Dommen, J., Flagan, R. C., Hansel, A., Kulmala, M., Volkamer, R., Winkler, P. M., Worsnop, D. R., Donahue, N. M., Kirkby, J. and Curtius, J.: Molecular understanding of new-particle formation from alpha-pinene between -50 and +25 °C, Atmos. Chem. Phys., 20(15), 9183–9207, doi:10.5194/acp-20-9183-2020, 2020.

Sipilä, M., Sarnela, N., Jokinen, T., Henschel, H., Junninen, H., Kontkanen, J., Richters, S., Kangasluoma, J., Franchin, A., Peräkylä, O., Rissanen, M. P., Ehn, M., Vehkamäki, H., Kurten, T., Berndt, T., Petäjä, T., Worsnop, D., Ceburnis, D., Kerminen, V.-M., Kulmala, M. and O'Dowd, C.: Molecular-scale evidence of aerosol particle formation via sequential addition of HIO3, Nature, 537, 532, doi:10.1038/nature19314, 2016.

Smolarkiewicz, P. K.: A fully multidimensional positive definite advection transport algorithm with small implicit diffusion, J. Comput. Phys., 54(2), 325–362, doi:https://doi.org/10.1016/0021-9991(84)90121-9, 1984.

Stolzenburg, D., Steiner, G. and Winkler, P. M.: A DMA-train for precision measurement of sub-10 nm aerosol dynamics, Atmos. Meas. Tech., 10(4), 1639–1651, doi:10.5194/amt-10-1639-2017, 2017.

Stolzenburg, D., Fischer, L., Vogel, A. L., Heinritzi, M., Schervish, M., Simon, M., Wagner, A. C., Dada, L., Ahonen, L. R., Amorim, A., Baccarini, A., Bauer, P. S., Baumgartner, B., Bergen, A., Bianchi, F., Breitenlechner, M., Brilke, S., Buenrostro Mazon, S., Chen, D., Dias, A., Draper, D. C., Duplissy, J., El Haddad, I., Finkenzeller, H., Frege, C., Fuchs, C., Garmash, O., Gordon, H., He, X., Helm, J., Hofbauer, V., Hoyle, C. R., Kim, C., Kirkby, J., Kontkanen, J., Kürten, A., Lampilahti, J., Lawler, M., Lehtipalo, K., Leiminger, M., Mai, H., Mathot, S., Mentler, B., Molteni, U., Nie, W., Nieminen, T., Nowak, J. B., Ojdanic, A., Onnela, A., Passananti, M., Petäjä, T., Quéléver, L. L. J., Rissanen, M. P., Sarnela, N., Schallhart, S., Tauber, C., Tomé, A., Wagner, R., Wang, M., Weitz, L., Wimmer, D., Xiao, M., Yan, C., Ye, P., Zha, Q., Baltensperger, U., Curtius, J., Dommen, J., Flagan, R. C., Kulmala, M., Smith, J. N., Worsnop, D. R., Hansel, A., Donahue, N. M. and Winkler, P. M.: Rapid growth of organic aerosol nanoparticles over a wide tropospheric temperature range, P. Nat. Acad. Sci. USA, 115(37), 9122–9127, doi:10.1073/pnas.1807604115, 2018.

Stolzenburg, D., Simon, M., Ranjithkumar, A., Kürten, A., Lehtipalo, K., Gordon, H., Ehrhart, S., Finkenzeller, H., Pichelstorfer, L., Nieminen, T., He, X.-C., Brilke, S., Xiao, M., Amorim, A., Baalbaki, R., Baccarini, A., Beck, L., Bräkling, S., Caudillo Murillo, L., Chen, D., Chu, B., Dada, L., Dias, A., Dommen, J., Duplissy, J., El Haddad, I., Fischer, L., Gonzalez Carracedo, L., Heinritzi, M., Kim, C., Koenig, T. K., Kong, W., Lamkaddam, H., Lee, C. P., Leiminger, M.,

Li, Z., Makhmutov, V., Manninen, H. E., Marie, G., Marten, R., Müller, T., Nie, W., Partoll, E., Petäjä, T., Pfeifer, J., Philippov, M., Rissanen, M. P., Rörup, B., Schobesberger, S., Schuchmann, S., Shen, J., Sipilä, M., Steiner, G., Stozhkov, Y., Tauber, C., Tham, Y. J., Tomé, A., Vazquez-Pufleau, M., Wagner, A. C., Wang, M., Wang, Y., Weber, S. K., Wimmer, D., Wlasits, P. J., Wu, Y., Ye, Q., Zauner-Wieczorek, M., Baltensperger, U., Carslaw, K. S., Curtius, J., Donahue, N. M., Flagan, R. C., Hansel, A., Kulmala, M., Lelieveld, J., Volkamer, R., Kirkby, J. and Winkler, P. M.: Enhanced growth rate of atmospheric particles from sulfuric acid, Atmos. Chem. Phys., 20(12), 7359–7372, doi:10.5194/acp-20-7359-2020, 2020.

Stolzenburg, M. R. and McMurry, P. H.: Equations Governing Single and Tandem DMA Configurations and a New Lognormal Approximation to the Transfer Function, Aerosol Sci. Tech., 42, 421–432, doi:10.1080/02786820802157823, 2008.

Tauber, C., Schmoll, D., Gruenwald, J., Brilke, S., Wlasits, P. J., Winkler, P. M. and Wimmer, D.: Characterization of a Non-Thermal Plasma Source for the Use as a Mass Spec Calibration Tool and Non-Radioactive Aerosol Charger, Atmos. Meas. Tech., 13(11), 5993–6006, doi:10.5194/amt-13-5993-2020, 2020.

Tröstl, J., Chuang, W. K., Gordon, H., Heinritzi, M., Yan, C., Molteni, U., Ahlm, L., Frege, C., Bianchi, F., Wagner, R., Simon, M., Lehtipalo, K., Williamson, C., Craven, J. S., Duplissy, J., Adamov, A., Almeida, J., Bernhammer, A.-K., Breitenlechner, M., Brilke, S., Dias, A., Ehrhart, S., Flagan, R. C., Franchin, A., Fuchs, C., Guida, R., Gysel, M., Hansel, A., Hoyle, C. R., Jokinen, T., Junninen, H., Kangasluoma, J., Keskinen, H., Kim, J., Krapf, M., Kürten, A., Laaksonen, A., Lawler, M., Leiminger, M., Mathot, S., Möhler, O., Nieminen, T., Onnela, A., Petäjä, T., Piel, F. M., Miettinen, P., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Sengupta, K., Sipilä, M., Smith, J. N., Steiner, G., Tomè, A., Virtanen, A., Wagner, A. C., Weingartner, E., Wimmer, D., Winkler, P. M., Ye, P., Carslaw, K. S., Curtius, J., Dommen, J., Kirkby, J., Kulmala, M., Riipinen, I., Worsnop, D. R., Donahue, N. M. and Baltensperger, U.: The role of low-volatility organic compounds in initial particle growth in the atmosphere, Nature, 533, 527–531, doi:10.1038/nature18271, 2016.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ., 8(12), 1251–1256, doi:https://doi.org/10.1016/0004-6981(74)90004-3, 1974.

Verheggen, B. and Mozurkewich, M.: An inverse modeling procedure to determine particle growth and nucleation rates from measured aerosol size distributions, Atmos. Chem. Phys., 6(10), 2927–2942, doi:10.5194/acp-6-2927-2006, 2006.

Viskari, T., Asmi, E., Kolmonen, P., Vuollekoski, H., Petäjä, T. and Järvinen, H.: Estimation of aerosol particle number distributions with Kalman Filtering – Part 1: Theory, general aspects and statistical validity, Atmos. Chem. Phys., 12(24), 11767–11779, doi:10.5194/acp-12-11767-2012, 2012.

Voutilainen, A. and Kaipio, J. P.: Estimation of non-stationary aerosol size distributions using the state-space approach, J. Aerosol Sci., 32(5), 631–648, doi:https://doi.org/10.1016/S0021-8502(00)00110-5, 2001.

Wiedensohler, A.: An approximation of the bipolar charge distribution for particles in the submicron size range, J. Aerosol Sci., 19(3), 387–389, doi:10.1016/0021-8502(88)90278-9, 1988.

Wlasits, P. J., Stolzenburg, D., Tauber, C., Brilke, S., Schmitt, S. H., Winkler, P. M. and Wimmer, D.: Counting on chemistry: laboratory evaluation of seed-material-dependent detection efficiencies of ultrafine condensation particle counters, Atmos. Meas. Tech., 13(7), 3787–3798, doi:10.5194/amt-13-3787-2020, 2020.

Wolfenbarger, J. K. and Seinfeld, J. H.: Inversion of aerosol size distribution data, J. Aerosol Sci., 21(2), 227–247, doi:https://doi.org/10.1016/0021-8502(90)90007-K, 1990.