Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-236-RC2, 2021

Geoscientific Model Development Discussions

Interactive comment on "Retrieval of process rate parameters in the general dynamic equation for aerosols using Bayesian state estimation" *by* Matthew Ozon et al.

Anonymous Referee #2

Received and published: 16 February 2021

The study proposes an inverse modeling approach, based on Bayesian state estimation, to assess the parameters controlling the evolution of an atmospheric aerosol particle size distribution during particle formation and growth, as well as the parameter uncertainties. The model is applicable to measured particle number concentration data, plenty of which is available from field and laboratory measurements of particle formation events.

The application of different state estimation techniques, namely the Extended Kalman filter (EKF) and the Fixed Interval Kalman Smoother (FIKS) is demonstrated using model-generated synthetic particle distribution data. The results seem very promising:

the "true" parameters used for the data generation are generally well recovered, and especially the FIKS smoother reduces the uncertainty ranges of the key parameters. This is an important step forward in aerosol data analysis, since traditionally the parameters are assessed applying simple assumptions with no sophisticated uncertainty estimates.

In addition to contributing to improving the understanding of aerosol microphysics and processes controlling atmospheric aerosol distributions, the information that the proposed methods give can also improve the description of aerosols in larger-scale transport and climate models. The work is thus in the scope of Geoscientific Model Development. I have a few comments that I would ask the authors to address before recommending the paper for publication:

1. If I understand correctly, no constraints for the fitted parameters are used in the preset work. Could a measurement of the gas-phase condensable vapors provide a useful upper-limit constraint for the condensation rate parameter g (corresponding to irreversible condensation)? Would this improve the results? Similarly, the upper limit of the nucleation rate could possibly be assessed if the identities of the nucleating vapors are reasonably well known (at least in well-controlled laboratory experiments).

2. Figure 1: A minor observation on the NE event: the pre-existing distribution of larger aerosols seems to persist throughout the event, although normally atmospheric observations show that it is diluted due to boundary layer growth around noon (thus also enhancing the particle formation efficiency). Can such a dilution effect be added to the Bayesian model equations?

3. P7, L173: It is reasoned that all the unknown parameters of the GDE are non-negative. However, isn't it possible that particles shrink at low vapor concentrations

(e.g. Salma et al., Atmos. Chem. Phys., 16, 7837-7851, 2016), thus making the condensational growth rate g negative? (Similarly, also the formation rate J may be negative in the case of shrinking particles.)

4. Is it controlled that the applied Euler's method for time integration does not contain (cumulative) errors? While the integration is stabilized by the CFL criterion, does it ensure that a shorter time step will not have a notable effect on the result? Short steps may become relevant at high particle concentrations in polluted atmospheric environments, such as very polluted urban conditions.

5. While previously introduced GDE-based methods to assess particle growth rates or other parameters (i.e. those cited on P3, L60-67, and also the more recent methods proposed by Pichelstorfer et al., Atmos. Chem. Phys., 18, 1307-1323, 2018) do not include parameter uncertainty estimates, can it still be useful to test also these methods against the Bayesian state estimation?

6. It is explained that the evolution of the aerosol size distribution is determined by the nucleation, condensation, coagulation, and sink parameters. I'm wondering if the aerosol particles can exist in different charging states, and if this can have effects on the distribution through e.g. modified growth and coagulation kernels?

7. P4, L93-94: It is stated that the formation of particles occurs typically at 1.5-2 nm. Yet, the nucleation size in the simulations is assumed to be only 0.87 nm, which is barely a molecular size and would generally be expected to correspond to "pre-nucleation" molecular clusters. How was this size chosen? Are the results affected if the size is erroneously assigned?

8. Furthermore, for the steady state (SS) simulation case, the modeled size range is divided to 1731 logarithmically distributed size bins in the range [0.87 nm, 10.00 nm] (P16, L383). For such a high number of bins in a narrow range, the bins are very dense especially at the smallest particle sizes.

Is it ensured that the width is still physically consistent, that is, corresponds to at least one-molecule increment in terms of particle growth? I believe that for large atmospheric molecules, such as oxidized organic species, 10 nm-particles may consist of significantly less than 1731 molecules, in which case the bin widths would be unphysical.

9. P4, L96-97: It is stated that the particle flux to the smallest measurable size class is driven by condensational growth, as defined by Eq. (2). Can coagulation also play a role in polluted environments with high concentrations of nucleated particles? Would this affect the model formulation?

10. P16, L361 onward: Why is a wall loss rate estimate discussed for the "nucleation event" (NE) case? It is explained that the NE case represents a typical particle formation event in the atmosphere. Which type of walls in the ambient atmosphere does the loss rate correspond to, or does it refer to the walls inside a measurement device?

Technical comments:

1. The colored solid lines, corresponding to the "true" and estimated parameters as functions of time or particle size, in most figures are somewhat difficult to see; can they be made e.g. thicker and brighter (and maybe have different line styles)?

Also, the shaded areas corresponding to the uncertainties in the estimates could preferably be lighter and/or more transparent in order to see the lines better. It may also not be obvious by first look that the brown shade is the overlap between the blue and yellow shades; the figure legends list blue and yellow shades, but in most figures the yellow is almost entirely missing and there is only brown / blue.

2. P2, L42: Please remove the comma in “Methods 1 and 2 are applicable to cases, in which. . .”

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-236, 2020.