# J-GAIN v1.0: A flexible tool to incorporate aerosol formation rates obtained by molecular models into large-scale models

Daniel Yazgi[1] and Tinja Olenius[1]

[1]Swedish Meteorological and Hydrological Institute (SMHI), SE-601 76 Norrköping, Sweden

**Correspondence:** Daniel Yazgi (daniel.yazgi@smhi.se), Tinja Olenius (tinja.olenius@alumni.helsinki.fi)

**Abstract.** New-particle formation from condensable gases is a common atmospheric process that has significant but uncertain effects on aerosol particle number concentrations and aerosol–cloud–climate interactions. Assessing the formation rates of nanometer-sized particles from different vapors is an active field of research within atmospheric sciences, with new data being constantly produced by molecular modeling and experimental studies. Such data can be used in large-scale climate and air quality models through parameterizations or look-up tables. Molecular cluster dynamics modeling, ideally benchmarked against measurements when available for the given precursor vapors, provides a straight-forward means to calculate high-resolution formation rate data over wide ranges of atmospheric conditions. Ideally, the incorporation of such data should be easy, efficient and flexible in the sense that same tools can be conveniently applied for different data sets in which the formation rate depends on different parameters. In this work, we present a tool to generate and interpolate look-up tables of formation rates for user-defined input parameters. The table generator primarily applies cluster dynamics modeling to calculate formation rates from an input quantum chemistry data set defined by the user, but the interpolator may be used also for tables generated by other models or data sources. The interpolation routine uses a multivariate interpolation algorithm, which is applicable to different numbers of independent parameters, and gives fast and accurate results with typical interpolation errors of up to a few percent. These routines facilitate the implementation and testing of different aerosol formation rate predictions in large-scale models, allowing straight-forward inclusion of new or updated data without the need to apply separate parameterizations or routines for different data sets that involve different chemical compounds or other parameters.

## 1 Introduction

Formation of secondary aerosol particles from condensable gases is a well-known and ubiquitous phenomenon in Earth's atmosphere (Kerminen et al., 2018; Kontkanen et al., 2017). Aerosols have significant effects on climate and air quality, with the indirect climate effect through aerosol–cloud interactions forming the largest uncertainty in radiative forcing assessments (Naik et al., 2021). New-particle formation, including the formation and condensational growth of nanoparticles from vapors, is an important factor affecting aerosol number concentrations and size distributions (Fountoukis et al., 2012; Makkonen et al., 2012; Gordon et al., 2017). The primary quantity characterizing the formation process is the initial particle formation rate $J$, often referred to as the nucleation rate, which gives the rate at which new particles of ca. 1–2 nm in diameter form per unit time and volume at given ambient conditions through clustering of vapor molecules. The chemical compounds that are able to drive

and enhance the initial formation process include acids, bases, and organic species (Glasoe et al., 2015; Lehtipalo et al., 2018; Xiao et al., 2021). The roles and effects of different potentially important precursor vapors are not, however, resolved. While it is established that sulfuric acid ($H_2SO_4$) initiates particle formation in the presence of stabilizing species such as ammonia ($NH_3$) or monoamines (Jen et al., 2014; Almeida et al., 2013), there exist various species, such as diamines (Jen et al., 2016), organic acids (Zhang et al., 2004) and complex highly oxidized organic molecules (Kirkby et al., 2016), that may contribute to particle formation with or without sulfuric acid. Furthermore, quantitative particle formation rates are very challenging to assess: both theoretically and experimentally deduced formation rates involve high uncertainties (Almeida et al., 2013; Kürten et al., 2018), and may be sensitive to ambient conditions such as temperature and scavenging sink of the initial molecular clusters (Olenius et al., 2017; Elm et al., 2020).

The combination of quantum chemistry and molecular cluster dynamics simulations is the state-of-the-art method to calculate theoretical particle formation rates (Elm et al., 2020). Quantum chemistry is applied to compute the thermodynamics of the formation of the initial molecular clusters from vapor molecules. The thermochemistry data are converted to cluster evaporation rate constants and used as input in cluster population dynamics simulations, which determine the formation rate by modeling cluster formation and growth considering molecular collisions and evaporations and other dynamic processes such as cluster sinks (Olenius et al., 2013). Such simulations are fast and easy to perform, thus enabling the assessment of formation rates over wide ranges of atmospheric conditions for given chemical compounds. Benchmarking the theoretical methods against experimental data, which often cover a limited set of conditions, provides quality control and uncertainty estimates for the predicted formation rates (Almeida et al., 2013; Kürten et al., 2016).

New quantum chemical data, calculated for different chemical species (Elm et al., 2020; Elm, 2019a) or by applying different physical and computational approaches (Besel et al., 2020), is constantly becoming available. Such data can be used to test the effects of different particle formation mechanisms or improved quantitative formation rate predictions in atmospheric large-scale models. To this end, the formation rate data must be implemented as parameterizations or look-up tables (Dunne et al., 2016; Ehrhart et al., 2018; Kazil et al., 2010; Yu et al., 2020; Baranizadeh et al., 2016), as applying formation rate simulations within a large-scale model is not feasible. The formation rate is a function of several parameters, the most obvious of which are the concentrations of the vapors participating in the formation process, and the temperature. In addition, formation rates obtained by cluster simulations include the effects of cluster sink caused by pre-existing larger aerosol particles and, depending on the chemical compounds and quantum chemical input data, possibly also atmospheric ions and relative humidity. In the presence of several cluster-forming vapors, particle formation can proceed through multicompound clustering, or parallel clustering pathways involving non-interacting chemical mechanisms (Elm, 2019b).

Deriving parameterizations quickly becomes cumbersome as the number of independent parameters increases: finding parameterization formulas and coefficients that reproduce the formation rate data with a reasonable accuracy throughout the parameter space is virtually impossible for arbitrary chemical systems. Furthermore, evaluation of very complex formulas within a computationally heavy large-scale model may not be optimal. The benefit of look-up tables is that values determined from a table of sufficient resolution are guaranteed to be close to the original data, and no pre-processing of the formation rate data is needed. The use of such tables, on the other hand, requires multivariate interpolation algorithms that should ideally be

applicable to tables of arbitrary dimensions, with no need for manual changes depending on the number of independent parameters. While interpolated values must be sufficiently accurate, the interpolation routine should be computationally efficient and preferably simple.

In this work, we introduce a tool to incorporate molecular modeling data in atmospheric models by flexible routines to generate and interpolate formation rate look-up tables. With this approach, we aim to cover the following aspects:

– flexible implementation of state-of-the-art molecular modeling results in large-scale models,

– inclusion of arbitrary chemical compounds,

– efficient multivariate interpolation, and

– user-friendly routines with no need for modifications depending on, for instance, included chemical compounds.

We construct a look-up table generator that embeds a publicly available cluster dynamics model to calculate formation rates from user-defined quantum chemistry input data, and provide a table interpolation routine that can be readily implemented in a large-scale model. The table generator enables easy application of the standard formation rate modeling approach, but the interpolator can be used also for tables saved from other models in the same, simple format. We present assessments of optimal table resolution and interpolation accuracy, and demonstrate the use and performance of the tool by application on sulfuric acid—ammonia cluster chemistry data. We generate tables suitable for global applications for these data, and assess the performance of the interpolation routine in typical computationally heavy model set-ups.

## 2   Methods

The proposed look-up table approach comprises two components: a generator routine to create tables, and an interpolation routine to be implemented in an atmospheric model. We refer to the routines as J-GAIN (Formation rate ($J$) look-up table Generator And INterpolator). The details of the generation and interpolation routines are presented in Sects. 2.1 and 2.2, respectively, and their application is further discussed in Sects. 2.3 and 2.4. The table generator primarily uses input data from quantum chemical calculations and calls a molecular cluster dynamics model to determine the formation rates, as this is the standard method in theoretical atmospheric particle formation studies. The interpolation routine can, however, be applied also to formation rate data obtained by other means by saving the tabulated data in the given table format. Fig. 1 shows a schematic presentation of the J-GAIN routines, summarizing the input, output and usage. J-GAIN is written in Fortran, with the exception that the cluster dynamics model used by the table generator applies also Perl, and licensed under GPL 3.0. The code repository includes detailed instructions for using the programs (Yazgi and Olenius, 2022a, b).

### 2.1   Table generation

By default, the table generator takes as input molecular cluster thermochemistry data for the given chemical compounds, and calculates the formation rates by molecular cluster dynamics simulations through coupling to the open-source cluster model
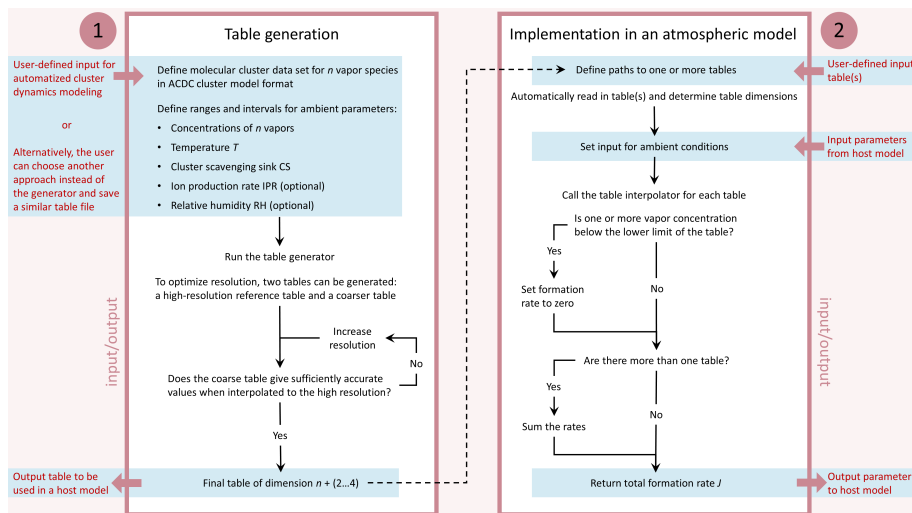
**Figure 1.** Flow chart illustrating the generation and application of particle formation rate tables by J-GAIN. The boxes summarize the steps for using the two parts: (1) table generation with automatized calculation of formation rates by cluster dynamics modeling, and (2) implementation of tables and the table interpolator in a host model. User-defined input and output are specified outside the boxes.

ACDC (Olenius, 2021). In practice, application of the generator on a new cluster data set consists of two steps: creating the formation rate equations by ACDC, and running the generator to obtain the table (see Yazgi and Olenius, 2022a). As the embedded ACDC application provides automatized and flexible treatment of arbitrary cluster data sets, we recommend the default table generator when applying the commonly used combination of quantum chemical input and cluster dynamics modeling to obtain formation rates. It can be noted that ACDC also includes a wide selection of options that increase the flexibility: for example, while cluster evaporation rate constants are by default obtained from the thermochemistry data, they can also be given as direct input if the user wishes to assess them in some other way. The advanced user may also modify other cluster simulation settings (see the ACDC manual; Olenius, 2021). However, the table interpolator is not limited to tables generated by the default table generator, and thus it is possible to use alternative approaches to determine formation rates. For this, the rates must be saved in the same file format as that produced by the default generator, as detailed in Appendix A.

In addition to the molecular cluster data, the table generator requires user-defined input for the ranges of the parameters that define the ambient conditions. These parameters include the concentrations of the vapor compounds, temperature ($T$), coagulational scavenging sink (CS; given as vapor condensation sink which is scaled for different cluster sizes within the cluster dynamics model as derived by Lehtinen et al. (2007)), and optionally also ion production rate (IPR; given as generic ion pairs per unit volume and time) and relative humidity (RH). This gives $n + (2 \ldots 4)$ independent parameters, where $n$ is the number of chemical compounds. While vapor concentrations, temperature and cluster sink always affect the particle formation rate and are thus included by default, the inclusion of IPR and RH depends on the chemical system, and also on data availability. For strongly clustering chemistries, ion effects may be negligible (Myllys et al., 2019), and also RH effects may

be minor (Henschel et al., 2016). As the addition of charged species and hydrates in a quantum chemistry data set requires a significant computational effort, these effects are not always included in available thermochemistry data sets.

The user sets the range for the values of each parameter by giving the lower and upper limits, and the number of values to be placed within the limits at even intervals. Parameters can be defined as "logarithmic", in which case the data points are placed evenly on a logarithmic scale. This is relevant for, for example, vapor concentrations. The formation rate table is then generated by running the generator which calls ACDC to obtain the formation rate $J$ for each combination of parameter values.

The tables are outputted as binary files, and a descriptor file is generated together with the table. The latter contains the essential information on the table, including the names and units of the independent parameters, the lower and upper limits of the parameter values, and the numbers of values along each dimension. In order to ensure sufficient accuracy, several tables can be generated at different resolutions. Then, possible errors in interpolated values can be assessed by interpolating a coarser-resolution table on the grid of a higher-resolution reference table (Sect. 3.1).

## 2.2 Table interpolation

The interpolation routine uses the descriptor file to obtain the number and identities of the independent parameters for the corresponding look-up table. After loading the table, the routine determines the formation rate by linear multivariate interpolation. In general, for an $N$-dimensional function $f = f(x_1, x_2, ..., x_N)$ with nearest known data points at locations $x_{j,0}$ and $x_{j,1}$ for each variable $x_j$, where $x_{j,0} < x_{j,1}$ and $j = 1, \ldots, N$, the interpolated value $\tilde{f}$ can be written as

$$\tilde{f}(x_1, x_2, ..., x_N) = \sum_{(i_1, i_2, ..., i_N) \in \{0,1\}^N} f\left(x_{1,i_1}, x_{2,i_2}, ..., x_{N,i_N}\right) \prod_{j=1}^{N} \begin{cases} x_j', & i_j = 1 \\ 1 - x_j', & i_j = 0 \end{cases}, \tag{1}$$

where the summation goes over the values of $f$ at all combinations of points $x_{j,0}$ and $x_{j,1}$ for $j = 1, \ldots, N$, and

$$x_j' = \frac{x_j - x_{j,0}}{x_{j,1} - x_{j,0}}. \tag{2}$$

Here, we perform linear interpolation by default for $f = \log J$ and either $x_j$ or $\log x_j$ depending on if parameter $j$ is defined as linear or logarithmic, respectively. This gives more accurate results than assuming purely linear relationships, as $J$ generally varies smoothly on a logarithmic scale. Moreover, a linear dependence between the logarithms of $J$ and the concentration $C$ of a contributing vapor within a short interval along the $C$-axis is expected also based on simplified cluster formation theories (see e.g. Li and Signorell, 2021). The choices of applying either the actual values or the logarithms of the dependent and independent parameters can also be changed by the user.

Treatment of input values that are outside the ranges covered by the table is parameter-dependent and defined by the user. For each parameter, the following options are included in the example routines:

1. Value below/above the table limits $\rightarrow$ the value is set to the lower/upper limit of the table (this is the default behavior)
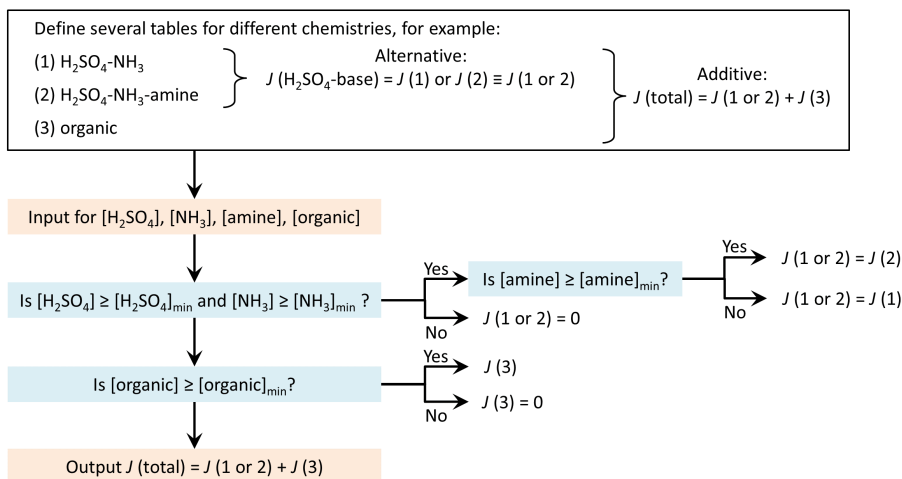
**Figure 2.** Schematic presentation of treatment of several tables: an example of a possible table combination that the user may construct according to their needs.

2. Value below the lower limit $\rightarrow J = 0$ is returned; input value above the upper limit $\rightarrow$ the value is set to the upper limit

The latter option is for vapor concentrations, provided that the lower limits of the table are chosen so that effectively no particle formation is expected at concentrations below the limits. The first option is by default used for other parameters.

## 2.3 Incorporation of tables in a host model

In the code repository, we provide an example of simple interfaces to load and interpolate look-up tables within a host model. The input parameters of the interpolation subroutine include all independent parameters that the particle formation rate may depend on, and the total formation rate is returned as output. Importantly, the interpolator is not limited to using a single table: separate particle formation pathways, corresponding to different chemical compounds, can be incorporated as separate tables. If more than one table are used, the interpolator is applied separately to each table, and the total formation rate can be obtained as the sum of the individual formation rates.

The repository includes a simple example of summing the rates interpolated from two separate tables. However, the user may construct different ways to treat several tables according to their needs and data availability. To give an example, a possible practical application could be as follows: separate tables are used for parallel formation mechanisms, for example, inorganic $H_2SO_4$–base and organics-driven pathways. There may also be alternative tables that are selected based on the presence of a given chemical species, that is, if the concentration of the species is high enough for the species to contribute. For instance, there may be data for particle formation from $H_2SO_4$ and $NH_3$ with or without an amine species. In the presence of the amine, a table of $H_2SO_4$—$NH_3$–amine formation rates is selected for the $H_2SO_4$–base pathway, while otherwise a $H_2SO_4$—$NH_3$ table is applied. This example of a potential table combination is schematically presented in Fig. 2.

## 2.4 Example case: sulfuric acid—ammonia particle formation

We demonstrate the application and performance of the J-GAIN table generator and interpolator using previously published quantum chemical data for sulfuric acid and ammonia (Olenius et al., 2013), which are common atmospheric particle formation precursor vapors. Here, the input molecular cluster data for the $H_2SO_4$-–$NH_3$ chemistry includes charged clusters but no hydration, and therefore the independent parameters that determine the formation rate are vapor concentrations $[H_2SO_4]$ and $[NH_3]$, $T$, CS and IPR.

We apply the $H_2SO_4$—$NH_3$ cluster data to generate tables of different resolution and coverage. First, we demonstrate the effect of table resolution by generating tables that cover subsets of independent parameter ranges where $J$ is sensitive to the parameter values, and compare the interpolated values of $J$ to accurate values given by a high-resolution reference table. Second, we generate extended tables suitable for global applications, where the ranges of all independent parameters cover the various environments from boundary layer to upper troposphere. The extended tables are applied to evaluate the accuracy of $J$ interpolated over the full set of parameters $[H_2SO_4]$, $[NH_3]$, $T$, CS and IPR that follow representative diurnal cycles, corresponding to practical model implementations. To assess the performance of the interpolation routine for given table sizes, we determine the interpolation speed for test tables of different numbers of data points and dimensions.

## 3 Results and discussion

### 3.1 Effect of table resolution on accuracy

Fig. 3 presents $J$ interpolated along the $[H_2SO_4]$-axis and the error in the interpolated values at representative atmospheric conditions for tables with $2^k + 1, k = 4...9$ data points on the axis. That is, for each subsequent table the number of points is doubled, resulting in numbers ranging from 17 to 513. For the lowest resolution of $2^4 + 1 = 17$ points, the relative error is less than $\pm 10\%$, and for $\geq 2^6 + 1 = 65$ points the error is well below $\pm 1\%$. Fig. 4 shows the error as a function of both $[H_2SO_4]$ and $[NH_3]$ at two different temperatures for the two lowest resolutions used here. For $2^4 + 1 = 17$ points along both axes, the error is up to 10 % but mostly below it, and doubling the resolution to $2^5 + 1 = 33$ points drops the error below a couple of percent.

In order to demonstrate the application of the interpolator for interpolation over all independent parameters at realistic ambient conditions, corresponding to implementation of the routine in an atmospheric model, $J$ is determined for a representative diurnal cycle as shown in Fig. 5. The independent parameters are set to follow 24-hour time profiles as described in Appendix B. Here, we apply extended tables that cover wide ranges of parameter values, suitable for larger-scale chemical transport or general circulation models: $[H_2SO_4] = 10^5 - 10^8$ cm$^{-3}$, $[NH_3] = 10^6 - 3 \times 10^{11}$ cm$^{-3}$, $T = 180 - 320$ K, CS $= 10^{-5} - 10^{-1}$ s$^{-1}$, and IPR $= 0.1 - 60$ cm$^{-3}$ s$^{-1}$. The errors in Fig. 5 are generally of the same order as in Figs. 3 and 4, although somewhat higher due to interpolation over all parameters. Doubling the table resolution from $2^4 + 1 = 17$ to $2^5 + 1 = 33$ points (1) only for vapor concentrations, and (2) for all parameters decreases the maximum errors down to ca. $\pm 15\%$ and well below $\pm 10\%$,
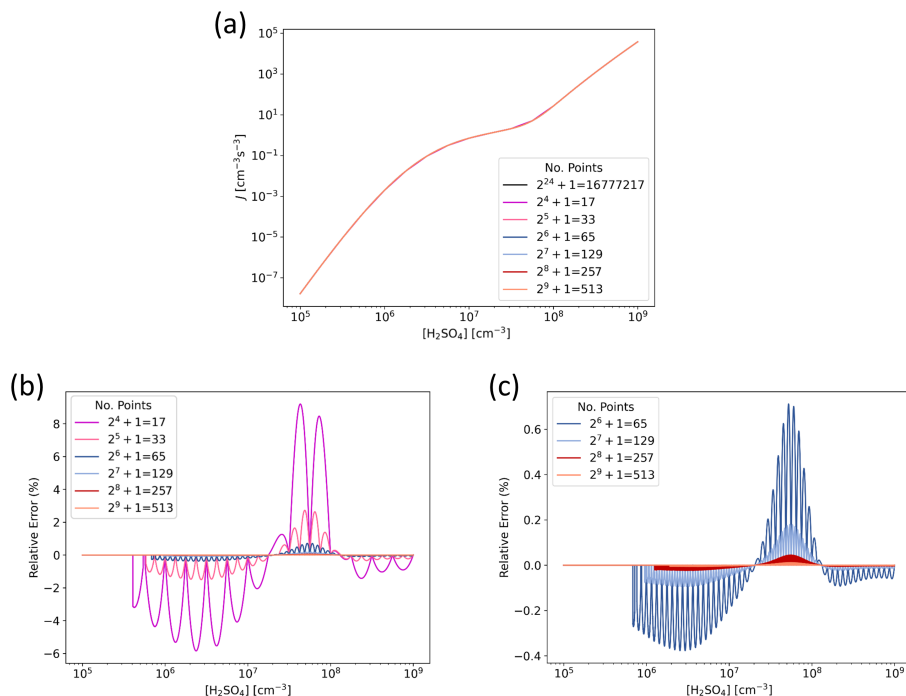
**Figure 3.** Particle formation rate $J$ as a function of $H_2SO_4$ concentration for $H_2SO_4$—$NH_3$ model chemistry ($[NH_3] = 10^{10}$ cm$^{-3}$, $T = 298.15$ K, CS $= 10^{-3}$ s$^{-1}$, IPR $= 3$ cm$^{-3}$s$^{-1}$), determined from look-up tables of different resolution. (a) Absolute $J$. Note that the different lines mostly fall on top of each other. (b) Relative error in interpolated $J$ compared to accurate values given by the highest-resolution reference table with $2^{24} + 1 \sim 2 \cdot 10^7$ points along the [$H_2SO_4$]-axis. (c) Relative error in interpolated $J$ for resolutions of $\geq 2^6 + 1 = 65$ points.

respectively. The errors are similar when using different absolute values of the independent parameters, tested by modifying the diurnal profiles by scaling the parameter values up or down (by e.g. an order of magnitude for vapor concentrations or CS).

In general, such errors are very small compared to typical uncertainty estimates for formation rate data: uncertainties in both theoretical and experimental $J$ generally span up to an order of magnitude and beyond (see e.g. Kürten et al., 2016). Based on the present evaluations, the interpolation approach is not expected to add notable uncertainties in the formation rate representation even at the lowest resolutions applied. Moreover, the accuracy is significantly increased by doubling the resolution. However, it can be noted that high-resolution tables quickly become very large, and the interpolation speed may be affected by the number of data points along the dimensions. Therefore, the choice of resolution depends on (1) the desired accuracy, (2) how large tables the user is willing to generate and store, and (3) the computational aspects of the model application considering the number of calls to the interpolator (Sect. 3.2).
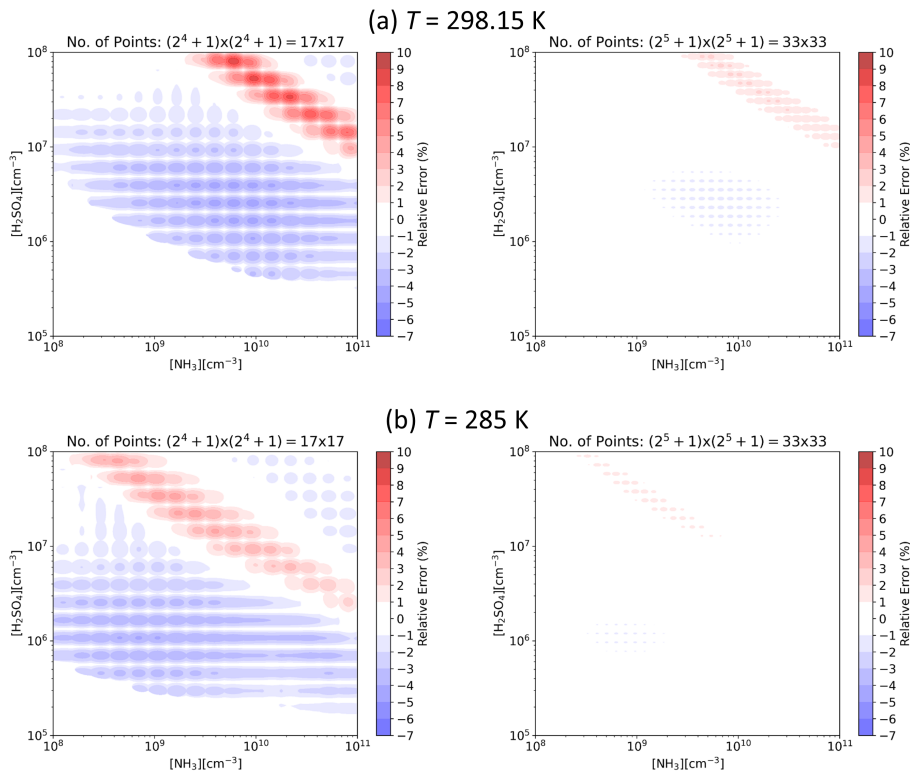
**Figure 4.** Relative error in interpolated particle formation rate $J$ as a function of $H_2SO_4$ and $NH_3$ concentrations for look-up tables with $2^4 + 1 = 17$ and $2^5 + 1 = 33$ data points along the [$H_2SO_4$]- and [$NH_3$]-axes at $CS = 10^{-3}$ s$^{-1}$, IPR $= 3$ cm$^{-3}$s$^{-1}$, and (a) $T = 298.15$ K, (b) $T = 285$ K.

## 3.2 Effect of table size and number of independent parameters on performance

In order to assess the interpolation speed with respect to the numbers of data points and independent parameters, we apply arbitrary test tables of different sizes and dimensions. Fig. 6 shows the mean run time for one million calls to the interpolator with randomly assigned input parameter values. Here, the $x$-axis is the base-2 logarithm of the table size (total number of values in the table). Therefore, for a table with equal numbers of points $2^k + 1$ along all dimensions, the $x$-axis corresponds to approximately $N \times k$, where $N$ is the number of dimensions.

The performance is primarily affected by the number of dimensions: the addition of a new dimension may increase the run time by up to a factor of $\sim 2$. In addition, the run time exhibits a major increase when the table size becomes very large, here beyond ca. $2^{28}$ (of the order of $\gtrsim 10^8$) data points. This is due to memory limits and management, and thus the exact threshold size depends on the computing system. For example, for the $H_2SO_4$–$NH_3$ table with $N = 5$ and the current simple test set-up with 64 GiB memory on the node, the threshold of $\sim 2^{28}$ points corresponds to $k > 5$. In the case that managing
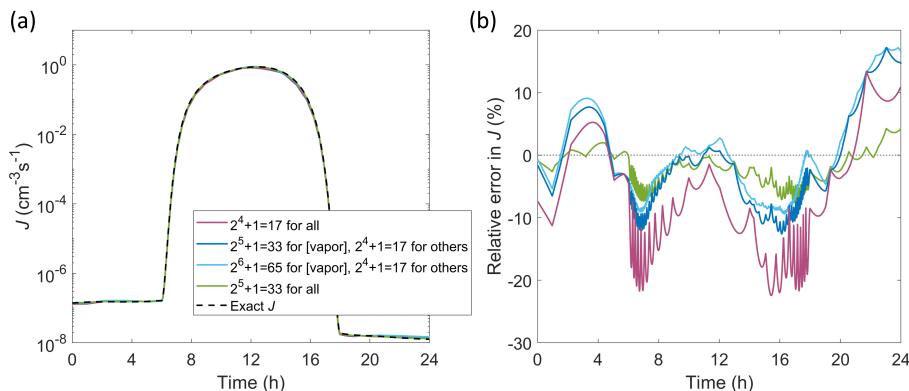
**Figure 5.** Particle formation rate $J$ for a diurnal cycle in which all independent parameters $[H_2SO_4]$, $[NH_3]$, $T$, CS and IPR vary with time (Appendix B). (a) Absolute $J$. (b) Relative error in interpolated $J$ compared to accurate values. Interpolated values are determined from the extended tables with increasing resolution either along only vapor concentration axes, or along all axes as indicated in the legend.

very large tables becomes slow on a given system, the performance can be optimized by splitting the table into subtables that cover different parts of the ambient conditions parameter space.

The approximative run times of Fig. 6 can be compared to typical overall times for e.g. global-scale applications. We have implemented the $H_2SO_4–NH_3$ table for first tests in the EC-Earth climate model (Döscher et al., 2022). In this EC-Earth set-up, the total number of grid points where the interpolation routine is called is 367200 ($120 \times 90$ horizontal grid points and 34 vertical layers), the time step for determining the formation rate is one hour, and the chemistry model uses 90 processes. This corresponds to approximately $\lesssim 1$ minute run time for the formation rate routine for a simulation of one year, assuming a mean time of $\lesssim 2$ seconds for $10^6$ look-ups (Fig. 6, resolutions of $2^4 + 1 = 17$ and $2^5 + 1 = 33$ for all axes). Such contribution can be considered acceptable compared to the overall model run time of $\gtrsim 10$ hours per year, or to the contribution of the atmospheric chemistry component which accounts for up to even $\sim 90\,\%$ of the total run time (van Noije et al., 2014).

For applications with e.g. higher spatial resolution or shorter time step, the performance of the formation rate routine can be optimized by reasonable choices of table dimensions and size. For example, for very strongly clustering chemical systems, the presence of atmospheric ions has only minor effects. The IPR parameter could thus be discarded in the interest of speed. In addition to the resolution of the independent parameter values (i.e. the absolute intervals), also their ranges (minimum and maximum values) can be chosen considering the modeled environment, so that redundant values are avoided. If the numbers and ranges of parameters cannot be optimized further, very large tables can be divided into separate subtables that cover different sets of ambient conditions, and selected within the host model application based on the input conditions.

It can also be noted that considering the high accuracy of the interpolated values, interpolation from pre-generated tables is superior compared to the time required to calculate the formation rates by the molecular cluster model. For example, generating the table of $(2^4 + 1)^5 = 1419857$ data points takes 27 hours on a single process, corresponding to a mean time of ca. 0.068 seconds per value. Such times would be infeasible for operational models.
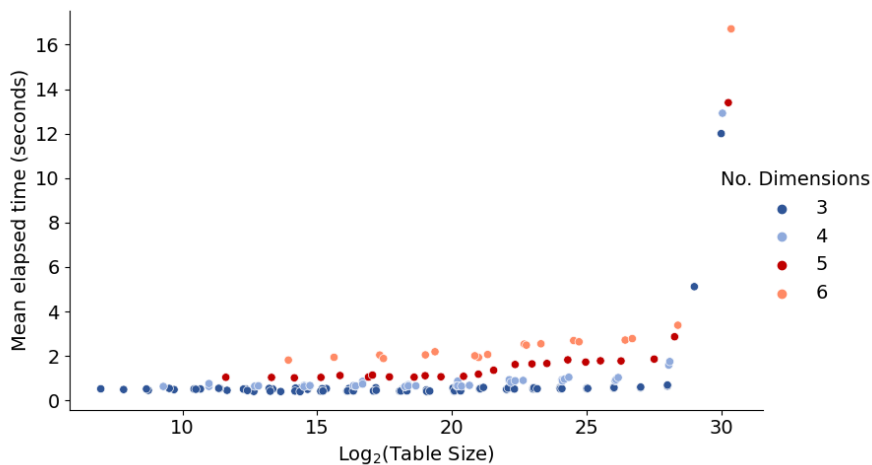
**Figure 6.** Time required to perform one million look-ups for tables of different sizes. The figure shows different combinations $(2^k + 1, 2^l + 1, 2^m + 1, ..)$ of data points along axes, keeping $k = l$ and the total number of points $< (2^7 + 1)^2 (2^4 + 1)^4 \sim 2^{30.37}$. For more details, see discussion in Sect. 3.2.

### 3.3 Potential limitations in applying formation rates in a host model

It must be noted that incorporating aerosol formation rates in an atmospheric model involves given limitations. These limitations are independent of the data source of implementation method, and apply equally to look-up tables, parameterizations or other approaches. An obvious restriction is the availability of tracers in the host model. While various chemical compounds may contribute to atmospheric particle formation, including large numbers of new species in a chemical transport model may be cumbersome. In addition, sources of individual chemical species, such as different types of amines or organic acids, may not be well quantified.

Here, we apply particle formation rates from $H_2SO_4$ and $NH_3$ which are common species included in atmospheric models, and comprise a central particle formation pathway according to current understanding (Dunne et al., 2016; Gordon et al., 2017). Tables generated in this work can thus be easily implemented in standard models that include aerosol formation processes. On the other hand, $H_2SO_4$–amine particle formation, for example, requires the inclusion of amine species which are not common tracers despite their potential importance to aerosol formation (Bergman et al., 2015). Different types of amines may exhibit different particle formation efficiencies, but some amine species are similar in terms of their effects on the formation rate (Jen et al., 2014; Olenius et al., 2017). In order to test the importance of such particle formation agents in a large-scale model, some simplifications are often needed.

A simplified option to include rough source–sink dynamics for species that can be assumed to have common sources and similar properties with respect to gas-phase chemistry and gas-to-particle partitioning is to implement a lumped or representative trace compound. For example, monoamines with similar properties—namely di- and trimethylamines—have been approximated as a single representative alkylamine species, the emissions of which are scaled from ammonia emissions by as-

sumed amine-to-ammonia ratio due to their common sources (Bergman et al., 2015). To take into account possible differences in the particle formation rates due to different amines, an option is to estimate the contributions of individual species based on available measurements (Schade and Crutzen, 1995). Oxidized organic species can be treated in a similar manner through a representative highly oxidized, ultra/extremely low-volatile compound (generally referred to as HOM or ULVOC/ELVOC; Kirkby et al., 2016; Schervish and Donahue, 2021), which is already included in some transport models (Gordon et al., 2017; Julin et al., 2018; Patoulias and Pandis, 2022).

In addition, it can be noted that the standard approach to implement formation rates involves certain approximations related to molecular cluster kinetics. Namely, formation rates are assumed to be determined solely by ambient conditions, applying the steady-state approximation for the cluster population and omitting time-dependent vapor–cluster–aerosol kinetics (Yu, 2003; Olenius and Riipinen, 2017; Olenius and Roldin, 2022). While such simplifications are needed for large-scale applications, they may affect the secondary particle concentrations. To minimize such effects, it can be recommended to extend the modeled aerosol size distribution to as small sizes as achievable instead of extrapolating the formation rate to a substantially larger size (Lee et al., 2013), or to apply a more detailed size distribution representation on the smallest particle sizes (Blichner et al., 2021). Nevertheless, the inherent steady-state assumption in the initial formation rate at ca. 1 nm is likely to cause overprediction in the case of strongly clustering, low-concentration trace species, and thus the rate should be considered as an upper-limit estimate for such chemistries (Olenius and Roldin, 2022).

## 4    Conclusions

Adequate representation of aerosol particle formation from vapors in atmospheric models is needed for assessing the climate and health effects of aerosols. The increasing amount of available computational molecular cluster chemistry data enable calculation of new-particle formation rates for different chemical compounds and computational chemistry methods. As formation rates are typically complicated functions of ambient conditions, a practical approach to apply them in an atmospheric model framework is through look-up tables. Here, we provide a tool to generate and interpolate formation rate tables, applicable to arbitrary sets of chemical species, for conveniently incorporating theoretical particle formation rate data in large-scale models.

Tests conducted using data for $H_2SO_4$–$NH_3$ particle formation show that the interpolation approach is efficient and accurate, with maximum interpolation errors typically ranging from negligible to ca. $\pm(10-20)$ % depending on table resolution. Interpolation errors can also be minimized by choosing to interpolate given parameters on a logarithmic instead of a linear scale. As interpolation speed is affected by the number of independent parameters and also by the table size for very large tables, the choice of parameters and the resolution should be optimized considering the computational burden of the application where the table is implemented. For heavy applications, redundant independent parameters that have only minor effects on the formation rate can be discarded, and the number of data points along each dimension can be set according to the desired accuracy. The flexibility of the provided tool makes the routines easy to apply to different data sets and choices of table dimensions and size. This design facilitates data transfer between the molecular modeling and the large-scale modeling communities.

## Appendix A: Data sources and table format

For J-GAIN, formation rate tables can be created (1) by the provided table generator when applying the molecular cluster dynamics modeling approach, or (2) by saving tabulated formation rates from another data source in a similar table. The details of the formats are listed below, and detailed instructions for the table generator are available in the J-GAIN repository (Yazgi and Olenius, 2022a).

For the J-GAIN table generator, the molecular cluster input data is given in the format of the ACDC model. Briefly, the set of input files include

1. Molecular compositions of the clusters

2. Cluster formation free energies as enthalpies and entropies

3. Cluster dipole moments and polarizabilities if also charged species are included

These files are summarized in the J-GAIN repository where also example files for the present test case can be found (Yazgi and Olenius, 2022a, ACDC subdirectory), and described in detail in the ACDC technical manual (Chapter 3) available in the ACDC repository (Olenius, 2021). Before applying the table generator, the user creates the ACDC formation rate equations for the given input by running the program provided in the J-GAIN repository. This gives the standard model set-up, although it can be noted that the advanced user may also adjust the cluster modeling settings within the program if wanting to, for example, modify rate constants (ACDC manual, Chapter 2.5). Lastly, for running the table generator, the user sets the ranges of the ambient conditions in the namelist file (template provided in the generator subdirectory).

For other data sources, the data need to be processed into the table format applied by the J-GAIN interpolator, including two files:

1. Binary file that contains the formation rate array

2. Descriptor file that lists the independent parameters and their ranges

The rates are saved in the binary file as a flattened 1-dimensional array that can be constructed by iterating over the independent parameter values in a nested loop. The order of the parameters in the nested loop must be the same as that in the descriptor file. That is, if the descriptor file lists parameters $x_1, x_2, ..., x_N$, in this order, the outermost loop must correspond to $x_1$ and the innermost to $x_N$:

$i = 0$

**for** $i_1 = 1$ to $i_{\max,1}$ **do**

   ...

   **for** $i_N = 1$ to $i_{\max,N}$ **do**

     $i = i + 1$

     Save $f(i) = f(x_{1,i_1}, x_{2,i_2}, ..., x_{N,i_N})$

where $i_j$ is the index corresponding to values given for parameter $x_j$, and $i_{\mathrm{max},j}$ is the total number of values for $x_j$. Example codes for such loop are included in the J-GAIN repository. The descriptor file can be created by following the instructions and
315   example files in the repository (table generator subdirectory).

### Appendix B: Parameter descriptions in the diurnal test case

Table B1 lists the functional forms of the independent parameters applied for the diurnal test case (Fig. 5). The time profiles, visualized in Fig. B1, are simply set to exhibit realistic magnitudes and shapes: the temperature peaks at daytime while the condensation sink drops as boundary layer grows, and $[H_2SO_4]$ follows a sinusoidal diurnal profile due to the photochemical
320   production of $H_2SO_4$. $[NH_3]$ and ion production rate generally do not show regular diurnal patterns, but are here set to vary in order to test the interpolation approach at different combinations of independent parameters.

**Table B1.** Parameter value as a function of time $t$ (in hours) for the diurnal test case.

| Parameter | Value $f = f(t), t = 0...24$ h |
|---|---|
| $[H_2SO_4]$ [cm$^{-3}$] | $\begin{cases} 0.5 \times (10^7 - 10^5) \times (\sin(2\pi(t-6)/12 - 0.5\pi) + 1) + 10^5, & t \in \{6...18\} \\ 10^5, & \text{otherwise} \end{cases}$ |
| $[NH_3]$ [cm$^{-3}$] | $(10^9 - 2 \cdot 10^{10}) \times \exp(-((t-18)/12)^2) + 2 \cdot 10^{10}$ |
| $T$ [K] | $10 \times \exp(-((t-10)/6)^2) + 283.15$ |
| CS [s$^{-1}$] | $(10^{-3} - 6 \cdot 10^{-3}) \times \exp(-((t-10)/12)^2) + 6 \cdot 10^{-3}$ |
| IPR [cm$^{-3}$s$^{-1}$] | $0.5 \times (4 - 1) \times (\sin(2\pi(t-6)/24 - 0.5\pi) + 1) + 1$ |
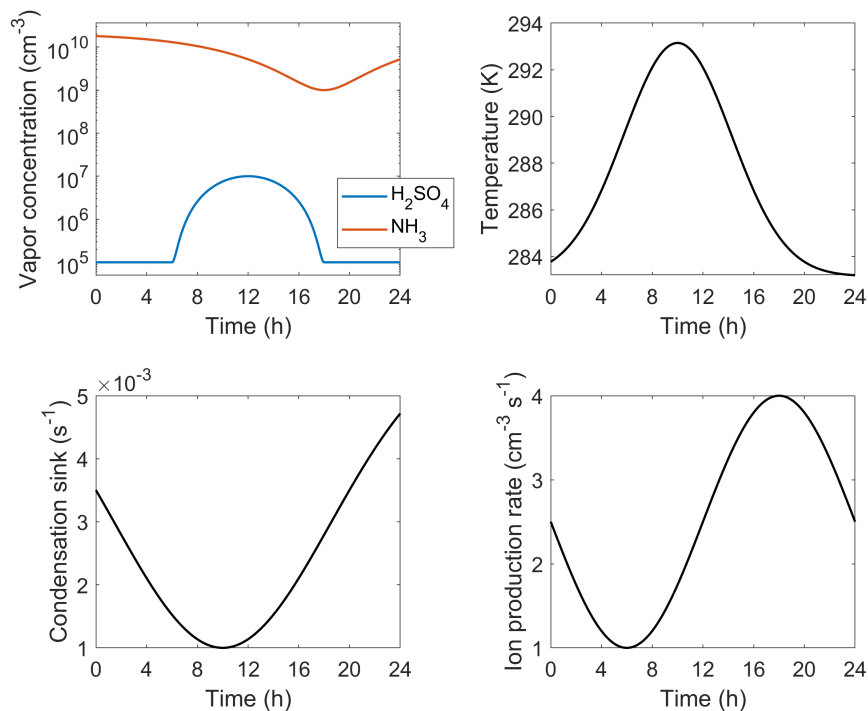
**Figure B1.** Time profiles of the independent parameters for the diurnal test case.

*Code and data availability.* J-GAIN is available at https://github.com/tolenius/J-GAIN (current version) and doi:10.5281/zenodo.7457152 (v1.0). Scripts and tables applied for the figures in this work are available from the corresponding authors.

*Author contributions.* DY designed, wrote and tested the programs, and created the figures. TO conceived the project and wrote the manuscript.
325 All authors contributed to discussing the results and revising the manuscript.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors gratefully acknowledge financial support from the Swedish Research Council VR (grant no. 2019-04853) and the Swedish Research Council for Sustainable Development FORMAS (grant no. 2019-01433). We thank Pontus Roldin and Carl Svenhag for discussions and help with testing the tool.

# References

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

Baranizadeh, E., Murphy, B. N., Julin, J., Falahat, S., Reddington, C. L., Arola, A., Ahlm, L., Mikkonen, S., Fountoukis, C., Patoulias, D., Minikin, A., Hamburger, T., Laaksonen, A., Pandis, S. N., Vehkamäki, H., Lehtinen, K. E. J., and Riipinen, I.: Implementation of state-of-the-art ternary new-particle formation scheme to the regional chemical transport model PMCAMx-UF in Europe, Geosci. Model Dev., 9, 2741–2754, https://doi.org/10.5194/gmd-9-2741-2016, 2016.

Bergman, T., Laaksonen, A., Korhonen, H., Malila, J., Dunne, E. M., Mielonen, T., Lehtinen, K. E. J., Kühn, T., Arola, A., and Kokkola, H.: Geographical and diurnal features of amine-enhanced boundary layer nucleation, J. Geophys. Res. Atmos., 120, 9606–9624, https://doi.org/10.1002/2015JD023181, 2015.

Besel, V., Kubečka, J., Kurtén, T., and Vehkamäki, H.: Impact of Quantum Chemistry Parameter Choices and Cluster Distribution Model Settings on Modeled Atmospheric Particle Formation Rates, J. Phys. Chem. A, 124, 5931–5943, https://doi.org/10.1021/acs.jpca.0c03984, 2020.

Blichner, S. M., Sporre, M. K., Makkonen, R., and Berntsen, T. K.: Implementing a sectional scheme for early aerosol growth from new particle formation in the Norwegian Earth System Model v2: comparison to observations and climate impacts, Geosci. Model Dev., 14, 3335–3359, https://doi.org/10.5194/gmd-14-3335-2021, 2021.

Dunne, E. M., Gordon, H., Kürten, A., Almeida, J., Duplissy, J., Williamson, C., Ortega, I. K., Pringle, K. J., Adamov, A., Baltensperger, U., Barmet, P., Benduhn, F., Bianchi, F., Breitenlechner, M., Clarke, A., Curtius, J., Dommen, J., Donahue, N. M., Ehrhart, S., Flagan, R. C., Franchin, A., Guida, R., Hakala, J., Hansel, A., Heinritzi, M., Jokinen, T., Kangasluoma, J., Kirkby, J., Kulmala, M., Kupc, A., Lawler, M. J., Lehtipalo, K., Makhmutov, V., Mann, G., Mathot, S., Merikanto, J., Miettinen, P., Nenes, A., Onnela, A., Rap, A., Reddington, C. L. S., Riccobono, F., Richards, N. A. D., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Sengupta, K., Simon, M., Sipilä, M., Smith, J. N., Stozhkov, Y., Tomé, A., Tröstl, J., Wagner, P. E., Wimmer, D., Winkler, P. M., Worsnop, D. R., and Carslaw, K. S.: Global atmospheric particle formation from CERN CLOUD measurements, Science, 354, 1119–1124, https://doi.org/10.1126/science.aaf2649, 2016.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F.,

Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

370  Ehrhart, S., Dunne, E. M., Manninen, H. E., Nieminen, T., Lelieveld, J., and Pozzer, A.: Two new submodels for the Modular Earth Submodel System (MESSy): New Aerosol Nucleation (NAN) and small ions (IONS) version 1.0, Geosci. Model Dev., 11, 4987–5001, https://doi.org/10.5194/gmd-11-4987-2018, 2018.

Elm, J.: An Atmospheric Cluster Database Consisting of Sulfuric Acid, Bases, Organics, and Water, ACS Omega, 4, 10 965–10 974, https://doi.org/10.1021/acsomega.9b00860, 2019a.

375  Elm, J.: Unexpected Growth Coordinate in Large Clusters Consisting of Sulfuric Acid and C8H12O6 Tricarboxylic Acid, J. Phys. Chem. A, 123, 3170–3175, https://doi.org/10.1021/acs.jpca.9b00428, 2019b.

Elm, J., Kubečka, J., Besel, V., Jääskeläinen, M. J., Halonen, R., Kurtén, T., and Vehkamäki, H.: Modeling the formation and growth of atmospheric molecular clusters: A review, J. Aerosol Sci., 149, 105 621, https://doi.org/10.1016/j.jaerosci.2020.105621, 2020.

Fountoukis, C., Riipinen, I., Denier van der Gon, H. a. C., Charalampidis, P. E., Pilinis, C., Wiedensohler, A., O'Dowd, C., Putaud, J. P., Moerman, M., and Pandis, S. N.: Simulating ultrafine particle formation in Europe using a regional CTM: contribution of primary emissions versus secondary formation to aerosol number concentrations, Atmos. Chem. Phys., 12, 8663–8677, https://doi.org/10.5194/acp-12-8663-2012, 2012.

Glasoe, W. A., Volz, K., Panta, B., Freshour, N., Bachman, R., Hanson, D. R., McMurry, P. H., and Jen, C.: Sulfuric acid nucleation: An experimental study of the effect of seven bases, J. Geophys. Res. Atmos., 120, 1933–1950, https://doi.org/10.1002/2014JD022730, 2015.

385  Gordon, H., Kirkby, J., Baltensperger, U., Bianchi, F., Breitenlechner, M., Curtius, J., Dias, A., Dommen, J., Donahue, N. M., Dunne, E. M., Duplissy, J., Ehrhart, S., Flagan, R. C., Frege, C., Fuchs, C., Hansel, A., Hoyle, C. R., Kulmala, M., Kürten, A., Lehtipalo, K., Makhmutov, V., Molteni, U., Rissanen, M. P., Stozhkov, Y., Tröstl, J., Tsagkogeorgas, G., Wagner, R., Williamson, C., Wimmer, D., Winkler, P. M., Yan, C., and Carslaw, K. S.: Causes and importance of new particle formation in the present-day and preindustrial atmospheres, J. Geophys. Res. Atmos., 122, 8739–8760, https://doi.org/10.1002/2017JD026844, 2017.

390  Henschel, H., Kurtén, T., and Vehkamäki, H.: Computational Study on the Effect of Hydration on New Particle Formation in the Sulfuric Acid/Ammonia and Sulfuric Acid/Dimethylamine Systems, J. Phys. Chem. A, 120, 1886–1896, https://doi.org/10.1021/acs.jpca.5b11366, 2016.

Jen, C. N., McMurry, P. H., and Hanson, D. R.: Stabilization of sulfuric acid dimers by ammonia, methylamine, dimethylamine, and trimethylamine, J. Geophys. Res. Atmos., 119, 7502–7514, https://doi.org/https://doi.org/10.1002/2014JD021592, 2014.

395  Jen, C. N., Bachman, R., Zhao, J., McMurry, P. H., and Hanson, D. R.: Diamine-sulfuric acid reactions are a potent source of new particle formation, Geophys. Res. Lett., 43, 867–873, https://doi.org/10.1002/2015GL066958, 2016.

Julin, J., Murphy, B. N., Patoulias, D., Fountoukis, C., Olenius, T., Pandis, S. N., and Riipinen, I.: Impacts of Future European Emission Reductions on Aerosol Particle Number Concentrations Accounting for Effects of Ammonia, Amines, and Organic Species, Environ. Sci. Technol., 52, 692–700, https://doi.org/10.1021/acs.est.7b05122, 2018.

400  Kazil, J., Stier, P., Zhang, K., Quaas, J., Kinne, S., O'Donnell, D., Rast, S., Esch, M., Ferrachat, S., Lohmann, U., and Feichter, J.: Aerosol nucleation and its role for clouds and Earth's radiative forcing in the aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 10, 10 733–10 752, https://doi.org/10.5194/acp-10-10733-2010, 2010.

Kerminen, V.-M., Chen, X., Vakkari, V., Petäjä, T., Kulmala, M., and Bianchi, F.: Atmospheric new particle formation and growth: review of field observations, Environ. Res. Lett., 13, 103 003, https://doi.org/10.1088/1748-9326/aadf3c, 2018.

405 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,
410 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, https://doi.org/10.1038/nature17953, 2016.

Kontkanen, J., Lehtipalo, K., Ahonen, L., Kangasluoma, J., Manninen, H. E., Hakala, J., Rose, C., Sellegri, K., Xiao, S., Wang, L., Qi,
415 X., Nie, W., Ding, A., Yu, H., Lee, S., Kerminen, V.-M., Petäjä, T., and Kulmala, M.: Measurements of sub-3 nm particles using a particle size magnifier in different environments: from clean mountain top to polluted megacities, Atmos. Chem. Phys., 17, 2163–2187, https://doi.org/10.5194/acp-17-2163-2017, 2017.

Kürten, A., Bianchi, F., Almeida, J., Kupiainen-Määttä, O., Dunne, E. M., Duplissy, J., Williamson, C., Barmet, P., Breitenlechner, M., Dommen, J., Donahue, N. M., Flagan, R. C., Franchin, A., Gordon, H., Hakala, J., Hansel, A., Heinritzi, M., Ickes, L., Jokinen, T.,
420 Kangasluoma, J., Kim, J., Kirkby, J., Kupc, A., Lehtipalo, K., Leiminger, M., Makhmutov, V., Onnela, A., Ortega, I. K., Petäjä, T., Praplan, A. P., Riccobono, F., Rissanen, M. P., Rondo, L., Schnitzhofer, R., Schobesberger, S., Smith, J. N., Steiner, G., Stozhkov, Y., Tomé, A., Tröstl, J., Tsagkogeorgas, G., Wagner, P. E., Wimmer, D., Ye, P., Baltensperger, U., Carslaw, K., Kulmala, M., and Curtius, J.: Experimental particle formation rates spanning tropospheric sulfuric acid and ammonia abundances, ion production rates, and temperatures, J. Geophys. Res. Atmos., 121, 12,377–12,400, https://doi.org/10.1002/2015JD023908, 2016.

425 Kürten, A., Li, C., Bianchi, F., Curtius, J., Dias, A., Donahue, N. M., Duplissy, J., Flagan, R. C., Hakala, J., Jokinen, T., Kirkby, J., Kulmala, M., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Onnela, A., Rissanen, M. P., Simon, M., Sipilä, M., Stozhkov, Y., Tröstl, J., Ye, P., and McMurry, P. H.: New particle formation in the sulfuric acid–dimethylamine–water system: reevaluation of CLOUD chamber measurements and comparison to an aerosol nucleation and growth model, Atmos. Chem. Phys., 18, 845–863, https://doi.org/10.5194/acp-18-845-2018, 2018.

430 Lee, Y. H., Pierce, J. R., and Adams, P. J.: Representation of nucleation mode microphysics in a global aerosol model with sectional microphysics, Geosci. Model Dev., 6, 1221–1232, https://doi.org/10.5194/gmd-6-1221-2013, 2013.

Lehtinen, K. E. J., Dal Maso, M., Kulmala, M., and Kerminen, V.-M.: Estimating nucleation rates from apparent particle formation rates and vice versa: Revised formulation of the Kerminen–Kulmala equation, J. Aerosol Sci., 38, 988–994, https://doi.org/10.1016/j.jaerosci.2007.06.009, 2007.

435 Lehtipalo, K., Yan, C., Dada, L., Bianchi, F., Xiao, M., Wagner, R., Stolzenburg, D., Ahonen, L. R., Amorim, A., Baccarini, A., Bauer, P. S., Baumgartner, B., Bergen, A., Bernhammer, A.-K., Breitenlechner, M., Brilke, S., Buchholz, A., Mazon, S. B., Chen, D., Chen, X., Dias, A., Dommen, J., Draper, D. C., Duplissy, J., Ehn, M., Finkenzeller, H., Fischer, L., Frege, C., Fuchs, C., Garmash, O., Gordon, H., Hakala, J., He, X., Heikkinen, L., Heinritzi, M., Helm, J. C., Hofbauer, V., Hoyle, C. R., Jokinen, T., Kangasluoma, J., Kerminen, V.-M., Kim, C., Kirkby, J., Kontkanen, J., Kürten, A., Lawler, M. J., Mai, H., Mathot, S., Mauldin, R. L., Molteni, U., Nichman, L., Nie, W., Nieminen,
440 T., Ojdanic, A., Onnela, A., Passananti, M., Petäjä, T., Piel, F., Pospisilova, V., Quéléver, L. L. J., Rissanen, M. P., Rose, C., Sarnela, N., Schallhart, S., Schuchmann, S., Sengupta, K., Simon, M., Sipilä, M., Tauber, C., Tomé, A., Tröstl, J., Väisänen, O., Vogel, A. L., Volkamer, R., Wagner, A. C., Wang, M., Weitz, L., Wimmer, D., Ye, P., Ylisirniö, A., Zha, Q., Carslaw, K. S., Curtius, J., Donahue, N. M.,

Flagan, R. C., Hansel, A., Riipinen, I., Virtanen, A., Winkler, P. M., Baltensperger, U., Kulmala, M., and Worsnop, D. R.: Multicomponent new particle formation from sulfuric acid, ammonia, and biogenic vapors, Sci. Adv., 4, eaau5363, https://doi.org/10.1126/sciadv.aau5363, 2018.

Li, C. and Signorell, R.: Understanding vapor nucleation on the molecular level: A review, J. Aerosol Sci., 153, 105 676, https://doi.org/10.1016/j.jaerosci.2020.105676, 2021.

Makkonen, R., Asmi, A., Kerminen, V.-M., Boy, M., Arneth, A., Guenther, A., and Kulmala, M.: BVOC-aerosol-climate interactions in the global aerosol-climate model ECHAM5.5-HAM2, Atmos. Chem. Phys., 12, 10 077–10 096, https://doi.org/10.5194/acp-12-10077-2012, 2012.

Myllys, N., Kubečka, J., Besel, V., Alfaouri, D., Olenius, T., Smith, J. N., and Passananti, M.: Role of base strength, cluster structure and charge in sulfuric-acid-driven particle formation, Atmos. Chem. Phys., 19, 9753–9768, https://doi.org/10.5194/acp-19-9753-2019, 2019.

Naik, V., Szopa, S., Adhikary, B., Artaxo, P., Berntsen, T., Collins, W. D., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A., Klimont, Z., Liao, H., Unger, N., and Zanis, P.: Short-Lived Climate Forcers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu R., and Zhou B. (eds.)], Cambridge University Press, in press, 2021.

Olenius, T.: Atmospheric Cluster Dynamics Code: Software repository, https://github.com/tolenius/ACDC, 2021.

Olenius, T. and Riipinen, I.: Molecular-resolution simulations of new particle formation: Evaluation of common assumptions made in describing nucleation in aerosol dynamics models, Aerosol Sci. Tech., 51, 397–408, https://doi.org/10.1080/02786826.2016.1262530, 2017.

Olenius, T. and Roldin, P.: Role of gas–molecular cluster–aerosol dynamics in atmospheric new-particle formation, Sci. Rep., 12, 10 135, https://doi.org/10.1038/s41598-022-14525-y, 2022.

Olenius, T., Kupiainen-Määttä, O., Ortega, I. K., Kurtén, T., and Vehkamäki, H.: Free energy barrier in the growth of sulfuric acid–ammonia and sulfuric acid–dimethylamine clusters, J. Chem. Phys., 139, 084 312, https://doi.org/10.1063/1.4819024, 2013.

Olenius, T., Halonen, R., Kurtén, T., Henschel, H., Kupiainen-Määttä, O., Ortega, I. K., Jen, C. N., Vehkamäki, H., and Riipinen, I.: New particle formation from sulfuric acid and amines: Comparison of monomethylamine, dimethylamine, and trimethylamine, J. Geophys. Res. Atmos., 122, 7103–7118, https://doi.org/10.1002/2017JD026501, 2017.

Patoulias, D. and Pandis, S. N.: Simulation of the effects of low-volatility organic compounds on aerosol number concentrations in Europe, Atmos. Chem. Phys., 22, 1689–1706, https://doi.org/10.5194/acp-22-1689-2022, 2022.

Schade, G. W. and Crutzen, P. J.: Emission of aliphatic amines from animal husbandry and their reactions: Potential source of N2O and HCN, J. Atmos. Chem., 22, 319–346, https://doi.org/10.1007/BF00696641, 1995.

Schervish, M. and Donahue, N. M.: Peroxy radical kinetics and new particle formation, Environ. Sci. Atmos., 1, 79–92, https://doi.org/10.1039/D0EA00017E, 2021.

van Noije, T. P. C., Le Sager, P., Segers, A. J., van Velthoven, P. F. J., Krol, M. C., Hazeleger, W., Williams, A. G., and Chambers, S. D.: Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth, Geosci. Model Dev., 7, 2435–2475, https://doi.org/10.5194/gmd-7-2435-2014, 2014.

Xiao, M., Hoyle, C. R., Dada, L., Stolzenburg, D., Kürten, A., Wang, M., Lamkaddam, H., Garmash, O., Mentler, B., Molteni, U., Baccarini, A., Simon, M., He, X.-C., Lehtipalo, K., Ahonen, L. R., Baalbaki, R., Bauer, P. S., Beck, L., Bell, D., Bianchi, F., Brilke, S., Chen, D., Chiu, R., Dias, A., Duplissy, J., Finkenzeller, H., Gordon, H., Hofbauer, V., Kim, C., Koenig, T. K., Lampilahti, J., Lee, C. P., Li, Z., Mai, H., Makhmutov, V., Manninen, H. E., Marten, R., Mathot, S., Mauldin, R. L., Nie, W., Onnela, A., Partoll, E., Petäjä, T., Pfeifer,

J., Pospisilova, V., Quéléver, L. L. J., Rissanen, M., Schobesberger, S., Schuchmann, S., Stozhkov, Y., Tauber, C., Tham, Y. J., Tomé, A., Vazquez-Pufleau, M., Wagner, A. C., Wagner, R., Wang, Y., Weitz, L., Wimmer, D., Wu, Y., Yan, C., Ye, P., Ye, Q., Zha, Q., Zhou, X., Amorim, A., Carslaw, K., Curtius, J., Hansel, A., Volkamer, R., Winkler, P. M., Flagan, R. C., Kulmala, M., Worsnop, D. R., Kirkby, J., Donahue, N. M., Baltensperger, U., El Haddad, I., and Dommen, J.: The driving factors of new particle formation and growth in the polluted boundary layer, Atmos. Chem. Phys., 21, 14 275–14 291, https://doi.org/10.5194/acp-21-14275-2021, 2021.

Yazgi, D. and Olenius, T.: J-GAIN: Aerosol particle formation rate look-up table generator and interpolator: Software repository, https://github.com/tolenius/J-GAIN, 2022a.

Yazgi, D. and Olenius, T.: J-GAIN v1.0: Aerosol particle formation rate look-up table generator and interpolator, Zenodo [code], https://doi.org/10.5281/zenodo.7457152, 2022b.

Yu, F.: Nucleation rate of particles in the lower atmosphere: Estimated time needed to reach pseudo-steady state and sensitivity to H2SO4 gas concentration, Geophys. Res. Lett., 30, https://doi.org/10.1029/2003GL017084, 2003.

Yu, F., Nadykto, A. B., Luo, G., and Herb, J.: $H_2SO_4$–$H_2O$ binary and $H_2SO_4$–$H_2O$–$NH_3$ ternary homogeneous and ion-mediated nucleation: lookup tables version 1.0 for 3-D modeling application, Geosci. Model Dev., 13, 2663–2670, https://doi.org/10.5194/gmd-13-2663-2020, 2020.

Zhang, R., Suh, I., Zhao, J., Zhang, D., Fortner, E. C., Tie, X., Molina, L. T., and Molina, M. J.: Atmospheric New Particle Formation Enhanced by Organic Acids, Science, 304, 1487–1490, https://doi.org/10.1126/science.1095139, 2004.