**Cluster Dynamics-based Parameterization for Sulfuric Acid-Dimethylamine** 

- **Nucleation: Comparison and Selection through Box- and Three-Dimensional-**
- **Modeling**
- 4 Jiewen Shen<sup>1,2</sup>, Bin Zhao<sup>1,2</sup>, Shuxiao Wang<sup>1,2,\*</sup>, An Ning<sup>3</sup>, Yuyang Li<sup>2</sup>, Runlong Cai<sup>4</sup>,
- 5 Da Gao<sup>1,2</sup>, Biwu Chu<sup>5,6</sup>, Yang Gao<sup>7</sup>, Manish Shrivastava<sup>8</sup>, Jingkun Jiang<sup>2</sup>, Xiuhui
- $Zhang<sup>3</sup>$ , Hong He<sup>5,6</sup>
- 
- 8 <sup>1</sup> State Key Joint Laboratory of Environment Simulation and Pollution Control, School
- *of Environment, Tsinghua University, Beijing, 100084, China*
- 10 <sup>2</sup> State Environmental Protection Key Laboratory of Sources and Control of Air
- *Pollution Complex, Beijing, 100084, China*
- 12 <sup>3</sup>Key Laboratory of Cluster Science, Ministry of Education of China, School of
- *Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, 100081, China*
- 15 <sup>4</sup>Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP<sup>3</sup>),
- *Department of Environmental Science & Engineering, Fudan University, Shanghai,*
- *200438, China*
- 5 *State Key Joint Laboratory of Environment Simulation and Pollution Control,*
- *Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences,*
- *Beijing 100085, China*
- 21 <sup>6</sup>College of Resources and Environment, University of Chinese Academy of Sciences,
- *Beijing 100049, China*
- 7 *Key Laboratory of Marine Environment and Ecology, Ministry of Education, Ocean*
- *University of China, Qingdao 266100, China*
- 8 *Pacific Northwest National Laboratory, Richland, Washington, USA*
- 
- 27 \* Correspondence to: Shuxiao Wang (shxwang@tsinghua.edu.cn)

#### **ABSTRACT**

 Clustering of gaseous sulfuric acid (SA) enhanced by dimethylamine (DMA) is a major mechanism for new particle formation (NPF) in polluted atmospheres. However, uncertainty remains regarding the SA-DMA nucleation parameterization that reasonably represents cluster dynamics and is applicable across various atmospheric conditions. This uncertainty hinders accurate three-dimensional (3-D) modeling of NPF and subsequent assessment of its environmental and climatic impacts. Here we extensively compare different cluster dynamics-based parameterizations for SA-DMA nucleation and identify the most reliable one through a combination of box-model simulations, 3-D modeling, and in-situ observations. Results show that the parameterization derived from Atmospheric Cluster Dynamic Code (ACDC) simulations, incorporating the latest theoretical insights (DLPNO-CCSD(T)/aug-cc- pVTZ//ωB97X-D/6-311++G(3df,3pd) level of theory) and adequate representation of cluster dynamics, exhibits dependable performance in 3-D NPF simulation for both winter and summer conditions in Beijing and shows promise for application in diverse atmospheric conditions. Another ACDC-derived parameterization, replacing the level of theory with RI-CC2/aug-cc-pV(T+d)Z//M06-2X/6–311++G(3df,3pd), also performs well in NPF modeling at relatively low temperatures around 280 K but exhibits limitations at higher temperatures due to inappropriate representation of SA-DMA cluster thermodynamics. Additionally, a previously reported parameterization incorporating simplifications is applicable for simulating NPF in polluted atmospheres but tends to overestimate particle formation rates under conditions of elevated 50 temperature ( $>$  ~300 K) and low condensation sink ( $<$  ~3×10<sup>-3</sup> s<sup>-1</sup>). Our findings highlight the applicability of the new ACDC-derived parameterization, which couples the latest SA-DMA nucleation theory and holistic cluster dynamics, in 3-D NPF modeling. The ACDC-derived parameterization framework provides valuable reference for developing parameterizations for other nucleation systems.



### **1 INTRODUCTION**

 Atmospheric aerosols have significant impacts on visibility, human health, and global climate (Gordon et al., 2016; Gao et al., 2024). New Particle Formation (NPF) is the predominant source of global aerosol population, with nucleation being the key stage of the gas-to-particle transformation (Zhao et al., 2020; Almeida et al., 2013). In polluted regions such as urban China, compelling evidence indicates that sulfuric acid (SA)-driven nucleation enhanced by dimethylamine (DMA) can generate thermodynamically stable SA-DMA clusters and lead to high particle formation rates close to kinetic limit of SA clustering, which is responsible for the observed intensive NPF events (Cai et al., 2021; Yao et al., 2018). Meanwhile, it has been demonstrated that variations in atmospheric conditions, including condensation sinks (CS) arising from background aerosols, along with temperature (*T*), can exert profound impacts on the cluster dynamics of SA-DMA nucleation by varying the particle formation rates across several orders of magnitude (Cai et al., 2021; Deng et al., 2020). Given that complex interactions exist among various gaseous precursors, molecular clusters, and pre-existing aerosols during nucleation, reasonable representation of the cluster dynamics of SA-DMA nucleation in three-dimensional (3-D) models is important for 3-D NPF modeling and subsequent assessment of its impacts on environment and climate.

 Empirical models in form of power law functions have been extensively utilized to examine how particle formation rates respond to precursor concentrations (Semeniuk and Dastoor, 2018). Through parameter fitting, these empirical models can effectively reproduce the particle formation rates observed in both laboratory experiments and field measurements (Kulmala et al., 2006; Riccobono et al., 2014; Semeniuk and Dastoor, 2018). Subsequently, they can be integrated into 3-D models for regional or global NPF simulations. Bergman et al. (2015) and Dunne et al. (2016) have simulated SA-DMA nucleation utilizing global models, which incorporate empirical equations derived from experimental data obtained from CLOUD chamber or flow tube experiments. These parameterization schemes successfully characterize the response of particle formation rates to precursor concentrations, however, they fail to account for dependencies on *T*  88 and CS due to the ignorance of explicit cluster dynamics. As a result, they are identified to be inadequate for accurately reproducing NPF events in winter Beijing (Li et al., 2023c).

 We recently developed an analytical equation for SA-DMA nucleation parameterization based on detailed cluster dynamics simulations (abbreviated as Dynamic\_Sim) (Li et al., 2023c). Previous theoretical insights into the SA-DMA system (Olenius et al., 2013, 2017; Ortega et al., 2012; Myllys et al., 2019) indicate that 95 (SA)<sub>k</sub>(DMA)<sub>k</sub> ( $k = 1-4$ ) and (SA)<sub>2</sub>(DMA)<sub>1</sub> clusters are considered the key clusters along the cluster formation pathways in SA-DMA nucleation. Under the polluted conditions 97 (CS > ~1.0×10<sup>-2</sup> s<sup>-1</sup>), the evaporation rates of clusters  $(SA)_k(DMA)_k$  ( $k = 2-4$ ) and  $(SA)<sub>2</sub>(DMA)<sub>1</sub> clusters are negligible compared to their coagulation sink. Accordingly,$ several simplifications have been made in Dynamic\_Sim, including 1) only

100 (SA)<sub>k</sub>(DMA)<sub>k</sub> ( $k = 1-4$ ) and (SA)<sub>2</sub>(DMA)<sub>1</sub> clusters are considered; 2) clusters larger 101 than  $(SA)_{1}(DMA)_{1}$  are regarded stable with no evaporation; and 3)  $(SA)_{4}(DMA)_{4}$  cluster is the only terminal cluster in calculating particle formation rates. Subsequent applications in 3-D modeling have demonstrated significantly improved performance of Dynamic\_Sim compared to previous data-fitting parameterizations in simulating the particle formation rates, the evolution of particle number size distributions (PNSDs), and NPF events in winter Beijing. However, the efficacy of Dynamic\_Sim in NPF simulation has yet to be assessed under varying atmospheric conditions, such as the summer season characterized by relatively higher *T* and lower CS compared to winter. Moreover, the impacts of simplifications made in the derivation of Dynamic\_Sim on 3- D NPF simulation under different atmospheric conditions remain unclear.

 In addition to the form of explicit formulations, integration of nucleation dynamics in 3-D models can also be realized using precomputed look-up tables generated by box models. Atmospheric Cluster Dynamics Code (ACDC) is a representative box model for simulating cluster dynamics and particle formation rates (Mcgrath et al., 2012; Olenius et al, 2013). In addition to representing *T*- and CS- dependencies for particle formation rate as Dynamic\_Sim, ACDC considers the source/sink terms of all given molecules/clusters within a nucleation system without simplifications of the clustering processes. By integrating quantum chemical calculations with ACDC, Almeida et al. (2013) discovered that the simulated SA-DMA nucleation provides valuable insights for interpreting the measurements from the CLOUD chamber experiments. Similarly, Lu et al. (2020) demonstrated that ACDC coupled with quantum chemistry calculations can effectively reproduce the particle formation rates observed in urban Shanghai. In addition to its extensive utilization in box modeling (Almeida et al., 2013; Lu et al., 2020; Yang et al., 2021), several studies have simulated nucleation pathways in chemical transport models using precomputed look-up tables generated by ACDC. For example, Baranizadeh et al. (2016) and Croft et al. (2016) used ACDC-derived look-up 127 tables as nucleation parameterizations to probe the impacts of SA–NH<sub>3</sub>–H<sub>2</sub>O nucleation on aerosol number concentration, cloud properties, and radiation balance. Olin et al. (2022) and Julin et al. (2018) evaluated the impact of new particle formation on aerosol number concentrations in Europe under historical and emission reduction scenarios, 131 respectively, using ACDC-derived parameterizations involving both SA–NH<sub>3</sub>–H<sub>2</sub>O and SA-DMA nucleation. It should be noted that ACDC program in modeling the nucleation process is highly reliant on specific thermodynamic data for the molecular clusters of interest, which are primarily obtained through quantum chemical calculations (Elm et al., 2020). A very recent study by Svenhag et al. (2024) compared the impact of two typical quantum calculation methods on 3-D modeling of SA-NH<sup>3</sup> nucleation using ACDC-derived parameterizations. However, it is still unclear how different quantum chemical methods affect the 3-D modeling of SA-DMA nucleation.

 This study aims to compare different cluster dynamic-based parameterizations for SA-DMA nucleation and identify the robust one applicable for 3-D models. We introduced parameterizations developed using the ACDC program, incorporating

 various quantum chemical calculations. Different cluster dynamic-based parameterizations, including ACDC-derived ones as well as Dynamic\_Sim, are comprehensively compared and evaluated through a combination of box-model simulations, 3-D modeling, and in-situ observational data. Our findings reveal that by incorporating the latest theoretical understanding and complete representation of cluster dynamics, ACDC-derived parameterization demonstrates reliable performance in 3-D NPF simulation for both winter and summer conditions in Beijing and exhibits potential applicability in diverse atmospheric conditions. The study sheds light on the impacts of employing various simplifications in cluster dynamics and different theoretical approaches in deriving parameterizations on NPF simulation. In addition to contributing to the precise simulation of SA-DMA nucleation and the quantification of its environmental and climatic effects, this study provides valuable references for simulating other nucleation mechanisms in 3-D models.

#### **2 METHODS**

### **2.1 Configurations of ACDC**

157 Here,  $(SA)<sub>m</sub>(DMA)<sub>n</sub>$  clusters ( $0 < n \le m \le 3$ , *m* and *n* represent the number of SA and DMA molecules in a cluster) are used to build the ACDC-derived parameterizations for SA-DMA nucleation due to their reported much higher stability compared to those containing more DMA molecules than SA molecules (Xie et al., 2017). The ACDC code is available at https://github.com/tolenius/ACDC. The conformations and 162 thermodynamics of SA-DMA clusters are taken from our other study (Ning et al., 2024). Briefly, the conformations of selected clusters are taken from the reported global minima from Li et al. (2020), and the key thermodynamic data for ACDC, Gibbs free energy change (Δ*G*), are recalculated at the DLPNO-CCSD(T)/aug-cc-pVTZ//ωB97X- D/6-311++G(3df,3pd) level of theory. Based on benchmark studies (Elm et al., 2020), this level of theory provides dependable thermodynamic insights into molecular clusters during nucleation and represents the latest theoretical approach. In addition, the rotational symmetry is consistently considered in quantum calculations following Besel et al. (2020). Following most previous ACDC simulation studies (Xie et al., 2017; Elm et al., 2020; Ning et al., 2020), (SA)4(DMA)<sup>3</sup> and (SA)4(DMA)<sup>4</sup> clusters are defined as the boundary conditions, i.e. the clusters fluxing out the simulated system and participating in subsequent growth in ACDC simulations, considering their high stability. Since clusters containing SA tetramers are estimated to have an electrical mobility diameter of 1.4 nm (Cai et al., 2023; Jen et al., 2014; Thomas et al., 2016), the 176 formation rates of  $(SA)_4(DMA)_3$  and  $(SA)_4(DMA)_4$  clusters are therefore deemed as the 177 particle formation rates at 1.4 nm  $(J_{1.4})$ . Size-dependent coagulation sink (CoagS) is counted for each SA-DMA cluster which is consistent with Dynamic\_Sim (Li et al., 2023c):

180 
$$
\text{Coags}_{i} = \text{CS} \left( \frac{V_{i}}{V_{1}} \right)^{\frac{1.7}{3}}
$$

181 where  $V_i$  and  $V_1$  (m<sup>3</sup>) represent the volume of cluster *i* and SA molecule, respectively.

The power-law exponent of -1.7 is selected according to typical range in the atmosphere

 (Lehtinen et al., 2007). In addition, enhancement for collision processes from Van de Waals forces is also considered. We refer to the ACDC-derived parameterization in coupling the DLPNO-CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311++G(3df,3pd) level of theory and adequate cluster dynamics as ACDC\_DB, which is established as the base-case for our discussion of other cluster dynamics-based parameterizations.

 In addition to the direct comparison of ACDC\_DB to Dynamic\_Sim, additional test parameterizations combining ACDC\_DB and three simplifications within Dynamic\_Sim are established and compared with ACDC\_DB to further probe the impacts of these simplifications on NPF simulations. According to our previous study, altering the simplifications within Dynamic\_Sim to explicit treatment would substantially escalate the computational demand by several orders of magnitude (Li et al. 2023c). Therefore, we utilize the ACDC-derived look-up tables to evaluate the impacts of the simplified treatments. The configurations of all parameterizations are detailed in Table 1. It should be noted that when all simplifications are applied on ACDC\_DB, Dynamic\_Sim still predicts higher *J*1.4 compared to ACDC\_DB (Figure 198 S1A). This is because the  $\Delta G$  value of the initial  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster at 298.15 K used in Dynamic\_Sim, which is taken from Myllys et al. (Myllys et al., 2019), is slightly 200 lower than that used in ACDC\_DB (-13.5 kcal mol<sup>-1</sup> for Dynamic\_Sim and -12.9 kcal 201 mol<sup>-1</sup> for ACDC\_DB) (Ning et al., 2024), even though both parameterizations employ the quantum chemical calculation method of DLPNO-CCSD(T). Possible reasons for 203 the discrepancy include the utilization of a larger basis set  $(3$ -zeta 6-311++G $(3df,3pd)$ ) and higher convergence criteria (Tight PNO + Tight SCF) in this study compared to 205 that in Myllys et al.. Aligning the  $\Delta G$  for  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster in Dynamic Sim with 206 that of ACDC leads to a high consistency in the predicted  $J_{1,4}$  between the two approaches (Figure S1B). The uncertainty of Δ*G* used in Dynamic\_Sim is discussed in our previous study (Li et al., 2023c) and here we mainly focus on the impacts of simplifications in Dynamic\_Sim.

210 While the DLPNO-CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311++G(3df,3pd) level of theory yields reasonable cluster thermodynamics, quantum chemistry calculations employing the RI-CC2 method predicting lower Δ*G* for cluster formation (stronger binding between molecules within clusters), has been widely used in conjunction with ACDC to interpret experimental and observed particle formation rates in previous studies (Almeida et al., 2013; Kürten et al., 2018; Ning et al., 2020). The prevalent combination used with the RI-CC2 method is RI-CC2/aug-cc-pV(T+d)Z//M06-2X/6- 311++G(3df,3pd) level of theory (Lu et al., 2020; Liu et al., 2021; Ning et al., 2022; Ning and Zhang, 2022; Liu et al., 2019). Based on Elm's work, compared to DLPNO-219 CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311++G(3df,3pd), the differences in predicted cluster binding energies primarily stem from discrepancies between DLPNO-CCSD(T) and RI-CC2 in single-point energy calculations, while the ωB97X-D and M06-2X functionals exhibit similar performance (Elm et al., 2013; Elm et al., 2020). Also, in previous studies the RI-CC2 method combined with ACDC was consistently accompanied by application of a sticking factor (SF) of 0.5 in treating collision  processes (Almeida et al., 2013; Lu et al., 2020). However, it is noteworthy that, according to Stolzenburg et al.'s work (Stolzenburg et al., 2020), the SF of the neutral SA-DMA cluster system should be unity. Here, we refer to the traditional theoretical 228 approach as employing the RI-CC2/aug-cc-pV(T+d)Z//M06-2X/6-311++G(3df,3pd) level of theory and incorporating the SF of 0.5 in collision processes. An ACDC-derived parameterization coupling the traditional theoretical approach is established to assess the effectiveness of the traditional method in NPF simulation (ACDC\_RM\_SF0.5). Except for the varied thermodynamic inputs and SF, the remaining configurations of 233 ACDC\_RM\_SF0.5 are identical to ACDC\_DB. Additionally, we establish a test parameterization coupling RI-CC2/aug-cc-pV(T+d)Z//M06-2X/6-311++G(3df,3pd) level of theory with an SF of unity (ACDC\_RM) to evaluate the impact solely arising 236 from the quantum chemical calculation method. Note that SF of unity is applied to all 237 parameterizations in this study except for the ACDC\_RM\_SF0.5.

238 To quantify the differences in simulating  $J_{1.4}$  among different cluster dynamics-239 based parameterizations compared to our base-case ACDC\_DB, we introduce a 240 parameter *R*:

$$
R_X = \frac{\sum_{i}^{n}(X_i/ACDC\_DB_i)}{n}
$$

242 where ACDC  $DB_i$  and  $X_i$  denote the simulated  $J_{1,4}$  by the base-case ACDC DB and another specific parameterization X, respectively, given the input scenarios of *i* (a set of input values for *T*, CS, concentration of SA ([SA]) and DMA ([DMA])), and *n* signifies the total number of input scenarios.

246







# **2.2 Incorporating the ACDC-derived Parameterizations into WRF-Chem/R2D-VBS Model**

 Various parameterizations are subsequently implemented in the Weather Research and Forecasting-Chemistry model (WRF-Chem) integrating an experimentally constrained Radical Two-Dimensional Volatility Basis Set (2D-VBS) (denoted as WRF-Chem/R2D-VBS) (Zhao et al., 2020). Incorporating the box-model ACDC into a 3-D model using the explicit mathematical formula, as Dynamic\_Sim, proves to be challenging. Here, we created a four-dimensional look-up table that delineates the response of *J*1.4 to four input variables (*T*, CS, [SA], and [DMA]) for each ACDC- derived parameterization (Yu, 2010). The table is derived based on multiple ACDC runs by varying input variables. The ranges for the input variables correspond to typical conditions of the atmosphere. Except for *T*, the ranges of variation for all other variables exceed at least one order of magnitude. Therefore, temperature is assumed to follow arithmetic uniform distribution, while the other variables are assumed to follow geometric uniform distribution. Details for the input variables are given in Table S1. In 265 WRF-Chem/R2D-VBS simulations,  $J_{1,4}$  are online calculated by interpolating values from a look-up table based on real-time input parameters. In our previous study, we have developed an emission inventory for China and its surrounding regions (Li et al., 2023c). Here [DMA] is calculated in WRF-Chem/R2D-VBS based on a comprehensive source-sink representation of DMA. More details of including DMA in WRF- Chem/R2D-VBS can be found in our previous study (Li et al., 2023c). In addition, a time-integrated-average [DMA] as well as [SA] of each time step were used to drive SA-DMA nucleation, since SA-DMA nucleation is accompanied with condensation of gaseous SA and DMA on pre-existing aerosols simultaneously in the atmosphere.

 Besides SA-DMA nucleation, seven other nucleation mechanisms have already been incorporated in WRF-Chem/R2D-VBS (Zhao et al., 2020), including neutral/ion- induced SA-H2O nucleation, neutral/ion-induced SA-NH3-H2O nucleation, neutral/ion- induced pure organics nucleation, and SA-organics nucleation. The organics involved in nucleation are ultralow- and extremely low-volatility organic compounds (ULVOC 279 and ELVOC) with O:C > 0.4. The formation chemistry of ULVOC and ELVOC from monoterpenes, including autoxidation and dimerization, is traced by the R2D-VBS framework (Zhao et al., 2020). Note that the impact of the other seven mechanisms on particle formation rates and particle number concentration is low compared to SA-DMA as revealed by our previous study (Li et al., 2023c). In WRF-Chem/R2D-VBS, the 284 evolution of PNSDs from 1nm to 10  $\mu$ m is treated by MOSAIC (Model for Simulating Aerosol Interactions and Chemistry) module. The newly formed 1.4 nm particles from SA-DMA nucleation are injected into the smallest size bin (1 - 1.5 nm) of the MOSAIC. **2.3 Configurations of WRF-Chem/R2D-VBS Model**

 The WRF-Chem/R2D-VBS model, incorporating various cluster dynamics-based SA-DMA nucleation parameterizations, was employed in a simulation over a domain with a spatial resolution of 27 km. This domain covers eastern Asia, with Beijing situated close to the center of the simulation area. Details of model configurations can be found in our previous study (Li et al., 2023c). Briefly, we use the ABaCAS-EI 2017 and IIASA 2015 emission inventories for mainland China and other areas in the domain, respectively, to represent the anthropogenic emissions (Zheng et al., 2019; Li et al., 2017; Li et al., 2023b); we use Model of Emissions of Gases and Aerosols from Nature (MEGAN) v2.04 to calculate the biogenic emissions (Guenther et al., 2006). To accurately represent the variation and distribution of chemical species concentrations during the simulation period, the chemical initial conditions, which represent the concentration field of chemical species at the initial simulation time, and the boundary conditions, which represent the flux or concentration around the simulation domain during the simulation period (Brasseur et al., 2017), are used in our WRF-Chem/R2D- VBS simulations. The simulation results from the National Center for Atmospheric Research's Community Atmosphere Model with Chemistry (https://www.acom.ucar.edu/cam-chem/cam-chem.shtml) is used for the chemical initial and boundary conditions in WRF-Chem/R2D-VBS simulations. In addition, we use a 5-day spin-up to minimize the impact of chemical initial conditions on simulation results.

 The simulation period consists of two parts: the winter period, which spans from January 14 to January 31, 2019, and the summer period, which is from August 18 to August 31, 2019. Previous observational studies have shown that the particle formation rates reach their highest and lowest levels during winter and summer in China, respectively (Deng et al., 2020; Chu et al., 2019). Therefore, periods from these two seasons are selected as representative simulation periods in this study and the specific time periods corresponded to those with relatively complete and continuous PNSDs and *J*1.4 observations. Since observational data for DMA concentration is only available for the period from January 1, 2019 to January 23, 2019, similar to our other study (Ning et al., 2024), we performed additional simulation for this period to compare observational and simulated DMA concentrations. For each season, all the SA-DMA parameterizations listed in Table 1 were employed for simulation. Among them, 320 ACDC\_DB, Dynamic\_Sim, and ACDC\_RM\_SF0.5 serve as three main

321 parameterizations, while ACDC DB CE, ACDC DB BC, ACDC DB CN, and ACDC\_RM are set as test cases to investigate the impact of individual simplification or theoretical approach on NPF simulations. In all comparisons, ACDC\_DB is set as a reference.

## **2.4 Ambient Measurements**

 In the 3-D simulations, we utilize measured concentrations of nucleation precursors and PNSDs as a criterion to discuss the model performance with various parameterizations. The duration of the observational data matches that of the simulations mentioned above. Detailed descriptions of the observation site and instruments can be found in our previous research (Deng et al., 2020; Zhu et al., 2022). Briefly, the observation site is located on the West Campus of the Beijing University of Chemical Technology. CI-TOF-MS (chemical ionization time-of-flight mass spectrometer; Aerodyne Research Inc.) were used to measure the concentrations of SA. 334 • Amine concentrations were measured with a modified TOF-MS using  $H_3O^+$  or its clusters as the reagent ions (Zhu et al., 2022). PNSDs from 1 nm to 10 µm were measured using a PSD (particle size distribution) system and a DEG-SMPS (diethyl glycol scanning mobility particle spectrometer). *J*1.4 derived from observation is calculated employing an improved aerosol population balance formula (Cai and Jiang, 2017).

### **3 RESULTS AND DISCUSSIONS**

# **3.1 Comparison of Different Parameterizations Based on Box-Model Simulations**

**3.1.1 Comparison between ACDC\_DB and Dynamic\_Sim**

 Figure 1 illustrates the comparison between the reported cluster dynamics-based parameterization with simplifications, Dynamic\_Sim, and the base-case parameterization ACDC\_DB. The comparison is based on a comprehensive dataset that includes over 40,000 box-model simulations for each parameterization, by varying parameters such as [SA] ( $1 \times 10^5 - 1 \times 10^8$  molec. cm<sup>-3</sup>), [DMA] ( $5 \times 10^6 - 5 \times 10^8$  348 molec. cm<sup>-3</sup>), CS (5  $\times$  10<sup>-4</sup> – 5  $\times$  10<sup>-1</sup> s<sup>-1</sup>), and *T* (250 – 320 K). In most scenarios, *J*<sub>1.4</sub> predicted by ACDC\_DB and Dynamic\_Sim demonstrates deviations within one order of magnitude, with the majority falling within a factor of 3. However, Dynamic\_Sim 351 predicts notably higher  $J_{1,4}$  than ACDC\_DB in scenarios where *T* exceeds ~300 K and 352 CS is below  $\sim 3 \times 10^{-3}$  s<sup>-1</sup>, characteristic of a clean atmosphere during summer. The 353 discrepancy in these scenarios elevates the overall *R*<sub>Dynamic</sub> s<sub>im</sub> up to 17.0. Furthermore, no clear correlation is observed between the differences of the two parameterizations and other input parameters such as [DMA] and [SA] (Figure S2). The differences between parameterizations are attributed to the combined effects of the three 357 simplifications and the lower  $\Delta G$  of  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster in Dynamic Sim. However, 358 the latter should not be the primary cause for the significant differences of  $J_{1,4}$  prediction under high *T* and low CS conditions, as it typically results in an overestimation within an order of magnitude (*R*=3.3) (Figure S1).



 **Figure 1.** Comparison of *J*1.4 predictions between ACDC\_DB and Dynamic\_Sim 364 correlated with *T* variation (A) and CS variation (B). Solid dots represent simulated  $J_{1,4}$  values, solid lines indicate a 1:1 line, dotted lines correspond to 1:3 and 3:1 lines, and dashed lines represent 1:10 and 10:1 lines.

 The impacts of the three simplifications made in Dynamic\_Sim are shown in Figure 2. Specifically, the simplification in cluster evaporations tends to elevate the predicted *J*1.4, whereas the simplifications in boundary conditions and cluster number tend to lower them. When applying the simplification in cluster evaporations (clusters larger than  $(SA)<sub>1</sub>(DMA)<sub>1</sub>$  are regarded stable with no evaporation) to ACDC DB, the 373 predicted  $J_{1,4}$  by ACDC\_DB\_CE only slightly exceed than that of ACDC\_DB within a

374 factor of 3 under conditions where  $T < -290$  K and CS  $> -0.1$  s<sup>-1</sup>. However, the 375 overestimation of  $J_{1,4}$  prediction by ACDC DB CE becomes much greater with increasing *T* and decreasing CS. The discrepancy between ACDC\_DB\_CE and ACDC\_DB should be primarily attributed to the pivotal role of *T* in influencing cluster evaporation rates (Ortega et al., 2012; Deng et al., 2020). At low *T*, the evaporation rates of clusters are low enough to allow efficient nucleation, thus whether setting the concerned SA-DMA clusters to evaporate based on the expected evaporation rates does not lead to a significant impact on *J*1.4 prediction. However, at high *T*, the evaporation rates of clusters significantly increase, therefore the simplification in cluster 383 evaporations within ACDC\_DB\_CE is likely to predict higher  $J_{1,4}$  than those with no simplification. The impact of simplification in cluster evaporations across varying *T* is also found in a nonbranched SA-DMA nucleation scheme from 280 K to 298 K reported by Li et al. (2023a). Note also that the overestimation of ACDC\_DB\_CE diminishes as CS increases (Figure 2D), with CS becoming the primary sink in the nucleation system and the impact of cluster evaporations becoming less pronounced. This underscores the connection between the specific deviation arising from simplification in cluster evaporations and the respective contributions of CS and cluster evaporations to the overall sink for clusters in nucleation. In addition, the relative independence of the differences between ACDC\_DB\_CE and ACDC\_DB from variations in precursor concentrations ([SA] and [DMA]) is similar to that between Dynamic\_Sim and ACDC\_DB (Figure S3). Overall, the scenarios where ACDC\_DB\_CE predicts higher *J*1.4 than ACDC\_DB only occurs under conditions of both high *T* and low CS (Figure 2A and Figure 2D). The averaged discrepancy between ACDC\_DB\_CE and 397 ACDC\_DB *R*ACDC\_DB\_CE is 22.3, closely resembling *R*Dynamic Sim, indicating that the simplification in cluster evaporations is a major factor contributing to the difference between Dynamic\_Sim and ACDC\_DB.





**Figure 2.** Comparison of *J*1.4 predictions between ACDC\_DB and test cases including

403 ACDC\_DB\_CE (A and D), ACDC\_DB\_BC (B and E), and ACDC\_DB\_CN (C and F). The first row in the panel (A, B and C) is correlated with *T* variation and the second row (D, E and F) is correlated with CS variation. Solid dots represent simulated *J*1.4 values, solid lines indicate a 1:1 line, dotted lines correspond to 1:3 and 3:1 lines, and dashed lines represent 1:10 and 10:1 lines.

408

409 The underestimations of ACDC\_DB\_BC and ACDC\_DB\_CN in *J*<sub>1.4</sub> prediction 410 compared to base-case ACDC\_DB are related to the growth pathways of SA-DMA 411 clusters. In the original scheme of ACDC\_DB, precursor molecules have the flexibility 412 to pass through any  $(SA)<sub>m</sub>(DMA)<sub>n</sub>$  clusters  $(0 < n \le m \le 3)$ , and terminal 1.4-nm 413 particles are formed when the clusters grow to  $(SA)_{4}(DMA)_{4}$  or  $(SA)_{4}(DMA)_{3}$ . As 414 expected, ACDC\_DB\_BC, which assumes  $(SA)_4(DMA)_4$  cluster as the only boundary 415 condition with an omission of  $(SA)_4(DMA)_3$  cluster, predicts lower  $J_{1,4}$  than ACDC DB. 416 (SA)<sub>4</sub>(DMA)<sub>3</sub> and  $(SA)_{4}$ (DMA)<sub>4</sub> clusters are primarily formed from  $(SA)_{3}$ (DMA)<sub>3</sub> 417 cluster by colliding with a SA molecule and a  $(SA)<sub>1</sub>(DMA)<sub>1</sub>$  cluster, respectively. As 418 the concentration of  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster is more sensitive to *T*, we further found that 419 the discrepancy between ACDC\_DB\_BC and ACDC\_DB becomes more pronounced 420 with increasing *T* (Figure 2B). Furthermore, we found no apparent correlation between 421 the variation of CS and the disparity between ACDC\_DB\_BC and ACDC\_DB (Figure 422 2E).

 In addition to ACDC\_DB\_BC, ACDC\_DB\_CN also underestimates *J*1.4 compared 424 to ACDC\_DB with a comparable value  $({\sim}0.5)$  of  $R_{ACDC}$  pB cN and  $R_{ACDC}$  pB BC. Under the simplification in cluster number, the formation of 1.4-nm clusters can only occur 426 through specific pathways, including  $(SA)_{1}(DMA)_{1} \rightarrow (SA)_{2}(DMA)_{2} \rightarrow (SA)_{3}(DMA)_{3}$  $427 \rightarrow (SA)<sub>4</sub>(DMA)<sub>4</sub>(SAA)<sub>4</sub>(DMA)<sub>3</sub>, (SA)<sub>1</sub>(DMA)<sub>1</sub> \rightarrow (SA)<sub>2</sub>(DMA)<sub>1</sub> \rightarrow (SA)<sub>2</sub>(DMA)<sub>2</sub> \rightarrow$  (SA)3(DMA)3  $\rightarrow$  (SA)4(DMA)4/(SA)4(DMA)3, or a combination thereof, while other pathways are restricted. Due to the variability in growth pathways and their 430 contributions to  $J_{1,4}$  under different atmospheric conditions, the difference between ACDC\_DB\_CN and ACDC\_DB is not strongly correlated with the variations of *T* and CS (Figure 2C and Figure 2F). Despite that, while the differences between the two tested parameterizations (ACDC\_DB\_BC and ACDC\_DB\_CN) involving cluster growth pathways and the original ACDC\_DB are not highly correlated with [DMA], there is a more pronounced correlation with [SA], which implies a more important role of SA in cluster growth (Figure S4 and Figure S5).

 In our previous study, we demonstrated improvements in computing CS- dependent *J*1.4 of SA-DMA nucleation with the Dynamic\_Sim compared to the previous power- law parameterizations under polluted atmospheric conditions (Li et al., 2023c). Here, we further show that, based on Dynamic\_Sim, the new ACDC\_DB with complete cluster dynamics can more reasonably simulate *J*1.4 under previously less studied 442 conditions of high  $T$  (> ~300 K) and low CS (< ~3×10<sup>-3</sup> s<sup>-1</sup>), where Dynamic\_Sim tends 443 to produce significant overestimation of  $J_{1,4}$ . This overestimation is primarily driven by the simplification in cluster evaporations within Dynamic\_Sim. Even though a  comparable performance in *J*1.4 prediction between ACDC\_DB and Dynamic\_Sim could be achieved under other ambient conditions, cautions should be made that the mutual offsetting effect between overestimation and underestimation resulting from different simplifications in Dynamic\_Sim when computing *J*1.4.

## **3.1.2 Comparison between ACDC\_DB and ACDC\_RM\_SF0.5**

 In Figure 3, ACDC\_DB is compared with another main ACDC-derived 451 parameterization, ACDC\_RM\_SF0.5, which uses the RI-CC2/aug-cc-pV(T+d)Z//M06- 2X/6-311++G(3df,3pd) level of theory and employs a SF of 0.5 in processing collision. 453 It can be observed that at lower temperatures (~280 K), ACDC RM SF0.5 and ACDC\_DB exhibit similar performance in predicting *J*1.4. However, with higher *T*  (accompanied by lower CS with a slight dependency), *J*1.4 predicted by 456 ACDC RM SF0.5 become higher than that predicted by ACDC DB, reaching even several orders of magnitude at the upper limit of the *T* range (320 K). Furthermore, we also observed that in scenarios close to the lower limit of the *T* range (250 K), the *J*1.4 predicted by ACDC\_RM\_SF0.5 shift from being higher to lower compared to ACDC\_DB.



 **Figure 3.** Comparison of *J*1.4 predictions between ACDC\_DB and ACDC\_RM\_SF0.5 correlated with *T* variation (A) and CS variation (B). Solid dots represent simulated *J*1.4 values, solid lines indicate a 1:1 line, dotted lines correspond to 1:3 and 3:1 lines, and dashed lines represent 1:10 and 10:1 lines.

 The distinction between ACDC\_RM\_SF0.5 and ACDC\_DB arises from the combined effects of variation in quantum chemical calculation method and the application of the 0.5 SF in collision processing. As depicted in Figure 4, when the SF 471 in ACDC\_RM\_SF0.5 is set to unity as in ACDC\_DB, the resulting ACDC\_RM parameterization predicts consistently higher *J*1.4 than ACDC\_DB. This implies that the modified quantum chemical calculation method, which results in lower evaporation 474 rates for clusters within the system compared to ACDC DB under the same condition, 475 leads to higher  $J_{1,4}$  predictions. The impact from varying quantum chemical calculation method is akin to that from simplification in cluster evaporations discussed earlier. The distinction between ACDC\_RM and ACDC\_DB\_CE lies in the fact that the modified quantum chemical calculation method affects all clusters within the system, whereas  the simplification in cluster evaporations is specific to limited clusters. This contributes 480 to a much higher  $R_{ACDC-RM}$  (614.5) compared to  $R_{ACDC-DB-CE}$  (22.3). Despite that, 481 compared to ACDC\_DB, the differences for both ACDC\_DB\_CE, ACDC\_RM, as well as ACDC\_RM\_SF0.5 demonstrate similar sensitivity to *T* (Figure 3A and Figure 4A) and CS (Figure 3B and Figure 4B) but independence on [SA] (Figure S6A and Figure S7A) and [DMA] (Figure S6B and Figure S7B). Comparing ACDC\_RM\_SF0.5 and 485 ACDC RM, it can be inferred that the application of a 0.5 SF in collision processes would result in an underestimation in *J*1.4 prediction. It can be noted that in most previous studies (Almeida et al., 2013; Kürten et al., 2018; Elm et al., 2020), comparisons of ACDC simulations using the traditional method and measured particle formation rates are conducted at around 280 K. At this temperature, all three main parameterizations of ACDC\_RM\_SF0.5, ACDC\_DB, and Dynamic\_Sim tends to yield similar *J*1.4 predictions and should have consistent applicability in NPF simulation.



 **Figure 4.** Comparison of *J*1.4 predictions between ACDC\_DB and ACDC\_RM 495 correlated with *T* variation (A) and CS variation (B). Solid dots represent simulated  $J_{1,4}$  values, solid lines indicate a 1:1 line, dotted lines correspond to 1:3 and 3:1 lines, and dashed lines represent 1:10 and 10:1 lines.

 In summary, based on our base-case parameterization ACDC\_DB, the extensive box-model simulations above demonstrate the characteristics of different parameterizations. Specifically, Dynamic\_Sim shows general consistency with 502 ACDC DB in simulating  $J_1$  a under most atmospheric conditions with  $T < \sim 300$  K or CS >  $\sim$ 3.0×10<sup>-3</sup> s<sup>-1</sup> while overestimating *J*<sub>1.4</sub> with *T* >  $\sim$ 300 K and CS >  $\sim$ 3.0×10<sup>-3</sup> s<sup>-1</sup> 504 compared to ACDC\_DB. ACDC\_RM\_SF0.5 performs similarly to ACDC\_DB under 505 conditions of  $\sim$  280 K but give different  $J_{1,4}$  predictions at other temperatures. We further use reported measurements from well-controlled CLOUD chamber experiments to examine the characteristics and applicability of these parameterizations (Xiao et al., 2021). As shown in Figure S8, simulated *J*1.4 using three main parameterizations, 509 ACDC\_DB, ACDC\_RM\_SF0.5, and Dynamic\_Sim, correspond well to measured  $J_{1.7}$ 510 at low temperature  $(T = 278 \text{ K})$ , proving the applicability of all three parameterizations 511 at this temperature. In the experiments with elevated temperature  $(T = 293 \text{ K})$ , ACDC\_DB and Dynamic\_Sim continues to exhibit similar performance, with slight  overestimation by approximately 2 factors. This may be because the much lower cluster concentrations at high temperatures compared to those at low temperatures lead to slower cluster growth and thus an enlarged gap between *J*1.4 and *J*1.7 (Figure S9). In contrast, ACDC\_RM\_SF0.5 only shows a slight *T*-dependence, which is deviated from the measurements. The comparison between controlled experiments and box-model simulations hence confirms our conclusions above, and provides a solid basis for further discussions on 3-D simulations using these parameterizations with constraint from field

observations.

## **3.2 Comparison of Different Parameterizations Based on 3-D Model Simulations**

 Various cluster dynamics-based parameterizations for SA-DMA nucleation were subsequently integrated into the WRF-Chem/R2D-VBS model. 3-D simulations using these parameterizations have been conducted for both wintertime and summertime conditions in Beijing. Given that the concentrations of precursors are crucial input variables for each parameterization, the simulated and observed concentrations of [DMA] and [SA] are compared. Figure S10, Figure S11 and Table S2 illustrates good consistencies in temporal variations and the mean values between simulations and observations in Beijing. This validates the reliability of our representation of sources and sinks for nucleating precursors and serves as a foundation for our discussions on the performances of various parameterizations. In the following sections, we discuss the results of 3-D NPF simulations in Beijing during winter and summer by employing different parameterizations. The evaluation of various parameterizations focuses on their ability to reproduce in situ NPF measurements across different seasons.

### **3.2.1 Wintertime Simulations**

 Figure 5A and Figure S12A primarily compare the simulated *J*1.4 values from different parameterizations with those derived from wintertime observations in Beijing, as *J*1.4 being a key parameter describing NPF events. The performance of Dynamic\_Sim in simulating *J*1.4 during wintertime Beijing has been discussed in our previous study 540 (Li et al., 2023c). The averaged  $J_{1,4}$  simulated by three main parameterizations (Dynamic\_Sim: 64.0 cm<sup>-3</sup> s<sup>-1</sup>; ACDC\_DB: 51.6 cm<sup>-3</sup> s<sup>-1</sup>; ACDC\_RM\_SF0.5: 54.5 cm<sup>-</sup> 542  $\frac{3}{5}$  s<sup>-1</sup>) approximate the observation (46.7 cm<sup>-3</sup> s<sup>-1</sup>). For test cases, however, only 543 ACDC\_DB\_CE (55.7 cm<sup>-3</sup> s<sup>-1</sup>) demonstrates a reasonable representation of  $J_{1.4}$ .  $J_{1.4}$ 544 simulated from ACDC\_DB\_BC (20.5 cm<sup>-3</sup> s<sup>-1</sup>) and ACDC\_DB\_CN (20.8 cm<sup>-3</sup> s<sup>-1</sup>) are approximately two times lower than the observed values, while ACDC\_RM (226.2 cm- <sup>3</sup> s<sup>-1</sup>) is approximately five times higher than the observations.

 The performances of different parameterizations on depicting *J*1.4 subsequently influences their representations of PNSDs evolution and NPF events, which are shown in Figure 5B. Generally, most parameterizations efficiently reproduce the observed time evolution of PNSDs and captures NPF events, such as those on 01/20, 01/21, 01/30, and 01/31, which are characterized by the burst of aerosol number concentrations in nanometer-sized range. Simulations using ACDC\_DB\_BC and ACDC\_DB\_CN result in lower particle concentrations in the low size range (1-10 nm) during the NPF period compared to three main parameterizations and the observations, while simulations with 555 ACDC RM show higher concentrations. This is consistent with the comparison of  $J_{1,4}$  among different parameterizations and further evident by the comparison of averaged PNSDs in Figure 5C. Notably, when compared to observations, all parameterizations consistently underestimate the averaged PNSDs within the 2-10 nm range but overestimate them in the 10-50 nm range. This discrepancy may stem from simplified assumptions in particle growth simulation, as discussed in our previous study (Li et al., 2023c).

- 
- 



 **Figure 5.** Comparison of simulated particle formation rates and particle number size distributions (PNSDs) with observations during January 13, 2019, to January 31, 2019, in Beijing. A represents the averaged particle formation rates during the period, the blue bars and orange bars represent observations and simulations, respectively, while the 569 blue dashed line represents the observed values. Daily maximum values of  $J_{1,4}$  are used following Deng et al. (2020); B for the time series of PNSDs; and C for the averaged PNSDs.

 The results show the applicability of all three main parameterizations in NPF modeling during wintertime periods. Importantly, the reliability of the new ACDC- derived parameterization based on the latest theoretical approach (ACDC\_DB) without simplifications in 3-D NPF simulation, is affirmed. The differences among various parameterizations can be explained by the comprehensive box-model simulations above at corresponding conditions. Compared to ACDC\_DB, the *J*1.4 and PNSDs 579 simulated by other two main parameterizations (Dynamic Sim and ACDC RM SF0.5)  agree similarly with observations, but for different reasons. In the case of Dynamic\_Sim, the simplification in cluster evaporations has minimal impact on NPF simulation since CS is the dominant sink for clusters under the wintertime conditions (averaged *T* and 583 CS is 274.7 K and  $3.3 \times 10^{-2}$  s<sup>-1</sup>, respectively). However, the simplifications in boundary 584 conditions and cluster number lead to the underestimation of the  $J_{1,4}$ , consequently lowering the simulated particle number concentrations in 1-100 nm size range due to the ignorance of clusters contributing to growth. As a result, the agreement of Dynamic\_sim to observations should result from a combination of underestimation due to simplifications in boundary conditions and cluster number, along with the 589 compensatory effect of the overestimation caused by lower  $\Delta G$  for  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster. For another main parameterization ACDC\_RM\_SF0.5, since the test parameterization ACDC\_RM considerably overestimates *J*1.4 and PNSDs compared to the observations, the general agreement between ACDC\_RM\_SF0.5 and observations should be attributed to a balance between reduced kinetic limit through the application of SF and the compensatory effect of the overestimation caused by inappropriate representation of cluster thermodynamics.

## **3.2.2 Summertime Simulations**

 Figure 6 provides additional insight into the performance of various parameterizations in NPF simulation during summer. It can be noted that there exists a significant difference in particle formation rates between winter and summer in Beijing. As shown in Figure 6 and Figure S12B, ACDC\_DB and Dynamic\_Sim continues to demonstrate consistent and effective performance in simulating *J*1.4 (within a factor of 2), PNSDs evolution as well as NPF events. However, distinct differences emerge in the NPF simulation for other parameterizations, including another main parameterization ACDC\_RM\_SF0.5. Specifically, in contrast to the good performance of ACDC\_DB and Dynamic\_Sim, ACDC\_RM\_SF0.5, along with the test case 606 ACDC RM, exhibits a significant overestimation of  $J_{1,4}$ , exceeding the observations by more than 15 times and over two orders of magnitude, respectively. This aligns with their overestimation of NPF occurrences and particle number concentration in the size range of 1-100 nm in comparison to observation, with a more pronounced overestimation for ACDC\_RM. Conversely, the test cases of ACDC\_DB\_BC and 611 ACDC DB CN show an underestimation of averaged  $J_{1,4}$  by approximately 4-5 times. They almost fail to depict NPF events, resulting in a significant underestimation of number concentrations in the 1-100 nm size range. Simulations using ACDC\_DB\_CE 614 notably overestimates  $J_{1,4}$  especially on  $08/28 - 08/31$  (Figure S11B), which results in 615 an overestimation of averaged  $J_{1,4}$  by approximately 4 times compared to the observations. However, apart from a moderate overestimation in the initial particle size, we can observe a closer alignment of particle number concentrations in the 2-50 nm range with observations for ACDC\_DB\_CE, which should result from a combination of surplus newly formed particles and fast particle growth from inadequate assumptions within the model. For the broader 2-100 nm size range, it can be observed that ACDC\_DB and Dynamic\_Sim are closer to the observations compared to 622 ACDC DB CE and another major parameterization ACDC RM SF0.5 (Figure S13). The latter two overestimate the average number concentrations during the simulation period by 1.6 times and 2.5 times, respectively. Given the more accurate representation of nucleation rates by ACDC\_DB and Dynamic\_Sim, the discrepancies in the 2-100 nm size range compared to the observed PNSDs should also stem from the simplified assumptions in particle growth simulations.



 **Figure 6.** Comparison of simulated particle formation rates and particle number size distributions (PNSDs) with observations during August 18, 2019, to August 31, 2019, in Beijing. A represents the averaged particle formation rates during the period, the blue bars and orange bars represent observations and simulations, respectively, while the 634 blue dashed line represents the observed values. Daily maximum values of  $J_1$  are used following Deng et al. (2020); B for the time series of PNSDs; and C for the averaged PNSDs.

 Most previous NPF studies combining experiments/observations with simulations are conducted under conditions biased towards winter (~280K) (Almeida et al., 2013; Lu et al., 2020). Under summer conditions with elevated *T*, there exists a deficiency in parameterization evaluations for simulating NPF. The 3-D simulation results during the summer period provide additional validation for the reliability of ACDC\_DB. For ACDC\_RM\_SF0.5, evidence from both box-model simulations and 3-D simulations suggests that it can accurately reproduce real SA-DMA nucleation at temperatures around 280 K, while it has limitations in higher temperatures. Another main parameterization Dynamic\_Sim consistently demonstrates good performance in NPF

 simulation, akin to its efficacy in winter conditions. With the increased temperature in summer (averaged *T* is 298.2 K), the influence of simplifications in cluster evaporations, cluster number, and boundary conditions becomes more profound, mirroring the trends observed in box-model simulations above. This leads to more significant overestimation for ACDC\_DB\_CE, and underestimation for ACDC\_DB\_CN and ACDC\_DB\_BC compared to the observation as well as the base-case ACDC\_DB. Note 653 that CS during the summer period (averaged CS is  $2.8 \times 10^{-2}$  s<sup>-1</sup>) decreases compared to winter but remains significantly higher than typical values in clean regions  $(\sim 3.0 \times 10^{-3}$  s<sup>-1</sup>) (Dal Maso et al., 2008). According to the limited conditions for Dynamic Sim 656 described above, although the overestimation of  $J_{1,4}$  prediction resulting from the simplification in cluster evaporations is more pronounced in summer compared to that in winter, impacts from diverse overestimations and underestimations from different 659 simplifications and varied thermodynamics for  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> cluster can still offset each other, thereby allowing Dynamic\_Sim to match observations. Based on previous comparisons using box-models, significant differences in *J*1.4 predictions between Dynamic\_Sim and ACDC\_DB only exist under conditions of high *T* > ~300 K and low  $CS < \sim 3 \times 10^{-3}$  s<sup>-1</sup>, thus similar performance of Dynamic\_Sim and ACDC\_DB can be 664 expected in the polluted atmosphere  $(CS > -1.0 \times 10^{-2} \text{ s}^{-1})$ . In clean atmosphere with high temperature, however, caution is advised when using Dynamic\_Sim for 3-D NPF simulations.

## **4. CONCLUSIONS and DISCUSSIONS**

 By integrating box modeling, 3-D simulations, also under the constraint from in situ measurements, this study conducts comprehensive comparison of different cluster dynamics-based parameterizations for SA-DMA nucleation. Among them, the ACDC- derived parameterization grounded in the latest molecular-level understanding and complete representation of cluster dynamics (ACDC\_DB), is identified to effectively model particle formation rates and PNSDs evolution in both winter and summer in Beijing within 3-D simulations. While a previously proposed simplified cluster dynamics-based parameterization (Dynamic\_Sim) performs comparably in modeling NPF in Beijing, analysis reveals that their similarity arises from a delicate balance between overestimation and underestimation due to simplifications in cluster dynamics processes and the difference in thermodynamics of initial cluster. Particularly, under 679 specific conditions of high temperature  $(> \sim 300 \text{ K})$  and low CS  $(< \sim 3 \times 10^{-3} \text{ s}^{-1})$ , Dynamic\_Sim tends to make significant overestimation of particle formation rates compared to the reality. Moreover, the study furnishes evidence that integrating ACDC- derived parameterizations with the traditional theoretical approach RI-CC2/aug-cc-683 pV(T+d)Z//M06-2X/6-311++G(3df,3pd) (ACDC RM SF0.5) effectively captures particle formation rates and the evolution of PNSDs around 280 K, a temperature range frequently explored in prior experiments and simulations investigating NPF (Kirkby et al., 2011; Almeida et al., 2013; Kirkby et al., 2016; Xie et al., 2017; He et al., 2021; Ma et al., 2019). Therefore, ACDC\_RM\_SF0.5 exhibits consistent applicability as other two parameterizations at around ~280 K. However, attributed to an inappropriate  representation of cluster thermodynamics, ACDC\_RM\_SF0.5 has limitations in predicting particle formation rates at elevated temperatures. Overall, considering all aspects, we recommend ACDC\_DB as a more reliable parameterization for simulating NPF across various atmospheric environments.

 In addition to contributing to a more reasonable 3-D modeling of NPF, our research further provides valuable references for the development of parameterizations for other nucleation systems. Firstly, we demonstrate the efficacy of the DLPNO-CCSD(T)/aug-696 cc-pVTZ// $\omega$ B97X-D/6-311++G(3df,3pd) level of theory in describing the thermodynamic properties of SA-DMA clusters through comprehensive evidence. This approach can thus be referenced when using quantum chemical calculations to obtain thermodynamic data for other nucleation clusters, especially for other alkylamines such as methylamine/trimethylamine-sulfuric acid clusters. Although DLPNO method still has uncertainties in accurately describing cluster thermodynamics (Besel et al., 2020), it is well recognized as the best available method currently (Elm et al., 2020). Besides, in some qualitative studies, e.g., comparing the enhancing potential or synergistic effects of different precursors in SA-driven nucleation, methods other than DLPNO- CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311++G(3df,3pd), such as RI-CC2/aug-cc-706 pV(T+d)Z//M06-2X/6-311++G(3df,3pd), are equally valid (Liu et al., 2019).

 Comprehensive modeling evidences are provided in this study that certain simplifications or assumptions in cluster dynamics, such as reducing the number of expected clusters, modifying boundary conditions, and assuming certain clusters to be non-evaporative, can significantly impact the prediction of particle formation rates and hence alter the 3-D NPF simulation under certain conditions. While applying certain simplifications concurrently under specific ambient conditions can offset different influences against each other, leading to a satisfactory model-observation comparison, there is a risk that certain simplifications may drive the model's outcomes away from reality when environmental conditions change. Therefore, caution should be exercised when applying these simplifications in derivation of nucleation parameterizations and subsequent application in 3-D models. In addition to the simplifications within the cluster dynamics regime, it should be noted that current standard treatments in 3-D models that ignore detailed gas-cluster-aerosol interactions may also lead to biases under certain atmospheric conditions (Olenius and Roldin, 2022). This applies not only to parameterizations involving explicit mathematical expressions but also to those using ACDC-derived look-up tables. Additional evaluations for the SA-DMA system indicate that the impacts of these treatments may be highest under a combination of low 724 temperature  $(< 270 \text{ K})$ , low CS  $(< 0.003 \text{ s}^{-1})$ , and low precursor concentrations, which leads to elevated time to reach steady state and a higher proportion of precursor consumption from cluster formation, as also indicated by Olenius and Roldin's study (Olenius and Roldin, 2022). Despite these impacts being generally limited under most atmospheric conditions in our modeling scenarios (see supporting information), further research, especially using computationally lightweight models, should aim to  circumvent the potential bias by linking the cluster and aerosol dynamics (Olenius and Roldin, 2022).

 It is recognized that the development of cluster dynamics-based nucleation parameterizations in the form of explicit mathematical expressions is subject to limitations, especially for systems involving multiple precursor species (Semeniuk and Dastoor, 2018). Given that the original ACDC has been extended to involve more than two precursor species, the ACDC-derived parameterization framework, in the form of a look-up table, is highly meaningful for establishing parameterizations for these multi- component nucleation systems. Given that multiple nucleation pathways may be simultaneously considered and simulated in 3-D modeling through ACDC-derived look-up tables, automized incorporation of tables are needed through useful tools such as J-GAIN developed recently (Yazgi and Olenius, 2023).

- **Appendix.** Abbreviations used in the main text.
- 
- **SA:** sulfuric acid
- **DMA:** dimethylamine
- **ACDC:** Atmospheric Cluster Dynamic Code
- **DB**: DLPNO-CCSD(T)/aug-cc-pVTZ//ωB97X-D/6-311++G(3df,3pd) level of theory
- **RM**: RI-CC2/aug-cc-pV(T+d)Z//M06-2X/6–311++G(3df,3pd) level of theory
- **CE:** simplification in cluster evaporations (only (SA)*k*(DMA)*<sup>k</sup>* (*k* = 1-4) and
- 750  $(SA)<sub>2</sub>(DMA)<sub>1</sub> clusters are considered)$
- 751 **CN**: simplification in cluster number (clusters larger than  $(SA)$ <sub>1</sub>(DMA)<sub>1</sub> are regarded
- stable with no evaporation)
- **BC:** simplification in boundary conditions ((SA)4(DMA)<sup>4</sup> cluster is set as the only
- terminal cluster in calculating particle formation rates)
- **SF**: sticking factor used in collision process
- **Dynamic\_Sim:** a reported cluster-dynamic based parameterization incorporating
- simplifications of CE, CN and BC.
- *J***1.4**: particle formation rate at 1.4 nm
- 759 **R**: a parameter to quantify the differences in simulating  $J_{1,4}$  among different cluster
- dynamics-based parameterizations compared to the base-case ACDC\_DB

 **Code and data availability.** The data and code used in this study are available upon request from the corresponding author.

 **Author contributions**. JS, BZ, and SW designed the research; AN and XZ collected the quantum chemistry calculation data; JS performed the ACDC and WRF- Chem/R2D-VBS simulations; YL, RC, and JJ collected the observational data. JS, BZ, and SW analyzed the data; RC, DG, JJ, YG, MS, BC, and HH presented important suggestions for the paper; JS, BZ, and SW wrote the paper with input from all co-authors.

- **Competing interests.** At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.
- **Acknowledgements.** This study was supported by the National Natural Science 775 Foundation of China (22188102 and 42275110) and Samsung PM<sub>2.5</sub> SRP.
- 
- **Financial support.** Financial support from National Natural Science Foundation of
- China (22188102 and 42275110) and Samsung PM2.5 SRP.

# **REFERENCES**





 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, Journal of Aerosol Science, 149, 10.1016/j.jaerosci.2020.105621, 2020. Gao, D., Zhao, B., Wang, S. X., Shen, J. W., Wang, Y., Zhou, C., Jiang, J. K., Wu, Q. R., Li, S. Y., Sun, Y. S., He, Y. C., Zhu, Y., and Jiang, Z.: Distinct PM2.5- Related Near-Term Climate Penalties Induced by Different Clean Air Measures in China, Geophysical Research Letters, 51, 10.1029/2024gl108204, 2024. Gordon, H., Sengupta, K., Rap, A., Duplissy, J., Frege, C., Williamson, C., Heinritzi, M., Simon, M., Yan, C., Almeida, J., Trostl, J., Nieminen, T., Ortega, I. K., Wagner, R., Dunne, E. M., Adamov, A., Amorim, A., Bernhammer, A. K., Bianchi, F., Breitenlechner, M., Brilke, S., Chen, X., Craven, J. S., Dias, A., Ehrhart, S., Fischer, L., Flagan, R. C., Franchin, A., Fuchs, C., Guida, R., Hakala, J., Hoyle, C. R., Jokinen, T., Junninen, H., Kangasluoma, J., Kim, J., Kirkby, J., Krapf, M., Kurten, A., Laaksonen, A., Lehtipalo, K., Makhmutov, V., Mathot, S., Molteni, U., Monks, S. A., Onnela, A., Perakyla, O., Piel, F., Petaja, T., Praplan, A. P., Pringle, K. J., Richards, N. A., Rissanen, M. P., Rondo, L., Sarnela, N., Schobesberger, S., Scott, C. E., Seinfeld, J. H., Sharma, S., Sipila, M., Steiner, G., Stozhkov, Y., Stratmann, F., Tome, 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., Curtius, J., and Carslaw, K. S.: Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation, Proc Natl Acad Sci U S A, 113, 12053- 12058, 10.1073/pnas.1602360113, 2016. Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181- 3210, 10.5194/acp-6-3181-2006, 2006. He, X. C., Tham, Y. J., Dada, L., Wang, M., Finkenzeller, H., Stolzenburg, D., Iyer, S., Simon, M., Kurten, A., Shen, J., Rorup, B., Rissanen, M., Schobesberger, S., Baalbaki, R., Wang, D. S., Koenig, T. K., Jokinen, T., Sarnela, N., Beck, L. J., Almeida, J., Amanatidis, S., Amorim, A., Ataei, F., Baccarini, A., Bertozzi, B., Bianchi, F., Brilke, S., Caudillo, L., Chen, D., Chiu, R., Chu, B., Dias, A., Ding, A., Dommen, J., Duplissy, J., El Haddad, I., Gonzalez Carracedo, L., Granzin, M., Hansel, A., Heinritzi, M., Hofbauer, V., Junninen, H., Kangasluoma, J., Kemppainen, D., Kim, C., Kong, W., Krechmer, J. E., Kvashin, A., Laitinen, T., Lamkaddam, H., Lee, C. P., Lehtipalo, K., Leiminger, M., Li, Z., Makhmutov, V., Manninen, H. E., Marie, G., Marten, R., Mathot, S., Mauldin, R. L., Mentler, B., Mohler, O., Muller, T., Nie, W., Onnela, A., Petaja, T., Pfeifer, J., Philippov, M., Ranjithkumar, A., Saiz-Lopez,











