Master's Thesis
Master's Programme in Atmospheric Sciences
Aerosol Physics

Atmospherically Relevant Chemistry and Aerosol Box Model – ARCA box

Petri Clusius
2020

Ohjaaja: Michael Boy
Tarkastaja: Hanna Vehkamäki

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE
PL 64 (Gustaf Hällströmin katu 2a)
00014 Helsingin yliopisto

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree programme
Matemaattis-luonnontieteellinen tiedekunta / Faculty of Science	Ilmakehätieteiden maisteriohjelma / Master's Programme in Atmospheric Sciences

Opintosuunta – Studieinrikting – Study track
Aerosolifysiikka / Aerosol physics

Tekijä – Författare – Author
Petri Sebastian Clusius

Työn nimi – Arbetets titel – Title
Atmospherically Relevant Chemistry and Aerosol Box Model – ARCA box

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Maisterin tutkielma / Master's Thesis	Marraskuu 2020 / November 2020	66

Tiivistelmä – Referat – Abstract

Työssä esitellään 0-ulottuvuuksinen laskennallinen malli Atmospherically Relevant Chemistry and Aerosol Box Model (ARCA box), joka soveltuu ilmakehän kemiallisten reaktioiden, molekulääristen klustereiden sekä aerosolihiukkasten synnyn ja kasvun simuloimiseen. Mallin käyttökohteita ovat muun muassa ulkoilman hivenkaasupitoisuuksien määrittäminen esiastekomponenteista, kaasureaktiokammiossa tehtävien kokeiden suunnittelu ja mallintaminen tai sisäilmanlaadun tutkimus. ARCA:ssa mallinnettujen kemiallisten reaktioiden peruskirjastona toimii Master Chemical Mechanism MCM, ja sitä on olennaisesti täydennetty monoterpeenien autoksidaatioreaktioilla (PRAM-mekanismi). Molekylääristen klustereiden simulaatiossa malli käyttää Atmospheric Cluster Dynamics Codea (ACDC). Aerosolien kokojakaumaa kuvataan mallissa kahdella vaihtoehtoisella diskreetillä mallilla, joiden hilaväli ja tiheys ovat käyttäjän valittavissa. Kokojakauman muutosta kaasumaisten hiilivetyjen tiivistymisen myötä, sekä aerosolien Brownin liikkeen johdosta tapahtuvaa koagulaatiota mallinnetaan käyttäen vakiintuneita kineettisen kaasuteorian ja termodynamiikan yhtälöitä.

ARCA:n käyttöliittymä poikkeaa huomattavasti aiemmista vastaavista malleista. Mallia pystyy käyttämään joustavasti graafisen käyttöliittymän avulla, helpottaen sekä simulaatioiden suunnittelua ja toistettavuutta. Käyttöliittymän myötä ARCA:lla on potentiaalia myös aerosolimallinnukseen erikoistuneiden ryhmien ulkopuolella, kuten kokeellisessa ilmakehätutkimuksessa tai viranomaiskäytössä.

This thesis presents the Atmospherically Relevant Chemistry and Aerosol Box Model (ARCA box), which is used for simulating atmospheric chemistry and the time evolution of aerosol particles and the formation of stable molecular clusters. The model can be used for example in solving of the concentrations of atmospheric trace gases formed from some predefined precursors, simulation and design of smog chamber experiments or indoor air quality estimation. The backbone of ARCAs chemical library comes from Master Chemical Mechanism (MCM), extended with Peroxy Radical Autoxidation Mechanism (PRAM), and is further extendable with any new reactions. Molecular clustering is simulated with the Atmospheric Cluster Dynamics Code (ACDC). The particle size distribution is represented with two alternative methods whose size and grid density are fully configurable. The evolution of the particle size distribution due to the condensation of low volatile organic vapours and the Brownian coagulation is simulated using established kinetic and thermodynamic theories.

The user interface of ARCA differs considerably from the previous comparable models. The model has a graphical user interface which improves its usability and repeatability of the simulations. The user interface increases the potential of ARCA being used also outside the modelling community, for example in the experimental atmospheric sciences or by authorities.

Avainsanat – Nyckelord – Keywords
ARCA box, ACDC, PRAM, aerosol physics, atmospheric modelling, new particle formation, HOM, SOA

Säilytyspaikka – Förvaringställe – Where deposited
E-thesis, Helsingin yliopisto / University of Helsinki

Muita tietoja – Övriga uppgifter – Additional information

Contents

1 Introduction

We are facing one of the most difficult problems we have ever had to tackle during the existence of *Homo sapiens*, certainly the most difficult during the historic era: human-inflicted climate change is going to change the living conditions on this planet fast, and there is overwhelming scientific evidence to show that the change is well underway.

The big picture seems clear; we need to cut down CO_2 emissions. However, this seems to be politically and economically quite difficult, and despite all the goodwill even modest cuts in greenhouse gas emissions seem to be extremely difficult to achieve. The scientific community must answer similar questions again and again: How much CO_2 can we live with? What options do we have? What will the world be like in 50 or 100 or 3000 years? To answer these questions, we can only turn to scientific theories and models derived from those theories.

Computational climate models have come a long way since Lewis Richardson published his "Weather prediction by numerical process" (Richardson 1922). The complexity and chaotic nature of the Earth's climate assures that there is always room for improvement in the models. Physical circulation of air and water aside, there is still some work left in properly describing the interchange of gaseous and particulate matter between land, sea and atmosphere. Earth receives practically all its energy from the Sun's radiation, and this radiation travels through the atmosphere on its way in and out. Gas composition, properties and distribution of cloud droplets and other suspended particles all have an impact on the radiative balance, and accurately modelling them is especially important when considering the large uncertainty involved in the effect of anthropogenic aerosols to global warming (IPCC 2009).

Central part in understanding the atmosphere is to know its chemistry. The flow of air masses, heat and water can be solved using the Navier-Stokes equations. But to scale this approach to a global model is well beyond the resources and power of any computer, and simplifications must be made. Same goes for atmospheric chemistry; there are tens of thousands chemical reactions currently known to happen in the atmosphere, and to model such set of equations is simply not feasible, or even necessary, if the insignificant reactions can be left out. Therefore, the essential processes need to be found in order to approximate the system in

a representative manner that can be scaled to a regional or global scale. One way to achieve this is to increase the level of detail with the expense of dimensions, use this model in small scale and derive simplifications from the results. This is where our Atmospherically Relevant Chemistry and Aerosol Box model (ARCA box) enters the stage.

A detailed atmospheric box model has many uses. In the urban environment, air quality is a constant subject of monitoring, and tools for modelling the air contents are invaluable, not only outdoors but indoors as well. Covid-19 pandemic has highlighted the necessity of understanding of air quality, aerosols and their physical and chemical properties.

In this work we present a box model, which combines the current knowledge of relevant chemical reactions and aerosol formation.

1.1 History of the ARCA box model

Dr. Michael Boy, the co-leader of the Multi-Scale Modelling group and the supervisor of this work, has during the last decade received requests on a monthly basis from international groups for access to the models created inside the group. Most of the requests were directed towards the chemistry-aerosol model MALTE-box, which was developed more than 15 years ago. The model was sent to different groups, but it soon became apparent that the complexity of the model and the minimal documentation caused many problems to their users. The strategy was therefore changed, and a short visit (1–2 months) by the students or postdocs to the group was required before handing over the model. In this way more than 15 young scientists have visited the group in the last decade, learned to use the model and apply it to different atmospheric research topics.

Approximately two years ago the group discussed the necessity and benefit of creating a new, easy-to-use atmospherically relevant chemistry and aerosol box model. The model should be based on the knowledge gained of the three existing models which are used in the group: MALTE-box (e.g. Xavier et al. 2019), ADCHAM (Roldin et al. 2014) and ADiC-model (Pichelstorfer and Hofmann 2015). It should contain "the best" science from all these three models but additionally have an easy-to-use interface, so that also researchers who are less involved in modelling would be able to apply it. The model design was started from scratch so that unnecessary and in some cases obsolete code, theory or tools could be

minimized. The main strategy from the beginning was to make the new model an open-source code after its release and in this way improve the ARCA box model continuously with new research results and with feedbacks from the users.

1.2 Objectives of the ARCA box

ARCA box was written with usability and extendibility in mind. A scientific research model requires constant development and improvement. Care has been taken in order to make this work as easy as possible, by writing a code which is as self-explanatory, commented and documented in the manual. The need to customize experiment settings with external user input means that hard coded options and variables were avoided except for the most established instances, like natural constants.

Setting every option outside the program soon amounts to a massive number of options and their combinations. This will create a very steep learning curve for using the model, and therefore a graphical user interface (GUI) has been written for the model. While GUIs have been the standard way to make usable programs since the 80's, they are still not common in the atmospheric modelling field.

The scientific objective of ARCA was to gather the established theory on aerosol chemistry and physics into one package. The level of detail of the theories applied within the model depend on the size of the system. ARCA contains ACDC (McGrath et al. 2012, Olenius et al. 2013), which solves the formation of stable molecular clusters using data obtained by quantum-chemical calculations, but on the other hand the aerosol dynamics for hundreds or thousands of compounds is solved using more general kinetic and thermodynamic approaches.

In Chapter 2 the theory behind the mechanisms of ARCA box is introduced. Chapter 3 covers the installation, and configuration of the model. Some variables, functions or modules will be mentioned already in Chapter 2 without detailed introduction, in which case a reference to the appropriate section in Chapter 3 is provided. In Chapter 4 ARCA box is used in a case study, which estimates the necessary ammonia and amine concentrations to explain new particle formation in Hyytiälä during a few event days.

2 Theory of the ARCA box

2.1 Chemistry

2.1.1 Tropospheric chemistry

The Earth's atmosphere is an oxidizing environment, and consequently vast majority of the important reactions in the lower troposphere are connected to oxidation of chemical species. The main ingredients of air, N_2, O_2, Ar and CO_2 are chemically very passive, and solar radiation in the lower troposphere is not energetic enough to break these molecules. The reactions are mostly driven by atmospheric oxidants such as hydroxyl and nitrate radicals (OH and NO_3) and ozone O_3. Higher up in the atmosphere the energy of sun's radiation increases, leading to increased photodissociation. The vertical mixing time of the troposphere is in the order of 1–2 days, and in the stratosphere vertical mixing takes months (Mohanakumar 2008) whereas the lifetime of the emitted volatile organic compounds range from minutes and hours (terpenes and isoprene), days and weeks (acetone, methanol, propane) to years, decades and centuries (methyl chloroform, HCFC, CFC) (Williams 2004). Therefore higher altitudes have decreasing number of chemical species, although quantitative comparisons are hard to find ("tens of thousands of organic compounds" in the boundary layer in Williams (2004), and "hundreds of different gases" in the stratosphere (Mohanakumar 2008).

From modelling point of view, the complexity of the tropospheric chemistry is a challenge. Even when the insignificant reaction pathways (either too slow or too unlikely reactions) are left out as an approximation, the number of reactions is many thousands. More pathways are constantly discovered, some of which may have far reaching consequences. Example of such pathways are the formation of stabilized Criegee Intermediates (sCI) which could affect the OH and H_2SO_2 concentrations (Taatjes 2017), and the autoxidation of organic molecules via hydrogen shift (Crounse et al. 2013, Ehn et al. 2014) which leads to formation of highly oxygenated organic molecules (HOMs).

Fundamentally, the complexity of atmospheric chemistry is due to life on earth; *flora* and *fauna* in land and sea emit a myriad of (mainly organic) compounds in the air due to their metabolism or as ways of coping with the environmental conditions. Many of the emissions are known, such as isoprene, mono- and sesquiterpenes, methane and many nitrous compounds, but even the most recent models have a large discrepancy compared to observations (e.g. Mogensen et al. 2015, Praplan et al. 2019), and there is still a lot to be discovered, either in the form of emissions or reactions.

2.1.2 Chemical models for atmospheric chemistry

In order to make any reasonable effort in atmospheric modelling, the chemical module needs to consider at least the most relevant reactions. Probably the most important collection of chemical mechanisms in the atmosphere is the Master Chemical Mechanism, developed initially at the University of Leeds (MCM v. 3.3, Jenkin et al. 1997, Saunders et al. 2003). In its current form it considers the degradation reaction chains of 143 organic compounds in roughly 17000 reactions and 6700 resulting species, and forms a backbone to many detailed atmospheric chemistry models, e.g. SOSAA (Boy et al. 2011), MALTE-box (Xavier et al. 2019), ADCHEM (Roldin et al. 2011).

Other important collection of chemical mechanisms is the Peroxy Radical Autoxidation Mechanism PRAM, which simulates the formation of HOMs from monoterpenes (Roldin et al. 2019). Since these reactions are not included in MCM, but lead to relatively heavy molecules, some of which can have very low saturation vapour pressures, PRAM is an invaluable and necessary addition to the chemistry scheme of an aerosol model.

These two sets of chemical mechanisms do not contain all known reactions involving atmospheric gases, they include merely the ones which are relevant from the kinetic point of view, in other words the reactions have a large enough probability to contribute to loss or production of a species. Additionally, as mentioned, many significant reactions are still missing or not incorporated to atmospheric chemistry models. One important, known shortcoming, is the HOM formation from benzene and toluene, which are major constituents of the anthropogenically emitted VOC cocktail in polluted cities.

2.1.3 Reaction kinetics

The atmosphere is always changing. If there ever is a momentary steady state, it is a dynamic one, depending on the concentration of every chemical compound, suspended particles, energy in the form of light and temperature as well as interactions with ocean and land. Reaction rates describe the speed of chemical reactions and enable us to find the rate of change of concentrations of compound x and consequently the momentary concentration [x]. Consider for example compound C which is produced in a second order reaction A + B $\rightarrow$ C + D with reaction rate k_l, then

$$-\frac{\mathrm{d[A]}}{\mathrm{d}t} = -\frac{\mathrm{d[B]}}{\mathrm{d}t} = \frac{\mathrm{d[C]}}{\mathrm{d}t} = \frac{\mathrm{d[D]}}{\mathrm{d}t} = k_1[\mathrm{A}][\mathrm{B}] = R_1. \tag{1}$$

However, [C] will be affected by any reaction where it is involved, so all N reactions producing or consuming [C] need to be combined to find any momentary concentration. This will lead to an ordinary differential equation (ODE)

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \sum_{i}^{N} R_i. \tag{2}$$

This ODE includes other compounds whose concentrations are governed by their respective ODEs, and this leads to a system of coupled ODEs, which can (usually) be numerically solved for all the participating compounds. Due to the large range of magnitudes of the reaction rates and concentrations, and the coupling between different species, the system of chemical ODEs is described as stiff, meaning that the integration time step is dictated by the stability and not the accuracy of the solution. This is something that the model user does not usually have to consider, but which might lead to slowing or stalling of the model if some extreme concentrations are either used as input or produced in the chemistry module.

In ARCA the reaction kinetics is solved using code produced by the Kinetic PreProcessor, introduced in section 3.1.1. Setting up a chemistry scheme is covered in section 3.1.2.

2.1.4 Variables used in the chemistry module

Initially, before the model main loop starts, all time-dependent variables have default values of zero. In the beginning of each time step those vari-

ables that are supplied by the user are updated. The necessary variables for the chemistry module are absolute humidity, temperature, pressure, photolysis rates and the concentrations of the precursors. The concentrations of the reaction products are those calculated in the chemistry module in the previous time step, and are initially zero.

2.1.5 Short wave radiation and actinic flux

The photochemical reactions need light as a driver. The wavelength dependent flux of photons towards a molecule from any direction is called actinic flux (Seinfeld and Pandis 2006). This is a necessary variable for solving the photolysis rate J, calculated at each time step. The measured total short wave radiation [W/m^2], which needs to be given as input, is transformed to wavelength dependent data by multiplying it with a vector of relative intensities of wavelengths between 280–650 nm. The default vector (loaded by the model from "ModelLib/General/swr_distribution.txt") has been optimized for Hyytiälä, and when the model is used in very different environment, for example in a chamber, it has to be updated with the appropriate spectrum. The light reflected from ground upwards is estimated using a constant surface albedo, together with the calculated solar elevation angle at the chosen location and time (for description of how to set up these in the model, see section 3.3.2).

2.1.6 Photolysis rates

With actinic flux the photolysis rates J_i for photochemically active species X_i are calculated using

$$J_i = \int_0^\infty q_{x,i}(\lambda)\sigma_{x,i}(\lambda)I(\lambda)\,\mathrm{d}\lambda \approx \sum_{l=280nm}^{650nm} q_{i,l}\sigma_{i,l}I_l\Delta l, \quad \Delta l = 5\mathrm{nm}, \tag{3}$$

where $q_{x,i}$, $\sigma_{x,i}$ and I are the quantum yield of the reaction, absorption cross section of the molecule, and the actinic flux, all wavelength λ dependent. In the model the integral is discretized to the summation form of eq. (3), valid only when the intensity of wavelength less than 280 nm is sufficiently low to be ignored, like in the boundary layer. On the other hand, light with wavelength above ~650 nm has no significant effect in most photochemical reactions, a notable exception being the photodissociation of NO_3 (Johnston et al. 1996). The wavelength dependent

quantum yields are stored in "ModelLib/General/Photolyse" directory, and come from MCM, except for HO_2NO_2 and N_2O_5 which are from Atkinson et al. (2004).

2.2 Formation of molecular clusters

ARCA has a couple of ways of handling new particle formation (NPF) rates, but the only explicit method is the Atmospheric Cluster Dynamics Code ACDC (McGrath et al. 2012, Olenius et al. 2013), which in ARCA is used to simulate the cluster formation of the two-component systems of sulfuric acid and ammonia, and sulfuric acid and dimethylamine (DMA).

2.2.1 ACDC

ACDC simulates stable cluster formation in a system of n compounds, where the molecules can be combined to form clusters in all possible ways. Assuming that maximum number of molecules of any one type i in the cluster is m_i, the number of possible clusters in the system is $\prod_{1=1}^{n} m_i$, where each cluster type is represented by the most energetically favourable geometry. The formation and stability of the clusters in the system are modelled by calculating the collision and evaporation coefficients, using kinetic gas theory and the Gibbs formation free energies of each cluster, respectively. With these coefficients known, the steady-state concentrations of each cluster type can be solved.

The calculation of evaporation coefficients is done by assuming a detailed balance of mass flux between each cluster pair in the system. This means that between any pair the net flux is zero. Therefore for any cluster pair i and j must be

$$\gamma_{(i+j)\rightarrow i} C_{i+j}^{\mathrm{eq}} = \beta_{i,j} C_i^{\mathrm{eq}} C_j^{\mathrm{eq}}, \tag{4}$$

where $\gamma_{(i+j)\rightarrow i}$ is the evaporation coefficient from (i+j) to i and $\beta_{i,j}$ is the collision coefficient between i and j. The equilibrium constant for clusters i and j can be written as

$$\frac{C_{i+j}^{eq}}{C_i^{eq} C_j^{eq}} = \frac{\exp\left(\frac{-\Delta G}{k_b T}\right)}{c_{ref}}, \tag{5}$$

where k_b is the Boltzmann constant and C^{eq}, ΔG and T are the equilibrium concentrations of different clusters, Gibbs formation free energy and temperature at reference concentration C_{ref}, respectively. Rearranging eq. (5) and inserting to (4), evaporation coefficient γ can be solved. Here the Gibbs formation free energies of the clusters come in use, and a numerical value for the evaporation coefficient can be obtained. The birth-death equation for the clusters can be written, and solved for C_i as a coupled system of ODEs:

$$\begin{aligned}\frac{\mathrm{d}C_i}{\mathrm{d}t} =& \frac{1}{2}\sum_{j=1}^{i-1}\beta_{j,i-j}C_jC_{i-j} + \sum_{j=i+1}^{n}\gamma_{j\to i,j-i}C_{i+j} \\ &- \sum_{j=1}^{n}\beta_{i,j}C_iC_j - \frac{1}{2}\sum_{j=1}^{i-1}\gamma_{i\to j,i-j}C_i - S_i,\end{aligned} \tag{6}$$

where C_i and S_i are the concentrations and sinks of cluster i, and $\beta_{x,y}$ is the collision coefficients of clusters x and y and $\gamma_{x\to y,y-x}$ is the evaporation coefficients of x to y and y-x. Finally, the flux out of the $\{m_i\}$ $i = \{1...n\}$ system is then considered as the formation rate. (McGrath et al. 2012, Olenius et al. 2013)

The molecules in any given cluster can be arranged in large number of ways, affecting its total energy. To find out which geometry is energetically the most favourable, and what is its Gibbs free energy, configurational sampling and quantum chemical methods are used. Initially a large list of possible candidates is narrowed down by filtering out the energetically least favourable configurations, starting with lower level quantum chemical theories, and increasing the theory level when the list of candidates gets shorter. This process is computationally very expensive and in practice limits the size of the systems considered in ACDC. The evaporation coefficients are quite sensitive to the Gibbs free energy values, and this sets high demands for the quantum chemical calculations. The energy input that is used in ARCA's ACDC are found to reproduce the CLOUD experiments with reasonable accuracy (Almeida et al. 2013).

There are some underlying assumptions in ACDC and in the way it is used in ARCA. First, it is assumed that the evaporation coefficients, which are derived in the detailed balance, are still valid at non-equilibrium conditions when there is net cluster flux in the system (the formation rate). Secondly, the size of the system has to be large enough, so that the outgrowing clusters really are stable. This condition is already

satisfied in ARCA for sulfuric acid – ammonia and sulfuric acid – DMA systems in reasonable concentrations, but not necessarily in hot environment or extremely low concentrations. On the other hand, in these cases the formation rates will anyway be negligible, so in practice the assumption does not cause major inaccuracies. Third point to consider relates to monomer concentrations, which are also affected by cluster formation. In ARCA's implementation of ACDC monomer concentrations are fixed at each time step, and the cluster formation will have no effect on them (see also discussion in section 3.3.8→"Run ACDC to steady state"). This is a reasonable assumption when there are much more molecules in the gas phase (monomers) than there are in the clusters. However, this might not be the case when the monomer concentrations of the clustering compounds differ by many orders of magnitude, and the user should make sure that the underlying condition is still met for all compounds. ARCA prints out the fraction of molecules in the ACDC clusters to monomers, and this should preferably be less than $\sim 10^{-3}$. (Vehkamäki 2006).

Since ACDC itself is not limited to the two systems used in ARCA, it is possible to define new systems and include them to the model, provided that the quantum chemical input data for the system in question exists. In case the user wants to rebuild ACDC Fortran files for example due to updated energy data or other modifications to the systems of H_2SO_4–NH3 or H_2SO_4–DMA, the ACDC Perl code is available in directory "src/ACDC".

2.2.2 Other methods to calculate the NPF rate

There is increasing evidence that new particle formation can also be initiated by clustering of sulfuric acid and organic molecules (Dada et al. 2017, Stolzenburg et al. 2018, Tröstl et al. 2016), although it is not clear if this is the result of homogeneous nucleation of these compounds or rapid condensational growth of formed initial clusters by organic vapours. Unfortunately, the vast variety of organic molecules makes it unpractical to simulate their clustering explicitly in ACDC. If the user wants to model sulfuric acid and organic clustering, a parametrization from Roldin et al. (2019) is included and can be enabled from the options (see section 3.3.8→"*Organic nucleation*"). The nucleating compounds are listed in the file "ModelLib/nucl_homs.txt" and can be edited by the user. Besides this, users might want to define their own mechanisms or implement pre-

viously published ones. The recommended way to do this is to write the function in to the subroutine **AFTER_NUCL**, found in the file "src/custom_functions.f90". There any additional formation rates can be added to the variable **J_TOTAL** (in unit $1/s/m^3$), as long as they represent the formation of the smallest modelled particles: if the formation rate is for some other diameter, it can be converted using the Kerminen-Kulmala equation from Lehtinen et al. (2007), utilizing the coagulation and growth rates from the aerosol module.

2.3 Particle size distribution (PSD) representations in ARCA

A non-uniform (polydisperse, i.e. of varying diameters) particle population can be represented mathematically in discrete or continuous form. In the latter case, the lognormal distribution, shown schematically in Figure 1, is one that at least coarsely fits the atmospheric observations and is thus most often used (Hinds 1990). Here the logarithm of particle sizes are assumed to follow a (Gaussian) normal distribution, and the descriptive variables are the geometric mean diameter, geometric standard deviation and total concentration. Due to its analytical nature, methods using the lognormal particle size distribution (PSD) have been used extensively in aerosol models, and for example many of the current global models use the multimodal M7 lognormal particle size representation (Vignati et al. 2004).

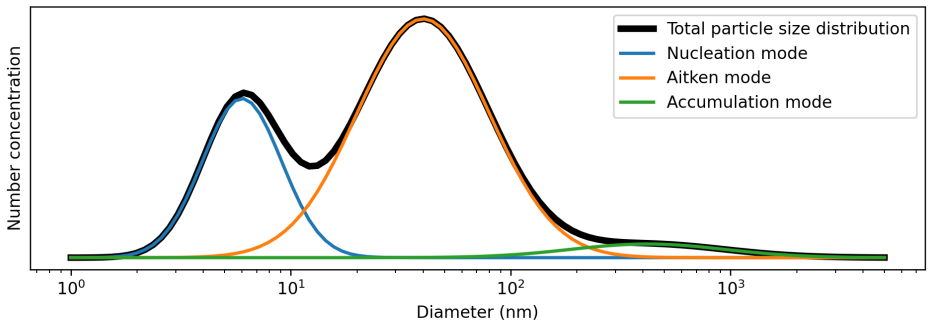

Figure 1: Schematic representation of log-normally distributed particle population (black line) with three modes (coloured lines).

Atmospheric particle distributions generally do not follow analytical functions and are formed of multiple compounds which affect the properties of the distribution such as growth by condensation. Therefore, when computational resources permit, discrete PSD representations are used to model ambient conditions with high fidelity. In principle any aerosol population can be described unambiguously using arrays where each component represents the number of molecules or some other building block of the particle. In atmospheric conditions the particles will be of different size, but also differ in chemical composition and physical properties. To represent such population exactly would be impractical due to computational burden, even if all the physical and chemical properties were known. The PSD is therefore in practice simplified in a computationally efficient way. In ARCA this is done in two alternative ways, the Fully Stationary (FS) or Moving Average (MA) representation. Common for both is that the particles are divided in bins which have a diameter range. Nominally the bins are referred with the diameter of the centre of the bin. Also, it is assumed that all the particles in any particular bin share the same chemical composition. Lastly the particles are assumed to be spherical and have the same bulk density throughout the whole distribution.

When discussing the differences between FS and MA, we are concentrating on what happens when the particles grow by condensation or coagulation, and how to handle the resulting volume change so that total mass of the particles is conserved.

2.3.1 Fully stationary PSD

The first method used in ARCA is the fully stationary method, described for example by Jacobson (1997) and shown schematically in Figure 2. There, the diameter grid stays constant (stationary) throughout the model run. All particles in a bin are always assumed to have the nominal diameter. When mass is changing by condensation or coagulation, the volume of a particle should grow or shrink to some new corresponding diameter, which generally is between two nominal diameters. Therefore the particles in these two adjacent bins are redistributed as follows. Assume additional volume G_i due to growth in bin i with volume V_i. V_{j+1} and V_j are the nominal volumes of bins j and $j+1$, where

$V_j < V_i + G_i < V_{j+1}$. Then the fraction M of the particles in bin i that are moved to the bin $j+1$ is defined by

$$M_{j+1} = \frac{G_j}{V_{j+1} - V_j} \tag{7}$$

and the fraction that is moved to bin j is then $1 - M$. With moderate growth typically $i = j$. After the particles have been distributed, their compositions are averaged within the bin so that all particles in the bin have same composition and matter is conserved.

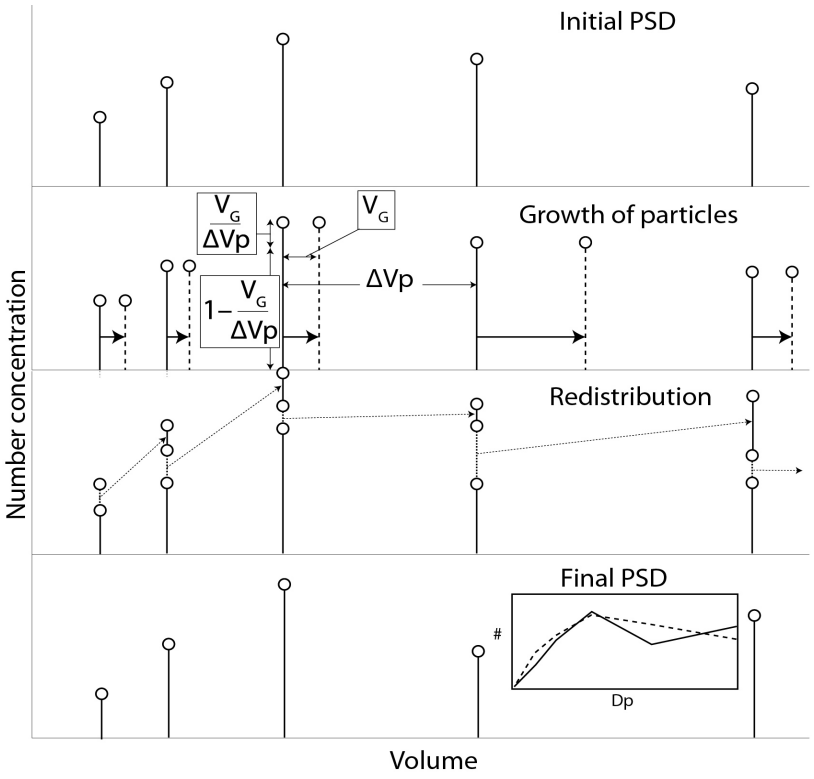

Figure 2: Schematics of the Fully Stationary (FS) particle size distribution (PSD) representation in ARCA. Inset shows initial (dashed line) and final (solid line) PSD. V_G=growth in volume, ΔV_p=nominal particle volume.

Fully stationary structure enables realistic handling of nucleation, transport, emissions and coagulation, but is weaker in describing condensation. This is because in FS growth is not continuous due to the fixed nature of the grid. Eventually this leads to a problem called "numerical diffusion".

It makes the particles spread in the diameter space, where the actual continuous growth would be more confined (Jacobson 1997). The effect of numerical diffusion is more pronounced with coarse diameter grid, but if dense enough grid is used, the problem becomes insignificant. Fully stationary representation is easy to implement especially when horizontal or vertical transport from other grids (in more than zero dimensional models) are considered. Another upside of the FS PSD is the smooth output it produces, making analysis and plotting easier than in the moving average PSD representation, introduced next.

2.3.2 Fixed grid, moving average PSD

Another PSD representation method in ARCA is using the moving average (MA), also described in Jacobson (1997) and depicted in Figure 3. This method or its variation is used in ADiC and ADCHEM models, among others, since it combines the practicality of fixed structure, while minimizing the numerical diffusion. In MA representation, instead of fixed diameter at the bin centre, the bin edges are fixed, but the particle mean diameter is let to continuously vary within the bin, eliminating numerical diffusion at this point. If the particles grow larger or shrink smaller than the bin edges, the whole bin population is moved to the neighbouring bin and redistributed with its current population. The merging of the added particles is done similarly than in the FS method, by averaging the particle compositions weighed by the particle concentration. The new diameter for the merged bin is obtained by calculating the single particle mass from the mass composition and using this mass and density to calculate the volume of the particle, assumed to be a sphere. Even if the averaging of the newly moved and old population is a source for numerical diffusion, in practice the effect is negligible (Jacobson 1997) because redistribution only happens when the particles cross the bin grid. Therefore, the moving average representation can be used with coarser grid and still get accurate results, while saving computational time especially when considering coagulation. Additional saving in computational time is achieved because redistribution is not done if the particles stay inside their current size bin.

In MA representation the varying diameter is used in all calculations involving condensation, coagulation or redistribution, but for practical reasons the diameters that are saved in the output of ARCA are the

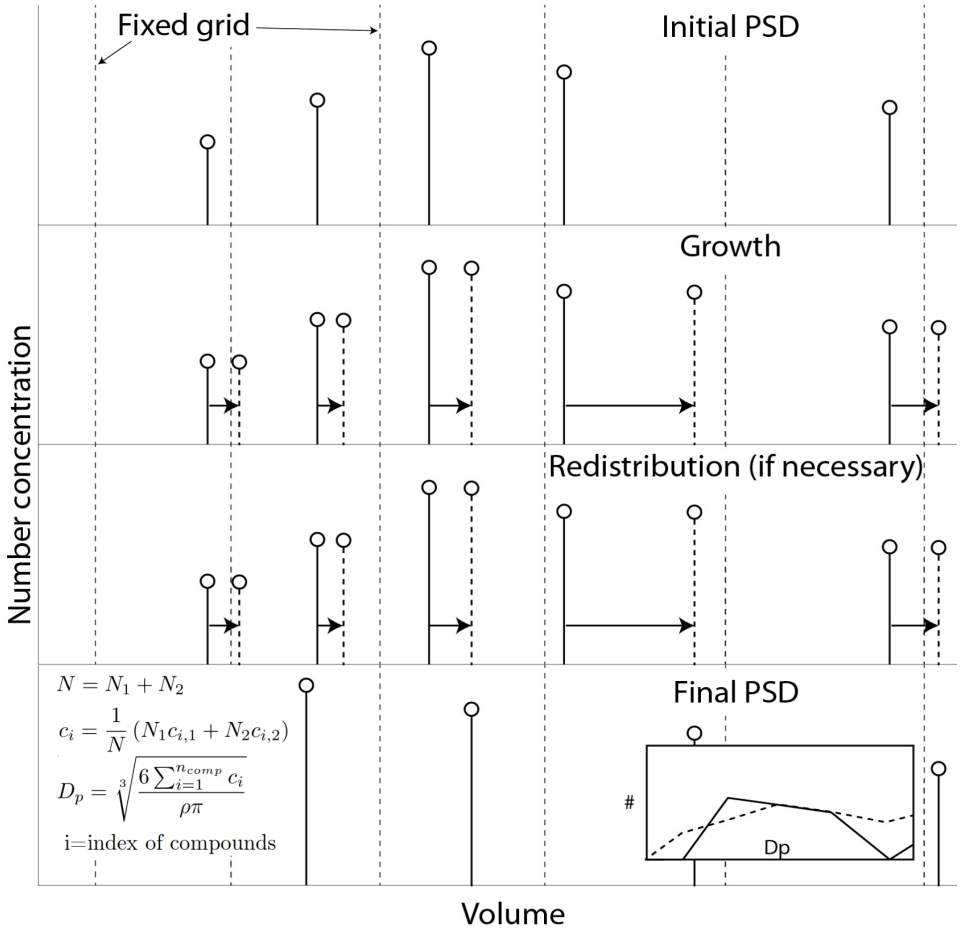

Figure 3: Schematic representation of the Moving Average (MA) particle size distribution (PSD) representation in ARCA. Lowest panel shows the formula for averaging two particle populations, where D_p is the particle diameter, N is particle count, C_i is the mass composition vector, and ρ is bulk density. Inset shows initial (dashed line) and final (solid line) PSD. Example also shows why pits and peaks are forming in the PSD.

nominal diameters, located in the geometric mean of each bin (centre of the bin edges in logarithmic scale). The information of the current (moving) diameter is still saved in the particle composition, similarly as was described earlier.

The output of moving average and fully stationary differ from each other somewhat. The way particle growth is handled in MA representation leads to pits and peaks in the PSD (Pichelstorfer and Hofmann 2015). These result from particles that have grown out of their bin, leaving it empty, thus also markedly increasing the concentration on the next

bin (Figure 4). Consequently the normalized number concentrations can differ substantially from bin to bin, and this is also demonstrated when the apparent maximum number concentration is compared with the FS method. However, as the total sum of particles in a lognormalized particle size distribution is proportional to the total area under the distribution curve, one can see from Figure 4 that the total particle counts are more or less the same in the two PSD representations. The pits and peaks are also clearly visible in the surface plots in Figure 5. Mohs and Bowman (2011) have developed methods to eliminate or deal with the artefacts, and some of these will be implemented in future ARCA versions. Roldin et al. (2011) used a hybrid version of FS and MA, where every 360th time step was done with FS and otherwise MA was used, eliminating both numerical diffusion and pitting.

Future versions of ARCA will also include other PSD representations like Fully Moving PSD and Fixed Grid, Moving Distribution PSD representations.

The model options for the aerosol structure are explained in section 3.3.3.

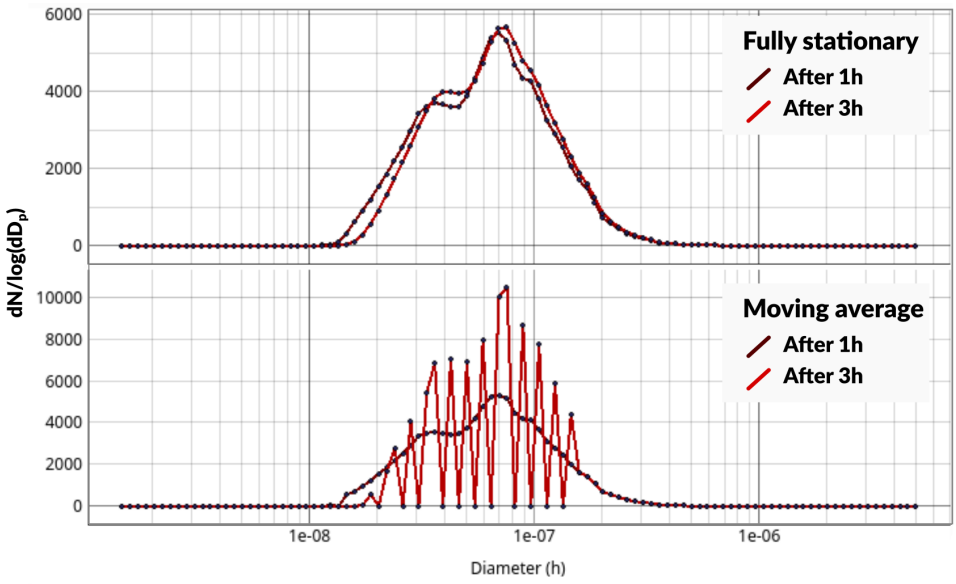

Figure 4: Particle size distribution of two runs using FS (upper) and MA (lower) PSD representation with 100 bins. Both runs used identical initialization and input data from SMEAR II in April 14th 2018. The figure shows the PSD after 1 hour (dark red curve) and 3 hours of run (red curve). The plots are from ARCA GUI with labels added.

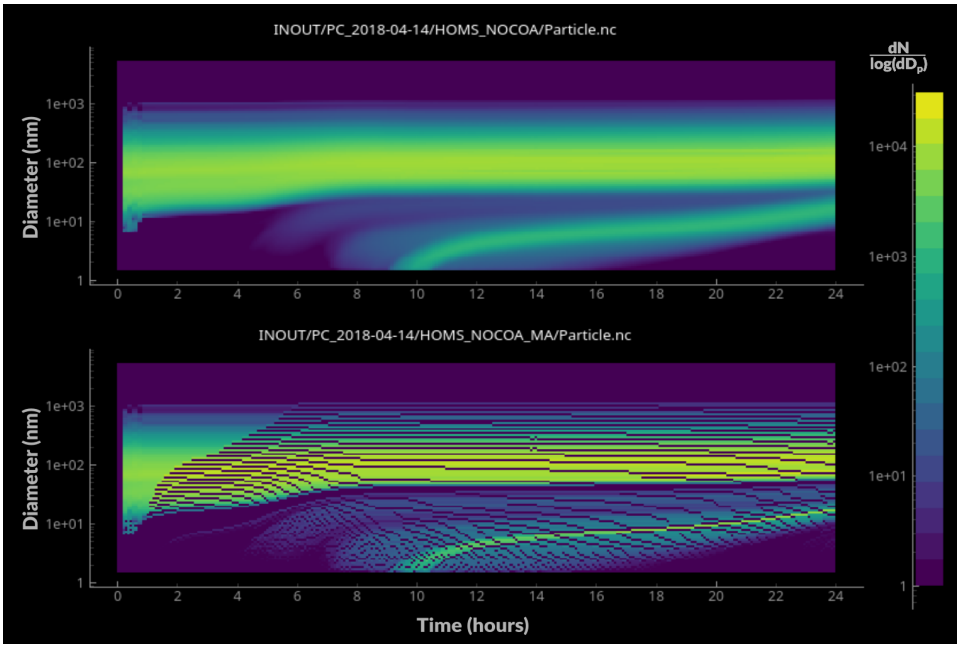

Figure 5: Surface plots of the runs shown in Figure 4. Upper panel shows FS and lower panel MA particle size distribution representation, both with 100 bins, using data from SMEAR II on April 14th, 2018. Coagulation module was turned off for this run. The plot is from ARCA GUI.

2.4 Condensation of vapours

ARCA calculates the condensational growth of particles due to organic vapours and sulfuric acid, defined by their saturation vapour pressure and gaseous concentration. The condensation mechanism in ARCA is based on the Analytical Predictor for Condensation (APC) from Jacobson (2002). There the change in composition $C_{c,i}$ for compound c in particle size bin i is defined as

$$\frac{\mathrm{d}C_{c,i,t+\Delta t}}{\mathrm{d}t} = k_{c,i,t}\left(C_{c,t} - S'_{c,i,t}C_{c,s,i,t}\right), \tag{8}$$

where $C_{c,t}$ is the gas phase concentration of compound c at time t, $S'_{c,i,t}$ is the equilibrium saturation ratio of the condensing gas and $C_{c,s,i,t}$ is the saturation vapour concentration over flat surface for pure compound c. Strictly speaking $S'_{c,i,t}$ should be calculated using the Köhler equation, which combines the Kelvin effect (accounts for the change in saturation vapour pressure due to droplet surface curvature) with the change of sat-

uration vapour pressure due to having a mixture of compounds in the particle droplet. However, in ARCA, Köhler effect is simplified by multiplying the Kelvin term with the molar fraction $x_{c,i}$ of compound c in particle phase:

$$S'_{c,i,t} = x_{c,i}\left(1 + \frac{2\sigma_{c,p}m_c}{R_g r_i T_t \rho_{c,p}}\right), \tag{9}$$

where $\sigma_{c,p}$ is the pure liquid surface tension, m_c is the molar mass and $\rho_{c,p}$ the liquid density of pure compound c, R_g is the universal gas constant, T_t temperature at time t and r_i the radius of particle in bin i.

Equation (8) is integrated over time step Δt (using Euler forward method) to get the change of concentration in the particle phase. Initially the change of concentration can lead to larger condensation than is available in gas phase. To preserve mass, the final concentrations are constrained by mass balance equation

$$C_{c,tot} = C_{c,t,gas} + \sum_{i=1}^{N_b} C_{c,i,t} = C_{c,t+\Delta t,gas} + \sum_{i=1}^{N_b} C_{c,i,t+\Delta t}, \tag{10}$$

where N_b is the number of particle bins. Finally the change in mass composition $\mathrm{d}C_{i,c}$ is calculated, and this is passed to the particle redistribution module described in section 2.3.

There are a number of important simplifications which have to be pointed out when the condensation scheme is considered. First, equations (8) and (9) rely on saturation vapour pressure and pure liquid surface tension of the compound c. Both of these variables have a sound base for a liquid droplet (Vehkamäki 2006), but the the applicability to a mixed phase droplet is not guaranteed. The surface tension of a compound in a droplet varies highly non-linearly depending on the droplet composition (Ekström et al. 2010); however, as in similar models (MALTE-box, ADCHEM, SOSAA), surface tension in ARCA is either constant or varies with temperature, but is the same for all compounds.

The most important factor governing growth in ACP scheme is the saturation vapour pressure for a compound, and there are many methods for deriving these values; Nannoolal et al. (2008) uses the molecular structure whereas the so called Volatility Basis Set (VBS) method uses the elemental composition, (e.g. Donahue et al. 2012). Since the structures of condensing vapours are known in ARCA, the former method is

preferred, but other methods can be used as well. It is important to realize the uncertainties brought in to the model by the large variability in saturation vapour pressure estimates (Kurtén et al. 2016, Peräkylä et al. 2020).

At present state the accumulation of ammonia or nitric acid to particles is not considered in the condensation scheme of ARCA. The saturation vapour pressures (P_{sat}) of these compounds over pure liquid exceed the modelled concentrations by many orders of magnitude (for ammonia, P_{sat} exceeds the atmospheric pressure already at 241 K) so they would not condense on particles in the model (Dean and Lange 1998, Duisman and Stern 1969). However, ammonia and nitric acid react forming ammonium nitrate, which can have impact to particle growth (Wang et al. 2020), and currently this reaction is not included in ARCA. Another missing feature is the dissolution of gaseous compounds to aqueous phase and back, affecting the direct accumulation of nitric acid and ammonia to particles. This mechanism is added in the near future when the particle phase chemistry module is implemented and the pH-values of the particles are calculated.

2.5 Coagulation

The Brownian motion of the molecules in air sets also the aerosol particles to random motion, leading to collisions and agglomeration between particles, in a process called Brownian (or thermal) coagulation. It leads to rapid decrease of the concentration of small particles, and due to the mass accumulation, slow diameter increase of larger particles. Overall, coagulation decreases particle concentration but does not affect total mass. The effect of coagulation to a particle size distribution is shown in Figure 6. In ARCA Brownian coagulation is solved by integrating (using Euler forward method) the discrete coagulation equation from Seinfeld and Pandis (2006):

$$\frac{\mathrm{d}N_k(t)}{\mathrm{d}t} = \frac{1}{2}\sum_{j=1}^{k-1} K_{j,k-j} N_j N_{k-j} - N_k \sum_{j=1}^{\infty} K_{k,j} N_j, \quad k \geq 2, \tag{11}$$

where N_x are the number concentrations and $K_{x,y}$ are the coagulation coefficients between particles in size bins x and y, respectively. Note the similarity between this equation and eq. (6) with the exception of the

summation limit in the second term in eq. (11). In practice the modelled particle distribution has an upper size limit, and to satisfy the validity constraints of the coagulation equation, upper limit of the modelled particle size range should be selected so that the last bin stays empty (or at least with negligible values) throughout the simulated time.

When deriving the coagulation coefficient, three size domains have to be considered. First is the free molecular regime, where the particle diameters are in the range or smaller than the mean free path of air ($\lambda_{\text{air}} \approx$70 nm), and the collisions between air molecules can be thought as separate events. Second is the continuum regime, where the particle sizes are well beyond λ_{air}, and the force due to the collision between the gas molecules and the particle can be thought as continuous drag. Between these is the transition regime. In 1964 N.A. Fuchs proposed a nowadays widely accepted formulation that applies to all three regimes (Seinfeld and Pandis 2006):

$$K_{j,k} = \; 2\pi \left(d_{p,j} + d_{p,k}\right)\left(D_j + D_k\right)\beta, \tag{12}$$

where $d_{p,x}$ are the diameters and D_x the diffusivities of particles x and β is a dimensionless correction factor:

$$\begin{aligned}
\beta &= \left(\frac{d_{p,j} + d_{p,k}}{d_{p,j} + d_{p,k} + 2\sqrt{g_j^2 + g_k^2}} + \frac{\frac{8}{\alpha}\left(D_j + D_k\right)}{\sqrt{\bar{c}_j^2 + \bar{c}_k^2}\left(d_{p,j} + d_{p,k}\right)} \right)^{-1} \\
g_i &= \frac{\left(d_{p,i} + \lambda_i\right)^3 - \sqrt{\left(d_{p,i} + \lambda^2\right)^3}}{3 d_{p,i} \lambda_i} - d_{p,i} \\
\lambda_i &= \frac{8 D_i}{\pi \bar{c}_i}, \bar{c}_i = \sqrt{\frac{8 k_b T}{\pi m_i}}, D_i = \frac{k_b T C_{c,i}}{3\pi\mu d_{p,i}} \\
C_{c,i} &= 1 + \frac{2\lambda_{\text{air}}}{d_{p,i}} \left[1.257 + 0.4 \exp\left(\frac{-1.1 d_{p,i}}{2\lambda_g} \right) \right],
\end{aligned} \tag{13}$$

where $\bar{c}_x$ is the average (thermal) speed of particle x (particles have the same kinetic energy as gas molecules) and g_x is a length factor, C_c is the Cunningham slip correction, λ_i is the mean free path of aerosol and α is the coagulation efficiency, i.e. fraction of the collisions that actually stick together. When β equals 1, the coagulation coefficient reduces to that of the continuum regime. In ARCA the coagulation coefficient is calculated using eq. (13) for all particle sizes, with coagulation efficiency of 1.

The collisions between particles happen also for other reasons than their thermal motion. One notable is the gravitational settling, where differently sized particles fall down with different velocities, resulting in collision and agglomeration. Also any wind condition which creates wind gradients, like turbulence or wind shear, will introduce velocity differences between particles and enhance coagulation. Since these kinds of processes would necessarily require dimensionality to be solved explicitly, they are omitted from a box model, and therefore not considered at all. Brownian coagulation is most dominant for particles of d_p<1 µm. For example, when considering coagulation between 100 nm and 1µm particle, the thermal coagulation is more than 100 times faster than gravitational or shear coagulation, and therefore can safely be ignored. If the model is to be used with supermicron particle sizes, these additional effects should be taken into account.

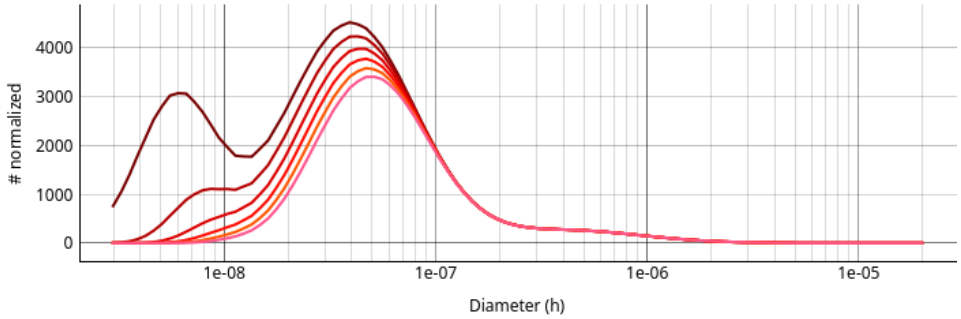

Figure 6: Effect of 10 hrs of coagulation to particle size distribution shown in Figure 1. Each line represents +2 h from previous, the darkest line is t=0. Coagulation quickly scavenges the smallest particles (here in the nucleation mode) but does not significantly affect particles above 100 nm of diameter (accumulation mode). The plot is from ARCA GUI.

2.6 Deposition and wall losses

Particles and chemicals do not stay indefinitely in the atmosphere but are removed by means of wet and dry deposition. Dry deposition to the canopy or ground is affected by vegetation type and meteorological conditions like wind and turbulence. Particles are affected differently based on their size, as largest particles will fall down faster than small particles

(due to their different terminal velocities), and also their trajectories with the air flow are different. Chemical deposition is affected by the solubility and reactivity of the compounds. Wet deposition removes gases and particles in water droplets, either in cloud or during rain. Wet deposition can be ignored in dry conditions, but dry deposition is a slow and continuous process, significantly affecting both gas phase and particle concentrations. Roldin et al. (2019) used a chemical transport model (ADCHEM) and found that inside a pine forest canopy dry deposition of HOMs was comparable in magnitude with their condensation to particles, accounting roughly half of the total losses. When the whole boundary layer was considered, dry deposition accounted for only $\sim$7 % of all HOMs losses. While deposition rate can be calculated in a model that has one or more dimensions, ARCA is a 0-dimensional model and in ambient simulations must resort to user supplied loss rates, which can be uploaded as size-resolved loss rate files (see section 3.3.7), in unit of 1/s. The chemical losses are calculated after chemistry, and aerosol losses are calculated after all other aerosol modules have finished.

In chamber experiments the wall losses of particles and vapours have significant impact to the results and cannot be ignored in model simulations. ARCA uses parametrization for aerosol wall loss similar to ADCHAM (Roldin et al. 2014), which is based on the work of Lai and Nazaroff (2000). There the user supplies dimensions of the chamber, and the model accounts for different deposition efficiencies in the upwards, downward and vertical facing surfaces. The tuning parameter for the deposition efficiency is the friction velocity $u*$, which is related to the turbulent mixing in the chamber.

Whichever method was used to derive the loss rates, the change in particle numbers is calculated with

$$\Delta N_{i,\Delta t} = N_{i,t} \left(1 - \exp(-k_{\mathrm{dep}} \Delta t)\right), \tag{14}$$

where k_{dep} is the loss rate and Δt is model time step. Currently the assumption in the model is that once deposited, the particles are not taking any part to the gas–particle phase partitioning. This is an oversimplification (Roldin et al. 2014), and future versions will address these mechanisms in more detail. Another future consideration for ARCA is the effect of particle charge, which has an effect to wall losses (and other processes).

The wall losses of condensing organic vapours are based on Matsunaga and Ziemann (2010). The parametrization accounts for reversible adsorption and desorption on Teflon coated walls. It calculates the first order rates for both adsorption and desorption of each vapour, and uses 0.01 second integration time to solve the gas phase and adsorbed concentrations. The parametrization is based on an assumption that there is a thin layer of stationary gas adjacent to the wall, followed by a laminar flow layer before the well mixed turbulent chamber. The tuning parameter for the gas phase wall loss is the coefficient of eddy diffusion, which describes the turbulent mixing intensity. Another parameter is the wall accommodation coefficient, which could vary from compound to compound, but is set to constant for all vapours. Effect of wall losses to gases is show in Figure 7, where the sum of concentrations for a set of HOMs is shown with and without wall loss. Configuration of wall losses in the model is introduced in section 3.3.7.

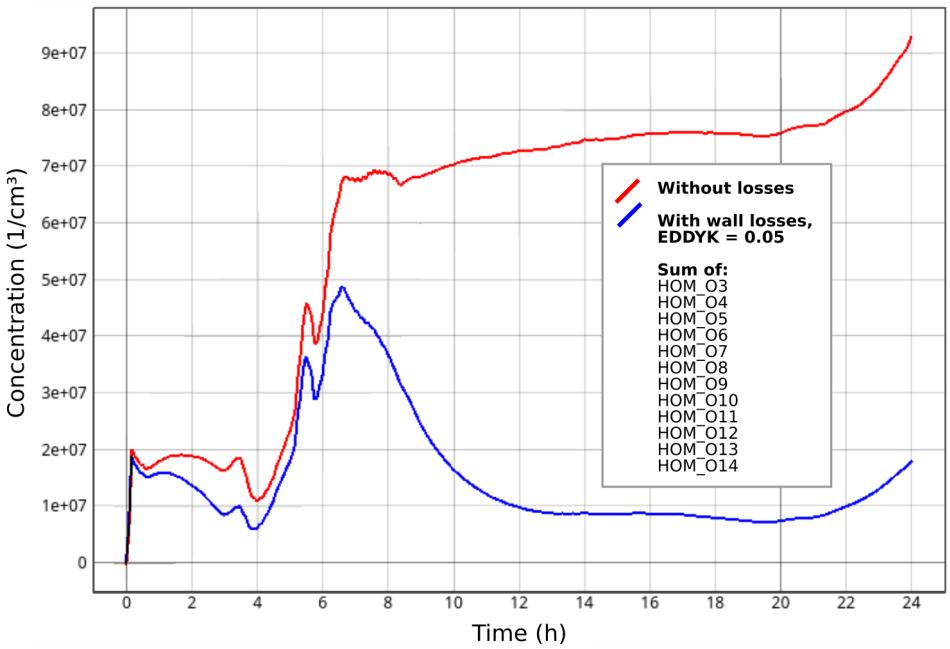

Figure 7: Effect of wall loss of gases for ambient simulation. EDDYK = coefficient of eddy diffusion, wall accommodation coefficient = 5×10^{-5}.

3 Using ARCA box

3.1 Requirements to run the model

The user will need to have Fortran compiler and NetCDF4-Fortran installed on the computer where the model is run. Note that this does not have to be the same computer where the GUI is run, permitting configuring the model on a local computer and running it from the terminal on a remote computer. The model is compiled from the command line by running "`make`" in the model root directory. To test the result, the numerical model can be run from the command line with: `arcabox.exe test_installation`, or using the GUI (in tab "Run ARCA").

Much of the hope that ARCA will be a useful tool for a broader audience than just modellers lies in its graphical user interface and the improved usability compared to the existing models. The GUI is written in Python 3, which is a versatile cross-platform open source programming language. The graphical interface relies on PyQt5, currently very widely used toolkit which is likely to be supported for a long time in the future. The plotting functions are based on PyQtGraph, which is a toolkit for scientific real-time plotting. Other necessary Python packages are NetCDF4 (for Python), NumPy, SciPy and Matplotlib.

The repository version of ARCA includes a skeleton chemistry for installation purposes, but for any real scientific work the user should compile a chemistry scheme which suits the simulation conditions. There the software Kinetic PreProcessor (KPP, http://people.cs.vt.edu/-asandu/Software/Kpp) is needed, and will therefore be introduced in the next section. The procedure for setting up a chemistry scheme is explained in section 3.1.2.

3.1.1 Kinetic PreProsessor (KPP)

Writing the system of chemical ODEs would be tedious and error prone work and on the other hand it is perfect task for a computer program. In ARCA, the code for the chemistry module is produced by the Kinetic PreProsessor (KPP). For any given set of reactions, their rates and the compounds involved, KPP creates the ODE matrices and the necessary

Fortran code to calculate the chemical concentrations at given time using one of the available integration methods.

When the chemistry scheme for ARCA is created, typically concentrations of precursor compounds that are input as measurements (or estimates), are defined to be fixed. These could be precursors or intermediate products. Fixed concentrations are not variables in the coupled ODEs but are used as parameters. If they are produced in some reaction, the concentration change is ignored.

It should be mentioned that although condensation and wall losses are sink terms that could be considered in the chemistry module, for simplicity they are considered later in the aerosol and deposition module. This approach can be justified with the assumption that the vapours which are condensing on the particles are for the most part at the end of the reaction chain they belong to, so ignoring their deposition sinks would not influence the production of some subsequent species. Another aspect is that the model can optimize the time step so that the relative changes in the concentrations in one time step are kept small, leading to smaller total error.

3.1.2 Defining a new chemistry scheme

The documentation of ARCA includes detailed instructions for creating a chemistry scheme, and it is covered here only superficially. The procedure starts with downloading the reactions and reaction rates from MCM website (mcm.york.ac.uk). There one can select the precursor VOCs and obtain the necessary rates as text file. Other chemical mechanisms (such as PRAM, see section 2.1.2) are added to chemistry at this point. Finally, running KPP will produce the chemistry module Fortran files, which should then be saved to a freely named directory "src/chemistry/ <my_chemistry>". It is important that KPP is run with `#DRIVER mcm_module` option, since this will create the modular version of KPP, interfaced to ARCA box by "Chemistry.f90". It is possible to keep multiple chemistry schemes in "src/chemistry", but the model has to be recompiled when the scheme is changed; it is easiest to do this in the GUI from the tab "Advanced → Recompile", where the list of available schemes is shown, and the desired one can be selected.

3.2 General structure of the model

ARCA box consists of two somewhat independent parts, the graphical user interface and the numerical model (Figure 8). Independence means that the numerical model can be run even if the GUI is not installed, on the other hand the simulations can be configured, or output examined with the GUI without the Fortran model. The main function of the GUI is to create the necessary initialization file (from here on referred as `INIT-FILE`), where the settings are passed to the model using Fortran *namelist* variables. GUI presents all the available model settings, contains tools to create parametric input, enables the user to save and load different model settings and inspect model output.

Broadly speaking the Fortran model consists of three main stages: 1) reading of input and initializing the program variables, 2) integration loops over the time of study, and 3) saving data from these loops.

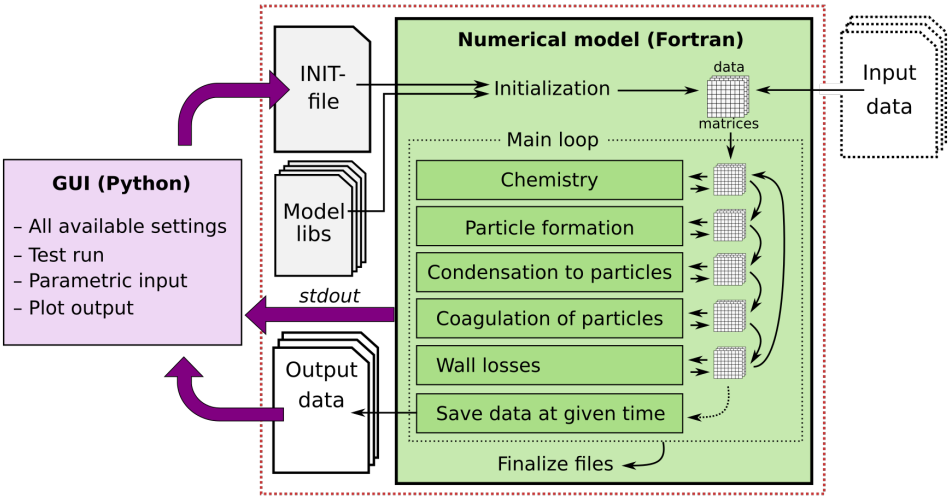

Figure 8: Schematic representation of ARCA box. The green rectangle contains the Fortran part of the model, darker shade green boxes contain the modules, the purple box the Python graphical user interface (GUI). Purple arrows show where the interaction between the GUI and Fortran executable takes place. GUI interacts with the Fortran model by writing the ***INITFILE*** *(top purple arrow), repeating the screen output of the model (middle purple arrow) and plotting the output data (bottom purple arrow). The dashed purple rectangle is the minimal configuration of ARCA, INPUT data is not strictly necessary as parametric input can also be used.*

3.3 User interface of ARCA box

In the following sections ARCA's graphical user interface (GUI) is explained in a way which also serves as a quick manual for ARCA box. A glossary of the input variable names in the Fortran model and their relation with the GUI is provided in Appendix 1.

3.3.1 Loading, printing and saving model settings

On the bottom of the GUI window are the buttons for printing, saving and loading an `INITFILE`, which is used by the Fortran model. "Print INIT" will print the contents of the `INITFILE` on the terminal screen which launched the GUI. "Save" and "Save as" are used to save the `INITFILE`. Any `INITFILE` created with the GUI can also be loaded in to the GUI, using "Load INIT". "Save as defaults" will save all current options as default values, loaded by ARCA GUI on start-up. The defaults are saved in directory "ModelLib/gui/defaults" as an `INITFILE`.

3.3.2 General options

Most important settings are found on the first tab "General options" (Figure 9).

Modules in use

These options turn the main modules on and off. Even if ACDC is off, the parametric organic formation rate (see 3.3.8) and input variable `NUC_RATE_IN` (see 3.3.5) can still be used.

Model run options

The simulation time is defined either in hours or seconds. "Case name" is a general name for the simulations, while "Run name" is a specific simulation for that "Day" or "Index". The naming logic of the output directories is shown in Figure 10. The choice of "Date" or "Index" has some consequences in the model run. If "Date" is used, the model will calculate the Sun's position on the sky based on the date, time and location (latitude is defined in "Advanced" tab, the default location is that of SMEAR II, Hyytiälä) and use this information together with the surface albedo (defined in the "Advanced" tab) to estimate the actinic flux.

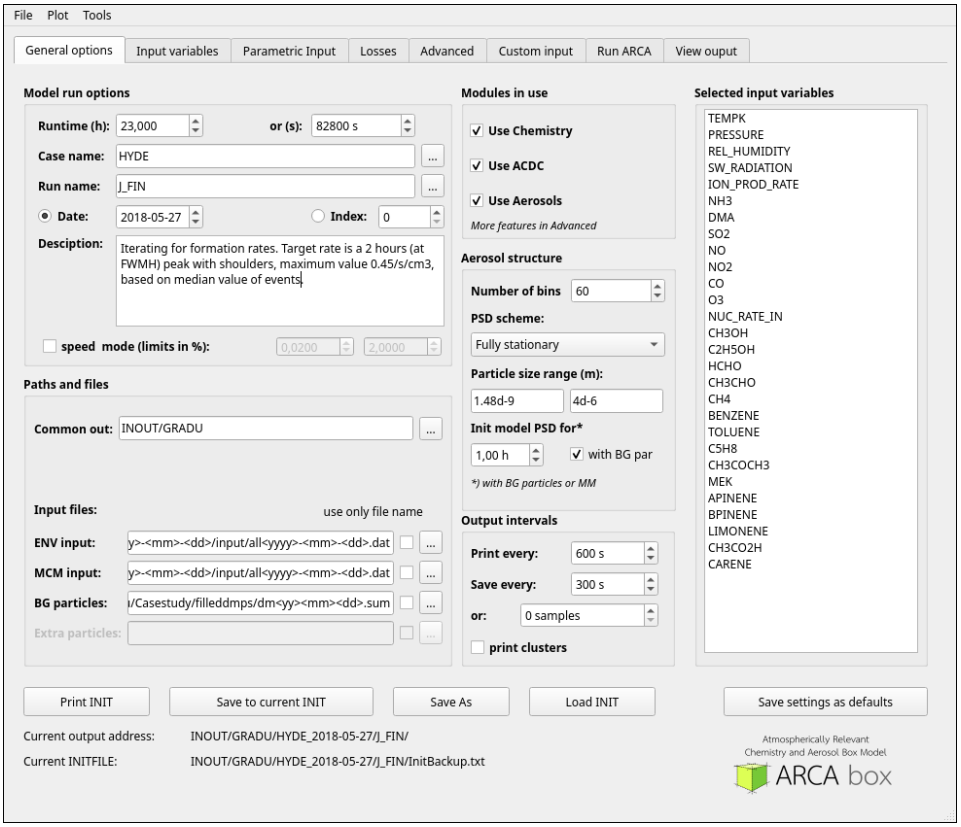

Figure 9: ARCA main settings.

The starting time of the model is always midnight, unless shifted with the option **`start_time_s`**, defined tab "Custom input". If "Index" is used instead of date, the location information is ignored, and the direction of the short wave radiation is always assumed to be from straight up (surface albedo is still used in the actinic flux calculation). Therefore "Index" is best suited for chamber simulations.

The "Description" text box is for saving notes about the settings. The text inserted here will be saved to all the output NetCDF files, introduced in section 3.3.11 (see also Table 2).

Speed mode

The integration time step can be dynamically optimized by the model. There the output of each module (the darker shaded green boxes in Figure 8) is compared with the initial values sent to the module. If the relative change is smaller than the target defined in the "limits in %", time step for that module is multiplied by 2, in other words the module is now

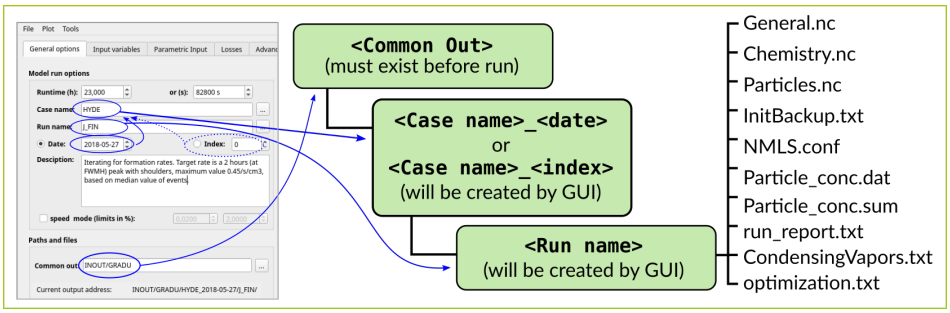

Figure 10: The output structure of ARCA. "Common out" can be outside the program directory, if absolute path is given, or relative to program location if relative path is given.

called once every two loops. If the values are changing very slowly, the time step will increase until the change will become larger than the accepted relative change. The model will then revert to the previous values, half the time step of the module which failed to meet the target precision, and repeat the loop. The nominal model time step Δt (defined in tab "Advanced") is then a minimum time step for all modules; individual modules can use values which are $\Delta t \times 2^n,\ n = \{0, 1, 2...\}$.

Paths and files

Here the root directory for output and paths for input files are defined. "Common out" is the root directory where all output is saved. Root path is relative to the main directory, and if it is outside the ARCA program directory, absolute path should be provided. The root directory must exist before the model can be run, neither the GUI nor the Fortran model will create it. The structure of the files in "ENV input", "MCM input" and "BG particles" will be explained in detail in sections 3.3.3 and 3.3.5. Here also the user can give either path relative to the main program directory, or absolute path. If the tick box "relative to case" is checked, only the filename is used and the file is always assumed to be found in directory "INPUT", located in directory "[Common out]/[CASE NAME]_[date or index]/". Wildcards can be used to name the file, practical for example when the input files are named with date or some index. Wildcards y=year, m=month, d=day and i=index are inserted in tags <>. For example "<yyyy>" means four-figure year, "<yy>" two figure year. When the cursor is hovered above the path text, the tool tip shows how the path will be parsed.

3.3.3 Aerosol structure

Options in tab General options→Aerosol structure define the number of bins, the particle size range and duration of the PSD initialisation. "PSD scheme" selects between the FS or MA PSD representation, presented in sections 2.3.1 and 2.3.2. As noted in section 2.5 regarding coagulation, selecting an upper size limit for the PSD is important from a model stability point of view. If the simulations are terminating in the middle of the run (not immediately upon start), the first troubleshooting option is to increase the particle upper size limit. A reasonable value for the number of bins is 100, but much lower values can give stable runs, especially if the Moving Average PSD is used.

The "with BG par" check box turns on and off the PSD initialization with the file defined in "BG particles." When measurements are used to initialize the model, they are read in from a text file ("BG particles", defined in tab "General"→"Paths and files"), which has to have the following structure:

	0	d_1	d_2	d_3	...	d_m
$time_0$	[ignored]	$conc_{0,1}$	$conc_{0,2}$	$conc_{0,3}$	...	$conc_{0,m}$
$time_1$	[ignored]	$conc_{1,1}$	$conc_{1,2}$	$conc_{1,3}$	...	$conc_{1,m}$
...	....	....	...		...	...
$time_n$	[ignored]	$conc_{n,1}$	$conc_{n,2}$	$conc_{n,3}$	...	$conc_{n,m}$

where d_i is the particle diameter in meters. If the second column starts with 0, it is ignored, otherwise it is treated as a size bin. This is because often in particle measurements their .sum-files (for example the ones from SMEAR stations) use the second column to store the total particle count, which is not used in ARCA. The actual time values in the input file are ignored, but the time interval must be same for the whole file (see also the next subsection "*Duration of the initialization*"). ARCA assumes that the second row (first measurement row) starts with the values at t=0 (model time) and proceeds from there with the given interval. The concentrations must be given as log-normalized concentrations, defined by

$$\text{Log-normalized conc} = \frac{dN}{\mathrm{d}\log_{10} d_p} \approx \frac{N_i}{\Delta\left(\log_{10} d_{p,i}\right)} = \frac{N_i}{\log_{10}\left(\frac{d_{p,i}}{d_{p,i-1}}\right)} \qquad (15)$$

where N_i is the number concentration in bin i and $\Delta\left(\log_{10} d_{p,i}\right)$ is the width of bin i. If the bins are with equal spacing in the logarithmic scale, the denominator in equation (15) is constant.

The particle size distribution can also be initialized using a (multimodal) lognormal distribution. The properties of such PSD are defined in tab "Advanced"→"Initialize PSD with (multi)modal". There the total particle count is first defined, and the PSD properties are given as one [(x_1 x_2 x_3)×number of modes] space separated vector where for each mode x_1 is the the count median diameter (in meters), x_2 the geometric standard deviation and x_3 a scaling factor for the concentration. By pressing the "Show PSD" text the resulting PSD is plotted interactively in a separate pop up window.

Duration of the initialization

The duration (in model hours) for replacing the modelled particle size distribution (PSD) with the measured PSD or modal lognormal distribution is defined with "Init model PSD for". When the model reaches a moment where measurements exist (by default every 10 minutes but this interval can be changed in tab "Custom input"), and the model time is less than what is defined in "Init model PSD for", ARCA reads the measured particle size distribution and redistributes it from the measured bins to model bins, preserving total number and mass. The modelled particle size distribution is then replaced with this redistributed PSD. In case (multi)modal lognormal distribution is defined, it will replace the model PSD on every time step as long as the model time is less than the time defined in "Init model PSD for". Since the composition of the particles in the initialization PSD is unknown, their composition is also initialized to a generic non-evaporating "compound" (called GENERIC in the output), which does not exist in the gas phase. This means that when the particles are initialized, it takes some time to see the effect of condensing organic vapours in the particle phase composition. But even after longer period, only some part of the particle mass will be due to condensation, and some part will have a GENERIC composition, however this is not the case with any particles that are formed after the initialization due to new particle formation. This must be kept in mind when interpreting the results of the model, especially when comparing modelled and measured particle composition.

3.3.4 Defining the condensible vapours

The particles in the model are composed of the nucleating and condensing vapours, and of the generic, unknown composition which comes from the initialization of the PSD. Together all these particle phase compounds form the particle composition. ARCA keeps track of the composition, and it is saved as time and diameter resolved array in the output file "Particles.nc" file in units of kg/particle.

When a chemistry scheme is defined for ARCA, the saturation vapour pressures for the condensible compounds are obtained (for a range of temperatures) using the Chemistry package (explained briefly in section 3.1.2), and the Antoine equation

$$\log_{10}\left(P_{sat,c}^{eq}\right) = a_c - \frac{b_c}{T} \tag{16}$$

is fitted to them, where temperature T is the independent variable and a_c and b_c the free parameters. For each compound its name, molar mass, a and b are stored in a single file which is later loaded by ARCA, and using eq. (16) the P_{sat} is calculated at every time step.

While any chemical compound produced in the chemistry module could in principle condense on the particle, provided that its saturation vapour pressure (P_{sat}) is known, only compounds with very low P_{sat} have any significance in the condensation. Computationally it makes sense to restrict the number of condensible vapours using a threshold value for P_{sat}. The threshold has to be considered with respect to the anticipated concentrations of the vapours; for example, if a compound has P_{sat} of 100 mPa, it would require mixing ratios of approximately 1 ppm to contribute to condensation in atmospheric pressure. If such concentrations are not realistic, the compound can be ignored in the condensation scheme.

The list of condensible vapours used in the model is defined in the tab "Advanced" with option "Vapour file", which is the path to the file containing names of the condensing vapours and their properties. The names must match exactly those in the chemistry module and are originally defined in KPP settings file. To limit the list of condensing vapours to n first compounds in the "Vapour file", "limit_vapours" option can be used in the tab "Custom input". In addition to the compounds found in the "Vapour file", sulfuric acid is always added to condensible vapours, with

extremely low saturation vapour pressure so that its condensation is only limited by its diffusivity.

3.3.5 Input of time-dependent variables

All one-dimensional input variables that can have time-dependent values are shown in the GUI on the right hand side list in tab ”Input variables”. These include gas concentrations, temperature, pressure, short wave radiation and other physical variables. The user must first select the variables that are considered in the current simulation, and by pressing ”Move to selected” they are selected to be uploaded to the model. In ARCA the input of these variables can be defined in two ways, via text files or parametric input (described in section 3.3.6). When text files are used, the values are provided as a time series. The variables are grouped in three categories based on the files (see section 3.3.2→”*Paths and files”*) from which the variables are uploaded to the model:

1. ”ENV input”: physical variables, concentrations of inorganic compounds, rates, see Table 1 for complete list of variables.
2. ”MCM input”: Concentrations of organic compounds (variables used in the chemistry module). The list consists of all the possible precursors in MCM (with some additions), but it can also be extended by the user if the chemistry requires some additional input. The naming follows that of the MCM, and for information about the structural properties and elemental composition of the compounds, visit mcm.york.ac.uk→Browse.
3. ”BG particles”: background particles, typically from measurements (described in section 3.3.3). Since these are two dimensional variables, they are not included in the tab ”Input variables but are instead uploaded from text file. Background particles can still be initialized parametrically using the (multi)modal PSD, also described in section 3.3.3.

The assumed structure for the ”ENV input” and ”MCM input” files is such that first column is the time in decimal days, and the next columns contain a variable each, separated by tab or space.

Table 1: Variables which can be uploaded to the model with the "ENV input" file (or alternatively with the parametric input). By default all values are zero.

Name in ENV input	Name, description	Available unit
`TEMPK`	Temperature (default value is 0°C)	C, K
`PRESSURE`	Air pressure	Pa, hPa, kPa, bar, mbar
`REL_HUMIDITY`	Relative humidity	%
`CONDENS_SINK`	Condensation sink of H_2SO_4. If not provided, will be calculated from particle size distribution (if it exists).	1/s
`CON_SIN_NITR`	Condensation sink of HNO_3	1/s
`SW_RADIATION`	Short wave radiation (280–650 nm)	W/m^2
`ION_PROD_RATE`	Ion production rate (used in ACDC)	Ion pairs/s/ cm^3
`H2SO4`	Concentration of sulfuric acid, H_2SO_4. Can also be modelled with proper chemistry.	$1/cm^3$, ppm_v, ppb_v, ppt_v,ppq_v
`NH3`	Ammonia, NH_3	
`DMA`	Dimethylamine, C_2H_6NH	
`SO2`	Sulfur dioxide, SO_2	
`NO`	Nitric oxide, NO	
`NO2`	Nitrogen dioxide, NO_2	
`CO`	Carbon monoxide, CO	
`H2`	Hydrogen gas, H_2	
`O3`	Ozone, O_3	
`NUC_RATE_IN`	Formation rate of the smallest particles in the model. This will always be added to what ACDC or other formation rate methods produce.	$1/s/cm^3$

The columns of the input files are assigned to the respective variable and unit in tab "Input variables"→"Column", using the column index. If the "ENV input and "MCM input" files have comment header (see Figure 11 for an example of an input file), the menu option "Tools→ "Print input headers with column numbers" shows the column index. For the "ENV input" and "MCM input" files the measurement points are linearly interpolated between measurement time points, and these do not have to be equally spaced. This means that if higher order interpolation is needed, it

has to be done outside the model, and the input files should have sufficiently small time intervals to get reasonable accuracy with linear interpolation. Same column can be used for multiple variables, and the "ENV input" and "MCM input" files can altogether be the same file. Note that any variable which is used by the model, but not selected as input in tab ”Input variables”, will by default be zero. For example, the formation rate due to clustering of ammonia and sulfuric acid depends on the ion production rate (a variable called `ION_PROD_RATE`). If module ACDC is used, but `ION_PROD_RATE` is not selected as input, the variable will be zero during the simulation.

The values that are read in from file can be manipulated using the modifiers ”Multiply” and ”Shift”, and their use is explained in the next section along with the parametric input.

Before the main loop starts, a screen print is provided with the list of all variables that are read in, the column from where they are read in, the units assumed and any modifications to nominal values. It is advisable to check the printed messages carefully before performing the actual simulations. If the chemistry module is used, the model will start the main loop only if all the organic compounds that are given as input (marked with grey colour in tab ”Input variables”) are actually found in the chemistry module.

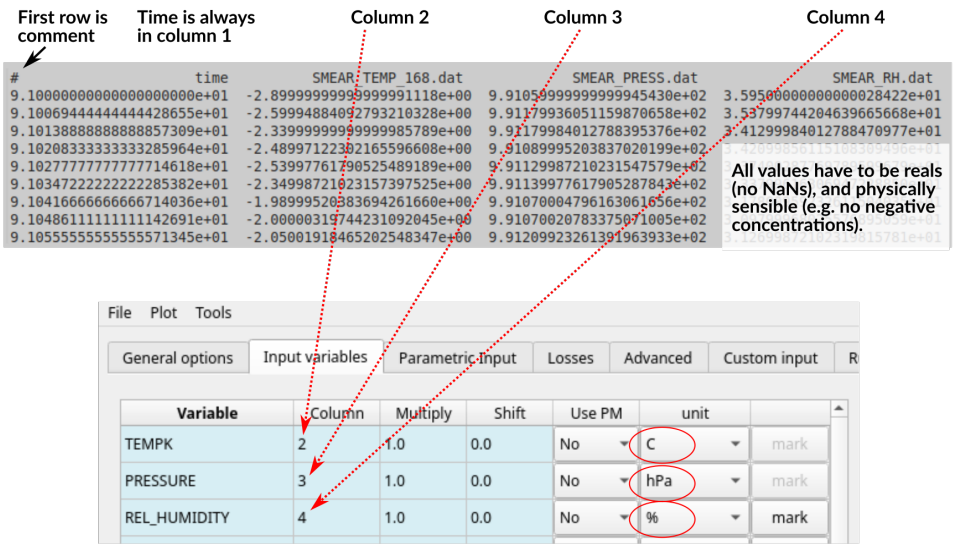

Figure 11: Example that shows the desired format for measurement input file, and the assignment of the file to the variables in the GUI.

3.3.6 Parametric input

In ARCA's parametric input, every time-dependent variable can be modified or even completely replaced by the parameters in the model settings. Couple of examples illustrate the idea: We might have a time series of measurement for NH_3, which would affect the molecular cluster formation rate. Using the modifier "Multiply", found in tab "Input variables", for NH_3, we could multiply the measurements by setting the parameter to 0.5, and then perform another run where it is set to 2. This would give a range of estimation of the formation rate calculated by ACDC. Another example would be when monoterpenes are measured, but we would like to assign the concentrations to alphapinene, betapinene, carene and limonene. Using the same input column for all of these but setting the multipliers for each so they add up to 1 would achieve this. We might also want to see what effect a 5°C temperature decrease has to the simulation. Using a Shift of -5 for `TEMPK` would achieve this. Even if no external input file for a given compound is used, a constant can be given by setting the Shift to that value, as the default value for any variable is zero.

In addition to multiplication and shifting, ARCA has a way to assign a smooth time series for any variable using the parametrised function. An example and instructions of this is shown in Figure 12, where the user wants to set a diurnal pattern for the nucleation rate. The parameters of the variable are saved in the `INITFILE` and these are used in the Fortran model to calculate a value for the variable in question at each time step.

3.3.7 Losses

The physical basis for calculating the wall loss rates is covered in section 2.6. In tab "Losses" the user can turn on and off the chemical and aerosol losses, define the chamber dimensions, modify the value for eddy diffusion coefficient, friction velocity and accommodation coefficients. Loss rates can also be uploaded to the model from a file, which should have similar structure to the "BG particles" file (shown in section 3.3.3), where the first row starts with 0, next n columns contain the diameters (in meters). Second row starts with time (in decimal days) followed by n columns containing the loss rates. If loss rates stay constant over time, no additional rows are needed. The rates are linearly interpolated both between the

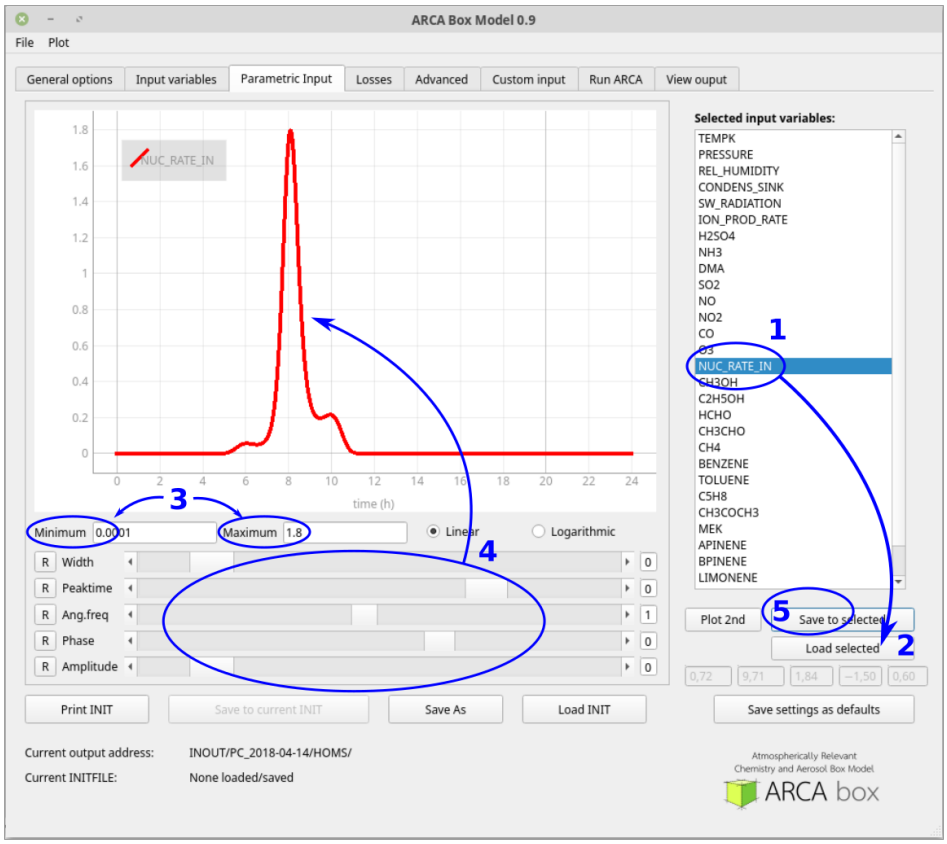

Figure 12: Parametric input explained. First select the variable in question (1). Press "Load selected" (2). Set the minimum and maximum for the function (3). Use the sliders to get the desired form for the time series (4). Save the parameters to the variable (5).

given timestamps, assumed to be in decimal days, and between the given diameters to fit the model particle diameters.

3.3.8 Advanced options

The tab ”Advanced”, shown in Figure 13, contains many important but somewhat miscellaneous options which have mostly been covered in the previous sections. The options that have not been previously explained, are introduced now.

Aerosols include condensation/coagulation

These options simply turn on and off the different aerosol dynamic processes.

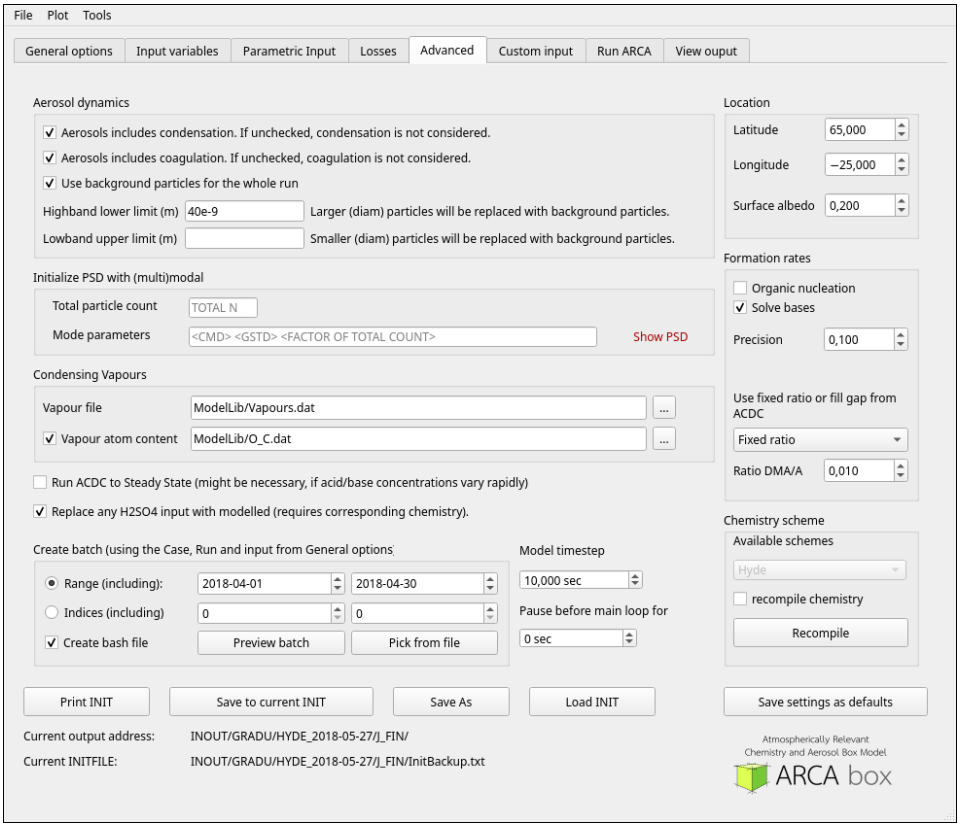

Figure 13: "Advanced" tab of the GUI.

Use background particles for the whole run

Choosing this option will result in replacing the PSD with the measured particle distribution even after the time defined in “Read BG particles for” has passed. The behaviour of the module changes so that only particles larger than diameter defined in “Highband lower limit” or smaller than diameter in “Lowband upper limit” are replaced with the measurement. This method is used when the focus of the simulation is in some particular size range, such as the nucleation mode. Using measured particles elsewhere produces more realistic condensation and coagulation sinks for the size range of interest. Leaving the **`Highband lower limit`** or **`Lowband upper limit`** empty has the same effect than setting infinitely large or small size limit, respectively. This option only applies when ”BG particles” are used and has no effect when (multi)modal PSD initialization is used.

Vapour atom content

This file contains the elemental composition of the condensing vapours (defined in "Vapour file"). It is used to calculate the diameter of the organic molecules using the method of Tang et al. (2015), affecting their diffusivity. If this option is not checked, diameter is calculated from the bulk density.

Run ACDC to Steady State

ACDC solves the cluster concentrations and the fluxes by running the simulation for a given maximum time, or as long a time-independent formation rate is reached. After this step the cluster population is saved in the memory and used as a starting point on the next time step. By default in ARCA, ACDC is only run for the duration of model time step (Δt), by which time a time-independent formation rate is not always reached, depending on the concentrations of the nucleating compounds and the cluster concentrations calculated at the previous time step. Optionally ACDC can be run to time-independent state by selecting the option "Run ACDC to Steady State". The simulated formation rates obtained with or without "Run ACDC to Steady State" generally differ by small amounts when the conditions affecting the cluster formation change slowly; then ACDC approaches time-independent formation rate even with shorter simulation time. However, when conditions change rapidly, like in the beginning of the run, when the initial cluster population is zero, or later if the measured monomer concentrations have sharp peaks, duration of the time step is not enough to reach time-independent formation rate.

The effect of simulating the formation for Δt and aiming for time-independence is illustrated in Figure 14, where the rates are modelled for one hour, using H_2SO_4 and NH_3 in concentrations often observed in ambient conditions but changing rapidly. When the formation rate is increasing due to increasing precursor concentrations, the Δt rates are lagging behind the time-independent rates and show lower values. In decreasing concentrations the lag leads to higher rates compared to time-independent rates. In these conditions the maximum formation rate in the time-independent case was approximately 4 times higher than the time-dependent case. Note that decreasing the ARCA main time step (and therefore,

also the time-dependent simulation time in ACDC) has no effect to the outcome. In Figure 14 the loss rate of the clusters, represented by the

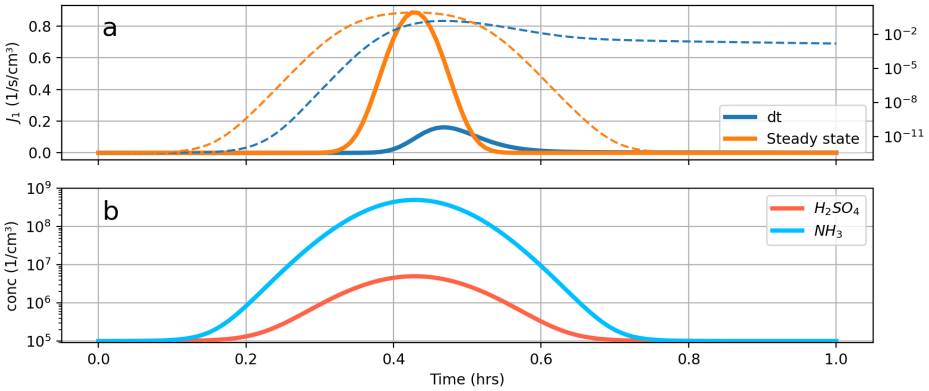

Figure 14: The effect of ACDC simulation time to obtained formation rates, with condensation sink = 0.001. In panel (a) the blue curve shows the rate when maximum simulation time was same as the model time step (10 s), orange curve shows the time-independent formation rate; the dashed lines, related to the right axis, show these in logarithmic scale. Panel (b) shows the concentrations used for H_2SO_4 (red curve) and NH_3 (blue).

condensation sink of sulfuric acid, was low (0.001/s), which explains why the formation rate in the Δt case declines so much slower than in the time-independent case.

A compromise between these approaches can be made with the parameter "`acdc_iterations`", defined in tab "Custom input", which is a multiplier for the ACDC simulation time step, and this way reach a middle ground between time-dependent and time-independent rates. It is up to the user to decide which method is used; it might be good to compare the extreme cases to see if the effect of the changing conditions is relevant. The effect of simulation time step should also be kept in mind when the formation rates from ARCA's ACDC are compared with results obtained with some other method, model or measurements.

Replace H2SO4 with modelled

By default ARCA does not use the modelled sulfuric acid concentration in ACDC or aerosol module, even if no external input for H_2SO_4 is provided. This option will use the modelled H_2SO_4 in ACDC and aerosol module, and even if the user supplies input data for H_2SO_4, this input will

not be used in the model. The input data will still be saved in the output file ”General.nc”, whereas the concentrations produced in the chemistry will always be saved in the output file ”Chemistry.nc”. Using this option requires a chemistry scheme which calculates the sulfuric acid production, and input of the precursor concentrations.

Create batch
This tool will create multiple copies of the current `INITFILE` settings and the corresponding directories to run a batch of simulations. Optionally, a *bash* file which queues all the runs is also saved. A date or index range is selected, and when pressing “Preview batch”, a pop up window is shown which lists the files and directories which would be created. If the user confirms the task, the files are created. By navigating to the batch file and running it, the runs will start immediately in a terminal window. It is up to the user to make sure that the input files exist in the corresponding input folders specified in the “General” tab. This way the runs could be performed on a remote cluster, as long as the output directory structure is copied upstream. Instead of a ”date” or ”index” range the run identifiers can also be loaded from a file using the button ”Pick from file”.

Pause before main loop for
This option has effect only when the model is run from the command line and will pause the program after the information of the initialization modules is printed and before the main loop starts. Negative value means ”Pause until Enter key”, 0 ”No pause” and positive value is pause for that many seconds.

Organic nucleation
Turns on the parametrization for organic nucleation, described in section 2.2.2.

Solve bases
Even if ammonia and dimethylamine are often referred as the key compounds when sulfuric acid is clustering with bases, other basic compounds are present in the atmosphere such as methyl- and trimethylamines, guanidine, which can have significance (Almeida et al. 2013, Kubečka et al. 2019, Myllys et al. 2019) even in concentrations that can-

not be reliably measured (Hemmilä et al. 2018, Sipilä et al. 2015). It might be interesting to see how much of DMA (or, as a first approximation, amine proxy) would be needed to fill the gap between observed and modelled formation rate. This can be done using the ”Solve bases” function, where either ammonia or DMA can be fixed and let the other vary, or tie them together with some defined fraction and solve both. When this function is used, the formation rate must be uploaded using the variable `NUC_RATE_IN`, defined in tab ”Input variables”. This is used as a target rate, which the ”Solve bases” function tries to match. The solved concentrations are saved in the output file ”General.nc”.

3.3.9 Custom input

This tab is mainly used for model development, which is always an ongoing work. There are two ways to use this tool. In the input module (in file ”src/input.f90”) of the Fortran model there is a *namelist* called `NML_CUSTOM`. By declaring variables in *input.f90* and adding them to `NML_CUSTOM` they can be modified through the GUI. The purpose of this is option to omit any need to manually tamper with the `INITFILE`s and still be able to easily transfer options to the model without a need for repeated recompiling.

The “Raw input” window prints everything directly “as is” to the `INITFILE`. Hash symbol (#) can be used to comment out lines.

3.3.10 Run ARCA

In the tab ”Run ARCA” the model can be run from inside the GUI, which is often the most practical way of running the model. The I/O between Fortran model and GUI is done via piping the STDOUT from the model to GUI using the Python module *subprocess*. The model is started by pressing ”Run with current settings”. The GUI then saves a temporary file ”ModelLib/gui/tmp/GUI_INIT.tmp” and calls the Fortran executable with this file. The printouts from the model appear on the GUI window (shown in Figure 15). While the model is running, the GUI is usable, since the two parts of ARCA are independent of each other. The model can be stopped with ”Stop” button (which sends a kill command to the model). During the run the screen update can be paused with ”Pause Scroll”.

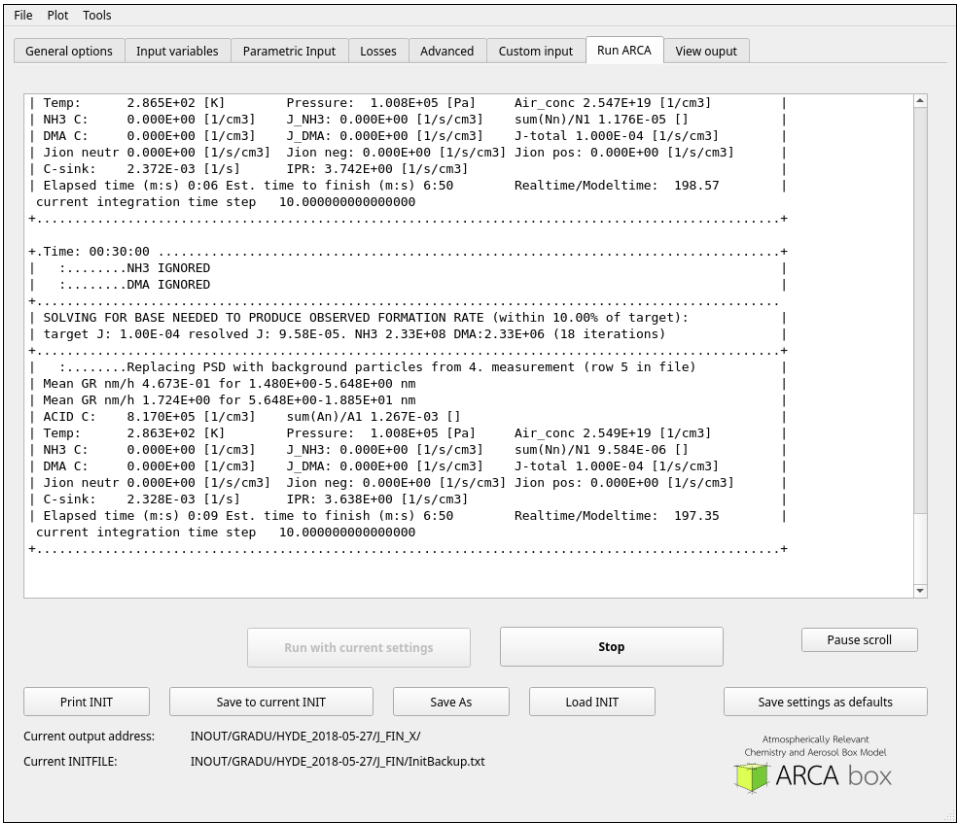

Figure 15: Test run tab from ARCA GUI.

3.3.11 Model output

ARCA box saves all its simulation output in a compressed NetCDF4 format. Additionally, the settings are documented in text files (see Table 2 for description of the output and Figure 10 for the output directory naming convention). The NetCDF files can be opened and the variables plotted in the GUI (see section 3.3.12), if the simulation has finished and the files have been closed correctly. Particle size distribution is additionally saved in two text files, where they are represented in number concentration (”particle_conc.dat”) and lognormal concentration, defined in eq.(15), similar to SMEAR sumfiles (”particle_conc.sum”, see section 3.3.3). NetCDF files can not be accessed when the model is running, but even then the particle size distribution can be plotted by loading one of the two ”particle_conc.*” files in tab ”View output”→”Surface plots”.

One of the challenges in modelling work is to keep good track of the settings used for any simulation. To tackle this problem, ARCA saves all the simulation settings in two ways. Any simulation can always be reproduced using the saved "InitBackup.txt" file, provided that the input files are the same. All ARCA *namelist* variables are documented in "NMLS.conf" file, also if the default values were not changed in the **INITFILE** (For list of the variables, see Appendix 1). If the model is run from the GUI (using "Run ARCA") the screen output (see Figure 15), which holds information of the model setup and its progression, is saved as text file "run_report.txt". In case the model is run from terminal, the print output can be saved for example by piping *tee* command: **arcabox.exe <INITFILE> |tee <path to screen print file>**.

Table 2: Description of the output from ARCA box

File name	**format**	**Description**	**Saved if**
General.nc	NetCDF4	Stores all ENV input variables and their units (whether the variables were supplied by the user or not), all supplied MCM variables and formation rates (0 if not calculated).	Always saved
Chemistry.nc	NetCDF4	Stores MCM output and their units.	Chemistry module is used
Particles.nc	NetCDF4	Stores condensible vapour concentrations, particle number and composition, particle diameter, units etc.	Aerosol module is used
particle_conc.dat	text	Particle size distribution in raw form, first row is diameter, first column is time, concentrations (in unit of $1/cm^3$) are not normalized.	Aerosol module is used
particle_conc.sum	text	Particle size distribution in log-normal form, first row is diameter, first column is time, second column is the sum of all particles, in unit of $1/cm^3$. The file is compatible with SMEAR sum-file.	Aerosol module is used
InitBackup.txt	text	A copy of the **INITFILE**	Always saved
NMLS.conf	text	List of all the *namelist* variables and their values	Always saved
runReport.txt	text	Contains the screen output from the model	If run from GUI
optimization.txt	text	Report from the time step optimization routine	Always saved

The intervals (in seconds) for saving data and printing the screen output are defined in the GUI from the tab ”General options”→”Output interval”. Alternatively, a number of samples to be saved can be selected from ”Samples” drop down menu. The model will find an interval which leads to even spacing of the saves and might part from the exact number of samples given to achieve this. “Print clusters” (available if ACDC is used) will print out the ACDC cluster population with the intervals defined in the “Print every” option.

3.3.12 View output

This tab contains three sub-tabs: ”Line graph”, ”Surface plots” and ”Particle mass and concentration”. The first is used plot time series of the variables saved in the NetCDF output files. The GUI opens the file, parses the variable names and shows them in the right hand window. Plotting scale can be changed between linear and log mode. Another file can be loaded from ”Load Similar”; the file has to be from a similar run than the initial file, since the variables and time are taken from the first file.

“Surface plot” tab is used to plot the file ”Particles.nc”, or alternatively one of the ”particle_conc” files. By pressing ”Use Current BG file” the file defined in ”BG particles” is plotted to the lower panel.

In tab “Particle mass and concentration” one can inspect the particle mass for each diameter or sum of selected diameters as a function of time, as well as particle size distributions at selected times. If ”BG particles” was defined in the model run settings which produced the ”Particles.nc” file, modelled masses and concentrations can be compared with the measurements by selecting option ”Show measurements”.

4 Case study with ARCA box

Next an example study with ARCA box is presented. New particle formation due to the clustering of ammonia and sulfuric acid or dimethylamine and sulfuric acid has been observed in a chamber experiment with vapour concentrations often found in polluted urban environment (Almeida et al. 2013). However, in a boreal type forest with clean air, ammonia concentrations are considerably lower. Continuous online measurements of ammonia or amines are difficult to conduct, and therefore reliable measurements from SMEAR II are scarce. Thus, it is hard to compare theoretical two-component nucleation rates with observed rates in Hyytiälä. It is quite likely that clustering is a multicomponent process, initially involving sulfuric acid and some base compounds, but as the clusters grow also organic molecules (Lehtipalo et al. 2016, Lehtipalo et al. 2018). The initial growth of particles, which is not yet a continuous condensation but a fluctuating, stochastic process, is a particular challenge for a model (Olenius et al. 2018). These nuances present a challenge for the ARCA box, where clustering is a two-component process and particle growth is always continuous. Bearing those limitations in mind, a theoretical experiment is performed, where the ”Solve bases” function is used to find out the base concentrations needed to explain the observed particle concentrations on seven days which showed new particle formation events. The exercise also involves modelling the sulfuric acid concentrations since these are not routinely measured at SMEAR II.

4.1 Model setup and results

All simulated days were classified as 1A event days with new particle formation (NPF), on criteria based on Dal Maso et al. (2005). Type 1 events show an emerging of a new mode of sub-25nm particles, and the continuous growth of the mode diameter. Type 1A includes the most clear and strongest events with very little existing particles in the nucleation mode before the actual event. These archetypical NPF events are not very common in Hyytiälä (only 12 type 1A events between 2016 and mid-2019), and on some occasions monoterpene measurements were lacking, making them unusable for this study. Other days showed signs of rapidly changing concentrations in the Aitken mode particles, not appar-

ently related to growth by condensation, and therefore comparison with box model would have made little sense. In the end seven days were chosen out of 12 (Figure 16), and ARCA box was used to simulate the atmospheric chemistry and aerosol formation starting from midnight until 23:00. An example `INITFILE` is shown in Appendix 2. Instead of specifically calculating the formation rates from particle measurements performed with Particle Size Magnifier (PSM), which existed only for two days in this set, they were estimated for the time of the particle formation as follows: the rates were tuned using the parametric input method in ARCA, until the model roughly reproduced the observed particle numbers in the 3–35 nm size range. The chemistry was modelled already at this stage, and therefore the HOMs formation affected the survival probability of the particle numbers in the size range of interest. The obtained formation rates are well within those reported by Dada et al. (2017) who showed median formation rates of 1.5 nm particles ($J_{1.5nm}$) between 1–3 particles/s/cm^3 during NPF events between March–October in Hyytiälä.

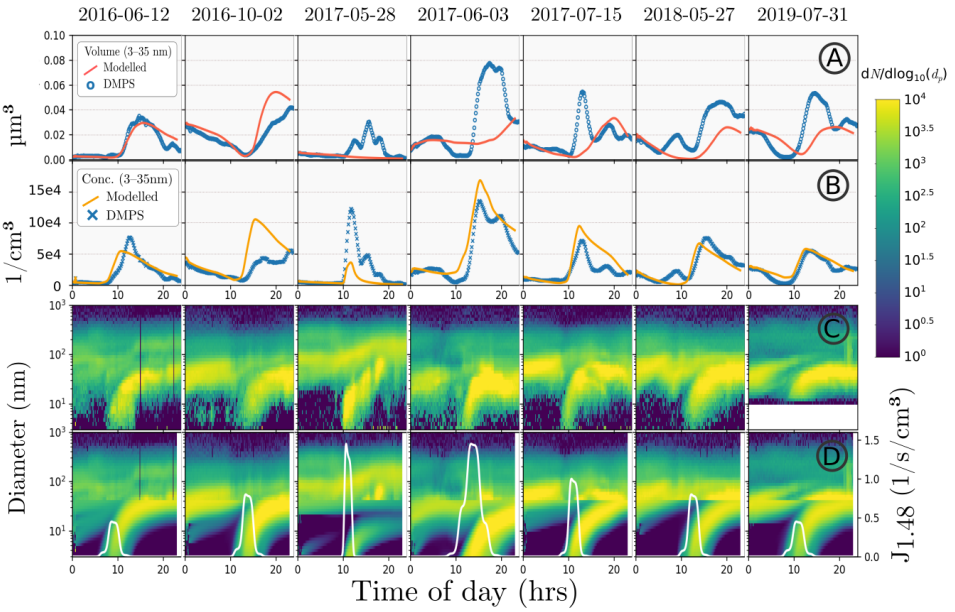

Figure 16: Seven class 1A event days modelled in ARCA. Panel A: Modelled (red curve) and measured (blue circles) summed particle volume in the 3–35 nm diameter range. Panel B: Modelled (yellow curve) and measured (blue crosses) summed particle concentration in the 3–5 nm diameter range. Panel C (measured) and D (modelled) show the surface plots for the 3–1000 nm size range. Formation rates used in the model are shown in panel D, right axis, as white curve.

The production of low volatility organic molecules was modelled using the measured monoterpene and other VOCs (in total, 15 organic compounds) as input. In the model sulfuric acid, necessary for NPF, was produced in the oxidation and subsequent reactions of the measured SO_2. Other input include measured concentrations of CO, O_3, NO_X and meteorological variables; temperature, relative humidity and intensity of short wave radiation. The precursor SO_2 and modelled sulfuric acid concentrations are shown in Figure 17. This figure also shows a problem with using SO_2 measurements, as the lower limit of detection is in the order of 2.5×10^9 molecules/cm^3 or 0.1 ppb (Dada et al. 2017), and only three days show values clearly greater than that, while the rest show noise. On the three days where the measurement is reliable, the modelled sulfuric acid maximum concentrations are between 5×10^6–2×10^7 molecules/cm^3. These are similar values to those estimated using the sulfuric acid proxy in Hyytiälä (Dada et al. 2017).

The chemistry module solved the time evolution for the concentrations of 1950 compounds, but in addition to sulfuric acid only the hydroxyl radical (OH) is shown here, due to its importance in atmospheric oxidation. Because its formation requires first ozone to be photolyzed, OH shows clear diurnal pattern, but also large differences between different days, with the daily maximum varying within one order of magnitude (Figure 18).

Figure 19a shows the base concentrations which are necessary to meet the target formation rate, given the modelled sulfuric acid concentration. Also shown is the gas phase NH_3 concentrations measured in Hyytiälä

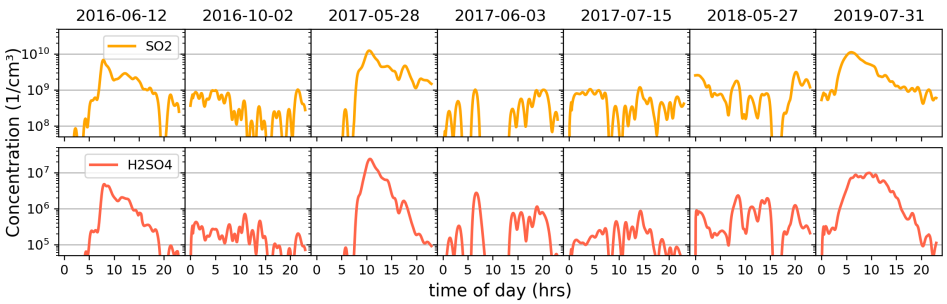

Figure 17: Measured SO_2 (top panel) and modelled sulfuric acid (bottom panel) concentrations. The SO_2 detection limit is about 2.5×10^9/cm^3, which is evident by the noisy signal below these values.

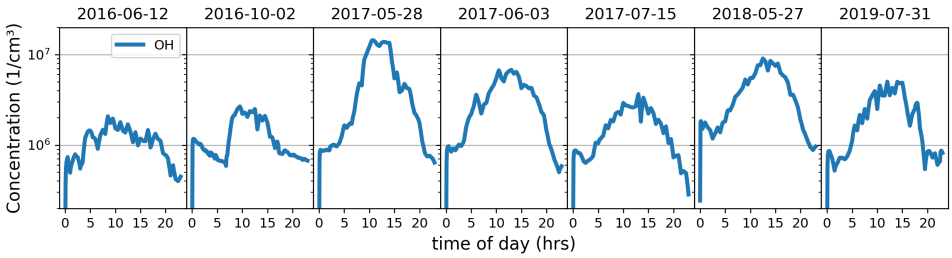

Figure 18: Modelled OH concentrations during the example days.

for the EMEP database (www.emep.int), using filter measurements with one week time resolution. Thus the ammonia data contains no information on the diurnal or inter-day variation of the concentration and could easily be influenced by transport from nearby point emissions (it is not known if these exist in the proximity of the Hyytiälä station). The ”Solve bases” was set so that the target formation rates were met with a base composition which contained 1 moles of DMA for 100 moles of ammonia. Figure 19b shows the fraction of the total formation caused by sulfuric acid – DMA clustering during the NPF event, while the rest of the particles are produced by sulfuric acid and ammonia.

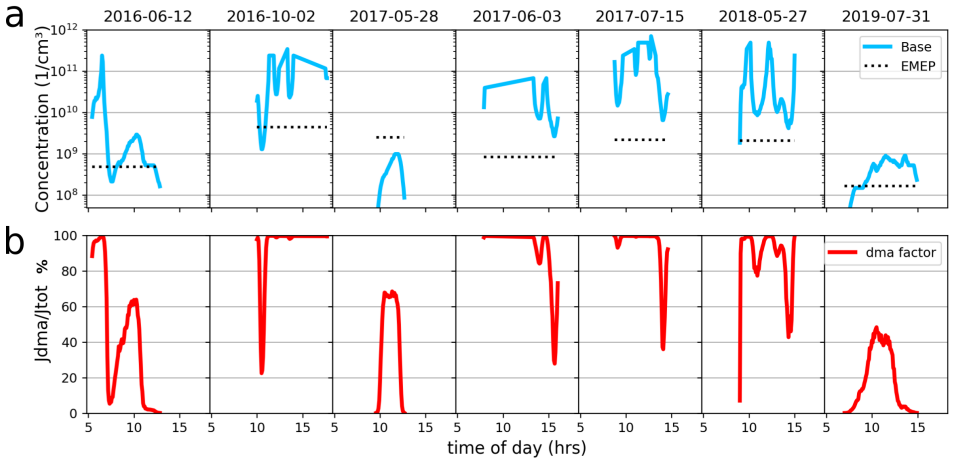

Figure 19: Panel a: Modelled base (99% NH_3, 1% DMA) levels during NPF (blue). Black dashed line shows the EMEP weekly mean concentration of gaseous ammonia. DMA concentrations follow those of NH_3 but are two orders of magnitude lower. Panel b shows the fraction (in %) of the total formation rate which is due to H_2SO_4–DMA clustering.

4.2 Discussion on the case study

The new particle formation rate was fitted so that the observed particle numbers in the 3–35 nm diameter range were approximately produced by the model. Even if the number concentrations of the model and the measurements agree, the particle volumes can still differ quite markedly (Figure 16a). For example, on 2016-06-12 the measured and modelled volumes match quite well, whereas on 2017-06-03 the observed total volume is not reproduced by the model. The particle numbers, even in this subclass of 3–35 nm size range, are dominated by the smallest particles, but the volume is governed by the larger end of the size range, and their number is dependent on the available vapours and their condensation, as well as the coagulation losses. One reason for the missing volume is the partitioning of nitric acid on the particles, which is currently not modelled in ARCA. Secondly, a box model is blind to transport by advection, fast vertical mixing or deposition, and these will alter the observed PSD in a way that the model cannot replicate. However, the model is likely still missing some low-volatile organics which contribute to the growth. The days from 2017 suggest this, as the growth rates (indicated by the steepness of the growing particle "plume") of the newly formed particles are clearly different in the observation and the model. On the other four days both the volumes and the apparent growth rate correspond much better. The formation rates deduced were realistic when compared with typical median formation rates ($J_{1.5nm}$) during events, indicating that the coagulation losses are reasonable. This suggests that more in-depth analysis of the growth of the particles in the model is needed to resolve the reasons for the occasional discrepancy between modelled and measured volume, ideally using results from more controlled environment like a smog chamber.

When considering the three days where sulfur dioxide was measured in reliable way, an interesting observation can be made, namely that the EMEP mean NH_3 concentrations are of the same order of magnitude or even higher that what would be required by the model. This is partly achieved because DMA was also included in the consideration, even when we have few measurements of the actual ambient gas phase DMA concentrations. Those made with MARGA-MS by Hemmilä et al. (2018) showed values above 1 ppt (2.5×10^7 molecules/cm^3), but due to the scarcity of the measurements we have little information on how representative these

results are. Therefore the results here are considered hypothetical, but it is interesting to note that the required concentration of DMA would have been less than or approximately 1 ppt during the event in all the three cases where SO_2 measurements were reliable. During these days the fraction of formation rate explained by H_2SO_4–DMA clustering was approximately 40–65%, when we assume DMA concentration to be 1% of the NH_3 concentration. This also means that DMA, or other amines with high clustering capability, would have to be available in the gaseous form to explain new particle formation. If, on the other hand, some other explanation beside the amines is considered, for example clustering of some organic molecules, they would have to be quite reactive to match the high clustering potential of DMA. As mentioned earlier, in the end multicomponent (with more than two components) clustering is very likely to explain the clustering in the atmosphere, but in ARCA it is not yet explicitly modelled.

The OH concentrations from the model were presented here because they showed quite large variation between the different days. The sample size is too small to draw conclusions but it would seem that the higher values were from early summer, and the lower values from mid and late summer. This could be attributed to the monoterpene emissions, which are temperature dependent and increase in the summer, thus increasing the OH loss rate.

5 Conclusions

In this work a novel atmospherically relevant aerosol box model ARCA box has been presented. It is versatile and flexible and should find many users in scientific community and other fields. Besides chamber experiment simulations ARCA box is suitable for ambient air sensitivity studies, or for example simulations of indoor air quality. The case study reported here is an example of how ARCA can be used for estimating unknown concentrations. The chemistry and nucleation modules offer more possibilities for in depth studies, for example by examining the chemical loss and production rates. It can also be a helpful tool in teaching of atmospheric chemistry and physics. The code is completely open source and does not use commercial software (in particular, it does not require Matlab), which will hopefully lead to its wider use.

ARCA box will be linked to the Multiscale Modelling group's website and can be downloaded from there. The model is currently in its initial state with version 0.9 available. The future development of the model will benefit from all user feedback and contribution, which the INAR Atmospheric Modelling Group is happy to receive, and is excited to develop ARCA further to meet demanding scientific tasks.

6 Acknowledgements

I would like to express my gratitude to the Multiscale Modelling Group in INAR, Helsinki, for letting me be part of the developers working with the ARCA box. I thank Carlton Xavier for writing the module for aerosol condensation and coagulation, Dr. Lukas Pichelstorfer for the particle size representation module and the speed optimization module. Dr. Michael Boy, who has always backed me, and been the best supervisor anyone could hope for, has written the chemistry interface and the photokinetic reactions, based on the work of Ditte Taipale (Mogensen) and others. Prof. Pontus Roldin has provided the PRAM mechanism, along with many methods that are, or will be, implemented to ARCA. The Computational Aerosol Physics group in INAR has provided the ACDC code, along with the data to use it, and prof. Hanna Vehkamäki, Dr. Tinja Olenius, Dr. Olli Pakarinen and Jakub Kubecka have always helped me with anything to do with molecular clustering, thermodynamics, or quantum chemistry. Especially I thank Hanna Vehkamäki for her patience and critical comments on this thesis. My part in ARCA has been to put all these components together to a working and usable program and document it.

Finally, I want to thank my family for their support and endurance towards my absent-mindedness at home during my whole studies.

7 References

Almeida, J., Schobesberger, S., Kürten, A., Ortega, I. K., Kupiainen-Määttä, O., Praplan, A. P., Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, 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., Tsagkogeorgas, 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. & Kirkby, J. (2013). Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere. *Nature*, 502, 359-363, doi:10.1038/nature12663.

Atkinson, R., Baulch, D. L., Cox, R. A., Crowley, J. N., Hampson, R. F., Hynes, R. G., Jenkin, M. E., Rossi, M. J. & Troe, J. (2004). Evaluated kinetic and photochemical data for atmospheric chemistry: Volume I - gas phase reactions of Ox, HOx, species. *Atmos. Chem. Phys.*, 4, 1461-1738, doi:10.5194/acp-4-1461-2004.

Boy, M., Sogachev, A., Lauros, J., Zhou, L., Guenther, A. & Smolander, S. (2011). SOSA – a new model to simulate the concentrations of organic vapours and sulphuric acid inside the ABL – Part 1: Model description and initial evaluation. *Atmos. Chem. Phys.*, 11, 43-51, doi:10.5194/acp-11-43-2011.

Crounse, J. D., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G. & Wennberg, P. O. (2013). Autoxidation of Organic Compounds in the Atmosphere. *The Journal of Physical Chemistry Letters*, 4, 3513-3520, doi:10.1021/jz4019207.

Dada, L., Paasonen, P., Nieminen, T., Mazon, S. B., Kontkanen, J., Peräkylä, O., Lehtipalo, K., Hussein, T., Petäjä, T., Kerminen, V.-M., Bäck, J. & Kulmala, M. (2017). Long-term analysis of clear-sky new particle formation events and nonevents in Hyytiälä. *Atmos. Chem. Phys.*, 17, 6227-6241, doi:10.5194/acp-17-6227-2017.

Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P. & Lehtinen, K. E. J. (2005). Formation and growth of fresh atmospheric aero-

sols: eight years of aerosol size distribution data from SMEAR II, Hyytiälä, Finland. *Boreal Environ. Res.*, 10, 323-336, doi:.

Dean, J. & Lange, N. (1998). *Lange's Handbook of Chemistry.* : McGraw-Hill Professional Publishing.

Donahue, N. M., Kroll, J. H., Pandis, S. N. & Robinson, A. L. (2012). A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution. *Atmos. Chem. Phys.*, 12, 615-634, doi:10.5194/acp-12-615-2012.

Duisman, J. A. & Stern, S. A. (1969). Vapor pressure and boiling point of pure nitric acid. *Journal of chemical and engineering data*, 14, 457-459, doi:.

Ehn, M., Thornton, J. A., Kleist, E., Sipilä, M., Junninen, H., Pullinen, I., Springer, M., Rubach, F., Tillmann, R., Lee, B., Lopez-Hilfiker, F., Andres, S., Acir, I.-H., Rissanen, M., Jokinen, T., Schobesberger, S., Kangasluoma, J., Kontkanen, J., Nieminen, T., Kurtén, T., Nielsen, L. B., Jørgensen, S., Kjaergaard, H. G., Canagaratna, M., Maso, M. D., Berndt, T., Petäjä, T., Wahner, A., Kerminen, V.-M., Kulmala, M., Worsnop, D. R., Wildt, J. & Mentel, T. F. (2014). A large source of low-volatility secondary organic aerosol. *Nature*, 506, 476-479, doi:10.1038/nature13032.

Ekström, S., Nozière, B., Hultberg, M., Alsberg, T., Magnér, J., Nilsson, E. D. & Artaxo, P. (2010). A possible role of ground-based microorganisms on cloud formation in the atmosphere. *Biogeosciences*, 7, 387-394, doi:10.5194/bg-7-387-2010.

Hemmilä, M., Hellén, H., Virkkula, A., Makkonen, U., Praplan, A. P., Kontkanen, J., Ahonen, L., Kulmala, M. & Hakola, H. (2018). Amines in boreal forest air at SMEAR II station in Finland. *Atmos. Chem. Phys.*, 18, 6367-6380, doi:10.5194/acp-18-6367-2018.

Hinds, W. C. (1990). *Aerosol technology : properties, behavior, and measurement of airborne particles.* New York, NY: Wiley-Interscience.

Jacobson, M. Z. (1997). Development and application of a new air pollution modeling system—II. Aerosol module structure and design. *Atmos. Environ.*, 31, 131-144, doi:10.1016/1352-2310(96)00202-6.

Jacobson, M. Z. (2002). Analysis of aerosol interactions with numerical techniques for solving coagulation, nucleation, condensation, dissolution, and reversible chemistry among multiple size distributions. *J. Geophys. Res.*, 107, doi:10.1029/2001jd002044.

Jenkin, M. E., Saunders, S. M. & Pilling, M. J. (1997). The tropospheric degradation of volatile organic compounds: a protocol for mechanism development. *Atmos. Environ.*, 31, 81-104, doi:10.1016/s1352-2310(96)00105-7.

Johnston, H. S., Davis, H. F. & Lee, Y. T. (1996). NO3 Photolysis Product Channels: Quantum Yields from Observed Energy Thresholds. *The Journal of Physical Chemistry*, 100, 4713-4723, doi:10.1021/jp952692x.

Kubečka, J., Besel, V., Kurtén, T., Myllys, N. & Vehkamäki, H. (2019). Configurational Sampling of Noncovalent (Atmospheric) Molecular Clusters: Sulfuric Acid and Guanidine. *J. Phys. Chem. A*, 123, 6022-6033, doi:10.1021/acs.jpca.9b03853.

Kurtén, T., Tiusanen, K., Roldin, P., Rissanen, M., Luy, J.-N., Boy, M., Ehn, M. & Donahue, N. (2016). α-Pinene Autoxidation Products May Not Have Extremely Low Saturation Vapor Pressures Despite High O:C Ratios. *J. Phys. Chem. A*, 120, 2569-2582, doi:10.1021/acs.jpca.6b02196.

Lai, A. C. K. & Nazaroff, W. W. (2000). MODELING INDOOR PARTICLE DEPOSITION FROM TURBULENT FLOW ONTO SMOOTH SURFACES. *J. Aerosol Sci.*, 31, 463 - 476, doi:https://doi.org/10.1016/S0021-8502(99)00536-4.

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

Lehtipalo, K., Rondo, L., Kontkanen, J., Schobesberger, S., Jokinen, T., Sarnela, N., Kürten, A., Ehrhart, S., Franchin, A., Nieminen, T., Riccobono, F., Sipilä, M., Yli-Juuti, T., Duplissy, J., Adamov, A., Ahlm, L., Almeida, J., Amorim, A., Bianchi, F., Breitenlechner, M., Dommen, J., Downard, A. J., Dunne, E. M., Flagan, R. C., Guida, R., Hakala, J., Hansel, A., Jud, W., Kangasluoma, J., Kerminen, V.-M., Keskinen, H., Kim, J., Kirkby, J., Kupc, A., Kupiainen-Määttä, O., Laaksonen, A., Lawler, M. J., Leiminger, M., Mathot, S., Olenius, T., Ortega, I. K., Onnela, A., Petäjä, T., Praplan, A., Rissanen, M. P., Ruuskanen, T., Santos, F. D., Schallhart, S., Schnitzhofer, R., Simon, M., Smith, J. N., Tröstl, J., Tsagkogeorgas, G., Tomé, A., Vaattovaara, P., Vehkamäki, H., Vrtala, A. E., Wagner, P. E., Williamson, C., Wimmer, D., Winkler, P. M., Virtanen, A., Donahue, N. M., Carslaw, K. S., Baltensperger, U., Riipinen, I., Curtius, J., Worsnop, D. R. & Kulmala, M. (2016). The effect of acid–base clustering and ions on the growth of atmospheric nano-particles. *Nat. Commun.*, 7, doi:10.1038/ncomms11594.

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., Kuerten, A., Lawler, M. J., Mai, H., Mathot, S., Mauldin, R. L., Molteni, U., Nichman, L., Nie, W., Nieminen, T., Ojdanic, A., Onnela, A., Passananti, M., Petäjä, T., Piel, F., Pospisilova, V., Quelever, L. L. J., Rissanen, M. P., Rose, C., Sarnela, N., Schallhart, S., Schuchmann, S., Sengupta, K., Simon, M., Sipilä, M., Tauber, C., Tome, A., Trostl, 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. & Worsnop, D. R. (2018). Multicomponent new particle formation from sulfuric acid, ammonia, and biogenic vapors. *Sci. Adv.*, 12, doi:10.1126/sciadv.aau5363.

McGrath, M. J., Olenius, T., Ortega, I. K., Loukonen, V., Paasonen, P., Kurtén, T., Kulmala, M. & Vehkamäki, H. (2012). Atmospheric Cluster Dynamics Code: a flexible method for solution of the birth-death equations. *Atmos. Chem. Phys.*, 12, 2345-2355, doi:10.5194/acp-12-2345-2012.

Mogensen, D., Gierens, R., Crowley, J. N., Keronen, P., Smolander, S., Sogachev, A., Nölscher, A. C., Zhou, L., Kulmala, M., Tang, M. J., Williams, J. & Boy, M. (2015). Simulations of atmospheric OH, O3 and NO_3 reactivities within and above the boreal forest. *Atmos. Chem. Phys.*, 15, 3909-3932, doi:10.5194/acp-15-3909-2015.

Mohanakumar, K., 2008. *Stratosphere--Troposphere Exchange*. In (Eds.), , Springer Netherlands, Dordrecht, pp. 331-355.

Mohs, A. J. & Bowman, F. M. (2011). Eliminating Numerical Artifacts When Presenting Moving Center Sectional Aerosol Size Distributions. *Aerosol Air Qual. Res.*, 11, 21-30, doi:10.4209/aaqr.2010.06.0046.

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

Nannoolal, Y., Rarey, J. & Ramjugernath, D. (2008). Estimation of pure component properties. *Fluid Phase Equilib.*, 269, 117-133, doi:10.1016/j.fluid.2008.04.020.

Olenius, T., Kupiainen-Määttä, O., Ortega, I. K., Kurtén, T. & Vehkamäki, H. (2013). Free energy barrier in the growth of sulfuric acid–ammonia and sul-

furic acid–dimethylamine clusters. *The Journal of Chemical Physics*, 139, 084312, doi:10.1063/1.4819024.

Olenius, T., Pichelstorfer, L., Stolzenburg, D., Winkler, P. M., Lehtinen, K. E. J. & Riipinen, I. (2018). Robust metric for quantifying the importance of stochastic effects on nanoparticle growth. *Sci. Rep.*, 8, doi:10.1038/s41598-018-32610-z.

Peräkylä, O., Riva, M., Heikkinen, L., Quéléver, L., Roldin, P. & Ehn, M. (2020). Experimental investigation into the volatilities of highly oxygenated organic molecules (HOMs). *Atmos. Chem. Phys.*, 20, 649-669, doi:10.5194/acp-20-649-2020.

Pichelstorfer, L. & Hofmann, W. (2015). Modeling aerosol dynamics of cigarette smoke in a denuder tube. *J. Aerosol Sci.*, 88, 72-89, doi:10.1016/j.jaerosci.2015.05.009.

Praplan, A. P., Tykkä, T., Chen, D., Boy, M., Taipale, D., Vakkari, V., Zhou, P., Petäjä, T. & Hellén, H. (2019). Long-term total OH reactivity measurements in a boreal forest. *Atmos. Chem. Phys.*, 19, 14431-14453, doi:10.5194/acp-19-14431-2019.

Richardson, L. F. (1922). *Weather prediction by numerical process.* : Cambridge University Press.

Roldin, P., Ehn, M., Kurtén, T., Olenius, T., Rissanen, M. P., Sarnela, N., Elm, J., Rantala, P., Hao, L., Hyttinen, N., Heikkinen, L., Worsnop, D. R., Pichelstorfer, L., Xavier, C., Clusius, P., Öström, E., Petäjä, T., Kulmala, M., Vehkamäki, H., Virtanen, A., Riipinen, I. & Boy, M. (2019). The role of highly oxygenated organic molecules in the Boreal aerosol-cloud-climate system. *Nat. Commun.*, 10, doi:10.1038/s41467-019-12338-8.

Roldin, P., Eriksson, A. C., Nordin, E. Z., Hermansson, E., Mogensen, D., Rusanen, A., Boy, M., Swietlicki, E., Svenningsson, B., Zelenyuk, A. & Pagels, J. (2014). Modelling non-equilibrium secondary organic aerosol formation and evaporation with the aerosol dynamics, gas- and particle-phase chemistry kinetic multilayer model ADCHAM. *Atmos. Chem. Phys.*, 14, 7953-7993, doi:10.5194/acp-14-7953-2014.

Roldin, P., Swietlicki, E., Schurgers, G., Arneth, A., Lehtinen, K. E. J., Boy, M. & Kulmala, M. (2011). Development and evaluation of the aerosol dynamics and gas phase chemistry model ADCHEM. *Atmospheric chemistry and physics*, 11, 5867-5896, doi:.

Saunders, S. M., Jenkin, M. E., Derwent, R. G. & Pilling, M. J. (2003). Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A):

tropospheric degradation of non-aromatic volatile organic compounds. *Atmos. Chem. Phys.*, 3, 161-180, doi:10.5194/acp-3-161-2003.

Seinfeld, J. H. & Pandis, S. N. (2006). *Atmospheric Chemistry and Physics - From Air Pollution to Climate Change.* : Wiley-Interscience.

Sipilä, M., Sarnela, N., Jokinen, T., Junninen, H., Hakala, J., Rissanen, M. P., Praplan, A., Simon, M., Kürten, A., Bianchi, F., Dommen, J., Curtius, J., Petäjä, T. & Worsnop, D. R. (2015). Bisulfate &ampmathsemicolonndash-mathsemicolon cluster based atmospheric pressure chemical ionization mass spectrometer for high-sensitivity (&ampmathsemicolonltmathsemicolon 100 ppqV) detection of atmospheric dimethyl amine: proof-of-concept and first ambient data from boreal forest. *Atmos. Meas. Tech.*, 8, 4001-4011, doi:10.5194/amt-8-4001-2015.

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

Taatjes, C. A. (2017). Criegee Intermediates: What Direct Production and Detection Can Teach Us About Reactions of Carbonyl Oxides. *Annu. Rev. Phys. Chem.*, 68, 183-207, doi:10.1146/annurev-physchem-052516-050739.

Tang, M. J., Shiraiwa, M., Pöschl, U., Cox, R. A. & Kalberer, M. (2015). Compilation and evaluation of gas phase diffusion coefficients of reactive trace gases in the atmosphere: Volume 2. Diffusivities of organic compounds, pressure-normalised mean free paths, and average Knudsen numbers for gas uptake calculations. *Atmos. Chem. Phys.*, 15, 5585-5598, doi:10.5194/acp-15-5585-2015.

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

Vehkamäki, H. (2006). *Classical nucleation theory in multicomponent systems*. United States: Springer.

Vignati, E., Wilson, J. & Stier, P. (2004). M7: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models. *Journal of Geophysical Research: Atmospheres*, 109, n/a-n/a, doi:10.1029/2003jd004485.

Wang, M., Kong, W., Marten, R., He, X.-C., Chen, D., Pfeifer, J., Heitto, A., Kontkanen, J., Dada, L., Kürten, A., Yli-Juuti, T., Manninen, H. E., Amanatidis, S., Amorim, A., Baalbaki, R., Baccarini, A., Bell, D. M., Bertozzi, B., Bräkling, S., Brilke, S., Murillo, L. C., Chiu, R., Chu, B., Menezes, L.-P. D., Duplissy, J., Finkenzeller, H., Carracedo, L. G., Granzin, M., Guida, R., Hansel, A., Hofbauer, V., Krechmer, J., Lehtipalo, K., Lamkaddam, H., Lampimäki, M., Lee, C. P., Makhmutov, V., Marie, G., Mathot, S., Mauldin, R. L., Mentler, B., Müller, T., Onnela, A., Partoll, E., Petäjä, T., Philippov, M., Pospisilova, V., Ranjithkumar, A., Rissanen, M., Rörup, B., Scholz, W., Shen, J., Simon, M., Sipilä, M., Steiner, G., Stolzenburg, D., Tham, Y. J., Tomé, A., Wagner, A. C., Wang, D. S., Wang, Y., Weber, S. K., Winkler, P. M., Wlasits, P. J., Wu, Y., Xiao, M., Ye, Q., Zauner-Wieczorek, M., Zhou, X., Volkamer, R., Riipinen, I., Dommen, J., Curtius, J., Baltensperger, U., Kulmala, M., Worsnop, D. R., Kirkby, J., Seinfeld, J. H., El-Haddad, I., Flagan, R. C. & Donahue, N. M. (2020). Rapid growth of new atmospheric particles by nitric acid and ammonia condensation. *Nature*, 581, 184-189, doi:10.1038/s41586-020-2270-4.

Williams, J. (2004). Organic Trace Gases in the Atmosphere: An Overview. *Environ. Chem.*, 1, 125, doi:10.1071/en04057.

Xavier, C., Rusanen, A., Zhou, P., Dean, C., Pichelstorfer, L., Roldin, P. & Boy, M. (2019). Aerosol mass yields of selected biogenic volatile organic compounds – a theoretical study with nearly explicit gas-phase chemistry. *Atmos. Chem. Phys.*, 19, 13741-13758, doi:10.5194/acp-19-13741-2019.

(2009). *Climate Change 2013 - The Physical Science Basis.* : Cambridge University Press.

Appendix 1: Glossary of input variables

The table relates the user definable variable names used in the Fortran model to the options in the GUI.

Namelist variable	type	Location in GUI
&NML_TIME		
RUNTIME	(real)	General options→Runtime (h) / (s)
DT	(real)	Advanced→Model timestep
FSAVE_INTERVAL	(real)	General options→Save every (s)
PRINT_INTERVAL	(real)	General options→Print every (s)
FSAVE_DIVISION	(integer)	General options→Or x samples
DATE	(character)	General options→Date
INDEX	(character)	General options→Index
&NML_FLAG		
CHEMISTRY_FLAG	(logical)	General options→Use Chemistry
AEROSOL_FLAG	(logical)	General options→Use Aerosols
ACDC_SOLVE_SS	(logical)	Advanced→Run ACDC to steady state
ACDC	(logical)	General options→Use ACDC
CONDENSATION	(logical)	Advanced→Aerosols include condensation
COAGULATION	(logical)	Advanced→Aerosols include coagulation
DEPOSITION	(logical)	Losses→Wall loss for aerosols
CHEM_DEPOSITION	(logical)	Losses→Wall loss for Gases
MODEL_H2SO4	(logical)	Advanced→Replace any H2SO4 input...
RESOLVE_BASE	(logical)	Advanced→Solve bases
PRINT_ACDC	(logical)	General options→print clusters
USE_SPEED	(logical)	General options→speed mode
ORG_NUCL	(logical)	Advanced→Organic nucleation
&NML_PATH		
INOUT_DIR	(character)	General options→Common out
CASE_NAME	(character)	General options→Case name
RUN_NAME	(character)	General options→Run name
&NML_MISC		
LAT	(real)	Advanced→Latitude
LON	(real)	Advanced→Longitude
WAIT_FOR	(integer)	Advanced→Pause before main loop for
DESCRIPTION	(character)	General options→Description
CH_ALBEDO	(real)	Advanced→Surface albedo
DMA_F	(real)	Advanced→Solve bases→Ratio DMA/A
RESOLVE_BASE_PRECISION	(real)	Advanced→Solve bases→Precision
FILL_FORMATION_WITH	(character)	Advanced→Solve bases→Use fixed or fill..
&NML_VAP		

USE_ATOMS	(logical)	Advanced→Vapour atom content	
VAP_NAMES	(character)	Advanced→Vapour file	
VAP_ATOMS	(character)	Advanced→Vapour atom file	
&NML_PARTICLE			
PSD_MODE	(integer)	General options→PSD scheme	
N_BINS_PARTICLE	(integer)	General options→Number of bins	
MIN_PARTICLE_DIAM	(real)	General options→Particle size range	
MAX_PARTICLE_DIAM	(real)	General options→Particle size range	
DMPS_FILE	(character)	General options→BG Particles	
EXTRA_PARTICLES	(character)	General options→Extra particles	
DMPS_READ_IN_TIME	(real)	General options→Init model PSD for	
DMPS_HIGHBAND_LOWER_LIMIT	(real)	Advanced→Highband lower limit	
DMPS_LOWBAND_UPPER_LIMIT	(real)	Advanced→lowband upper limit	
USE_DMPS	(logical)	General options→with BG par	
USE_DMPS_PARTIAL	(logical)	Advanced→Use background particles fo the whole run	
MMODAL_INPUT	(character)	Advanced→Initialize with (multi)modal→Mode parameters	
N_MODAL	(real)	Advanced→Initialize with (multi)modal→Total particle count	
&NML_ENV			
ENV_FILE	(character)	General options→ENV input	
LOSSES_FILE	(character)	Losses→First order loss rate file	
CHAMBER_FLOOR_AREA	(real)	Losses→Chamber dimensions→Floor area	
CHAMBER_CIRCUMFENCE	(real)	Losses→Chamber dimensions→Circumfence	
CHAMBER_HEIGHT	(real)	Losses→Chamber dimensions→Height	
EDDYK	(real)	Losses→Coefficient of eddy diffusion	
USTAR	(real)	Losses→Friction velocity	
ALPHAWALL	(real)	Losses→Wall accommodation coefficient	
&NML_MCM			
MCM_FILE	(character)	General options→MCM input	
&NML_CUSTOM (with their default values)			
USE_RAOULT	(logical)	Custom input	.true.
SKIP_ACDC	(logical)	Custom input	.true.
ACDC_ITERATIONS	(integer)	Custom input	1
VARIABLE_DENSITY	(logical)	Custom input	.false.
DMPS_TRES_MIN	(real)	Custom input	10
START_TIME_S	(real)	Custom input	0
DMPS_MULTI	(real)	Custom input	1.00E+06
INITIALIZE_WITH	(character)	Custom input	
INITIALIZE_FROM	(integer)	Custom input	0
VP_MULTI	(real)	Custom input	1

	DONT_SAVE_CONDENS-IBLES	(logical)	Custom input	.false.
	LIMIT_VAPOURS	(integer)	Custom input	999999
	END_DMPS_SPECIAL	(real)	Custom input	1.00E+100
	NO2_IS_NOX	(logical)	Custom input	.false.
	NO_NEGATIVE_CONCEN-TRATIONS	(logical)	Custom input	.false.
	FLOAT_CHEMISTRY_AFT ER_HRS	(real)	Custom input	1.00E+100
&NML_MODS				
	MODS(n)%MODE	(integer)	Input variables (defined by GUI)	
	MODS(n)%COL	(integer)	Input variables→Column	
	MODS(n)%MULTI	(real)	Input variables→Multiply	
	MODS(n)%SHIFT	(real)	Input variables→Shift	
	MODS(n)%MIN	(real)	Parametric input→Minimum	
	MODS(n)%MAX	(real)	Parametric input→Maximum	
	MODS(n)%SIG	(real)	Parametric input→Width	
	MODS(n)%MJU	(real)	Parametric input→Peaktime	
	MODS(n)%FV	(real)	Parametric input→Ang.freq	
	MODS(n)%PH	(real)	Parametric input→Phase	
	MODS(n)%AM	(real)	Parametric input→Amplitude	
	MODS(n)%UNIT	(character)	Input variables→Unit	
	MODS(n)%NAME	(character)	Input variable (defined by GUI)	
	MODS(n)%ISPROVIDED	(logical)	Input variable (defined by Fortfran model)	

Appendix 2: Model settings used in the case study runs

```
#-------------------------------------------------
#   ARCA box setting file: INOUT/GRADU/HYDE_2019-07-31/input/
#                          HYDE_2019-07-31_J_CHASE.conf
#    Created at: Oct 29 2020, 11:41:14
#-------------------------------------------------

&NML_PATH
 INOUT_DIR = 'INOUT/GRADU'
 CASE_NAME = 'HYDE'
 RUN_NAME = 'J_CHASE'
/

&NML_FLAG
 CHEMISTRY_FLAG = .TRUE.
 AEROSOL_FLAG = .TRUE.
 ACDC_SOLVE_SS = .FALSE.
 ACDC = .TRUE.
 CONDENSATION = .TRUE.
 COAGULATION = .TRUE.
 DEPOSITION = .FALSE.
 CHEM_DEPOSITION = .FALSE.
 MODEL_H2SO4 = .TRUE.
 RESOLVE_BASE = .TRUE.
 ORG_NUCL = .FALSE.
 PRINT_ACDC = .FALSE.
 USE_SPEED = .FALSE.
/

&NML_TIME
 RUNTIME = 23.0
 DT = 10
 FSAVE_INTERVAL = 300
 PRINT_INTERVAL = 600
 FSAVE_DIVISION = 0
 DATE = '2019-07-31'
 INDEX = ''
/

&NML_PARTICLE
 PSD_MODE = 1
 N_BINS_PARTICLE = 60
 MIN_PARTICLE_DIAM = 1.48d-9
 MAX_PARTICLE_DIAM = 4d-6
 DMPS_FILE = '/home/pecl/01-TUTKIMUS/Gradu/Casestudy/filleddmps/dm190731.sum'
 EXTRA_PARTICLES = ''
 DMPS_READ_IN_TIME = 1.0
 DMPS_HIGHBAND_LOWER_LIMIT = 40e-9
 DMPS_LOWBAND_UPPER_LIMIT =
 USE_DMPS = .TRUE.
 USE_DMPS_SPECIAL = .TRUE.
/

&NML_ENV
 ENV_FILE = '/home/pecl/01-TUTKIMUS/Gradu/Casestudy/2019-07-31/input/all2019-07-
31.dat'
 LOSSES_FILE = ''
 CHAMBER_FLOOR_AREA = 779.8
 CHAMBER_CIRCUMFENCE = 99.0
 CHAMBER_HEIGHT = 50.0
 EDDYK = 0.05
 USTAR = 0.05
/
```

```
&NML_MCM
 MCM_FILE = '/home/pecl/01-TUTKIMUS/Gradu/Casestudy/2019-07-31/input/all2019-07-
31.dat'
/

&NML_MODS
 MODS(1)  = 0 20 1.0d+0 0.0d+0 1.0d+01 1.0d+05 1.d0 12.d0 0.d0 0.d0 1.d0 'C'              ! TEMPK
 MODS(2)  = 0 16 1.0d+0 0.0d+0 1.0d+01 1.0d+05 1.d0 12.d0 0.d0 0.d0 1.d0 'hPa'            ! PRESSURE
 MODS(3)  = 0 17 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 '#'            ! REL_HUMIDITY
 MODS(6)  = 0 19 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 '#'            ! SW_RADIATION
 MODS(7)  = 1 -1 1.0d+0 3.0d+0 4.0d+0 3.0d+0 6.d0 12.000240d0 1.96d0 -5.40d0 4.70d0 '#'   ! ION_PROD_RATE
 MODS(9)  = 0 -1 1.0d+0 0.0d+0 1.0d+01 2.0d+02 4.68d0 9.000180d0 0.d0 0.d0 1.d0 'ppt'     ! NH3
 MODS(10) = 0 -1 1.0d+0 0.0d+0 1.0d+01 2.0d+02 2.71d0 12.d0 0.d0 0.d0 1.d0 'ppt'          ! DMA
 MODS(11) = 0 18 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! SO2
 MODS(12) = 0 13 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! NO
 MODS(13) = 0 14 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! NO2
 MODS(14) = 0  6 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! CO
 MODS(16) = 0 15 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! O3
 MODS(17) = 1 -1 1.0d+0 0.0d+0 1.0d-04 4.5d-01 1.08d0 11.714520d0 2.56d0 -2.70d0 0.35d0 '#' ! NUC_RATE_IN
 MODS(19) = 0 10 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! CH3OH
 MODS(20) = 0  7 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! C2H5OH
 MODS(37) = 0 -1 1.0d+0 4.d-01 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppt'          ! HCHO
 MODS(38) = 0  2 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! CH3CHO
 MODS(46) = 0 -1 1.0d+0 1.95d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppm'         ! CH4
 MODS(85) = 0  5 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! BENZENE
 MODS(86) = 0 21 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! TOLUENE
 MODS(120) = 0  8 1.0d+0 0.0d+0 1.0d-01 3.50d+02 4.08d0 12.000240d0 0.56d0 0.d0 1.d0 'ppb' ! C5H8
 MODS(141) = 0  4 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! CH3COCH3
 MODS(142) = 0  9 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! MEK
 MODS(151) = 0 11 5.1d-01 0.0d+0 1.0d+01 2.0d+02 2.71d0 12.d0 0.d0 0.d0 1.d0 'ppb'        ! APINENE
 MODS(152) = 0 11 1.2d-01 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'        ! BPINENE
 MODS(153) = 0 11 9.0d-02 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'        ! LIMONENE
 MODS(156) = 0  3 1.0d+0 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'          ! CH3CO2H
 MODS(162) = 0 11 2.8d-01 0.0d+0 1.0d+01 1.0d+05 2.34d0 12.d0 0.d0 0.d0 1.d0 'ppb'        ! CARENE
/

&NML_MISC
 LAT = 65.0
 LON = -25.0
 WAIT_FOR = 0
 DESCRIPTION = 'Iterating for formation rates. Target rate is a 2 hours wide (at FWMH) peak
with shoulders, maximum value 0.45/s/cm3, based on mean value of the days where formation
rates were calculated from PSM data.'
 CH_ALBEDO = 0.2
 DMA_F = 0.01
 RESOLVE_BASE_PRECISION = 0.1
 FILL_FORMATION_WITH = ''
/

&NML_VAP
 USE_ATOMS = .TRUE.
 VAP_NAMES = 'ModelLib/Vapour_names.dat'
 VAP_ATOMS = 'ModelLib/O_C.dat'
/

&NML_CUSTOM
 NO_NEGATIVE_CONCENTRATIONS = .true.
 NO2_IS_NOX = .true.
/
```