New trajectory driven aerosol and chemical process model: chemical and aerosol Lagrangian model (CALM)

P. Tunved¹, D. G. Partridge¹, and H. Korhonen²

¹Department of Applied Environmental Science Stockholm University, 10691 Stockholm, Sweden
²Finnish Meteorological Institute, Kuopio Unit P.O. Box 1627, 70211 Kuopio, Finland

Received: 24 May 2010 – Accepted: 5 June 2010 – Published: 21 June 2010

Correspondence to: P. Tunved (peter.tunved@itm.su.se)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

A new Chemical and Aerosol Lagrangian Model (CALM) have been developed and tested. The model incorporates all central aerosol dynamical processes, from nucleation, condensation, coagulation and deposition to cloud formation and in-cloud processing. The model is tested and evaluated against observations performed at the SMEAR II station located at Hyytiälä (61°51′ N, 24°17′ E) over a time period of two years, 2000–2001. The model shows good agreement with measurements throughout most of the year, but fails in reproducing the aerosol properties during the winter season, resulting in poor agreement between model and measurements especially during December-January. Nevertheless, through the rest of the year both trends and magnitude of modal concentrations show good agreement with observation, as do the monthly average size distribution properties. The model is also shown to capture individual nucleation events to a certain degree. This indicates that nucleation largely is controlled by the availability of nucleating material (as prescribed by the [H$_2$SO$_4$]), availability of condensing material (in this model 15% of primary reactions of monoterpenes (MT) are assumed to produce low volatile species) and the properties of the size distribution (more specifically, the condensation sink). This is further demonstrated by the fact that the model captures the annual trend in nuclei mode concentration. The model is also used, alongside sensitivity tests, to examine which processes dominate the aerosol size distribution physical properties. It is shown, in agreement with previous studies, that nucleation governs the number concentration while transport from clean areas takes place. It is also shown that primary number emissions almost exclusively govern the CN concentration when air from Central Europe is advected north over Scandinavia. We also show that biogenic emissions have a large influence on the amount of potential CCN observed over the boreal region, as shown by the agreement between observations and modeled results for the receptor SMEAR II, Hyytiälä, during the studied period.
1 Introduction

The representation of particles in the atmosphere remains one of the largest uncertainties in predicting our future climate (IPCC, 2007). Knowledge of the particle abundance, chemistry and size is of crucial importance to determine both indirect and direct climate effects of particles in the atmosphere. In order to accurately describe the aerosol properties on large spatial scales better, more efficient ways to parameterize important aerosol processes are needed. While regional and global transport models incorporating aerosol schemes are quite numerous, they often include quite coarse parameterizations of the dynamical processes relating to aerosols in the atmosphere. Typically, this is due to computational limitations.

A more detailed description of aerosol processes is possible in coupled chemical-aerosol box models, which have been previously developed and tested in different environments. These models have been of varying level detail, and often used to address specific problems or questions in different types of environments (e.g., Capaldo et al., 1999; Grini et al., 2005; Chameides and Stelson, 1992; Danilin et al., 1994). Box models are computational efficient since they do not solve the actual dispersion in space, giving the opportunity to investigate the usually computationally demanding aerosol dynamic processes with a higher level of detail than possible in large scale regional or global models.

In this study we present a new box model framework adopting state of the art aerosol dynamic description that will aid in the understanding of the different processes affecting the aerosol over the boreal regions, and in the future also at other sites. The model is a two layer box model that is intended to run along trajectories.

The aim with the study is to test this trajectory driven Lagrangian process model that seeks to capture and describe the processes that govern the evolution of aerosol chemical and physical properties. In this study we have focused on the processes dominating the aerosol as observed at Hyytiälä SMEAR station, Southern Finland (61°51′ N, 24°17′ E, 181 m a.s.l.). The SMEAR station holds today the longest record of aerosol
number size distribution observations, dating back to 1996, and also facilitates numerous measurements of other aerosol parameters and trace gases. The surroundings are dominated by a flora mainly consisting of pine forests of an age of roughly forty years. The closest large city is the city of Tampere some 60 km away from the station. The location of Hyytiälä, in the southern rim of the Scandinavian boreal region, makes it an excellent site to study both the role of anthropogenic as well as natural emissions. The location allows studies of natural sources when the air masses originating from the marine environment transport over the forest, as well as of aged anthropogenic air downstream of the large pollution sources in continental Europe. This fact has been demonstrated in a number of studies, and northerly and southerly transport of air is associated with distinct features. As has been shown in previous studies, marine air transport over the forested regions is associated with a rapid increase in aerosol number and, to a somewhat smaller extent, in mass (e.g., Tunved et al., 2006a; Tunved et al., 2006b). This increase in number appears to be controlled by nucleation mediated by sulfuric acid while the growth mainly seems to be facilitated by organics originating from the forest itself, most likely monoterpenes (MT) or similar compounds (e.g., Tunved et al., 2006a, Laaksonen et al., 2008). However, when polluted air arrives from the south it is seen that nucleation is largely absent over the forest, and number concentrations typically decrease when transport further northwards takes place (Tunved et al., 2005). Although located far from major pollution sources, observations at Hyytiälä do show that continental influence may occasionally dominate the aerosol properties (Dal Maso et al., 2007; Tunved et al., 2003) providing the site with a fairly large concentration of accumulation mode-sized particles.

In this study we use the newly developed model CALM to study the integrated effect of both natural and anthropogenic sources, as well as primary and secondary aerosol production. Using a number of sensitivity tests, we identify the processes dominating the appearance and abundance of particles in the different transport sectors. The difference between e.g. marine and continental air advection will be discussed in terms of the role of primary and secondary emissions. Several other model studies have been...
performed at Hyytiälä, covering a wide range of complexity (Eulerian process studies, Lagrangian process studies (Tunved et al., 2004), global model studies (Spracklen et al., 2006). Besides on the general features of the aerosol, many of these studies have focused on the role and mechanisms of nucleation over the forest and on the importance of primary emissions.

2 General model design

The basic model design consists of a trajectory driven box model. The model is fed single particle trajectories along which the process model simulates the evolution of both aerosol and gas species. The current model setup adopts a two layer structure, a residual and mixing layer (RL and ML, respectively). Both compartments are assumed to be well mixed internally. The trajectory dictates the transport of the model space described by these internally well mixed boxes. The trajectory itself however is not assigned specifically to either of the boxes, but instead describes the movement of this simplified model system along the geographical coordinates of the trajectory. Exchange between the layers is allowed, and this exchange is governed by the variation of the mixing layer height (MLH). The MLH typically follows a diurnal cycle with a maximum around noon. During morning hours, when the mixing layer starts to grow into the residual layer, the mixing layer gases and aerosols are mixed with their counterparts in the residual layer, a process that most of the time leads to dilution of the ML quantities. When the MLH starts to decrease, the ML aerosol and gases get partially trapped within the residual layer. The height of the mixing layer is provided by the trajectory model and typically varies between 250 m (which is the lower limit from the trajectory model) and some thousand meters above the ground level. Throughout each simulation the residual layer upper limit remains constant. This upper boundary of the model compartment is defined as the maximum MLH of each simulated trajectory and typically reaches altitudes around 1500–2500 m.
The trajectories were calculated using the HYSPLIT4 model on FNL data (Draxler and Hess, 1998). Trajectories were calculated arriving 100 m above ground level (m a.g.l.). Along each trajectory, surface temperature and surface wind speed were extracted from the meteorological data fields. This data is needed for the approximation of temperature dependent BVOC emissions and of wind speed dependent emissions from water surfaces. Other relevant meteorological parameters, such as relative humidity and ML height, were extracted from trajectory calculations as well.

In both layers we allow for general aerosol dynamics (e.g., condensation, nucleation, coagulation) as well as photochemistry. However, deposition is only considered in the ML and all ground based emissions are initially confined to the ML, although may be transported into the RL due to variation of the MLH. Two types of clouds are considered: stratus type and cumulus type clouds. The model setup only allows for clouds in the mixing layer. The frequency of clouds is described using available statistics of cloudiness over the model domain. Precipitation scavenging is accounted for in both ML and RL. The treatment of these processes will be described in detail below.

2.1 Aerosol dynamic model

The aerosol particle dynamics is described using the University of Helsinki Multi-component Aerosol model (UHMA) described in detail by (Korhonen et al., 2004). The model incorporates the major microphysical processes that affect the aerosol under clear sky conditions, namely nucleation, coagulation, multi-component condensation and dry deposition. In the current model setup we assume that the aerosol is distributed over 45 log-normally distributed bins over the size range 0.2 nm–1.2 µm, a size range and resolution that we consider sufficient for this study. The particles are assumed internally mixed in every size bin. In each bin we allow a composition defined by three different components: sulphuric acid, soluble organics and insoluble species. In the setup we assume that component one represents sulphur and sea salt, component two represents both primary and secondary formed organic aerosol constituents and component three represents insoluble species (i.e. soot in the current setup).
The nucleation is represented by the activation theory, where the nucleation rate is directly proportional to the concentration of sulphuric acid, $[\text{H}_2\text{SO}_4]$, with an empirically defined correlation coefficient, $A$ (Kulmala et al., 2006). The value of $A$ has been set to $2 \times 10^{-6}$ following empirical findings at Hyytiälä (Riipinen et al., 2007). It is worth mentioning that $A$ varies between different locations and seasons, reflecting the other unknown components necessary to accurately describe the rate of nucleation. Thus $A$ may be described as an environment-specific constant, but since the information available on the value of $A$ is still somewhat limited we assume that $A$ remains constant through the simulations. As we will see, this will not substantially affect the final simulated size distribution at the receptor location.

The dry deposition for particles and gases is calculated for the 17 classes of land cover in the International Geosphere-Biosphere Programme (IGBP; Loveland and Belward, 1997) global vegetation classification scheme. The Land Cover Classification product (MOD12Q1) includes 11 natural vegetation classes, 3 developed land classes (one of which is a mosaic with natural vegetation), permanent snow or ice, barren or sparsely vegetated land, and water.

Dry deposition is calculated following the dry deposition procedure by the EMEP/MSC-E regional model of heavy metals airborne pollution (MSCE-HM), using variations on the resistance analogy approach (Wesely et al., 1989) for each surface type (as documented by Travnikov and Ilyin, 2005).

The dry deposition of gases is calculated following the gaseous deposition model by Zhang et al. (2003). This model utilizes a “big-leaf” resistance approach model for representing the process of gaseous deposition. It is very important to describe in some manner the deposition of terpenes within CALM. It is difficult to assign the gas properties, e.g. molecular diffusivities, for these organic compounds since there are very few measurements in the literature. Therefore for consistency the carbonyl groups presented by Zhang et al. (2003) are used as a proxy for these gases in CALM.
2.2 Chemical model

The gas-phase chemistry is solved using the quasi-stationary state assumption (QSSA). The chemistry module solves the chemistry for every time step of the simulation, thereby updating the concentration of the included species. Besides chemical reactions, the trace constituents are also subject to deposition (wet and dry) and removal from the gas phase via condensation onto existing particles (and naturally also nucleation, as in the case for sulfuric acid, see Sect. 2.2). In this study we estimate the chemistry for 76 different compounds and intermediates. The module solves the chemistry of the most important oxidants; hydroxyl radicals (OH, HO₂), nitrate radicals (NO₃) and ozone (O₃).

In the model, the photolysis constants are taken from a lookup table where the constants have been calculated using the Tropospheric Ultraviolet-Visible Model (TUV, Madronich et al., 2002) under different atmospheric conditions. Additional modules to process the lookup table is also used (JVAL1, Madronich et al., 2002). The photolysis constants for 56 different reactions are extracted from the look-up table. The selection is based on solar zenith angle, azimuth and atmospheric properties such as optical depth and concentration of absorbing gases. The solar zenith angle and azimuth constants are computed from the geographical information supplied by the HYSPLIT4 output, i.e. latitude, longitude, time and date. In our set up we have fixed the optical depth to 0.1 and the single scattering albedo to 0.85. Most likely, these values will vary significantly over the modeled domain. However, we do not have the possibility to specifically address these properties more in detail. The ozone column aloft is assumed to be 330 Dobson units. The surface albedo is set to 0.3 throughout the simulations. For the purpose of this study we do not think that this assumption will affect the results in a major way. If clouds are present in the model run, TUV photolysis constants are adjusted accordingly, assuming a cloud albedo of 20 (corresponding to reasonable cloudiness), modifying photolysis constants above (i.e. in the residual layer) and below the cloud column (i.e. in the mixing layer).
The terpene chemistry is represented by the relatively well-known reaction chain of \( \alpha \)-pinene, and the reaction scheme used is the one presented in (Andersson-Skold and Simpson, 2001) and references therein. This reaction scheme includes the reactions of \( \alpha \)-pinene and its products considering oxidative reactions including nitrate radicals (\( \text{NO}_3 \)), hydroxyl radicals (\( \text{OH} \)) and ozone (\( \text{O}_3 \)). While the reactions described in this scheme certainly are important for the gas phase chemistry, e.g. radical abundance and ozone production, we do not use the products in this scheme explicitly when estimating the production of condensable species. Instead, we assume a fixed stochiometric yield of 15% condensable compounds from the primary reactions between \( \alpha \)-pinene and \( \text{NO}_3 \), \( \text{OH} \) and \( \text{O}_3 \). This compound is assumed to have saturation vapor pressure of \( 3 \times 10^{12} \) molecules per cubic meter, and molar mass of 186 g mol\(^{-1}\) (Kulmala et al., 1998). The yield agrees well with earlier estimates presented by Tunved et al. (2006b) considering what is required to reproduce the observed growth rates of particles over the boreal forest.

Isoprene chemistry follows the scheme suggested by Simpson et al. (EMEP, MSC-W, 1993), including 18 different reactions of isoprene and its products. Based on findings over the Amazonian region we also over the boreal region assume a fixed stochiometric yield of condensable gases of 0.2% of primary reactions between isoprene and major oxidants (\( \text{OH} \), \( \text{NO}_3 \), and \( \text{O}_3 \)).

Sesquiterpenes are assumed to be immediately oxidized once emitted, and thus we do not describe their chemistry at all. Sesquiterpenes are highly reactive, and are quickly oxidized under atmospheric conditions. Chamber studies indicate that the aerosol mass yield from its oxidation by common oxidants is between 17–67% (Griffin et al., 1999). In our setup 20% of the emitted sesquiterpenes are assumed to form a condensable product with the same properties as that formed from mono-terpene oxidation.

In the current setup, ethane is assumed to represent all anthropogenic NMVOC emissions. The chemistry of ethane is described using the reaction sequence presented by Simpson et al. (EMEP, MSC-W, 1993). The same goes for the methane chemistry.
With the reactions outlined above, the chemistry of NO\textsubscript{x} is solved and the ozone HO\textsubscript{x} (OH, HO\textsubscript{2}) and nitrate radical concentrations are calculated.

### 2.3 Cloud description

The occurrence of clouds is prescribed randomly in the model. The cloud frequency, however, is constrained by observational seasonal data. The cloud frequency for different seasons and locations (December–February, March–May, June–August, September–November) are adopted from the online Climatic Atlas of Clouds over Land and Ocean (http://www.atmos.washington.edu/~ignatius/CloudMap/, Warren et al., 2006). This approach allows for both seasonal and spatial constraints with respect to cloud frequency over the simulated domain. The cloud types considered are stratus type clouds and cumulus type clouds. The stratus type clouds fill the upper 150 m of the ML. The typical coverage of the stratus clouds is set to 6/8. The cumulus type clouds are assumed to form from the middle of the mixing layer up to the top of the mixing layer. Furthermore cumulus clouds are only present when the mixing layer height is above 600 m. The typical coverage of the cumulus type clouds is set to 4/8 or 50%. This means that only a fraction of the aerosol will be processed by a cloud, and this fraction is based on the horizontal coverage and vertical extent of the cloud as described above.

When a cloud is assumed to form, the ML aerosol size distribution and associated properties are fed to the cloud module, where the dynamics of the aerosol population in the clouds is described using the common growth equations (e.g., Seinfeld and Pandis, 1998), using a constant updraft through the cloud. The updraft in turn depends on the cloud type. Stratus type clouds are allowed to have a prescribed updraft between 0.025–0.125 ms\textsuperscript{-1} and the cumulus type clouds are allowed to have an updraft between 1–3 ms\textsuperscript{-1}. Although fixed for the individual clouds simulated, the updraft for every cloud cycle is determined randomly, allowing the updraft to vary within the above-mentioned limits from case to case. The growth is calculated based on the size and composition of the aerosol particles. Sulfates are assumed to be completely soluble,
while the solubility of the organics is assumed to be 10%. Variable aerosol properties (such as size and chemistry) and variable updraft will thus govern a variation in lowest activation radius from case to case. The droplet growth in the cloud is described using the moving center approach, in contrast to the fixed sectional approach adopted for the “dry” aerosol dynamics.

Once a cloud is formed, cloudy conditions are assumed to apply for 4 h. However, since clouds are dynamic phenomena, their physical properties and environmental parameters change through time. This means that a static cloud with constant properties during four hours would be unrealistic. We bypass this issue by allowing the cloud properties (base-height, thickness (as for cumulus clouds), LWC and droplet distribution) to remain constant for only 15 min. After 15 min the cloud dissipates and the cloud-processed aerosol is mixed with the rest of the mixing layer aerosol before it is recalculated from an updated aerosol size distribution and updraft. In this way, for every cloud period, the cloud is recalculated 16 times. This 15 min cycle is applied in order to prevent an unrealistic behavior of the cloud, and allows the cloud to be replenished with gases and aerosol, thereby trying to imitate the actual dynamics of clouds. Due to the overall simplicity of the model structure, we do not think that we are able to describe convective clouds in a detailed way. Thus we do not consider convective clouds or associated precipitation events in the model.

The current setup of clouds allows for no precipitation, and thus no in-cloud scavenging. This is admittedly a compromise. Several attempts were made to mimic the in-cloud scavenging by allowing the shallow clouds to precipitate. However, the results soon became unrealistic since the aerosol was removed too quickly from the modeled two layer column. Thus, it was concluded that the current setup could not provide realistic precipitation description. Instead, precipitation data (mm h\(^{-1}\)) are taken from the HYSPLIT4 model output. This precipitation is assumed to fall through both the RL and ML, but the precipitating cloud itself is formed above the modeled column. In practice, this means that explicit treatment of in-cloud scavenging is omitted in the model since the cloud is formed on an aerosol outside the modeled column. This approach
is likely valid for frontal type precipitation (i.e. nimbostratus) considering that (1) the front itself serves as a boundary between two air-masses and (2) the air flows behind and ahead of the front are not the same since they occur in different air-masses. In practice, the movement of a warm front includes lifting of the warm air on top of the colder air-mass. Since the trajectories are calculated arriving at 100 m and given the fact that the trajectories rather seldom subside close to the station, the likelihood for a trajectory having spent time inside a NB-type cloud during the last couple of days is typically small. This means that rainout processes (i.e. in-cloud scavenging) are less likely to affect the mixing layer aerosol as compared to washout processes (i.e. below cloud scavenging) considering timescales of a couple of days. Thus, in the model we assume washout (i.e. below cloud scavenging) of aerosols only. For this purpose we make use of the parameterization presented by (Laakso et al., 2003). This empirically derived parameterization is based on 6 years of ground level size distribution data measured at Hyytiälä (61°51′ N, 24°17′ E) and provides size dependent scavenging coefficients between 10–500 nm. In this study, the parameterization is extrapolated up to 1 µm. Since the approach of the study presented by Laakso et al. incorporates size distribution data and precipitation data measured on the ground level, the resulting parameterization will largely describe below cloud scavenging, but since the method relies on the measured rate of change in aerosol concentration versus precipitation intensity, also other processes will by necessity affect the rate of change in aerosol properties (i.e. size dependent number concentration). This means that the parameterization will also indirectly take into account in-cloud scavenging if the precipitation takes place in a cloud confined within a well mixed boundary layer.

The CALM cloud module further takes into account in-cloud cloud oxidation of sulfate by ozone and hydrogen peroxide. We here use a bulk-water approach by summing up the total volume of water in the cloud. The bulk-water of the cloud is assumed to be infinitely dilute and is iteratively equilibrated to the surrounding gases (NH$_3$, SO$_2$, H$_2$O$_2$, O$_3$, CO$_2$) and pH is evaluated based on the amount of dissolved gases and aerosol bulk composition (considering the sulfate fraction only). The soluble gases
(SO₂, NH₃, O₃ and H₂O₂) are partitioned between gas and liquid phase based on thermo-dynamical limitations. The in-cloud oxidation of S(IV) to S(VI) is initialised by calculating the pH of the cloud bulk water, taking into account the liquid water content (LWC), concentration of surrounding gases and the particle content of sulphate. Once the pH has been established and the equilibrium of reacting gases has been reached, the liquid phase oxidation is calculated. The pH and concentration of reacting gases in the liquid phase are recalculated every time step.

\[ \frac{d[S(IV)O_3]}{dt} = [O_3, aq] \left( k_0 [H_2SO_3] + k_1 [HSO_3^-] + k_2 [SO_3^{2-}] \right) \]  
\[ \frac{d[S(IV)H_2O_2]}{dt} = k_4 [H_3O^+] [HSO_3^-] [H_2O_2, aq] / (1 + K [H_3O^+]) \]  

Reaction constants used are those from Seinfeld and Pandis (1998). When the cloud dissipates, the produced sulphate (equivalent amount of sulphuric acid) is distributed over the size range of activated particles following the sectional method. Each particle gains sulphate proportional to the individual water content of the droplet. This causes activated particles to leave the cloud with a larger size than they entered the cloud, and clouds thus provide a source of sulphate in the model.

2.4 Emissions

2.4.1 BVOC

In the model setup we consider biogenic emissions from different land use types according to the IGBP land use maps (Loveland and Belward, 1997): Needle leaf forest, mixed forest, deciduous needle leaf forest, open shrubs, closed shrubs, grasslands, permanent wetlands and croplands. The land use type is defined as a fraction per grid of the land use map (0–1 on a quarter degree grid). To apply a seasonal dependence on the foliar biomass density we adopt the Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI, Tucker et al., 2004), from which we derive a foliar biomass density using the equations as presented in (Guenther et al., 1995) and references therein. The NDVI data is used as follows:

\[ ... \]
The monthly average of the Global vegetative indices (G) is calculated as:

\[ G = 100(1 + \text{NDVI}) \]  

(3)

And the foliar biomass \(D_m\) density is calculated as:

\[ D_m = D_p(\exp(\ln(2)((G - G_2)/(G_{\text{max}} - G_2)))) \]  

(4)

\(D_p\) represents annual peak foliar density. \(D_m\) is only calculated if \(G\) is above a certain threshold \(G_2\). Representative values of \(G_2\) are taken from (Guenther et al., 1995). Otherwise, \(D_m\) is set to zero. \(D_p\) is in turn calculated as

\[ D_p = D_r \times \text{NPP} \]  

(5)

Where \(D_r\) is an empirical coefficient and NPP represent the net primary production taken from tabulated values for different ecosystem types corresponding to our IGBP land use classification.

The emission potential for spruce and pine is allowed to vary depending on season (Tarvainen et al., 2007). For deciduous forests we assume an emission potential of monoterpenes of \(\mu g\, g(\text{dry weight})^{-1}\, h^{-1}\) (i.e. emissions per hour and gram of leaf dry weight).

Biogenic Volatile Organic Compounds (BVOC) are assumed to consist of monoterpenes (represented by \(\alpha\)-pinene), isoprene and sesquiterpenes. Concerning the mono- and sesquiterpenes emissions we assume pool dependent temperature controlled emissions (Guenther et al., 1993, 1995). Monoterpene emissions are strongly dependent on temperature and the total flux of terpenes can be calculated using the relation

\[ F = \epsilon D \gamma \]  

where \(F\) represent the total flux of monoterpenes from the forest in \(\mu g\, m^{-2}\, h^{-1}\), \(\epsilon\) is the emission potential (\(\mu g\, g(\text{dry weight})^{-1}\, h^{-1}\)), \(D\) is the foliar biomass density in \(g(\text{dry weight})\, m^{-2}\), and \(\gamma\) is an environmental correction taking into account the temperature dependency of the emission rate \((\gamma(\text{pool})=\exp(\beta(T-T_s)), \beta=0.09C^{-1}\) and \(T_s=303.15\)). The temperature used is the surface temperature as derived from the FNL data set used for trajectory calculations. The sesquiterpenes are calculated
in a similar manner, but with a seasonal dependence of the emission potential associated with conifer trees with a zero emission potential through November–March (Hakola et al., 2006). No light dependence is considered for monoterpane emissions. Isoprene is calculated following Guenther et al. (1995). This isoprene emission estimates adopt both light and temperature dependence. The photosynthetic photon flux density (PPFD) is calculated following (Alados-Arboledas et al., 2000). The clear sky values of the PPFD are used regardless of cloudiness in the model. The emission potentials used are $0.1 \mu g (\text{dry weight})^{-1} h^{-1}$ for conifers and $1 \mu g (\text{dry weight})^{-1} h^{-1}$ for deciduous trees.

2.4.2 Other trace gases

In its current form, the model allows for gridded emissions of the most common inorganic trace gases as well as anthropogenic emissions of non-methane volatile organic compounds (NMVOC). These emissions are taken from the EMEP data base (50×50 km grid, Vestreng et al., 2006) and include besides NMVOC (in our study represented by ethane) also CO, anthropogenic SO$_2$, natural SO$_2$, and NO. The emissions are adjusted to comply with the seasonal emission pattern (David Simpson, personal communication). DMS emissions are calculated according to (Kettle et al., 1999), making use of monthly mean sea surface DMS concentrations. Actual DMS fluxes are calculated from these surface water concentrations by using surface temperature and wind speed.

2.4.3 Primary emissions

The model also incorporates primary aerosol emissions. For this purpose, the model includes anthropogenic emissions of anthropogenic primary organic particles (POA and primary EC) as well as POA and EC emissions from forest fires. These emissions are taken from the AEROCOM database (Dentener et al., 2006) using year 2000 as reference. The emissions are mass based, but remapped to standard size distributions
as shown in Table 1. The seasonal dependence for the emissions is accounted for by applying the previously discussed seasonal scaling factors also for these data sets. The forest fire emissions are however given as monthly values throughout the year. Sea spray emissions are estimated assuming dependence on wind speed and temperature following (Martensson et al., 2003).

3 Results

In the following the general model performance will be discussed followed by an investigation of the model’s sensitivity to certain key processes/parameters. The model is initialized with either a marine, rural or polluted continental size distribution, depending on the starting location. The model parameters of these initial size distributions are given in Table 2. The model is further initialized, regardless of starting location and/or time of year with 35 ppb ozone, 40 ppt SO$_2$ and 500 ppt NO$_x$. CO is a rather long lived species and also plays an important role in regulating the abundance of oxidants in the atmosphere. Therefore, initial CO concentrations are selected based on both season and location, i.e. either summer or winter values for continental and marine starting locations (100 and 220 ppb for summer and winter time continental locations and 100 and 150 ppb for summer and winter marine locations, thereby adopting typical values to initialize the model). All BVOC are initially set to zero. Methane is set to 1750 ppb.

3.1 A single trajectory run: simplified description of processes

In order to show the evolution of some selected key parameters along a trajectory run we selected a case where marine air is advected over Scandinavia (Fig. 1). Time and date of this simulation ends at 13 April 2000, 06:00 UTC. The model was simulated along the trajectory shown in Fig. 1 for 216 h, i.e. 9 days, starting in the marine Arctic environment. The evolution of selected parameters is displayed in Fig. 2, where the concentration of SO$_2$, O$_3$, H$_2$SO$_4$, [condensible organics], mixing layer height and
cloudiness are shown. This plot is accompanied by Fig. 3 showing the evolution of aerosol number size distribution along the trajectory. As can be seen from the figures, as long as the air resides over the marine environment, some nucleation is taking place, but the magnitude of both nucleation and growth is too low to support production of particles remaining as individuals for a longer period of time. However, as soon as the air arrives at the coast, the onset of biogenic emissions (represented by MT, third panel in Fig. 2) provides the amount of organic gases to support the growth. The air parcel reaches the coast after approximately 144 h of simulation, and during the following days, three consecutive nucleation events is suggested by the model, of which each one is contributing to increasing number concentration. The simulated day is also associated with nucleation at the receptor site Hyytiala, a typical feature observed at this station when marine polar and arctic air masses arrive at the station. The idea that the forest supports the growth has been pointed out in numerous studies as discussed in the introduction and the current model result suggests that several consecutive, nucleation events provide the number concentration that is observed at the station when this type of transport takes place. As a comparison, the resulting simulated and observed aerosol number size distributions are shown in Fig. 4. Although the growth is slightly smaller of the simulated distribution, there is a good agreement between the modeled and measured distributions. It should however be mentioned that while the agreement in this specific case is comparably good, some of the other simulations result in size distributions that on short time scales are quite different from the observations.

The influence of clouds is also displayed in the figures: as can be seen in the bottom panel of Fig. 2 five cloud events are prescribed by the model. The role of clouds becomes clear by comparing the onset of cloudiness after around 150 h of simulation. When the cloud forms at this time of simulation, the sulfur dioxide concentration is sufficiently high to provide an important source of particle sulfate via in cloud production. This is evident looking at the corresponding time in Fig. 3. The population of activated CCN evaporates and leaving significantly larger particles than those entering the cloud. This is clearly shown by the evolution of the minimum developing around
110 nm concurrent with the onset of cloudiness, shifting these particles into a larger size range.

3.2 Comparison with observational data

Both aerosol properties and trace gases will be discussed under two separate paragraphs under this section. As a base case reference years we have chosen years 2000–2001. For this period, 4 trajectories have been calculated each day (00:00, 06:00, 12:00 and 18:00). The model has been run along these trajectories, thus supplying the following model-measurement comparison with four endpoints each day throughout the two years. This provides the necessary model output required to perform both seasonal and monthly averaging of the data. This modeling approach also allows for Eulerian interpretation of the simulated data and thus straightforward comparison with observations.

3.2.1 Aerosol properties

In the following we show the results from the modeled and measured aerosol physical properties. In order to allow a direct comparison between the measured and modeled results, the modeled data was remapped to the bin width and size of the measured data using Matlab Spline fitting. This gives the measured and modeled data identical structure, while still introducing a minimum of bias to the data.

Shown in Figs. 5 and 6 are the comparisons between measured and observed aerosol number size distributions for years 2000 and 2001, respectively. The figures display the monthly average measured size distributions and their corresponding 25–75 percentile range. The modeled data is shown by the thick black line. The model is capable of reproducing the typical features of the observational data, as monthly average, with a relatively good accuracy for 9 of 12 months for year 2000. The modeled data typically falls within the modeled 25–75 percentiles of the measured results, and occasionally the modeled monthly median agrees very well with observed values.
(e.g., February, September, July and August). Not only the overall magnitude of the measured data is captured by the model during these months but also the shape of the size distribution is properly represented, and follows the changes in observational data, i.e. towards mono- or bimodal structure. Also for year 2001 model and measurements show a good agreement, apart from January–December. This indicates that the model in general is capable of reproducing the influence from relevant processes and sources. It is however obvious that the model performs worst during the winter months, especially December and January. In these cases the model overestimates the observational data and further suggests a typically uni-modal size distribution, while the measured size distribution is bi-modal.

It may also be informative to study how the model performs when the air arrives from different source regions, since such an analysis could aid in finding areas where sources and/or processes are less well captured. Therefore, the trajectories for years 2000–2001 have been clustered according to their geographical coordinates using the MATLAB kmeans.m routine on the ensemble of trajectories. Clustering is a method that captures, in the case of trajectory clustering, typical and distinct transport paths that the air follows before arriving to the receptor, maximizing the difference between the different clusters, while minimizing the difference between trajectories belonging to a certain cluster. The calculated clusters are shown in Fig. 7 as the cluster centroids (i.e. “average” trajectory of the different clusters). 10 clusters have been considered in this analysis. The associated number size distributions for the period February–November for both years are shown in Fig. 8. During this period, the model performs well in most transport directions. Especially some of the marine trajectory clusters are associated with very good agreement between model and measurements (cf. Fig. 7, clusters 3, 7, 8, 9). The reason for just choosing February–November is that during December and January model performance is poor for most clusters, and would thus bias the otherwise satisfying agreement between model and measurements for the rest of the year. The cause of this generally poor representation of aerosol data during winter months is not well understood, but may be linked to more complicated real world
winter time meteorology which CALM is not able to capture (e.g. stability, effect of clouds etc.) or inadequate representation of the wintertime sources.

In Figs. 9 and 10 the seasonal variation of the three dominating size ranges (or modes) of modeled and measured data for the receptor Hyytiälä during year 2000 and 2001 is shown. This comparison considers the accumulation mode size range \((D_p > 100 \text{ nm})
\), the Aitken mode size range \((30 \text{ nm} < D_p < 100 \text{ nm})\)
, and the nuclei size range \((D_p < 30 \text{ nm})\).

Starting with the accumulation mode it is clear that the model captures the magnitude and trends of the accumulation mode in an excellent way during both years. Model and measurements follow each other very well for long periods of time. As with the monthly average size distributions previously described, the model representation does get worse during the end of the year. The modeled accumulation mode number concentration very of the higher than measured during the winter period. However, since both magnitude and trends are captured by the model, it is concluded that the way we represent sources and dominating processes is sufficiently accurate for the purpose of this study. This is an important finding since accumulation mode provides the largest fraction of potential CCN, and thus an accurate description of this size range is of crucial importance for a proper determination of the important indirect aerosol effect on climate.

The seasonal variation of the Aitken mode \((30 < D_p < 100 \text{ nm})\) number concentration through year 2000 is shown in the second frame of Figs. 9 and 10. As previously recognized, the model performs well especially during months 2–10, while the winter time agreement gets poorer as the model typically overestimates the Aitken mode concentration. The annual averages of modeled and measured Aitken mode concentration are \(1190 \text{ cm}^{-3}\) and \(760 \text{ cm}^{-3}\), respectively. The model result is far patchier compared to the measured data, with several high peaks throughout the year. The general trend is however satisfying. One cause that could result in the discrepancy between the model and measurements is the fact that our current single trajectory approach will be very sensitive for local point sources. As described in the method part of this study, the mass based primary emissions are transformed into a fixed size distribution (Table 1),
something that also could introduce an erroneous representation of the actual size distribution emitted, resulting in disagreement between the model and measurements.

The third frame of Figs. 9 and 10 shows the modeled and measured nuclei mode ($D_p<30\,\text{nm}$) concentration. This figure is indicative of several interesting features related to the processes governing nucleation over Scandinavian boreal forests. Firstly, the model overestimates the average nuclei mode size range concentration (i.e. $D_p<30\,\text{nm}$). Annual modeled average is $390\,\text{cm}^{-3}$ and measured annual average is $270\,\text{cm}^{-3}$. However, as reported in several previous studies (e.g., Dal Maso et al., 2005, 2007) nucleation observed at Hyytiälä is generally associated with a pronounced seasonal pattern. Nucleation occurs rarely during winter months, followed by a spring maximum, and then again a summer minimum followed by another increase in nucleation during autumn. Previous discussions have invoked several possible explanations for this behavior. It is therefore highly interesting to notice that this seasonal trend is also captured by the model (in terms of abundance of nuclei mode particles), as is especially evident during year 2001 (Fig. 10). As shown in the figures both modeled and measured sub-30 nm particles show a good agreement with respect to the annual trend. Evident are the spring maximum, summer minimum and finally a smaller but significant increase during the autumn. In our model nucleation subsequent growth to detectable size is controlled by the balance between available condensing and nucleating vapors and the condensation and coagulation sink of the particles. Growth is favored during the summer month when the forest emits high amounts of BVOC (in our case represented by isoprene, monoterpenes and sesquiterpenes) while nucleation is governed by the concentration of $\text{H}_2\text{SO}_4$, which in turn is controlled by $\text{SO}_2$ emissions, condensation sink and insolation. During year 2000 also the modeled and measured trends of nuclei mode concentrations agree very well qualitatively.

It is further evident that CALM is able to reproduce the aerosol number and size with a good accuracy using only primary emissions and nucleation in the lower atmosphere and neglecting the influence of free troposphere (FT) nucleation and following entrainment. This contradicts to some extent the findings from earlier global model studies
which have suggested that FT can account for up to 25% of the boundary layer particle number and CCN concentration (Merikanto et al., 2009).

3.2.2 Two case studies

Two highly resolved runs were performed to investigate how the model performs in air masses arriving from the clean marine sector and the polluted continental sector. The clean period is represented by a period of 10 days in May 2000, associated with trajectories of marine origin that were advected to Hyytiälä. Nucleation is commonly observed over the boreal forest during spring when this type of air-mass transport dominates. As clean marine air is advected over Scandinavia, an abrupt change in sources occurs when going from the marine environment to forested areas which are proven to emit large amounts of BVOC (e.g., Simpson et al., 1995). As previously shown in Tunved et al. (2006a,b), nucleation followed by condensation of low or semi-volatile species is the process that most likely governs the evolution of number over the boreal forest during this type of transport. During the period investigated here, 1–10 May 2000, several nucleation events are indeed observed at the receptor site (Fig. 11). Pro- nounced nucleation events are observed 1, 2, 4, 6, and 7 May, corresponding to DOY 122, 123, 125, 127, 128. Traces of nucleation events are also noticed an all other days except 3-May. In order to investigate how the model captures this type of transport in detail, CALM was run for the corresponding ten day period with a resolution of 1 h (i.e., the model was run along one individual trajectory every 1 h through the studied period). The median modeled and 25–75 percentile ranges of the number size distribution are shown in Fig. 13 (left frame). With respect to the general shape and magnitude, the modeled and measured size distributions agree to a large degree. The model is able to reproduce the type of size distribution observed during this type of ad- vection. However, as can be seen in Fig. 13, the modeled growth is slightly slower than the observed one, which is obvious from the shift in the Aitken mode size range. The explanation closest at hand is that the model underestimates the amount of condensing gases, which slows the growth relative to the measurements. The model number and size distribution observed 1, 2, 4, 6, and 7 May, corresponding to DOY 122, 123, 125, 127, 128. Traces of nucleation events are also noticed an all other days except 3-May. In order to investigate how the model captures this type of transport in detail, CALM was run for the corresponding ten day period with a resolution of 1 h (i.e., the model was run along one individual trajectory every 1 h through the studied period). The median modeled and 25–75 percentile ranges of the number size distribution are shown in Fig. 13 (left frame). With respect to the general shape and magnitude, the modeled and measured size distributions agree to a large degree. The model is able to reproduce the type of size distribution observed during this type of ad- vection. However, as can be seen in Fig. 13, the modeled growth is slightly slower than the observed one, which is obvious from the shift in the Aitken mode size range. The explanation closest at hand is that the model underestimates the amount of condensing gases, which slows the growth relative to the measurements. The model number and size distribution observed 1, 2, 4, 6, and 7 May, corresponding to DOY 122, 123, 125, 127, 128. Traces of nucleation events are also noticed an all other days except 3-May. In order to investigate how the model captures this type of transport in detail, CALM was run for the corresponding ten day period with a resolution of 1 h (i.e., the model was run along one individual trajectory every 1 h through the studied period). The median modeled and 25–75 percentile ranges of the number size distribution are shown in Fig. 13 (left frame). With respect to the general shape and magnitude, the modeled and measured size distributions agree to a large degree. The model is able to reproduce the type of size distribution observed during this type of ad- vection. However, as can be seen in Fig. 13, the modeled growth is slightly slower than the observed one, which is obvious from the shift in the Aitken mode size range. The explanation closest at hand is that the model underestimates the amount of condensing gases, which slows the growth relative to the measurements. The model
predicts a distribution peaking around 25 nm, while in the measurements the maximum is located around 35 nm. Agreement between measured and modeled number concentrations above 100 nm is good.

The lower frame of Fig. 11 shows the modeled time evolution of the aerosol size distribution at Hyytiälä during the corresponding period. As can be seen, the model largely captures the temporal evolution of nucleation events during most of the days. The model does however suggest nucleation taking place also 3 May (although a rather weak nucleation), which is not seen in the observational data. Especially well captured are the particle formation and growth events during 1–2 May, 4–5 May and 6–7 May, although the model seems to overestimate the number of particles.

The polluted continental case was represented by 7 days in July 2001 (12–19 July), where trajectories arrived at Hyytiälä from SW in the beginning of the period, with a shift towards continental sources to the S–SE during the end of the period. The modeled and observed evolution of the size distribution at Hyytiälä is shown in Fig. 12. The period is characterized by a persistent accumulation mode located around 100 nm, with a rather small variation in size and concentration (Fig. 12, upper frame). This is a typical feature in continental air as observed at Hyytiälä. Furthermore, the data show low activity of new particle formation close to the receptor.

From Fig. 12, lower frame, it is seen that the model captures the general properties of the variation of the size distribution observed during the selected period. The modeled accumulation mode is however even more stable than observations, and some peaks indicating recent new particle formation are present in the observational data. This feature is lacking in the modeled data. The median modeled data during this case period is shown in the right frame of Fig. 13 together with measured median and 25–75 percentile range of the data.

Comparing the two case studies, it is shown that the model produces widely separate size distributions comparing the two different extremes in source areas. The marine case is associated with a high number concentration and a size distribution shifted towards smaller sizes, i.e. traces of recent nucleation. The continental case is
associated with lower number of particles, but the distribution is shifted towards the ac-
accumulation mode size range. The fact that these results agree well with observational
data is encouraging.

3.2.3 Trace gases

In the following we show the agreement between measured and modeled sulfur diox-
ide (SO$_2$) and ozone (O$_3$). These gases are highly important controllers of both the
nucleation rate and condensation growth (via H$_2$SO$_4$). Ozone in turn is an important
oxidant itself, and is furthermore intimately linked to the production and cycling of reactive
radicals (e.g. NO$_3$, OH, HO$_2$). Shown in (Fig. 14) is a comparison of measured and
modeled [SO$_2$] and [O$_3$] at the SMEAR II station in Hyytiälä. In order to better see the
general annual behavior, the data is presented as a 24 h running average to smooth
out the presence of intermittent high peaks.

The annual average of measured SO$_2$ was found to be 130 ppt for year 2000. The
modeled annual average was found to be 120 ppt. During 2001 the annual average of
measured SO$_2$ was found to be 120 ppt, but the modeled annual average 160 ppt.

The annual average of measured O$_3$ was found to be 28 ppb for year 2000. The
modeled annual average was found to be 27 ppb. During 2001 the annual average of
measured O$_3$ was found to be 27 ppb, but with a modeled annual average of 26 ppb.

As can be seen in the figures both modeled and measured SO$_2$ show a pronounced
seasonal variation, with maximum concentration during winter months. This is most
likely the result of the combined effect of higher emissions, but also lower rate of pho-
tolyltic degradation of SO$_2$ towards sulfuric acid. It is worth mentioning that the model
appears to overestimate SO$_2$ during the winter. The seasonal variation of O$_3$ follows
an opposite pattern, with maximum concentration during summer months, and mini-
mum during winter, reflecting the opposite photochemical dependence of SO$_2$. The
agreement between the model and measurement on the finer scale is not as good as
for SO$_2$, but as shown by the annual average the magnitude of modeled and measured
ozone agrees perfectly.
The modeled concentration of monoterpenes was found to be \( \sim 80 \) ppt as an annual average, which is lower, but in the same range as measured literature values (e.g., Hakola et al., 2009), which are typically below 100 ppt during winter and above 200 ppt during the middle of the summer. However, these measurements are often performed close to the canopy, whereas our modeled concentrations represent the average of the whole ML. Since the emissions are temperature dependent, the highest emissions occur during summer time. Annual average for modeled isoprene was found to be 110 ppt, which is in the range, although slightly lower, compared to observed values.

Figure 15 shows the annual variation of different oxidants and condensable species in the model. Hydroxyl radical (OH) concentration reaches maximum during summer, with maximum concentrations around \( 10^7 \) cm\(^{-3} \). Typical noon concentrations were around \( 2 \times 10^6 \). This is well in the range, although slightly higher, of observations at Hyytiälä (Petaja et al., 2009), who observed \( 3–6 \times 10^5 \) OH cm\(^{-3} \) during March–June).

Seasonal variation of nitrate radical is shown in the second frame of Fig. 15. The concentration is typically in the range of \( 10^7 \) cm\(^{-3} \), but occasionally reaches above \( 10^8 \) cm\(^{-3} \).

Sulfuric acid exhibits a typical seasonal variation, with maximum during summer months. Maximum concentrations is in the range of \( 1.0–1.5 \times 10^7 \) cm\(^{-3} \) with an annual noon average of \( 1.8 \times 10^6 \) cm\(^{-3} \). This agrees well with observed concentrations at the Hyytiälä measurement station (Petaja et al., 2009).

The modeled concentration of condensable species at Hyytiälä resulting from monoterpane oxidation varies between \( 0.2–15 \times 10^7 \) cm\(^{-3} \). There are no available measurements to confront this result, but according to e.g., Kulmala et al. (2001); Spracklen et al. (2006) and references therein, investigations of new particle formation events at Hyytiälä indicate that the required concentration of condensable species to sustain the observed growth must be around \( 2 \times 10^7 – 1.3 \times 10^8 \) cm\(^{-3} \). Thus, our results are in the lower range of these estimates.
3.3 Sensitivity tests

In this section we present the results for the sensitivity tests performed for the model. For these comparisons we have chosen year 2000. The main results are tabulated in Table 4, and will be referenced accordingly in the following text. The results of sensitivity tests with respect to nucleation, primary emissions, monoterpenes, clouds and precipitation and model initialization will be shown. The results will be discussed in terms of seasons and source regions of the air parcels simulated. In order to resolve the dependence of source regions we performed a clustering of the trajectories according to their typical advection paths. This will aid in the understanding of the importance of different processes at different locations/environments. The clusters for 2000 are shown in Fig. 16.

3.3.1 Nucleation and primary emissions

In the model, addition of aerosols by number is controlled by either primary emitted particles (as described under Sect. 2.4.3) or nucleation and consecutive growth. The primary emissions are confined to predefined size ranges, while nucleation always start with sub-nm particles formed from the gas phase. Since the number is relevant for both health and climate issues related to the atmospheric aerosol, it is of interest to investigate how the model responds to perturbations of these number concentration controlling processes. In this section we present results derived from three different sensitivity tests, one with nucleation completely disabled, one with an activation coefficient representing the minimum A derived at Hyytiälä ($4 \times 10^{-7}$ s$^{-1}$) and one with a maximum empirically derived A ($6 \times 10^{-6}$ s$^{-1}$). These numbers are taken from Rippinen et al. (2007) who compared the first order activation nucleation coefficient for different locations. In the base case, as mentioned previously, we assume a flat coefficient A of $2 \times 10^{-6}$ s$^{-1}$. Additional to these tests we also investigate the response to a complete cancellation of primary emissions.
By applying the range of values of A to two different simulations, only very small changes in the final average size distributions were noticed (not shown). This suggests that the exact value of the nucleation coefficient is insignificant for modeling on this scale. This conclusion is valid even with an order of magnitude difference in A between the two runs. Instead, beside the nucleation itself, the ability of the particles to grow in the environment in which they are formed is of crucial importance. This ability in turn is related to the condensation sink of pre-existing particles as well as the rate of coagulation of the freshly formed particles.

Figure 17 shows the winter (October–March) and summer (April–September) average size distributions for the run with nucleation disabled and the base case distribution. Both are for year 2000. As can be clearly seen, nucleation provides the largest addition of aerosol number during the summer months, while the runs with no nucleation during the winter period result in virtually the same aerosol number size distribution as the base case. This is well in agreement with previous findings, showing nucleation to occur preferentially during the spring-autumn period, indicating the dependence on photochemical processes governing nucleation. As shown in Table 4, with nucleation disabled, the annual average of nuclei mode particles is decreased from 390 to 180 cm$^{-3}$, the Aitken mode is reduced from 1190 to 820 cm$^{-3}$ and the accumulation mode is reduced from 400 to 360 cm$^{-3}$. The accumulation mode typically represents the amount of available CCN, and since nucleation provides $\sim$10% of these particles it is clear that nucleation provides a substantial contribution to potential CCN’s, also as an annual average. As shown, this contribution is most pronounced during the summer period (April–September). The presence of nuclei mode particles also in the simulation where nucleation is turned off is explained by the primary emissions, contributing to sub 30 nm particles.

More interestingly, as shown in Fig. 18, the role of nucleation differs largely between different clusters (Fig. 16). This figure shows the summer time base case runs, together with runs with nucleation disabled. As can be seen, nucleation has a large impact on sub-100 nm particles in clusters 2, 4, 6, 9 and 10. Clusters 4, 6, 9 and 10
are typically associated with marine air advection from north and cluster 2 centroid is directed over continental sources oriented NE of Hyytiälä. The least dependence of final size distribution on nucleation is observed in clusters 1, 5 and 8. These clusters arrive from continental sources, and both base case and the run with nucleation disabled are nearly identical. Remaining clusters share both continental and marine sources, and consequently there are some differences between no nucleation case and base case. These findings indicate that nucleation is an important contributor to the total aerosol number at Hyytiälä only when advection occurs from clean regions. The fact that nucleation preferentially occurs in clean, polar air masses has been known for several years (Boy et al., 2005; Sogacheva et al., 2008). What is interesting with these results is however, that even if nucleation is disabled over the continental region, this does not have a large impact on the resulting size distribution observed in Hyytiälä. These findings qualitatively agree with the results and discussions presented by Spracklen et al. (2006), using the global model GLOMAP. In clean regions, nucleation is a significant contributor to aerosol number, while in the outflow of large pollution sources it is not.

In Fig. 19 we show the summer and winter distributions resulting from base case simulations and simulations with all primary sources disabled. As can be seen, the lack of primary emissions affects both the summer and winter periods, although the difference is most pronounced for the winter period. This is opposite to the dependence on nucleation. From Fig. 20 it is also obvious that the clusters with the largest continental influence (i.e. cluster 1, 5 and 8) are most sensitive to the primary emissions (although insensitive to nucleation). This means that primary emissions have a much larger impact than nucleation during these transport conditions. Also interesting to notice is that the clusters that were most prone to be influenced by nucleation, respond quite differently compared to their continental counterparts. This is especially evident in clusters 9 and 10, where the reduced primary emissions apparently lead to an overall higher concentration in the nuclei mode size range (here defined as particle with $D_p<30\text{nm}$). This is without a doubt the result of lower condensation sink associated with the runs neglecting primary emissions, something that favors nucleation.
and subsequent detectable growth. During the winter period, each cluster is subject to reduction in particle number due to the absence of nucleation.

In summary this means that particle number is governed by different processes in polluted and clean environments, i.e. primary and secondary formation, respectively.

### 3.3.2 BVOC

In the model we apply a stochiometric yield of 15% condensable species from primary oxidation of MT (as represented by α-pinene). By cancelling out monoterpene emissions we now investigate the role of the dominating source of BVOC in the model. As a response in terms of annual average of the three different modes we notice a reduction in accumulation mode (400 and 270 cm\(^{-3}\) for base case and sensitivity, respectively), a slight decrease in Aitken mode (1190 and 1120 cm\(^{-3}\)) and an increase in nuclei mode particles (390 and 510 cm\(^{-3}\)). The reduction of accumulation mode particles reflects the fact that less condensable species are available to support growth into larger sizes. The insignificant change in the Aitken mode however is argued to be the combined result of more frequent nucleation due to reduction of the condensation sink combined with the fact that primary emissions largely contribute to this size range, and also that particles emitted in this size range do not grow into the accumulation mode size range. The increase of nuclei mode particles is explained by the reduced CS. More particles are formed on average via nucleation, although they grow much slower.

Figure 21 confirms this, by showing winter and summer average size distributions for base case and runs with no MT emissions. During both seasons, the number size distribution is shifted towards smaller sizes. This effect is most pronounced during the summer time when MT emissions are expected to be high. The shift towards smaller sizes is evident in all transport directions, but most pronounced in cluster 9, where the sizes are not only smaller, but the number of nuclei mode particles is far higher without MT emissions as compared to the base case (not shown). MT emissions are thus argued to be an important contributor to the growth of both nucleated and primarily emitted particles into CCN sized particles.
3.3.3 Initial conditions

In order to assure that the model result isn’t biased by the initial conditions, the role of the starting size distribution was investigated. This is the most probable parameter to be of importance during initialization of the model since gaseous components have relatively short lifetimes in relation to particles. Therefore, two different setups were chosen: one where all simulations are initialized with the continental type size distribution and one where all runs were initialized with the marine type size distribution. The results are shown in Figs. 22–23 for continental and marine starting distributions, respectively. As can be seen, if the model is initialized with only continental size distributions, a shift of both wintertime and summer time average size distributions towards larger sizes is obvious. However, on closer inspection with respect to the cluster orientations, it was concluded that the difference is only pronounced in the clean, marine clusters 4, 9, 10 (not shown). This is interpreted as the lifetime of the starting distribution in these cases being too long for the model to equilibrate to marine conditions. However, if all initial size distributions were substituted with their marine counterparts, negligible change compared to base case conditions was observed. This means that the source strength in polluted areas in principle is strong enough to allow for equilibration to continental conditions rather rapidly.

3.3.4 Clouds and precipitation

The role of clouds and precipitation was investigated by simply cancelling out clouds and rain in the model scheme. The largest relative effect of this was an increase of the accumulation mode number concentration (Table 4) which increases from 400 to 580 cm$^{-3}$ while the Aitken mode increases from 1190 to 1360 cm$^{-3}$. Somehow, the nuclei mode annual average concentration remained unaffected. This probably reflects the fact that although the condensations sink increases, SO$_2$ will be more abundant due to less scavenging by the clouds and precipitation, thus providing more nucleating material in terms of H$_2$SO$_4$. This indicates that clouds do play an important role in
the model, and thus are necessary to achieve the observed good agreement between model and measurements during the studied period. The change is most pronounced during the summer months, although of similar magnitude during winter. There was no clear (relative) dependence on source regions when performing the analysis per cluster, but instead all clusters were associated with a substantial increase in the Aitken-accumulation mode size range.

4 Summary and conclusions

In this study we have presented a new Chemical and Aerosol Lagrangian Model (CALM) that describe the evolution of particle distribution and chemical key species along trajectories and selected receptor sites. The model performs quickly enough to be run on standard PC units, and supplies users with an easy tool to investigate the relative role of dynamic processes controlling the appearance and fate of particles in the atmosphere. The model incorporates the most central aspects of the aerosol dynamics in both the dry and wet phase of the atmosphere and does as shown provide an excellent tool for determining dominating processes with a large degree of transparency and accessibility.

Considering the model performance, CALM is able to capture the most prominent aspects of the observed aerosol at the Hyytiälä measurement station. This statement is valid, with the exception of the winter period, when the model performs poorly. The explanation for this remains open, but may relate to more complicated meteorology and/or behavior of the sources. Overall, there is generally quite a good agreement between measured and modeled accumulation mode number concentrations with respect to both magnitude and seasonal trends. This applies also to the modeled and measured nuclei mode number concentration. Investigation of the model results versus actual observations provided a mechanistically sound explanation/description of the seasonal variation of nuclei mode particles, highlighting the importance of balance between the generation of nucleating and condensable species and their corresponding
condensation sink. There is however only fairly good agreement between modeled and measured Aitken mode number concentration. The poorer agreement between the measured and modeled Aitken mode particle number may be related to the way primary emissions are described. The model also performs satisfactorily in comparing either size distribution in relation to different months or different source regions and transport routes. This indicates that the model can cope with different environments and source regions and more importantly, that the model provides a qualitative and quantitative balance between the governing processes, something that is shown by the satisfying agreement between observed and measured sulfur dioxide and the fair agreement between modeled and measured ozone, as well as reasonable representation of the concentrations of monoterpenes, isoprene, radicals and sulfuric acid.

In addition, the sensitivity of the model to different processes has been investigated in detail. The overall findings suggest that the processes are well balanced. The sensitivity tests have also provided insight into the processes governing the aerosol as observed at the Hyytiälä measurement station. These conclusions include, but are not limited to:

- Nucleation is important for the provision of particle number in clean air masses only. Under continental, polluted conditions primary emissions provide most particle number

- The model result (as represented by the receptor site Hyytiälä) proved virtually insensitive over an order of magnitude range of nucleation coefficient \( (4 \times 10^{-7} - 6 \times 10^{-6} \text{s}^{-1}, \text{activation theory, } J \sim [\text{H}_2\text{SO}_4]) \). This suggests that other processes than nucleation rate itself limit the provision of stable particles via this nucleation mechanism within this range of nucleation coefficients. This could for example be availability of condensable species and/or rate of coagulation of freshly formed particles. This finding eases the selection of \( A \) for modelers (The nucleation activation theory coefficient \( A \) has been shown to vary between different environments (Riipinen et al., 2007).
Monoterpenes are an important contributor to particles above 100 nm since their oxidation products facilitates growth of particles over the entire size range. This indicates that monoterpenes may indirectly influence the radiation budget of the atmosphere. The role of monoterpenes was found to be most pronounced in the clean transport sectors of Hyytiälä.

Clouds, as described in CALM, do play an important role in regulating especially the accumulation mode concentration. By cancelling clouds and precipitation, the number of accumulation mode particles is substantially increased.

The model has been extensively tested and the results of these tests have proved the model to perform reasonably well. It has not been within the scope of this study to extend the testing to other sites than the Hyytiälä station. As with most model approaches, also the presented approach is associated with good and less good approximations. Thus, the model approach is associated with a number of advantages and disadvantages which are outlined below.

The model can with reasonable computational demand cope with a high level of detail of the aerosol dynamic processes in the dry atmosphere. CALM can be used with an arbitrary number of sections/compositions without significant loss in computational efficiency (within reasonable limits).

The model is associated with a high level of transparency with respect to the influence of different processes and is easily adoptable to different sites and locations. CALM is furthermore easily expandable to test different chemical schemes and dynamic process descriptions and is thus suitable for use as an evaluation tool for aerosol dynamic process parameterizations in regional and global models.

However, Lagrangian box-models obviously suffer, by necessity, from several more or less coarse parameterizations. In the case of CALM, the most central disadvantages are the rudimentary description of vertical and horizontal mixing and exchange. Two layer representation of model space could be unsuitable for certain locations. This may lead to unrealistic results, especially if the modeled receptor is located close to polluted
regions with large density of point sources. However, the average result of simulations over longer periods of time (weeks-months) likely reflects the actual source field experienced over the same period. This is evident from the good agreement between the model and measurements, especially looking at monthly and annual averages. Furthermore, the method is likely to be suitable for regions surrounded by homogeneous source fields similar to the boreal region itself. It is undeniably doubtful that the often good agreement of hourly simulated and measurement data (especially considering the accumulation mode concentration) would be coincidental only. Also, the validity of the simplified cloud description/treatment in the model must be tested more extensively, e.g. for other sites and locations. The model only considers washout processes (i.e. below cloud scavenging), with rain coming from clouds formed above the modeled column. This approximation is likely valid for rain from frontal type clouds (e.g. nimbostratus clouds), while not being the preferred solution for convective type precipitation with boundary layer air actively partaking in the cloud and precipitation forming processes. Nevertheless, the semi-empirical parameterization used implicitly takes into account both washout and rainout processes since it is based on the observed rate of change of boundary layer aerosol concentration and its relation to precipitation rate.

It remains a future task to better validate the presented process description, especially considering the model performance over longer time periods and over larger spatial scales. Ongoing measurement initiatives will support this future work by supplying relevant data required to validate the model (especially the EUCAARI- and EUSAAR-programs).

Acknowledgement. This work has been supported by Swedish Research Council (VR 2006-5360), EUCAARI (European Integrated project on Aerosol Cloud Climate and Air Quality interactions) No. 036833-2 and EUSAAR and the European Community – Research Infrastructure Action under the FP6 “Structuring the European Research Area” Program, EUSAAR Contract No. RI3-CT-2006-026140. University of Helsinki is acknowledged for providing the data from the SMEAR II station, Hyytiälä, Finland.
References


Spracklen, D. V., Carslaw, K. S., Kulmala, M., Kerminen, V.-M., Mann, G. W., and Sihto, S.-L.: The contribution of boundary layer nucleation events to total particle concentrations on re-


Table 1. Adopted size distribution properties of the various AEROCOM primary emissions. Shown is the sigma and geometric mean for the emission size distribution and emission year used for each sector.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$D_g$</th>
<th>Sigma ($\sigma_g$)</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biofuel Soot</td>
<td>80 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Bond and Streets, 1996</td>
</tr>
<tr>
<td>Biofuel POM</td>
<td>80 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Bond and Streets, 1996</td>
</tr>
<tr>
<td>Fossil fuel Soot</td>
<td>30 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Bond and Streets, 1996</td>
</tr>
<tr>
<td>Fossil fuel POM</td>
<td>30 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Bond and Streets, 1996</td>
</tr>
<tr>
<td>Forest Fires Soot</td>
<td>80 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Global Fire Emissions Database (GFED) version 1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Forest fires POM</td>
<td>80 nm</td>
<td>1.8</td>
<td>2000</td>
<td>Global Fire Emissions Database (GFED) version 1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> [http://ess1.ess.uci.edu/~jranders/data/GFED2/](http://ess1.ess.uci.edu/~jranders/data/GFED2/)
Table 2. Size distribution properties of the input size distributions.

<table>
<thead>
<tr>
<th>Source region</th>
<th>N1</th>
<th>GSD1</th>
<th>Dg1</th>
<th>N2</th>
<th>GSD2</th>
<th>Dg2</th>
<th>N3</th>
<th>GSD3</th>
<th>Dg3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marine</td>
<td>60</td>
<td>1.19</td>
<td>14</td>
<td>235</td>
<td>1.43</td>
<td>37</td>
<td>110</td>
<td>1.44</td>
<td>170</td>
</tr>
<tr>
<td>Rural</td>
<td>6650</td>
<td>1.252</td>
<td>15</td>
<td>147</td>
<td>1.745</td>
<td>54</td>
<td>1990</td>
<td>1.305</td>
<td>84</td>
</tr>
<tr>
<td>Remote continental</td>
<td>111</td>
<td>1.52</td>
<td>35.6</td>
<td>326</td>
<td>1.45</td>
<td>67.6</td>
<td>204</td>
<td>1.6</td>
<td>167</td>
</tr>
</tbody>
</table>
### Table 3. Number of trajectories belonging to each one of the clusters.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td># trajectories</td>
<td>94</td>
<td>132</td>
<td>58</td>
<td>207</td>
<td>144</td>
<td>140</td>
<td>89</td>
<td>71</td>
<td>63</td>
<td>97</td>
</tr>
</tbody>
</table>
Table 4. Comparison of annual average modal concentration for different sensitivity tests and base case simulation. Shown is the median and 25–75 percentile ranges for each sensitivity test. All units in cm$^{-3}$. Year 2000, Hyytiälä as receptor.

<table>
<thead>
<tr>
<th>Mode</th>
<th>BASE</th>
<th>No MT</th>
<th>No NUC</th>
<th>No clouds</th>
<th>No primary</th>
<th>All continental</th>
<th>All marine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleation mode ($D_p&lt;30$ nm)</td>
<td>390</td>
<td>510</td>
<td>180</td>
<td>390</td>
<td>110</td>
<td>310</td>
<td>390</td>
</tr>
<tr>
<td>Aitken mode ($30&lt;D_p&lt;100$ nm)</td>
<td>1190</td>
<td>1120</td>
<td>820</td>
<td>1360</td>
<td>465</td>
<td>970</td>
<td>1230</td>
</tr>
<tr>
<td>Accumulation mode ($D_p&gt;100$ nm)</td>
<td>400</td>
<td>270</td>
<td>360</td>
<td>580</td>
<td>218</td>
<td>560</td>
<td>400</td>
</tr>
</tbody>
</table>
Fig. 1. Trajectory arriving Hyytiälä at 13 April 2001 06:00 UTC.
Fig. 2. Evolution of trace gases and meteorological parameters during the 216 h simulation along the trajectory shown in Fig. 1. Sulfuric acid and condensable organic species are in units of molecules cm$^{-3}$. Ozone and SO$_2$ in ppb’s and ppt’s, respectively. Clouds are shown as either on or off, giving values of 1 or 0, respectively.
Fig. 3. Evolution of the mixing layer aerosol number size distribution along the trajectory ending in Hyytiälä 13 April 2001 06:00 UTC.
Fig. 4. Comparison between modeled and measured size distribution as of 13 April 2001 06:00 UTC. Mixing layer, Hyytiälä.
Fig. 5. Monthly median modeled size distributions and 25–75 percentile for corresponding observational aerosol number size distribution data, Hyttiyälä 2000.
Fig. 6. Monthly median modeled size distributions and 25–75 percentile for corresponding observational aerosol number size distribution data, Hyytiälä 2001.
Fig. 7. The ten trajectory clusters resulting from the clustering. Cluster number is indicated at the end of each centroid.
Fig. 8. Modelled and measured aerosol number size distribution data belonging to each of the clusters in Fig. AP3. Actual measurements at the receptor site Hyytiälä are presented as median and 25–75 percentile ranges as indicated by the error bars. The thick black line corresponds to modelled median.
Fig. 9. Modeled and observed number concentration of accumulation mode particles ($D_p > 100$ nm, top frame), Aitken mode particles ($30 < D_p < 100$ nm, middle frame) and nuclei mode ($D_p < 30$ nm, bottom frame), cm$^{-3}$. Blue line show modeled data and red line shows measured concentration. Hyytiälä, 2000.
Fig. 10. Modeled and observed number concentration of accumulation mode particles ($D_p>100$ nm, top frame), Aitken mode particles ($30<D_p<100$ nm, middle frame) and nuclei mode ($D_p<30$ nm, bottom frame), cm$^{-3}$. Blue line show modeled data and red line shows measured concentration. Hyytiälä, 2001.
**Fig. 11.** Observed (top frame) and modeled (bottom frame) size distribution evolution. Hyytiälä, 1–10 May 2000.
Fig. 12. Observed (top frame) and modeled (bottom frame) size distribution evolution. Hyytiälä, 12–19 July 2001.
Fig. 13. Modelled and measured number size distribution. Measurement data shown as median and 25–75 percentile range. Model data represented by the thick solid line. Left frame, marine transport, Hyytiälä 1–10 May 2000. Right frame, continental transport, Hyytiälä 12–19 July 2001.
Fig. 14. Measured and modeled SO$_2$ concentrations in ppt (top frames) and measured and modeled O$_3$ concentration in ppb through years 2000–2001 (bottom frames). Hyytiälä 2000–2001.
Fig. 15. Modeled concentrations of hydroxyl radical (OH, top left), nitrate radical (NO$_3$, top right), sulfuric acid (H$_2$SO$_4$, lower left) and condensable organics (lower right). All units as cm$^{-3}$. Hyytiälä 2001.
Fig. 16. Trajectory clusters centroids calculated for year 2000. Cluster number is indicated at the end of each centroid.
Fig. 17. Comparison between model base case and runs with nucleation disabled. Two periods considered, summer (April–September, right frame) and winter (October–March, left frame). Hyyttiälä 2000.
Fig. 18. Comparison between model base case and runs with nucleation disabled per cluster. Dashed line indicate the runs with nucleation disabled, full lines represent base case conditions. Hyytiälä 2000.
Fig. 19. Comparison between model base case and runs with primary emissions disabled. Two periods considered, summer (April–September, right frame) and winter (October–March, left frame). Hyytiälä 2000.
Fig. 20. Comparison between model base case and runs with primary emissions disabled per cluster. Dashed line indicate the duns with primary emissions disabled, full lines represent base case conditions. Hyytiäliä 2000.
Fig. 21. Comparison between model base case and runs with monoterpane emissions disabled. Two periods considered, summer (April–September, right frame) and winter (October–March, left frame). Hyytiälä 2000.
Fig. 22. Comparison between model base case and runs initialized with continental distributions only. Two periods considered, summer (April–September, right frame) and winter (October–March, left frame). Hyytiälä 2000.
Fig. 23. Comparison between model base case and runs initialized with marine distributions only. Two periods considered, summer (April–September, right frame) and winter (October–March, left frame). Hyytiälä 2000.