Zhu Jialei (Orcid ID: 0000-0003-2057-8996)
Penner Joyce E. (Orcid ID: 0000-0001-5577-452X)

Global modelling of secondary organic aerosol (SOA) with organic nucleation

Jialei Zhu[1] and Joyce E. Penner[1]

1 Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA

Key Points:

- A new version of the CESM/IMPACT atmospheric model is developed to include three organic nucleation schemes
- Including organic nucleation improves the model's ability to simulate aerosol number concentration
- Ion-induced pure organic nucleation is the largest contributor to the global new organic particle formation.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019JD030414

Abstract

Organic nucleation has been identified as an important way to form secondary organic aerosol (SOA) and change the number concentration of aerosol and thus its climate effect. A global atmospheric chemistry model is developed to include a comprehensive organic nucleation scheme that includes heteromolecular nucleation of sulfuric acid and organics (HET), neutral (NON) and ion-induced (ION) pure organic nucleation. Our model simulation shows reasonable agreement with the seasonal as well as spatial pattern of organic carbon (OC) concentration in America, while it fails to predict the seasonal pattern of OC in Europe due to the lack of sharp increases in primary organic aerosol emissions in the winter. Including organic nucleation decreases the bias of the annual average particle number concentration at 54% of the available observation sites and increases the temporal correlation coefficients at 58% of the sites. ION contributes the most to the total organic nucleation rate, which peaks around 400hPa in the tropics. HET dominates the total organic nucleation rate in the summer and mostly occurs in the lower troposphere. The number concentration of particles formed from organic nucleation (newSOA) in the nucleation and Aitken modes is highest in the tropics, while accumulation mode newSOA is highest in the Northern Hemisphere due to growth as a result of the condensation of sulfate. Three sensitivity experiments suggest that more studies are needed to investigate the formation mechanism of newSOA, so that a more accurate simulation of the spatial and size distribution of newSOA can be developed.

1. Introduction

Secondary organic aerosol (SOA) is a dominant constituent of submicrometer atmospheric particles and the largest component of global organic aerosol (*Jimenez et al.*, 2009). SOA forms from both biogenic and anthropogenic precursors, comprising a large number of structurally different organic oxygenates (*Zhang et al.*, 2015). SOA is ubiquitous in various atmospheric environments and on average accounts for 60~90% of the total organic aerosol (*Jimenez et al.*, 2009; *Zhang et al.*, 2007). SOA makes a considerable contribution to air pollution as well as to climate both directly by absorbing and scattering radiation and indirectly by altering the albedo of clouds (*Hallquist et al.*, 2009; *Shrivastava et al.*, 2017; *Zhang et al.*, 2011). Many of these effects depend on the particle size distribution, which is partly governed by the formation mechanism of new particles (*Zhu et al.*, 2019).

Nucleation of atmospheric vapors contribute most to the atmospheric aerosol number concentration (*Kulmala et al.*, 2004), and is thought to be responsible for up to half of the global cloud condensation nuclei (CCN) (*Wang and Penner*, 2009). Nucleation has a major influence on the microphysical properties of clouds and the radiative balance of the global climate system (*Gordon et al.*, 2016; *Wang and Penner*, 2009). Observations in the planetary boundary layer (PBL) revealed a consistent correlation between sulfuric acid and the concentration of newly formed particles at many sites (*Kuang et al.*, 2008; *Sihto et al.*, 2006; *Weber et al.*, 1997), so sulfuric acid was thought to be essential to initiate most particle formation in the atmosphere (*Kulmala et al.*, 2013). Many models only use a particle formation

rate based on the concentration of sulfuric acid (*Makkonen et al.*, 2012; *Mann et al.*, 2014). In sulfur-rich environments, nucleation can be explained by a simplified acid-base model through the formation of dimers and trimers involving the sulfate molecular and amines (*Chen et al.*, 2012). Although sulfuric acid is almost always involved in the initial cluster formation, its concentration cannot account for the observed early rapid growth and seasonal variations of atmospheric nanoparticles, especially in the forested areas with minimal influence from anthropogenic pollution (*Boy et al.*, 2008; *Paasonen et al.*, 2010). The nucleation rate is underestimated and the sensitivity of the nucleation rate to the sulfuric acid concentration is overestimated in most models with only classical nucleation theories of sulfuric acid (*Mann et al.*, 2014; *Scott et al.*, 2014).

Recently, highly oxygenated molecules (HOMs) with extremely low volatility were detected both in laboratory studies and in the ambient atmosphere (*Ehn et al.*, 2012). HOMs probably provide an initial organic medium to form molecular clusters (*Ehn et al.*, 2014). Multiple field studies have shown that biogenic SOA formation and growth from the oxidation of biogenic volatile organic compounds (BVOC) emitted by terrestrial ecosystems play an importantrole in atmospheric CCN production (*Jokinen et al.*, 2015; *Kulmala et al.*, 2013; *Paasonen et al.*, 2013). Comprehensive measurements performed as part of the CLOUD (Cosmics Leaving Outdoor Droplets) project detected the ternary nucleation of sulfuric acid with water and oxidized organics, and a model using this mechanism was able to reproduce the observed seasonal cycle of particle number concentrations (*Riccobono et al.*, 2014). The latest CLOUD experiments

observed high rates of new particle formation from HOMs even in the absence of sulfuric acid when ions were present (*Kirkby et al.*, 2016). This pure organic nucleation scheme might be able to explain the high concentration of nucleation mode particles and new particle formation observed in the upper troposphere over the Amazon where SO_2 levels are extremely low (*Zhu et al.*, 2019). This new scheme is also important for estimating the historical radiative forcing since it provides a way to form new particles in the pristine pre-industrial atmosphere, which may change the baseline CCN concentrations, depending on the formation mechanism for the nucleating organics that is used (*Gordon et al.*, 2016; *Zhu et al.*, 2019).

New organic particle formation and growth has been considered as an important contributor to global aerosol number concentrations and climate (*Gordon et al.*, 2016; *Jokinen et al.*, 2015), but models of SOA formation remain highly uncertain because of the chemical complexity associated with volatile organic compounds (VOC) and their oxidation (*Chen et al.*, 2017; *Pierce and Adams*, 2009). HOMs with low volatility are regarded as necessary for organic nucleation (*Kirkby et al.*, 2016), and the oxidation of α-pinene by ozone has become a canonical way to form HOMs (*Claeys et al.*, 2009; *Kristensen et al.*, 2014; *Yasmeen et al.*, 2010). However, the molecular structures and formation pathways of HOMs remain unclear (*Hall and Johnston*, 2011; *Zhang et al.*, 2015). As a result, most model studies assume empirical or semi-empirical fixed HOM yields after the first oxidation step of α-pinene to activate organic nucleation (*Gordon et al.*, 2016; *Jokinen et al.*, 2015; *Scott et al.*, 2018), which affects the

time scales for SOA formation as well as the spatial distribution of newly formed SOA (*Zhu et al.*, 2019). In addition to α-pinene, cyclohexene, limonene and β-pinene may also form HOMs (*Ehn et al.*, 2014; *Shrivastava et al.*, 2015; *Shrivastava et al.*, 2008; *Shrivastava et al.*, 2013), but their role in organic nucleation has not been studied as much as that of α-pinene. Previous studies have shown that the growth of the organic aerosol in the nucleation mode is largely via partitioning of semi-volatile organic vapor, but the growth of organic particles from nucleation mode to CCN sizes requires additional unidentified compounds and reactions in the particle phase. (*Apsokardu and Johnston*, 2018; *Burkart et al.*, 2017; *Bzdek and Johnston*, 2010).

New particle formation (NPF) (i.e. the production of clusters which then grow to > 3 nm diameter) has been observed over forests with large sources of BVOC (*Held et al.*, 2004; *Pierce et al.*, 2014; *Yu et al.*, 2014). Isoprene makes up to nearly half of the total global BVOC budget (*Guenther et al.*, 2012). However, the role of isoprene in NPF is controversial (*Kiendler-Scharr et al.*, 2012). Isoprene expoxydiols (IEPOX) comprise a substantial fraction (6%~36%) of total organic aerosol mass (*Hu et al.*, 2015), which is important for SOA formation (*Jokinen et al.*, 2015; *Lin et al.*, 2012). However, some field measurements credit isoprene with the suppression of new particle formation events, even though there are sufficient concentrations of monoterpenes to lead to NPF events (*Kanawade et al.*, 2011; *Kiendler-Scharr et al.*, 2012; *Kiendler-Scharr et al.*, 2009; *Lee et al.*, 2016). A plant chamber study indicated that such suppression effects are dependent on the concentration ratio of isoprene carbon to the monoterpene carbon (*Kiendler-Scharr*

et al., 2009). Isoprene and its oxidation products may change the chemical reactions of monoterpene ozonolysis, but a detailed chemical mechanism leading to the absence of NPF in isoprene rich forests is still not available (*Lee et al.*, 2016). Moreover, a recent study detected that NO_x strongly suppresses not only NPF but also SOA mass yields, revealing the importance of interactions of SOA formation with anthropogenic emissions (*Zhao et al.*, 2018). However, NO_x may also increase SOA formation by increasing oxidants like ozone and OH radicals which then increase the amount of VOC reacted (*Shrivastava et al.*, 2019). The specific chemical mechanism leading to the suppression of NPF is also not known.

In this study, we simulate SOA formation using a global model with a comprehensive nucleation scheme. The spatial and temporal distribution of SOA from this model is evaluated by comparison with observations of organic carbon (OC) and condensation nuclei (CN) concentrations over the world. The contribution of multiple nucleation schemes and the size distribution of SOA are analyzed using the model results. Three sensitivity experiments with different SOA formation mechanisms are designed to investigate uncertainties and the possible influence of missing mechanisms in the model on the number concentration and size distribution of newSOA.

2. Methods

2.1 Model description

We used the Community Earth System Model (CESM) version 1.2.2.1 coupled with the University of Michigan Integrated Massively Parallel Atmospheric Chemical Transport (IMPACT) aerosol model for this study. The IMPACT aerosol module gets the necessary meteorology field from the CESM model at each time step, while changes to the aerosols in IMPACT don't provide feedback to the CESM. The aerosol scheme in the model is partially modal (newSOA and sulfate) and partially sectional (dust and sea salt). In addition to SOA, fifteen other aerosol species in five types were simulated in the model, including: (1) sulfate in three modes (i.e. nucleation (0-5 nm radius), Aitken (5-50nm radius)and accumulation (50-630nm radius)), (2) soot from open biomass burning (bSoot, i.e. primary organic aerosol and black carbon from biomass burning), (3) soot from fossil fuel and biofuel burning (fSoot, i.e. primary organic carbon and black carbon), and (4) dust and (5) sea salt, the latter two of which were each carried in four separate bins with varying radii. All the nonsulfate aerosols are internally mixed with sulfate (and organics, see below) through condensation and coagulation processes or through sulfate formation in cloud droplets. The IPCC emission dataset for 2000 is applied for the source of precursors of aerosol as well as BC and OC. The emission of precursors for biogenic SOA (i.e. isoprene, α-pinene and limonene) are estimated by the MEGAN model (*Guenther et al.*, 2012) coupled to our model. The basic model description and setup can be found in *Zhou and Penner* (2014).

For the SOA simulation, four species of gaseous precursors were included: isoprene, α-pinene, limonene and aromatics. An explicit gas phase chemical mechanism was applied to predict the formation of semi-volatile organic compounds (SVOC) which mainly consist of organic nitrates and peroxides from the oxidation of precursors by ozone and the hydroxyl radical (*Ito et al.*, 2007; *Lin et al.*, 2012). Lower volatility compounds like oligomers form as a result of a simple e-folding time-scale of 24 hours when semi-volatile organics are incorporated into the aerosol phase (smaller than the low end of the range quoted for SOA from the reactions of O_3 with β-pinene or OH with dimethylsiloxane from *Apsokardu and Johnston* (2018)). The explicit aqueous phase reactions of glyoxal and methyglyoxal and heterogeneous reactions of IEPOX are also included to form SOA (*Lin et al.*, 2014). IEPOX, glyoxal and methylglyoxal form low-volatility products on immediate kinetic uptake by aerosols that are formed from the nucleation of sulfate with a rate that is proportional to the available surface area of these aerosols, which will be further discussed in section 2.2. Thermodynamic formation from partitioning of semi-volatile species followed by oligomer formation within aerosols contributes SOA to all preexisting aerosols based on the amount of organics that are present within the preexisting aerosols. The formation of SOA via aqueous phase formation within drops is distributed to the particles that act as CCN based on the number of CCN. All of the SOA that forms (other than HOMs that nucleate new particles) is internally mixed with sulfate, bSoot, fSoot, dust and sea salt based on the mechanism of formation (*Zhu et al.*, 2017). The

HOMs that nucleate new particles also become mixed with sulfate (*Zhu et al.*, 2019).

Low volatility HOMs from α-pinene oxidation were used to drive organic nucleation in the model. The HOMs are formed by the oxidation of α-pinene through explicit gaseous and particle phase reactions based on the experimental results of *X Zhang et al.* (2015) and the Master Chemical Mechanism (MCM) (*Jenkin et al.*, 2003; *Jenkin et al.*, 2015; *Saunders et al.*, 2003). New particles form from both binary sulfuric acid-water nucleation and organic nucleation. The organic nucleation mechanisms used in this model include heteromolecular nucleation of sulfuric acid and organics (HET), neutral organic nucleation (NON) and ion-induced organic nucleation (ION). The rate of HET was parameterized following equation (2) in *Riccobono et al.* (2014):

$$J = k_m[H_2SO_4]^2[HOMs]$$

where k_m is the multicomponent prefactor which is set to 3.27×10^{-21} cm^6 s^{-1}.

The rate of NON (J_n) and ION (J_{iin}) were parameterized following equation (4) in *Kirkby et al.* (2016):

$$J_n = a_1[HOMs]^{a_2+a_5/[HOMs]}$$

$$J_{iin} = 2[n_{\pm}]a_3[HOMs]^{a_4+a_5/[HOMs]}$$

Where ($n_{\pm}$) is the ion concentration. The parameters a_n are determined from fits to the experimental data and have the values a_1=0.04001, a_2=1.848, a_3=0.001366, a_4=1.566 and a_5=0.1863.

The ionization rate in the model was estimated based on the look-up table in *Yu et al.* (2008). In the look-up table, the global ionization rate due to cosmic rays is calculated based on the schemes given in *Usoskin and Kovaltsov* (2006), and the contribution of radioactive materials form soil to ionization rates is parameterized based on the profiles given in *Reiter* (1992). The ions do not condense on preexisting particles in this model, which may lead to overestimation of the ion concentrations used calculate the rate of ION due to the lack of these sinks. The organic nucleation rates were adjusted from the temperature of the CLOUD experiments (278 K) to other temperatures by multiplying by exp(-(T-278)/10) as suggested in (*Dunne et al.*, 2016). The new particles formed from organic nucleation (newSOA) can grow to Aitken and accumulation mode by coagulation with each other and by the condensation of sulfuric acid and semi-volatile organics based on thermodynamic principles. The newSOA includes three modes with the same size range as that used for new sulfate particles. Gas phase HOMs and other low-volatility organics that do not nucleate new particles can also lead to the growth of newSOA particles as well as other aerosols by thermal partitioning. NewSOA is removed by coagulation with other species of preexisting aerosols as well as by dry and wet deposition. The total CN concentration includes the number concentration of all new particles formed from both sulfuric acid and organic nucleation. SOA and sulfate can become internally mixed in the model as a result of condensation or coagulation, but we separately follow the method of nucleation in order to distinguish which process dominates for new particle formation in the model. The detailed chemical reactions that form HOMs and the organic

nucleation mechanisms appear in the Supplementary Information in *Zhu et al.* (2019).

2.2 Simulation setup

The BASE case and three sensitivity cases were designed to investigate the characteristics of SOA formed from organic nucleation in the present day and its uncertainty. Each simulation was performed for 5 years after a one year spin-up using present-day climate conditions and anthropogenic emissions. The sea surface temperature is fixed to observations and the model is run in a climate mode. The BASE case was set up as described in Section 2.1.

The first sensitivity case (EX1) was designed to examine the effects of uncertainties in the growth of newSOA. IEPOX is usually taken up by acidic aerosols (*Gaston et al.*, 2014). Sulfate is strongly correlated with the acidity of an aerosol, which is important for the chemistry of IEPOX (*Surratt et al.*, 2007). Glyoxal and methylglyoxal can dissolve in the aqueous phase and be further oxidized by OH and NO_3 radicals to form products with lower volatility (e.g., dicarboxylic acids and oligomers) (*Ervens et al.*, 2011; *Lim et al.*, 2010). Sulfate always provides a wet surface to take up glyoxal and methylglyoxal because of its high hygroscopicity. As a result, the low-volatility products formed from IEPOX, glyoxal and methylglyoxal are only taken up by sulfate in the BASE case. However, sulfuric acid takes part in HET organic nucleation and is thereby part of the initial composition in newSOA. It also makes a large contribution to the growth of newSOA by condensation. Therefore, the sulfuric acid involved in newSOA

may generate an acidic and wet surface to take up IEPOX, glyoxal and methylglyoxal similar to that which occurs on new sulfate particles. In EX1, we assumed that the low-volatility products formed from IEPOX, glyoxal and methylglyoxal reacting with newSOA and sulfate become internally mixed with both newSOA and new sulfate particles, while these low-volatility products are only internally mixed with new sulfate in the BASE case. There were no other differences from the BASE case.

The second sensitivity case (EX2) was designed to examine the uncertainty of organic nucleation under an isoprene-rich environment. Some field measurements indicated that isoprene suppresses organic nucleation (*Kiendler-Scharr et al.*, 2012; *Kiendler-Scharr et al.*, 2009; *Lee et al.*, 2016). It is clear that NPF rarely takes place when emissions of isoprene are high, although the chemical and physical mechanism of suppression is not clear. A plant chamber study indicated such suppression effects are dependent on the concentration ratio of isoprene carbon to the monoterpene carbon (*Kiendler-Scharr et al.*, 2009). We assumed that organic nucleation is shut off when the concentration ratio of isoprene carbon to the α-pinene carbon is higher than 1 in EX2. The specific test follows that applied in *Lee et al.* (2016), who state that it is not intended to represent the mechanism of new particle formation suppression, but rather to fit observations.

The third sensitive case (EX3) was designed to examine uncertainties in HOM formation from species other than α-pinene and their influence on organic nucleation. Laboratory experiments have shown that cyclohexene, limonene and β-

pinene are able to produce HOMs with extremely low volatility similar to α-pinene, which may also drive organic nucleation and the growth of newSOA (*Ehn et al.*, 2014). Based on the experiments conducted in *Ehn et al.* (2014), the HOM yield of cyclohexene is slightly lower than that found for α-pinene because of the more complex structure of α-pinene. The annual emission of cyclohexene over the world is less than that of α-pinene. The HOM yield of β-pinene is about 2 orders of magnitude less than that of α-pinene, which is attributed to the exocylic double bond in β-pinene instead of the endocyclic double bond in α-pinene. By contrast, the HOM yield of limonene is roughly twice that of α-pinene because of the two double bonds in limonene, which can thus react twice with ozone. As the result, in EX3 we assumed limonene also generates HOMs but ignore the contribution from cyclohexene and β-pinene due to their lower yields of HOMs. However, the chemical pathway to form HOMs from limonene is still unclear. We used the same pathway for limonene as the HOMs generated form α-pinene which can be found in Supplementary Information in *Zhu et al.* (2019), but with twice the yield.

3. Results and discussion

3.1 Comparison with measurements

Our model's ability to simulate the formation of SOA has previously been evaluated by comparison with both surface and vertical measurements (*Lin et al.*, 2012; *Lin et al.*, 2016; *Lin et al.*, 2014). Moreover, the model results for number concentration in the Amazon have been evaluated by comparing with aircraft

measurements taken by *Andreae et al.* (2018) as detailed in our previous publication (*Zhu et al.*, 2019). Those studies examined the annual average OC concentration, SOA concentration, organic aerosol composition and O/C ratio in comparison to observations which were available at that time. In total, the bias of the model was within the range of the model-observation comparisons shown in *Tsigaridis et al.* (2014) for the AeroCOM organic aerosol model intercomparison project. However, previous studies did not evaluate the temporal variation of organic aerosol, which we show here. The inclusion of organic nucleation in this study does not change the burden or the surface concentration of organic aerosol by very much. As indicated in *Zhu et al.* (2019), inclusion of organic nucleation only changes the global SOA burden by 7.3%. However, it does change the CN concentrations significantly. We only compare the OC mass concentrations for the BASE case here, since the OC mass concentrations in the sensitivity cases are very similar to that in the BASE case. However, the CN concentrations are very different in BASE case and sensitivity cases. Therefore, we compare the simulated CN concentration with observations over the world for all cases to evaluate the ability of the model to predict CN concentrations after including organic nucleation. Section 3.4 discusses the evaluation of CN concentrations for sensitivity cases.

3.1.1 Surface OC concentration

The monthly averaged surface OC concentration in 48 regions as determined from 196 sites of Interagency Monitoring of Protected Visual Environments (IMPROVE) network in Unite States during 2005-2008 (*Hand et al.*, 2011) are

used here to evaluate the model. The high local levels of OC in urban regions are not expected to be predicted in the model due to its coarse resolution (1.9° by 2.5°). Therefore, all the measurements sites in the IMPROVE network which are urban in the IMPROVE report (*Hand et al.*, 2011) were not included in the 196 sites and thus have been left out in our comparison. The IMPROVE monitor collects 24-hour samples every third day with the version II sampler. The detailed sampling and technical information can been found in *Hand et al.* (2011). The original data is given as OC in micrograms of carbon per cubic meter, while our model predicts the mass of organic aerosols (OA). To convert OA to OC, a factor of 1.4 for primary OA (POA) and 1.8 for SOA are used as suggested in *Turpin and Lim* (2001). Table S1 shows a comparison between the annual average concentration of OC, the monthly average normalized mean bias, and the temporal correlation coefficient between the simulation and measurements. The BASE case model (as well as the other sensitivity experiements) always underestimates surface OC concentrations but is generally within a factor of 2 of the observed concentration within most of the IMPROVE regions. The mean predicted OC concentration across all IMPROVE regions used in the comparison is 28% lower than observations (Table 1). Figure S1 shows the monthly variation of simulated and observed regional average OC concentrations with the simulated concentrations of POC and SOC. Including SOA formation improves the model enabling it to reproduce the temporal pattern and OC levels significantly, especially in the summer. However, the model still underestimates the peak of OC concentrations in the summer in most regions while it predicts OC levels that are similar to

observations in winter. Thus, there are possibly missing mechanisms for the formation of SOA in the model leading to lower peak concentrations in summer than in observations. The model underestimates the OC concentration in the winter by 71% at the Columbia River Gorge, Hells Canyon, Northern Great Plains and Northern Rockies, which probably indicates an underestimated source of POC in the northwestern United States. In contrast, the model overestimates the OC concentration by 41% in the summer in the Appalachia, Northeast, Ohio River Valley and Ontario regions, while the simulated OC concentration is close to observations in the winter there. It is possible that the source of BVOC is overestimated in the region around the Great Lakes. The temporal correlation coefficients between the simulated and observed OC concentrations are higher than 0.7 in 74% of the IMPROVE regions with the average of 0.73 across all regions (Figure 1a and Table 1). The model is able to predict the temporal variation of OC concentrations in most regions across the United States except for South Arizona which has low temporal correlation coefficients. There could be an underestimate of sources or SOA formation processes in the model for this region. The model also captures the spatial pattern of OC concentration in United States with a spatial correlation coefficient of 0.70 (Table 1).

In comparison with the IMPROVE network, the model shows a poorer ability to reproduce OC concentrations at the European Monitoring and Evaluation Program (EMEP) network sites. The particles were sampled daily at the EMEP network sites and the OC concentration was analyzed using a thermal desorption method. The detailed technical information can be found in the EMEP manual for

sampling and analysis (https://projects.nilu.no//ccc/manual/index.html). The simulated OC concentration is 67% lower than that observed on average (Table 1, Table S2). However, the average concentrations observed at EMEP sites are higher by a factor of ~3 than those at IMPROVE sites (Table 1). One reason is that 70% of EMEP sites collected PM_{10} while all IMPROVE sites collected $PM_{2.5}$. The large particles observed in EMEP sites are not captured in the model because we assumed that all organics are only present as submicron particles (*Lin et al.*, 2012). However, the mass concentration is expected to decrease substantially with aerosol size. The emissions of organic aerosol with diameter < 1μm are expected to explain most of the mass concentration of all organic aerosols (*Bond et al.*, 2004). This assumption might cause some underestimation of total OC concentration, but could not be the dominate reason to explain the disagreement with EMEP measurements since the mass concentration normally decreases substantially with aerosol size. In addition, the measurements at most of the EMEP sites have a peak concentration of OC in winter (Figure S2) which probably comes from biomass burning primary emissions (*Gelencser et al.*, 2007). Domestic combustion in winter is not fully represented in our emission database and the emission of POC in the model does not have any significant seasonal variation due to the lack of data to support estimates of the temporal variation. As a result, the large discrepancy between observed and simulated OC concentrations in winter in Europe was also found in other model comparisons (*Gilardoni et al.*, 2011; *Szidat et al.*, 2007). Our model underestimates the mean OC concentration by 83% in winter due to the lack of increased emissions of POC in winter. The difference in the OC concentrations

between the simulation and observations in the summer is reduced to 46% due to the effect of SOA formation and increased POC emissions from wildfires, so there may still be some missing mechanism of SOA formation or missing POC emissions. The model is not able to represent the temporal variation at most EMEP sites and, on average, has a negative correlation coefficient because of the observed large increase in the OC concentration in winter (Figure 1b). Reasonable temporal variations are reproduced by the model with correlation coefficients of >0.5 at sites such as Rigi, Campisabalos, and Birkennes, where, perhaps, since they are either mountainous or coastal locations, the emissions of POC are not as large in winter (Figure S2). The model captures the spatial pattern of OC concentration relatively well at the EMEP sites with a spatial correlation coefficient of 0.43 (Table 1).

3.1.2 Particle number concentration

Organic nucleation makes a large contribution to the CN concentration. Here, we evaluate the model ability to represent CN concentrations by comparison with observations at 27 sites over the world for observations performed during 2000-2010 (data from EMEP dataset online at http://ebas.nilu.no). The observations of CN were conducted using condensation particle counters (CPC) with nominal cut-off diameters of 3nm or 10nm. We used the number concentration of all aerosol species in the Aitken and accumulation modes (diameter > 5nm) calculated in the model to compare with the observations. The model overestimates the CN concentration by 19% on average compared to the observations with a normalized mean error of less than 50% at most of sites (Table 2, Table S3). The spatial

pattern of CN concentration is reproduced well by the model with a spatial correlation coefficient of 0.85 (Table 2). Close agreement in the CN concentrations between the simulation and observations is not expected due to the low model resolution and uncertainties associated with the nucleation parameterizations, but the model that includes organic nucleation exhibits an improvement in the simulation of CN concentrations compared to a simulation that does not include organic nucleation. We have shown a summary of the comparison of aerosol burden and column number between with and without organic nucleation in a previous publication (*Zhu et al.*, 2019). After including organic nucleation, the bias in the annual average CN concentration between the simulation and observations is decreased at 54% of the sites and the temporal correlation coefficients were increased at 58% of the sites (Table S3). In a previous model study which excluded organic nucleation, the simulated CN concentrations in the continental boundary layer were underestimated by 74% with a spatial correlation coefficient of 0.68 (*Spracklen et al.*, 2010). *Gordon et al.* (2016) improved the model bias in the CN concentration to -41% after including pure organic nucleation. As noted above, our model has a bias of 19%.

The model is unable to explain the observed monthly variation for all sites in the EMEP data base (Figure S3). The average temporal correlation coefficient is only 0.13, and has a significant spatial difference (Figure 1c). 31% of the sites have a negative temporal correlation between the simulation and observations, while the correlation coefficient is >0.6 at a few sites in remote areas (Neumayer, Mace Head, Zeppelin Mountain and South Pole). The monthly variation in the CN

concentration is determined by the number concentration of sulfate at most of sites (Figure S3). However, there is no monthly variation of SO_2 (the precursor of sulfate) emissions considered in the model, which might influence the correlation between simulation and observation. We note that there is a larger average correlation coefficient (0.277) between the simulated number concentration of newSOA and the observed CN concentrations than that for the total number concentration (0.134, see Table 2). This indicates that the monthly variation of the CN concentration at some sites may dominated by newSOA rather than by the new sulfate particles, as in the model. For example, the simulated CN concentration in two sites in Finland (i.e. Varrio and Hyytiälä) exhibits a poor correlation with the observations (Figure S3). However, the correlation coefficient between the simulated number concentration of newSOA and the observed CN concentration are all higher than 0.6 in these sites. Because the simulated number concentrations of new sulfate particles are much higher than those of newSOA at these sites, the newSOA contributes little to the simulated temporal variation of total CN. Since the measurements of aerosol composition in Hyytiälä, Finland showed a smaller mass concentration of sulfate (0.57 μg m^{-3}) than the model result (0.99 μg m^{-3}) as well as a larger mass concentration of organics (1.47 μg m^{-3}) compared to the model result (1.05 μg m^{-3})(*Häkkinen et al.*, 2012), the relative importance of sulfate in forming new particles may be overestimated. Speculatively, the underestimation of newSOA may be due to the uncertainties in the nucleation scheme and emissions of precursors. In addition, the overestimation of sulfate in Finland in the model may be caused by the poor grid resolution of the model,

which would include anthropogenic emissions of SO_2 from sources near the Hyytiälä observation site. These issues might explain the negative correlation between simulation and observation here.

3.2 Organic nucleation rate

Figure 2 shows the vertically integrated distribution of the total organic nucleation rate (the sum of the HET, NON, and ION rates) in the troposphere in boreal spring (March to May), summer (June to August), fall (September to November) and winter (December to February). Peak organic nucleation rates are found in the Amazon year round because of the high emissions of α-pinene and the high production of HOMs from α-pinene. The emission of α-pinene in the Amazon is highest in the fall, which is the reason for the highest nucleation rate occurring in the fall (Figure 2e). By comparison, the organic nucleation rate in the middle latitudes of the Northern Hemisphere (NH) (30° N~50° N) such as over the United States and China, make a significant contribution to the high global average organic nucleation rate in the summer (Figure 2c). The global average organic nucleation rate is lowest in boreal winter (Figure 2g) owing to the low organic nucleation rate in the NH in winter. The global organic nucleation rate in the winter is 25% smaller than that in the fall.

Figure 2 also shows the vertical structure of the zonal average organic nucleation rate in each season. The organic nucleation rate always peaks around 400hPa in the tropics. Organic nucleation mainly occurs throughout the middle and

upper troposphere from 600hPa to 150hPa with much smaller rates occurring in the lower troposphere and near surface in the tropics for all seasons except in summer. The highest organic nucleation rate in the upper troposphere takes place in the fall. Observations in the Amazon also indicated that there was a high number concentration of small organic particles in the upper troposphere in the fall with little nucleation within the PBL (*Andreae et al.*, 2015; *Martin et al.*, 2010). The small particles that are present in the boundary layer are sustained by vertical transport from the free troposphere (*Wang et al.*, 2016). *Andreae et al.* (2018) found that due to the frequency of deep convection, large amounts of BVOC can be brought up to upper troposphere. This is not the case in the lower troposphere in the middle latitudes of the NH in the summer where organic nucleation occurs frequently (Figure 2d). The nucleation in the PBL at this location offsets a decrease of the nucleation rate in the upper troposphere compared to that in the tropics. Apparently, convection is not as strong in the mid-latitudes of the NH as that in the tropics, which can be inferred from the much lower convective precipitation in rate in the mid-latitudes than that in the tropics (Figure S5), explaining both the increased PBL nucleation rates and reduced upper tropospheric nucleation rates.

Although organic nucleation contributes significantly to NPF, the total organic nucleation rate is small compared to total the sulfuric nucleation rate (Fig S4). The global annual average vertically integrated sulfuric nucleation rate is higher by a factor of 18.5 than the organic nucleation rate. As a result, sulfuric acid nucleation still dominates NPF in the atmosphere. However, the annual and vertically-average sulfuric acid nucleation rate is only 41% of the organic

nucleation rate in the Amazon, which is the reason for the much higher number concentration of newSOA than new sulfate in the Amazon. Importantly, the sulfuric acid nucleation rate in the Amazon in the fall, the season with highest organic nucleation rate, is 79% lower than the organic nucleation rate. Thus, organic nucleation is the most important source for new particle formation in the Amazon.

The simulated concentration of organic aerosol in the Amazon (1.699 μg m^{-3} in the wet season and 4.685 μg m^{-3} in the dry season) is much higher than the concentration of sulfate (0.236 μg m^{-3} in the wet season and 0.280 μg m^{-3} in the dry season). The simulation is close to the wet season observation found during the Green Ocean Amazon (GoAmazon 2014/5) field campaign (Sulfate: 0.268±0.136 μg m^{-3} and organic aerosol: 1.671±0.579 μg m^{-3}) (*Glasius et al.*, 2018). However, the simulation is lower than the observation in the dry season (Sulfate: 1.600±0.917 μg m^{-3} and organic aerosol: 8.783±2.642 μg m^{-3}), probably because of the influence of nearby emissions from biomass burning and the plume from a nearby city during the field campaign in the dry season, which are not captured in our model with its 1.9 x 2.5 degree horizontal resolution (*Glasius et al.*, 2018).

The horizontal and vertical spatial distributions as well as seasonal variations of organic nucleation rates are determined by the different organic nucleation mechanisms. Our model synthesizes the HET, NON and ION mechanisms to estimate the total organic nucleation rate as described in Section 2. The vertically integrated annual average organic nucleation rates in the troposphere for the three nucleation mechanisms along with their longitudinally-averaged

vertical distributions are shown in Figure 3. Ions are known to enhance nucleation as a result of charge effects on cluster stability (*Kirkby et al.*, 2016; *Yu*, 2010). As a result, ION is the largest contributor (60%) to the global annual average organic nucleation rate (Figure 3e). The ION rate is much higher in the middle and upper troposphere than in the lower troposphere because of the high ionization rates and low temperatures there (Figure 3f). HET contributes only 39% of the global average organic nucleation rate, but it dominates the organic nucleation rate over most continental regions outside of the tropics (Figure 3a). The HET rate is highest at latitudes between 20°N and 50°N, which is also the region with the highest industrial emissions of sulfur. The HET rate peaks in the lower troposphere of the middle latitudes of the NH which differs from the vertical distribution of the ION rate (compare Figures 3b and 3f). This is because of the high concentration of gaseous sulfuric acid in the NH mid-latitudes which is formed from anthropogenic surface emissions of its precursor SO_2. There are also high nucleation rates associated with HET in the tropical upper troposphere around 150hPa. This location has both high concentrations of gaseous sulfuric acid from transport and high concentrations of HOMs. NON is mainly found in the middle and lower troposphere in the tropics and has a rate that is one or two orders of magnitude lower than the other two organic nucleation mechanisms (Figure 3c, 3d). NON is responsible for only 1% of the total global organic nucleation because the initial neutral cluster has weak bonds as indicated by experiments in the laboratory (*Kirkby et al.*, 2016).

The variation of HET and ION rates also determines the seasonal pattern of total organic nucleation. HET contributes the most (56%) to organic nucleation in the summer because of the high emissions of α-pinene in the NH (Table 3) and the high formation rate of gaseous H_2SO_4. The HET nucleation rate in the summer is larger by a factor of 2 than that in the winter, which explains the importance of organic nucleation in the NH summer. The high ION rate in the fall determines the highest global average organic nucleation rate, and is responsible for 68% of newSOA formation in the fall. However, ION contributes the most to nucleation in the boreal winter (70%) due to the lowest HET rate in this season, resulting in a large contribution to newSOA formation in the Southern Hemisphere (SH) (Figure 2g).

New particle formation in the PBL is important for climate and air quality. Figure 4 shows the annual average organic nucleation integrated within the PBL. Unlike the vertically averaged organic nucleation in the troposphere, the organic nucleation within PBL is dominated by HET. HET is responsible for 85% of newSOA formation in the PBL, while ION only contributes 14% to the global annual average (Table 3). This is because 19% of the vertically averaged HET occurs within the PBL while only 2% of ION takes place in the PBL. Organic nucleation in the PBL mainly occurs in the middle latitudes of the NH, while there is some organic nucleation in the PBL of the Amazon from ION and NON (not shown). Owing to the high HET rate in the summer, the organic nucleation within the PBL contributes more to newSOA formation in the summer than that in the other seasons (Figure 4). Because of the large contribution to the global organic

nucleation from ION in the free troposphere, organic nucleation within the PBL is only 9% of that in the whole troposphere with the largest contribution (15%) in the summer and the smallest (5%) in the winter (Table 3).

3.3 SOA formed from nucleation

The organic nucleation rate is one of the most important factors contributing to the number of newSOA, but the concentration of newSOA does not spatially follow that of the nucleation rate. The number concentration of newSOA in the nucleation mode is highest in the Amazon in all seasons (Figure 5a, Figure 6) due to the high emission rates of α-pinene and resulting high nucleation rates there. However, the high organic nucleation rate in middle latitudes of the NH does not result in a high column number concentration of newSOA in the nucleation mode at these latitudes even in the summer (Figure 6c). This is because organic nucleation mostly occurs within lower troposphere in the NH, where there is high concentration of pre-existing aerosol. Most of newSOA are coagulated with other particles immediately after formation and internally mixed with them, which is an important sink to remove newSOA in the nucleation mode. The number concentration of newSOA in the nucleation mode always peaks around 200hPa in the extratropics, while the peak is around 400hPa in the tropics (Figure 5b). This is because the nucleation rate is highest at 400hPa in the tropics due to the peak in the concentration of HOMs, but the nucleation rate is higher at 200-300hPa than at 400hPa in the extratropics in part because the concentration of HOMs are nearly constant between 500 hPa and 200 hPa so that the low temperature and high ion

concentration at 200-300hPa determine the peak nucleation rate. There is very few (<1%) newSOA in the nucleation mode found within the PBL. Although organic nucleation occurs within the PBL, most of the newSOA in the nucleation mode within the PBL will be removed by coagulation due to the high background aerosol number concentration. Observational studies have difficulty in distinguishing whether organic particles are formed from organic nucleation or condensation on other particles or by primary emissions. As a result, it is hard to evaluate the mass and number concentration of newSOA from the model with observations.

The newSOA in nucleation mode grows to the Aitken and accumulation modes primarily as a result of the condensation of SVOC and sulfuric acid. The column number concentration of newSOA in the Aitken mode is highest in the tropics (Figure 5c). Many newSOA are not able to grow to the accumulation mode due to the very high number concentration of newSOA without sufficient condensable gases. Although the concentration of SVOC is high in the Amazon, there is little sulfuric acid in the lower troposphere there. Moreover, the large amount of isoprene that is emitted in the Amazon is oxidized to IEPOX, which is taken up only by new sulfate particles in the BASE version of the model and does not contribute to the growth of newSOA. In contrast, there is enough condensable SVOC and sulfuric acid to grow a limited number of newSOA particles to the accumulation mode in most ocean areas of middle latitudes (Figure 5e). The newSOA in the Aitken mode peaks in the upper troposphere above 200hPa in the extratropics at similar altitudes as the peak in the nucleation mode, whereas the Aitken mode peaks above that level in the tropics while the nucleation mode peaks

near 400 hPa (Figure 5d). In the lower troposphere of the middle latitudes, especially in the NH, sulfuric acid forms from anthropogenic emissions near the surface and the SVOC causes newSOA to grow to the accumulation mode, resulting in the high number concentration of newSOA in the accumulation mode in the NH while the Aitken mode is depleted at these altitudes (Figure 5f). The SVOC formed from the oxidation of BVOC can be vertically transported to the upper troposphere in the tropics. As a result, the number concentration of newSOA in the accumulation mode is high from the surface up to around 100hPa in tropics (Figure 5f).

The global average column number concentration shows a large seasonal variation (Table 4).The organic nucleation rates in the Amazon and SH are lowest in the boreal summer, resulting in the lowest global average column number concentration of newSOA in the nucleation mode in the boreal summer (Figure 6c). In contrast, newSOA in the boreal winter is mostly generated in the SH, where the atmosphere is clean and without a high number concentration of pre-existing particles. More newSOA survives in the boreal winter than in other seasons, resulting in the highest column number concentration of newSOA in the nucleation mode (Figure 6g), which is 67% higher than that in the boreal summer (Table 4). The column number concentration of newSOA in the Aitken mode is lowest in the summer (Figure 7c) while the column number concentration of newSOA in the accumulation mode is largest in the summer (Figure 8c). Thus, newSOA is more able to grow to the larger particles in the summer in the NH. Vertical convection and in-situ growth in the summer in the NH leads to high number concentrations of

newSOA in the accumulation mode in the upper troposphere (Figure 8d). The column number concentration of newSOA in the Aitken and accumulation modes is high in the boreal winter. There are a larger number of newSOA particles in the nucleation mode surviving removal by coagulation in the SH in the boreal winter. These newSOA grow to the Aitken and accumulation modes in the SH, resulting in the high concentration of newSOA in these modes there (Figure 7h, 8h). The high concentration of newSOA in the accumulation mode in the SH in the boreal winter offsets the very low concentration in the NH in the winter. The peak of the number concentration of newSOA in the accumulation mode changes between the NH and SH depending on the season (Figure 8).

NewSOA and new sulfate particles are only two species formed from nucleation in the model, but they have very different size, seasonal and spatial distributions. The column number concentration of newSOA in the nucleation mode is much smaller than that of new sulfate particles in the nucleation mode (Figure 9a, b) because of the lower global average organic nucleation rates than the sulfuric acid particle nucleation rates. The column number concentration of new sulfate is high in industrial regions such as East Asia, Southeast Asia, Europe and North America while it is very low in the Amazon, which is opposite to the spatial distribution of newSOA in the nucleation mode. The column number concentration of newSOA in the Aitken mode is higher than that of new sulfate (Figure 9c, d) while the column number concentration of newSOA in the accumulation mode is much lower than that of new sulfate in the accumulation mode (Figure 9e, f). Thus, new sulfate contributes more to the number concentration of accumulation mode

aerosol while newSOA dominates the number concentration of Aitken mode aerosol.

In comparison, the distribution of the mass burden of newSOA does not totally correspond with the distribution of the column number. The mass of newSOA is defined as the mass of HOMs that form newSOA combined with the organics that condense on newSOA. The horizontal distribution of the burden of newSOA in nucleation mode as well as vertical distribution of the mass concentration are similar to the distribution of number concentration of newSOA in the nucleation mode, which is highest in the upper troposphere of the Amazon (Figure 10a, 10b). This is determined by the high organic nucleation rate and small sink to remove new particles. However, the distribution of the burden of newSOA in the Aitken and accumulation modes are different from those of the column number concentration. The mass burden of newSOA in the Aitken and accumulation modes are both highest in the Amazon and the mass concentration peaks in the upper troposphere around 200hPa (Figure 10), which is similar to the mass and number distribution of newSOA in nucleation mode, but differs from the number distribution of newSOA in the Aitken and accumulation modes. The mass concentrations peak in the Amazon because the largest source of BVOC is in the Amazon. BVOC is oxidized to form SVOC there, which condenses on the newSOA and causes them to grow.

The spatial distribution patterns of the mass burden of newSOA in all three modes do not have any significant seasonal variation (Figure S6), while the mass burdens in the three modes have significant seasonal variation (Table 4). The mass

burden of newSOA is lowest in the boreal summer for all modes. The total mass burden of newSOA in the sum of the three modes in the summer is 55% lower than that in the spring (Table 4) . There are many newSOA formed in industrial areas with their high concentration of POC and sulfate during the summer. A lot of SVOC formed in the lower troposphere of the NH would mix internally with the POC and sulfate instead of with newSOA because of the small particle mass and surface area of the newSOA, leading to the 31% larger burden of SOA mixed with other aerosol (mixSOA) in the boreal summer than in spring (Table 4). Although the burden of newSOA is low in the summer and fall, the burden of total SOA (newSOA+mixSOA) is much higher in the summer and fall than that in the other two seasons due to the higher emission rates of BVOC (e.g. isoprene, α-pinene and limonene in the model). The burden of mixSOA dominates the burden of total SOA, which explains 90% of total SOA in the boreal summer.

3.4 Sensitivity experiments

3.4.1 Sensitivity to the inclusion of IEPOX, glyoxal and methylglyoxal to condensed organics

As described in the section of methods, we assumed IEPOX, glyoxal and methylglyoxal are able to be taken up by newSOA in EX1 because sulfuric acid is an important constituent of newSOA which would provide an acidic and wet environment. Therefore, the low-volatility products formed from IEPOX, glyoxal and methylglyoxal could take part in the growth of newSOA, so that newSOA

would quickly grow after its initial formation. In addition, the increased organic mass in the newSOA promotes increased partitioning of SVOC to newSOA and further increases the size of newSOA. As a result, the column number concentration of newSOA is decreased by 30% in the nucleation mode while increased by 47% and 259% in the Aitken and accumulation modes, respectively (Table 5). Figure 11 shows the difference in the column number concentration and the profile of the number concentration of newSOA between EX1 and BASE in each mode. In addition, the burden of organics condensed on sulfate and soot decreases by 77% and 21%, respectively, while the burden of newSOA increases by 277%, although the total burden of SOA changes very little (Table 5). Because there are more newSOA with a large concentration of organics that act to absorb HOMs, especially in the lower troposphere, the concentration of HOMs decreases significantly. The HET rate is limited by the concentration of HOMs in the NH where the concentration of sulfuric acid is high. The annual global average HET rate decreases by 36% with the most significant decrease in the boreal summer in the NH. The ION and NON new particle production rates are only changed by a small amount due to the formation of sufficient HOMs in the tropics where ION and NON rates are highest. Overall, the total annual global average organic nucleation rate decreases by 16% with the largest influence in the NH. This explains the significant decrease of newSOA in the nucleation mode in the NH (Figure 11a). The concentration of low-volatility products formed from IEPOX, glyoxal and methylglyoxal are always high in the lower troposphere of the tropics, so that the number concentrations of newSOA in the Aitken and accumulation

modes increase significantly there. As a result, the surface number concentration of aerosol is significantly changed in EX1. The average temporal correlation coefficient between simulated and observed number concentration of aerosol increases from 0.134 in BASE to 0.349 in EX1 (Table 2) with improvements seen at 73% of observation sites (Table S4). This is basically because of the larger contribution of SOA to the total aerosol number concentration. However, the mean number concentration is overestimated by 29% more in EX1 compared to BASE due to the increased number of newSOA in the Aitken and accumulation modes.

3.4.2 Sensitivity to the suppression of organic nucleation by isoprene

We examined the sensitivity of the formation of newSOA to the inclusion of the suppression by isoprene in EX2 as described in Section 2 based on field measurements that indicated that NPF rarely takes place when the concentration of isoprene is high (*Kiendler-Scharr et al.*, 2012; *Lee et al.*, 2016). The assumption of isoprene suppression in the model causes organic nucleation to shut off in many places, so that the annual global average organic nucleation rate is decreased by 67% (Table 5). The inclusion of isoprene suppression has a large influence on all the mechanisms of organic nucleation. The annual global average ION rate is influenced the most because of the high emission rate of isoprene in the tropics where the ION nucleation rate is most important. As the result, isoprene suppression makes the largest contribution to the decrease of the number concentration of newSOA in the tropics. However, the column number concentration of newSOA decreases over the entire world. The global average

column number concentration of newSOA in the nucleation and Aitken modes are decreased by 59%, while they decrease by 24% in the accumulation mode (Table 5). Isoprene suppression also changes the burden of newSOA though not as significantly as the number concentration. The global average burden of total newSOA is decreased by 8%. Owing to the decrease in the total number concentration of newSOA, more SVOC is distributed to the newSOA in the accumulation mode so that in this mode, the burden of newSOA is 53% larger in EX2 than in BASE, even though the column number concentration of newSOA in the accumulation mode is 24% lower in EX2. The bias in the number concentration of aerosol between the simulation and observations decreases from 19% to -0.1% because of the reduction in the number concentration of newSOA (Table 2). However, the average temporal correlation coefficient is decreased from 0.134 to 0.112 (Table S4). In general, isoprene suppression does not change the seasonal pattern of the aerosol number concentration significantly at most of sites used here (Figure S7).

3.4.3 Sensitivity of newSOA to the inclusion of HOMs from the oxidation of limonene

In EX3, we added HOMs formed from the oxidation of limonene to those from α-pinene to examine the uncertainty associated with additional sources of HOMs. The yield of HOMs from limonene is double of that from α-pinene, but the annual emissions of limonene are only 18% of the emissions of α-pinene. Limonene is primarily emitted in the tropics. Moreover, since the lifetime

of limonene is much shorter than that of α-pinene, the additional HOMs are mainly in the lower troposphere of tropics. The high concentration of preexisting aerosol within the lower troposphere is a large sink for HOMs. Including the mechanism of HOMs from limonene increases annual global average HOM concentration by 13%. The annual global average organic nucleation rate is increased by 2.3% in EX3, which is mostly explained by a 5.0% increase in the ION rate in the tropics (Table 5). The NON rate is increased by 6.3% because NON mostly occurs in the lower troposphere of the Amazon, while the HET rate does not change significantly. As a result, the newSOA in the Aitken mode is increased by 5.2% and the newSOA in the accumulation mode is increased by 2.9% (Table 5). In addition, there is more sulfuric acid condensed on newSOA in EX3 than in BASE due to the higher number concentration of newSOA, resulting in a decrease in the number concentration of new sulfate. The change of total aerosol number concentration is determined by the increased newSOA number and decreased new sulfate number. The average CN concentration in EX3 is overestimated by 19% compared with the observation at the 27 sites compared here, which is very similar to the result for BASE (Table 2). Including HOMs from limonene increases the average temporal correlation coefficient of the aerosol number concentration between the simulation and observation at 65% of sites, but primarily only improves the temporal correlation coefficient at sites with a correlation coefficients of <0.1 (Table S4, Figure S7). In total, the average temporal correlation coefficient is increased from 0.134 to 0.163 after including HOMs from limonene in EX3. The

additional mechanism of HOMs formed from the oxidation of limonene doesn't change the evaluation of the model by very much.

4 Conclusions

In this study, the number concentration and burden of SOA is simulated using the latest CESM/IMPACT model including three organic nucleation mechanisms: HET, NON and ION. NewSOA is formed from HOMs which are the product of the ozone and OH oxidation of α-pinene through explicit chemical reactions. After the inclusion of organic nucleation, the aerosol number concentration and the size distribution of SOA are changed. The simulated monthly average OC and CN concentration were evaluated by comparing with observations from the IMPROVE and EMEP networks as well as with several other sites over the world. The BASE version of the model underestimates the average surface OC concentration by 29% in the IMPROVE network (United States) and 67% in the EMEP (Europe) network. The comparison of the sensitivity tests with these data is similar to that of the base model. This large underestimation is possibly caused by the coarse resolution of the model and the underestimation of POC emissions. The model was able to capture the temporal variation of the OC concentration in most regions of United States with an average correlation coefficient of 0.73, but failed to predict the seasonal and temporal pattern at EMEP sites because of the high OC concentration observed in the winter in Europe. The lack of any significant seasonal variation in the POC emission rates explains much

of the negative correlation coefficients in Europe. The model predicted the spatial pattern of OC concentration well in both the United States and Europe. Organic nucleation contributes more to the aerosol number concentration than to the OC concentration. The average CN concentrations are close to the observations and the temporal correlation coefficients are increased at most of sites after including organic nucleation in the model. The temporal correlations of the CN concentrations between the simulation and observations are poor at the sites where the predicted CN concentration is dominated by sulfate. Thus, the simulated number concentration of newSOA shows a better correlation with the observed CN concentration than does the simulated total CN concentration.

Our model indicates that the ION mechanism generates 60% of the total newSOA, and is thus the most important organic nucleation mechanism. ION always occurs in the middle and upper troposphere of the tropics so that the upper troposphere in the Amazon is the largest source of newSOA. The organic nucleation rate is highest in the boreal fall, which is determined by the high ION rate in the fall in the Amazon. HET is the most widely present organic nucleation mechanism and is high in the middle latitudes of the NH. HET dominates the global organic nucleation rate in the boreal summer. The HET nucleation rate also explains most of the organic nucleation rate within the PBL, while ION explains most of the organic nucleation rate in the middle and upper troposphere. Due to the high concentration of preexisting aerosol in the lower troposphere of NH, the high HET nucleation rate does not lead to a high number concentration of newSOA in the nucleation mode there. In comparison, ION contributes to the large

concentration of newSOA in the nucleation mode in the upper troposphere of the Amazon. The column number concentration of newSOA in the Aitken mode is highest in the tropics because of high concentration of nucleated newSOA. However, the large number of nucleated newSOA are hard to grow to the accumulation mode in the tropics because of insufficient condensable sulfuric acid and SVOC. The column number concentration of newSOA in the accumulation mode is highest in the middle latitudes even though the number concentration of newSOA in the nucleation mode is not very high due to the high concentrations of sulfuric acid and SVOC in the lower troposphere. However, the distribution of the mass burden of newSOA in the Aitken and accumulation modes differs from that of the number concentration and is highest in the middle and upper troposphere of Amazon because of the much higher concentrations of SVOC that are able to condense on newSOA formed in the Amazon.

Three sensitivity experiments were conducted to investigate uncertainties in the nucleation and growth mechanisms in the model. When the low-volatility products formed from IEPOX, glyoxal and methylglyoxal take part in the growth of newSOA, a larger number of newSOA grow from the nucleation mode to the Aitken and accumulation modes, while the burden of SOA associated with sulfate and soot decrease significantly. The temporal correlation coefficient of the number concentration of surface aerosol between the simulation and observations is increased significantly but the overestimation of the average number concentration is larger. When the model includes the suppression of organic nucleation by isoprene, the annual global average organic nucleation rate is decreased by 67%,

resulting in a sharp decrease in the number concentration of newSOA in all modes. Isoprene suppression narrows the gap between the observations and the simulation of the average number concentration of aerosols significantly, but the temporal correlation becomes some worse. When we added the HOMs formed from the oxidation of limonene, the annual average organic nucleation rate is increased by 2.3% which is primarily determined by the increased ION rate in the tropics. The additional HOMs from limonene only have a small influence on the evaluation of the number concentration and temporal correlation coefficient.

Since the knowledge of organic nucleation is still very limited, large uncertainties in the simulation of newSOA remain. Future work is needed to investigate the mechanism to form HOMs and newSOA as well as the properties of newSOA. The chemical mechanism to form HOMs used in our model is one of the most detailed process models published to date. However, there likely are many other chemical mechanisms and different precursors to form HOMs that nucleate with different reaction times that need to be investigated in laboratory. The mechanism of HOM formation could have an influence on the number of newSOA and its spatial distribution. We only include four HOM species in the model which are formed from α-pinene and have been shown to contribute to nucleation in previous publications (*Kulmala et al.*, 2013; *Ortega et al.*, 2016; *Schobesberger et al.*, 2013; *Zhang et al.*, 2015; *Ziemann*, 2002). Semi-volatile organics only help to grow pre-existing aerosol, although some may have very low volatility at low temperatures and high altitudes. This treatment may cause an underestimation of organic nucleation rate. Further studies to determine whether more organics lead to

nucleation and should be included. Organic nucleation was described by a set of parameterizations based on experimental data, but these parameterizations may not fit all situations. A detailed dynamic mechanism of organic nucleation is needed to reduce the uncertainty of predicting the organic nucleation rate. Also, more studies are needed to determine the acidity and viscosity of newSOA in different modes. The growth mechanism of newSOA depends on the acidity and viscosity, which could change the predicted size distribution of newSOA. An accurate size distribution is needed to accurately estimate the radiative effects of aerosol. Moreover, work is needed to improve our understanding of the sources and sinks of BVOC as well as the seasonal variation of POC emissions. The correct description of emissions in the model is important to improve the prediction of the burden of aerosol and its seasonal pattern.

Acknowledgments

We are grateful for funding from the NASA ACMAP program under grant number NNX15AE34G as well as an NSF-GEO grant number 1540954. Computer time was provided by the NCAR CISL. We thank Ann Hjellbrekke and the EMEP network (http://ebas.nilu.no) for the surface OC concentration and surface aerosol number concentration in Europe. We thank the IMPROVE network (http://vista.cira.colostate.edu/Improve/improve-data/) for the surface OC concentration in the US. The updated CESM/IMPACT model and all model results used in this paper is archived at http://doi.org/10.7302/1c53-zy98.

References

Andreae, M. O., et al. (2015), The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols, *Atmospheric Chemistry and Physics*, *15*(18), 10723-10776, doi:10.5194/acp-15-10723-2015.

Andreae, M. O., et al. (2018), Aerosol characteristics and particle production in the upper troposphere over the Amazon Basin, *Atmospheric Chemistry and Physics*, *18*(2), 921-961, doi:10.5194/acp-18-921-2018.

Apsokardu, M. J., and M. V. Johnston (2018), Nanoparticle growth by particle-phase chemistry, *Atmospheric Chemistry and Physics*, *18*(3), 1895-1907, doi:10.5194/acp-18-1895-2018.

Bond, T. C., D. G. Streets, K. F. Yarber, S. M. Nelson, J. H. Woo, and Z. Klimont (2004), A technology -based global inventory of black and organic carbon emissions from combustion, *Journal of Geophysical Research: Atmospheres*, *109*(D14).

Boy, M., et al. (2008), New particle formation in the front range of the colorado rocky mountains, *Atmospheric Chemistry and Physics*, *8*(6), 1577-1590, doi:10.5194/acp-8-1577-2008.

Burkart, J., et al. (2017), Organic Condensation and Particle Growth to CCN Sizes in the Summertime Marine Arctic Is Driven by Materials More Semivolatile Than at Continental Sites, *Geophysical Research Letters*, *44*(20), 10,725-710,734, doi:10.1002/2017gl075671.

Bzdek, B. R., and M. V. Johnston (2010), New Particle Formation and Growth in the Troposphere, *Anal. Chem.*, *82*(19), 7871-7878, doi:10.1021/ac100856j.

Chen, M., et al. (2012), Acid-base chemical reaction model for nucleation rates in the polluted atmospheric boundary layer, *Proc. Natl. Acad. Sci. U. S. A.*, *109*(46), 18713-18718, doi:10.1073/pnas.1210285109.

Chen, Q., T. M. Fu, J. L. Hu, Q. Ying, and L. Zhang (2017), Modelling secondary organic aerosols in China, *Natl. Sci. Rev.*, *4*(6), 806-809, doi:10.1093/nsr/nwx143.

Claeys, M., et al. (2009), Terpenylic Acid and Related Compounds from the Oxidation of alpha-Pinene: Implications for New Particle Formation and Growth above Forests, *Environ. Sci. Technol.*, *43*(18), 6976-6982, doi:10.1021/es9007596.

Dunne, E. M., et al. (2016), Global atmospheric particle formation from CERN CLOUD measurements, *Science*, *354*(6316), 1119-1124, doi:10.1126/science.aaf2649.

Ehn, M., et al. (2012), Gas phase formation of extremely oxidized pinene reaction products in chamber and ambient air, *Atmospheric Chemistry and Physics*, *12*(11), 5113-5127, doi:10.5194/acp-12-5113-2012.

Ehn, M., et al. (2014), A large source of low-volatility secondary organic aerosol, *Nature*, *506*(7489), 476-479, doi:10.1038/nature13032.

Ervens, B., B. J. Turpin, and R. J. Weber (2011), Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field and model studies, *Atmospheric Chemistry and Physics*, *11*(21), 11069-11102, doi:10.5194/acp-11-11069-2011.

Gaston, C. J., T. P. Riedel, Z. F. Zhang, A. Gold, J. D. Surratt, and J. A. Thornton (2014), Reactive Uptake of an Isoprene-Derived Epoxydiol to Submicron Aerosol Particles, *Environ. Sci. Technol.*, *48*(19), 11178-11186, doi:10.1021/es5034266.

Gelencser, A., B. May, D. Simpson, A. Sanchez-Ochoa, A. Kasper-Giebl, H. Puxbaum, A. Caseiro, C. Pio, and M. Legrand (2007), Source apportionment of PM2.5 organic aerosol over Europe: Primary/secondary, natural/anthropogenic,

and fossil/biogenic origin, *Journal of Geophysical Research-Atmospheres*, *112*(D23), 12, doi:10.1029/2006jd008094.

Gilardoni, S., E. Vignati, F. Cavalli, J. P. Putaud, B. R. Larsen, M. Karl, K. Stenstrom, J. Genberg, S. Henne, and F. Dentener (2011), Better constraints on sources of carbonaceous aerosols using a combined C-14 - macro tracer analysis in a European rural background site, *Atmospheric Chemistry and Physics*, *11*(12), 5685-5700, doi:10.5194/acp-11-5685-2011.

Glasius, M., et al. (2018), Organosulfates in aerosols downwind of an urban region in central Amazon, *Environ Sci Process Impacts*, *20*(11), 1546-1558, doi:10.1039/c8em00413g.

Gordon, H., et al. (2016), Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation, *Proc Natl Acad Sci U S A*, *113*(43), 12053-12058, doi:10.1073/pnas.1602360113.

Guenther, A. B., X. Jiang, C. L. Heald, T. Sakulyanontvittaya, T. Duhl, L. K. Emmons, and X. Wang (2012), The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, *Geoscientific Model Development*, *5*(6), 1471-1492, doi:10.5194/gmd-5-1471-2012.

Häkkinen, S. A. K., et al. (2012), Long-term volatility measurements of submicron atmospheric aerosol in Hyytiälä, Finland, *Atmospheric Chemistry and Physics*, *12*(22), 10771-10786, doi:10.5194/acp-12-10771-2012.

Hall, W. A., and M. V. Johnston (2011), Oligomer Content of alpha-Pinene Secondary Organic Aerosol, *Aerosol Sci. Technol.*, *45*(1), 37-45, doi:10.1080/02786826.2010.517580.

Hallquist, M., et al. (2009), The formation, properties and impact of secondary organic aerosol: current and emerging issues, *Atmospheric Chemistry and Physics*, *9*(14), 5155-5236.

Hand, J., S. Copeland, D. Day, A. Dillner, H. Indresand, W. Malm, C. McDade, C. Moore, M. Pitchford, and B. Schichtel (2011), Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States, Report V, *Colorado State University, Fort Collins CO*.

Held, A., A. Nowak, W. Birmili, A. Wiedensohler, R. Forkel, and O. Klemm (2004), Observations of particle formation and growth in a mountainous forest region in central Europe, *Journal of Geophysical Research-Atmospheres*, *109*(D23), doi:10.1029/2004jd005346.

Hu, W. W., et al. (2015), Characterization of a real-time tracer for isoprene epoxydiols-derived secondary organic aerosol (IEPOX-SOA) from aerosol mass spectrometer measurements, *Atmospheric Chemistry and Physics*, *15*(20), 11807-11833, doi:10.5194/acp-15-11807-2015.

Ito, A., S. Sillman, and J. E. Penner (2007), Effects of additional nonmethane volatile organic compounds, organic nitrates, and direct emissions of oxygenated organic species on global tropospheric chemistry, *Journal of Geophysical Research-Atmospheres*, *112*(D6), 21, doi:10.1029/2005jd006556.

Jenkin, M. E., S. M. Saunders, V. Wagner, and M. J. Pilling (2003), Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part B): tropospheric degradation of aromatic volatile organic compounds, *Atmospheric Chemistry and Physics*, *3*, 181-193, doi:10.5194/acp-3-181-2003.

Jenkin, M. E., J. C. Young, and A. R. Rickard (2015), The MCM v3.3.1 degradation scheme for isoprene, *Atmospheric Chemistry and Physics*, *15*(20), 11433-11459, doi:10.5194/acp-15-11433-2015.

Jimenez, J. L., et al. (2009), Evolution of Organic Aerosols in the Atmosphere, *Science*, *326*(5959), 1525-1529, doi:10.1126/science.1180353.

Jokinen, T., et al. (2015), Production of extremely low volatile organic compounds from biogenic emissions: Measured yields and atmospheric implications, *Proc Natl Acad Sci U S A*, *112*(23), 7123-7128, doi:10.1073/pnas.1423977112.

Kanawade, V. P., B. T. Jobson, A. B. Guenther, M. E. Erupe, S. N. Pressley, S. N. Tripathi, and S. H. Lee (2011), Isoprene suppression of new particle formation in a mixed deciduous forest, *Atmospheric Chemistry and Physics*, *11*(12), 6013-6027, doi:10.5194/acp-11-6013-2011.

Kiendler-Scharr, A., et al. (2012), Isoprene in poplar emissions: effects on new particle formation and OH concentrations, *Atmospheric Chemistry and Physics*, *12*(2), 1021-1030, doi:10.5194/acp-12-1021-2012.

Kiendler-Scharr, A., J. Wildt, M. Dal Maso, T. Hohaus, E. Kleist, T. F. Mentel, R. Tillmann, R. Uerlings, U. Schurr, and A. Wahner (2009), New particle formation in forests inhibited by isoprene emissions, *Nature*, *461*(7262), 381-384, doi:10.1038/nature08292.

Kirkby, J., et al. (2016), Ion-induced nucleation of pure biogenic particles, *Nature*, *533*(7604), 521-526, doi:10.1038/nature17953.

Kristensen, K., T. Cui, H. Zhang, A. Gold, M. Glasius, and J. D. Surratt (2014), Dimers in α-pinene secondary organic aerosol: effect of hydroxyl radical, ozone, relative humidity and aerosol acidity, *Atmospheric Chemistry and Physics*, *14*(8), 4201-4218, doi:10.5194/acp-14-4201-2014.

Kuang, C., P. H. McMurry, A. V. McCormick, and F. L. Eisele (2008), Dependence of nucleation rates on sulfuric acid vapor concentration in diverse atmospheric locations, *Journal of Geophysical Research-Atmospheres*, *113*(D10), doi:10.1029/2007jd009253.

Kulmala, M., et al. (2013), Direct Observations of Atmospheric Aerosol Nucleation, *Science*, *339*(6122), 943-946, doi:10.1126/science.1227385.

Kulmala, M., H. Vehkamaki, T. Petaja, M. Dal Maso, A. Lauri, V. M. Kerminen, W. Birmili, and P. H. McMurry (2004), Formation and growth rates of ultrafine atmospheric particles: a review of observations, *J. Aerosol. Sci.*, *35*(2), 143-176, doi:10.1016/j.jaerosci.2003.10.003.

Lee, S. H., et al. (2016), Isoprene suppression of new particle formation: Potential mechanisms and implications, *Journal of Geophysical Research-Atmospheres*, *121*(24), 14621-14635, doi:10.1002/2016jd024844.

Lim, Y. B., Y. Tan, M. J. Perri, S. P. Seitzinger, and B. J. Turpin (2010), Aqueous chemistry and its role in secondary organic aerosol (SOA) formation, *Atmospheric Chemistry and Physics*, *10*(21), 10521-10539, doi:10.5194/acp-10-10521-2010.

Lin, G., J. E. Penner, S. Sillman, D. Taraborrelli, and J. Lelieveld (2012), Global modeling of SOA formation from dicarbonyls, epoxides, organic nitrates and peroxides, *Atmospheric Chemistry and Physics*, *12*(10), 4743-4774, doi:10.5194/acp-12-4743-2012.

Lin, G., J. E. Penner, and C. Zhou (2016), How will SOA change in the future?, *Geophysical Research Letters*, *43*(4), 1718-1726, doi:10.1002/2015gl067137.

Lin, G., S. Sillman, J. E. Penner, and A. Ito (2014), Global modeling of SOA: the use of different mechanisms for aqueous-phase formation, *Atmospheric Chemistry and Physics*, *14*(11), 5451-5475, doi:10.5194/acp-14-5451-2014.

Makkonen, R., A. Asmi, V. M. Kerminen, M. Boy, A. Arneth, P. Hari, and M. Kulmala (2012), Air pollution control and decreasing new particle formation lead to strong climate warming, *Atmospheric Chemistry and Physics*, *12*(3), 1515-1524, doi:10.5194/acp-12-1515-2012.

Mann, G. W., et al. (2014), Intercomparison and evaluation of global aerosol microphysical properties among AeroCom models of a range of complexity, *Atmospheric Chemistry and Physics*, *14*(9), 4679-4713, doi:10.5194/acp-14-4679-2014.

Martin, S. T., et al. (2010), SOURCES AND PROPERTIES OF AMAZONIAN AEROSOL PARTICLES, *Reviews of Geophysics*, *48*, 42, doi:10.1029/2008rg000280.

Ortega, I. K., N. M. Donahue, T. Kurten, M. Kulmala, C. Focsa, and H. Vehkamaki (2016), Can Highly Oxidized Organics Contribute to Atmospheric New Particle Formation?, *J Phys Chem A*, *120*(9), 1452-1458, doi:10.1021/acs.jpca.5b07427.

Paasonen, P., et al. (2013), Warming-induced increase in aerosol number concentration likely to moderate climate change, *Nature Geoscience*, *6*(6), 438-442, doi:10.1038/ngeo1800.

Paasonen, P., et al. (2010), On the roles of sulphuric acid and low-volatility organic vapours in the initial steps of atmospheric new particle formation, *Atmospheric Chemistry and Physics*, *10*(22), 11223-11242, doi:10.5194/acp-10-11223-2010.

Pierce, J. R., and P. J. Adams (2009), Uncertainty in global CCN concentrations from uncertain aerosol nucleation and primary emission rates, *Atmospheric Chemistry and Physics*, *9*(4), 1339-1356, doi:10.5194/acp-9-1339-2009.

Pierce, J. R., D. M. Westervelt, S. A. Atwood, E. A. Barnes, and W. R. Leaitch (2014), New-particle formation, growth and climate-relevant particle production in Egbert, Canada: analysis from 1 year of size-distribution observations, *Atmospheric Chemistry and Physics*, *14*(16), 8647-8663, doi:10.5194/acp-14-8647-2014.

Reiter, R. (1992), *Phenomena in atmospheric and environmental electricity*.

Riccobono, F., et al. (2014), Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles, *Science*, *344*(6185), 717-721, doi:10.1126/science.1243527.

Saunders, S. M., M. E. Jenkin, R. G. Derwent, and M. J. Pilling (2003), Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, *Atmospheric Chemistry and Physics*, *3*, 161-180, doi:10.5194/acp-3-161-2003.

Schobesberger, S., et al. (2013), Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules, *Proc Natl Acad Sci U S A*, *110*(43), 17223-17228, doi:10.1073/pnas.1306973110.

Scott, C. E., S. R. Arnold, S. A. Monks, A. Asmi, P. Paasonen, and D. V. Spracklen (2018), Substantial large-scale feedbacks between natural aerosols and climate, *Nature Geoscience*, *11*(1), 44-48, doi:10.1038/s41561-017-0020-5.

Scott, C. E., et al. (2014), The direct and indirect radiative effects of biogenic secondary organic aerosol, *Atmospheric Chemistry and Physics*, *14*(1), 447-470, doi:10.5194/acp-14-447-2014.

Shrivastava, M., M. O. Andreae, P. Artaxo, H. M. Barbosa, L. K. Berg, J. Brito, J. Ching, R. C. Easter, J. Fan, and J. D. Fast (2019), Urban pollution greatly enhances formation of natural aerosols over the Amazon rainforest, *Nature communications*, *10*(1), 1046.

Shrivastava, M., et al. (2017), Recent advances in understanding secondary organic aerosol: Implications for global climate forcing, *Reviews of Geophysics*, *55*(2), 509-559, doi:10.1002/2016rg000540.

Shrivastava, M., et al. (2015), Global transformation and fate of SOA: Implications of low-volatility SOA and gas-phase fragmentation reactions, *Journal of Geophysical Research: Atmospheres*, *120*(9), 4169-4195, doi:10.1002/2014jd022563.

Shrivastava, M., T. E. Lane, N. M. Donahue, S. N. Pandis, and A. L. Robinson (2008), Effects of gas particle partitioning and aging of primary emissions on

urban and regional organic aerosol concentrations, *Journal of Geophysical Research: Atmospheres*, *113*(D18).

Shrivastava, M., A. Zelenyuk, D. Imre, R. Easter, J. Beranek, R. A. Zaveri, and J. Fast (2013), Implications of low volatility SOA and gas -phase fragmenta reactions on SOA loadings and their spatial and temporal evolution in the atmosphere, *Journal of Geophysical Research: Atmospheres*, *118*(8), 3328-3342.

Sihto, S. L., et al. (2006), Atmospheric sulphuric acid and aerosol formation: implications from atmospheric measurements for nucleation and early growth mechanisms, *Atmospheric Chemistry and Physics*, *6*, 4079-4091, doi:10.5194/acp-6-4079-2006.

Spracklen, D. V., et al. (2010), Explaining global surface aerosol number concentrations in terms of primary emissions and particle formation, *Atmospheric Chemistry and Physics*, *10*(10), 4775-4793, doi:10.5194/acp-10-4775-2010.

Surratt, J. D., M. Lewandowski, J. H. Offenberg, M. Jaoui, T. E. Kleindienst, E. O. Edney, and J. H. Seinfeld (2007), Effect of acidity on secondary organic aerosol formation from isoprene, *Environ. Sci. Technol.*, *41*(15), 5363-5369, doi:10.1021/es0704176.

Szidat, S., A. S. H. Prevot, J. Sandradewi, M. R. Alfarra, H. A. Synal, L. Wacker, and U. Baltensperger (2007), Dominant impact of residential wood burning on particulate matter in Alpine valleys during winter, *Geophysical Research Letters*, *34*(5), 6, doi:10.1029/2006gl028325.

Tsigaridis, K., et al. (2014), The AeroCom evaluation and intercomparison of organic aerosol in global models, *Atmospheric Chemistry and Physics*, *14*(19), 10845-10895, doi:10.5194/acp-14-10845-2014.

Turpin, B. J., and H. J. Lim (2001), Species contributions to PM2.5 mass concentrations: Revisiting common assumptions for estimating organic mass, *Aerosol Sci. Technol.*, *35*(1), 602-610, doi:10.1080/02786820119445.

Usoskin, I. G., and G. A. Kovaltsov (2006), Cosmic ray induced ionization in the atmosphere: Full modeling and practical applications, *Journal of Geophysical Research: Atmospheres*, *111*(D21).

Wang, J., et al. (2016), Amazon boundary layer aerosol concentration sustained by vertical transport during rainfall, *Nature*, *539*(7629), 416-419, doi:10.1038/nature19819.

Wang, M., and J. E. Penner (2009), Aerosol indirect forcing in a global model with particle nucleation, *Atmospheric Chemistry and Physics*, *9*(1), 239-260.

Weber, R. J., J. J. Marti, P. H. McMurry, F. L. Eisele, D. J. Tanner, and A. Jefferson (1997), Measurements of new particle formation and ultrafine particle growth rates at a clean continental site, *Journal of Geophysical Research-Atmospheres*, *102*(D4), 4375-4385, doi:10.1029/96jd03656.

Yasmeen, F., R. Vermeylen, R. Szmigielski, Y. Iinuma, O. Boge, H. Herrmann, W. Maenhaut, and M. Claeys (2010), Terpenylic acid and related compounds: precursors for dimers in secondary organic aerosol from the ozonolysis of alpha- and beta-pinene, *Atmospheric Chemistry and Physics*, *10*(19), 9383-9392, doi:10.5194/acp-10-9383-2010.

Yu, F., Z. Wang, G. Luo, and R. Turco (2008), Ion-mediated nucleation as an important global source of tropospheric aerosols, *Atmospheric Chemistry and Physics*, *8*(9), 2537-2554, doi:10.5194/acp-8-2537-2008.

Yu, F. Q. (2010), Ion-mediated nucleation in the atmosphere: Key controlling parameters, implications, and look-up table, *Journal of Geophysical Research-Atmospheres*, *115*, 12, doi:10.1029/2009jd012630.

Yu, H., et al. (2014), New Particle Formation and Growth in an Isoprene-Dominated Ozark Forest: From Sub-5 nm to CCN-Active Sizes, *Aerosol Sci. Technol.*, *48*(12), 1285-1298, doi:10.1080/02786826.2014.984801.

Zhang, Q., et al. (2007), Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes, *Geophysical Research Letters*, *34*(13), n/a-n/a, doi:10.1029/2007gl029979.

Zhang, Q., J. L. Jimenez, M. R. Canagaratna, I. M. Ulbrich, N. L. Ng, D. R. Worsnop, and Y. L. Sun (2011), Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, *Anal. Bioanal. Chem.*, *401*(10), 3045-3067, doi:10.1007/s00216-011-5355-y.

Zhang, X., R. C. McVay, D. D. Huang, N. F. Dalleska, B. Aumont, R. C. Flagan, and J. H. Seinfeld (2015), Formation and evolution of molecular products in alpha-pinene secondary organic aerosol, *Proc. Natl. Acad. Sci. U. S. A.*, *112*(46), 14168-14173, doi:10.1073/pnas.1517742112.

Zhao, D. F., et al. (2018), Effects of NOx and SO2 on the secondary organic aerosol formation from photooxidation of alpha-pinene and limonene, *Atmospheric Chemistry and Physics*, *18*(3), 1611-1628, doi:10.5194/acp-18-1611-2018.

Zhou, C., and J. E. Penner (2014), Aircraft soot indirect effect on large-scale cirrus clouds: Is the indirect forcing by aircraft soot positive or negative?, *Journal of Geophysical Research-Atmospheres*, *119*(19), 11303-11320, doi:10.1002/2014jd021914.

Zhu, J., J. E. Penner, G. Lin, C. Zhou, L. Xu, and B. Zhuang (2017), Mechanism of SOA formation determines magnitude of radiative effects, *Proc Natl Acad Sci U S A*, *114*(48), 12685-12690, doi:10.1073/pnas.1712273114.

Zhu, J., J. E. Penner, F. Yu, S. Sillman, M. O. Andreae, and H. Coe (2019), Decrease in radiative forcing by organic aerosol nucleation, climate, and land use change, *Nature Communications*, *10*(1), 423, doi:10.1038/s41467-019-08407-7.

Ziemann, P. J. (2002), Evidence for low-volatility diacyl peroxides as a nucleating agent and major component of aerosol formed from reactions of O-3 with cyclohexene and homologous compounds, *J. Phys. Chem. A*, *106*(17), 4390-4402, doi:10.1021/jp012925m.

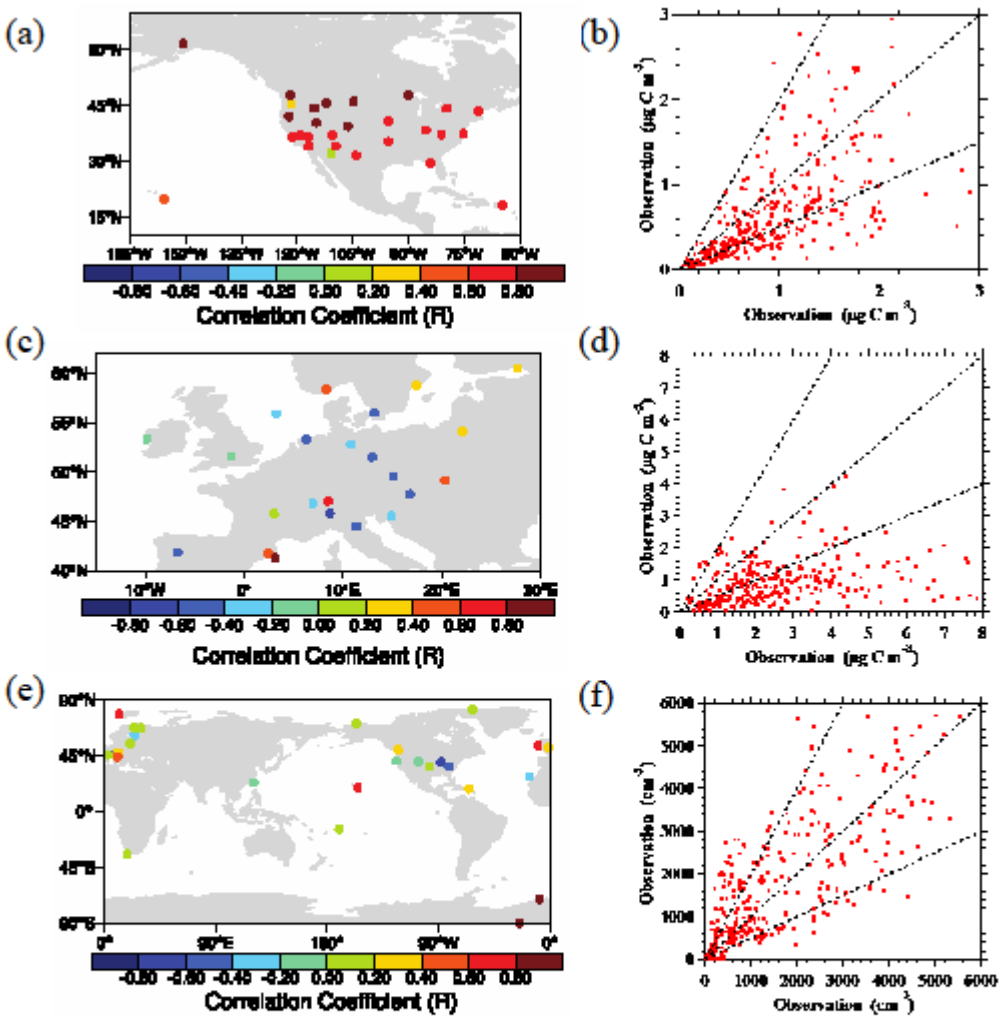


Figure 1. The temporal correlation coefficients between simulated and observed OC concentrations in the IMPROVE (a) and EMEP (b) networks, as well as the particle number concentrations (c).

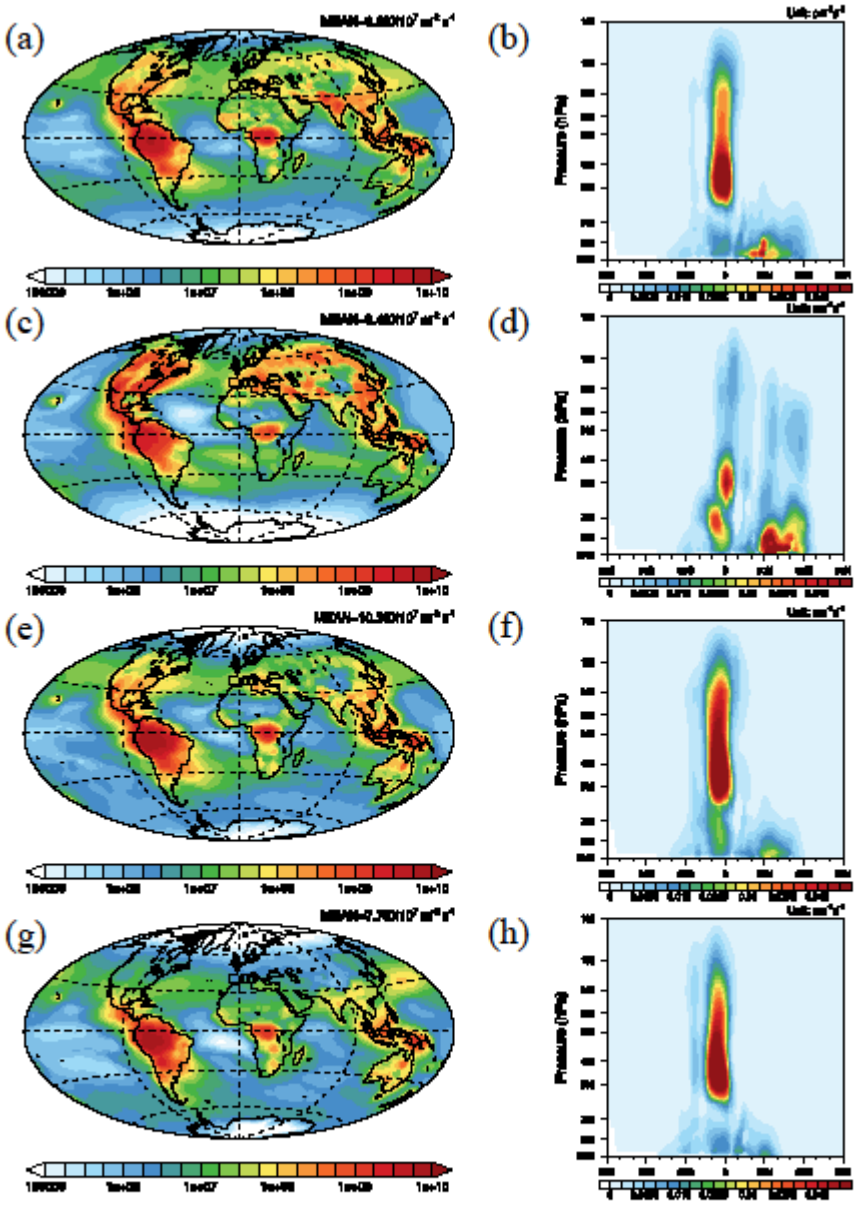


Figure 2. The vertically integrated organic nucleation rate (a,c,e,g) and zonal average of the organic nucleation rate (b,d,f,h) in boreal spring (a,b), summer (c,d), fall (e,f) and winter (g,h) for BASE case.

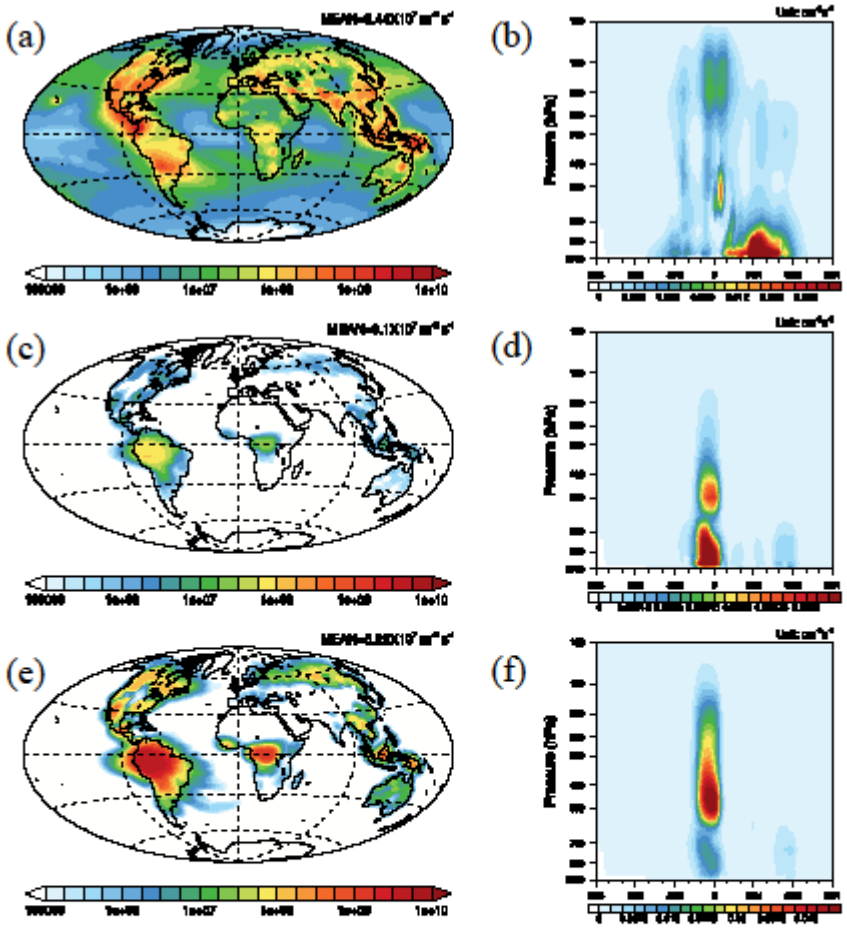


Figure 3. The annual average vertically integrated organic nucleation rate in the troposphere (a,c,e) and the zonal average organic nucleation rate (b,d,f) for HET (a,b), NON (c,d) and ION (e,f) for BASE case.

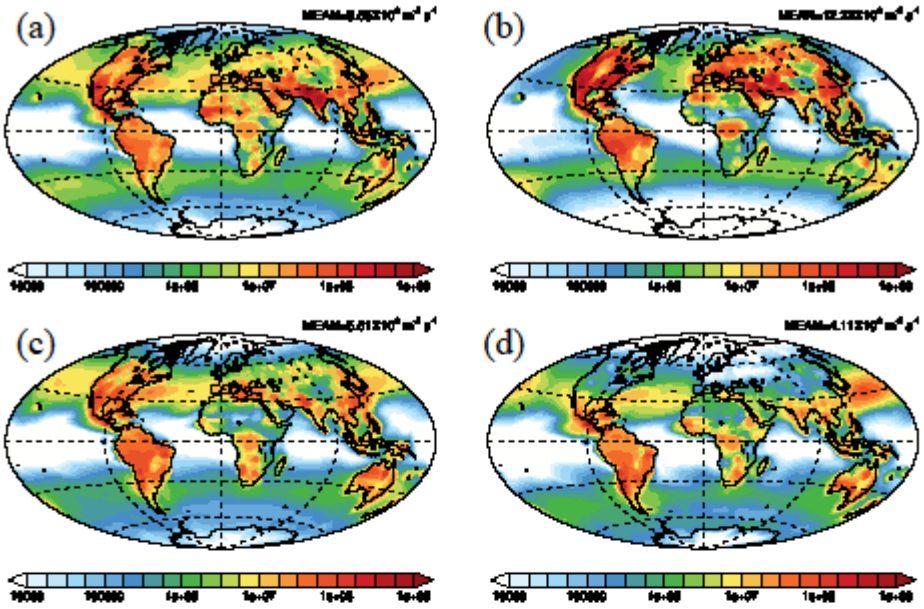


Figure 4. The organic nucleation rate integrated within the PBL in spring (a), summer (b), fall (c) and winter (d) for BASE case.

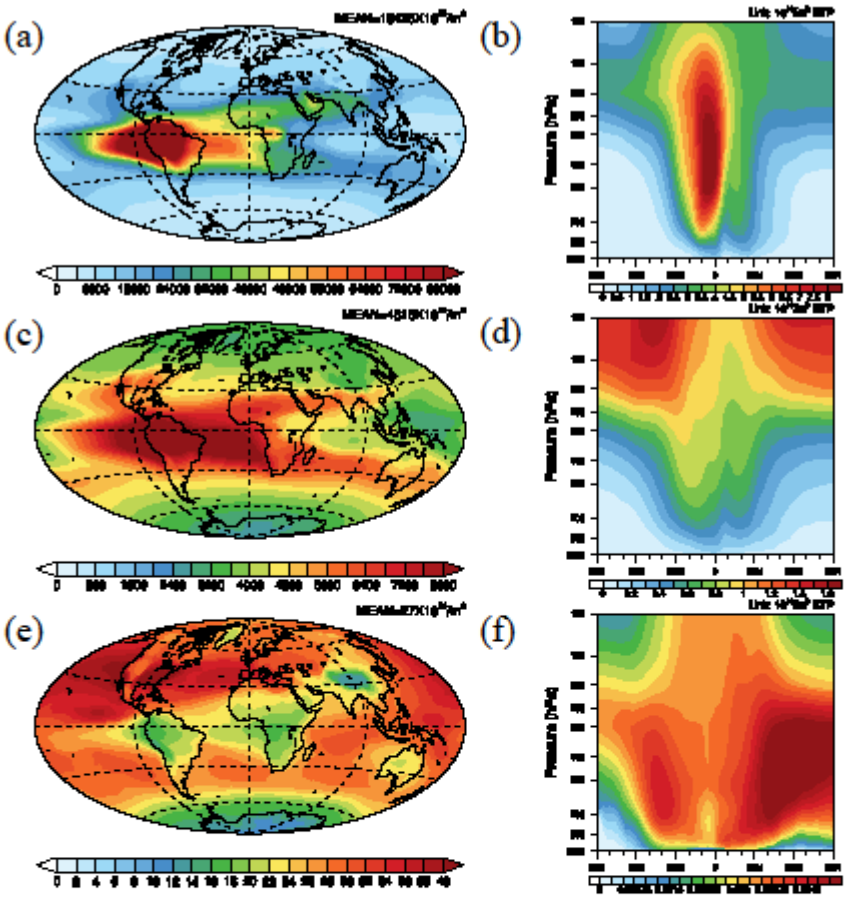


Figure 5. The annual average column number concentration (a,c,e) and zonal average number concentration (b,d,f) of newSOA in the nucleation mode (a,b), Aitken mode (b,d) and accumulation mode (e,f) for the BASE case. Note that the color bars limits vary for each mode.

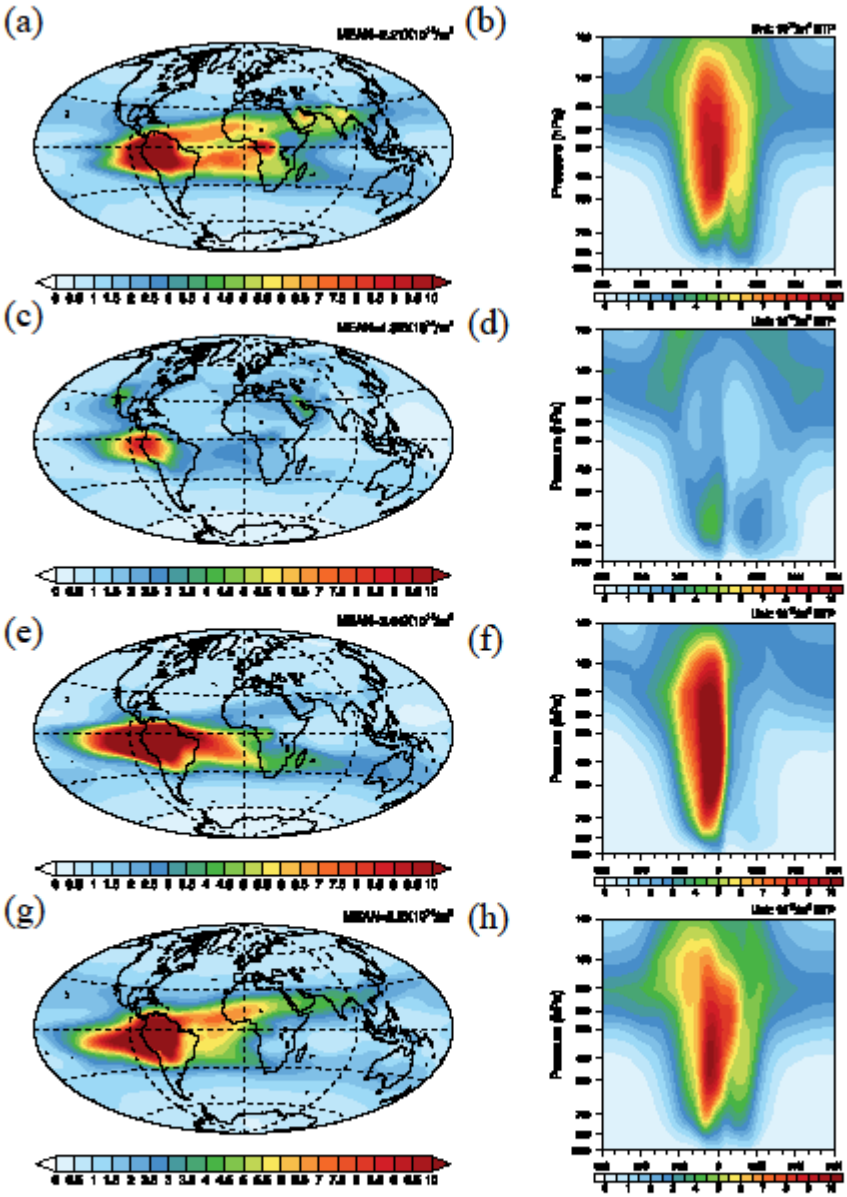


Figure 6. The column number concentration (a,c,e,g) and zonal average number concentration (b,d,f,h) of newSOA in the nucleation mode for the BASE case in boreal spring (a,b), summer (c,d), fall (e,f) and winter (g,h).

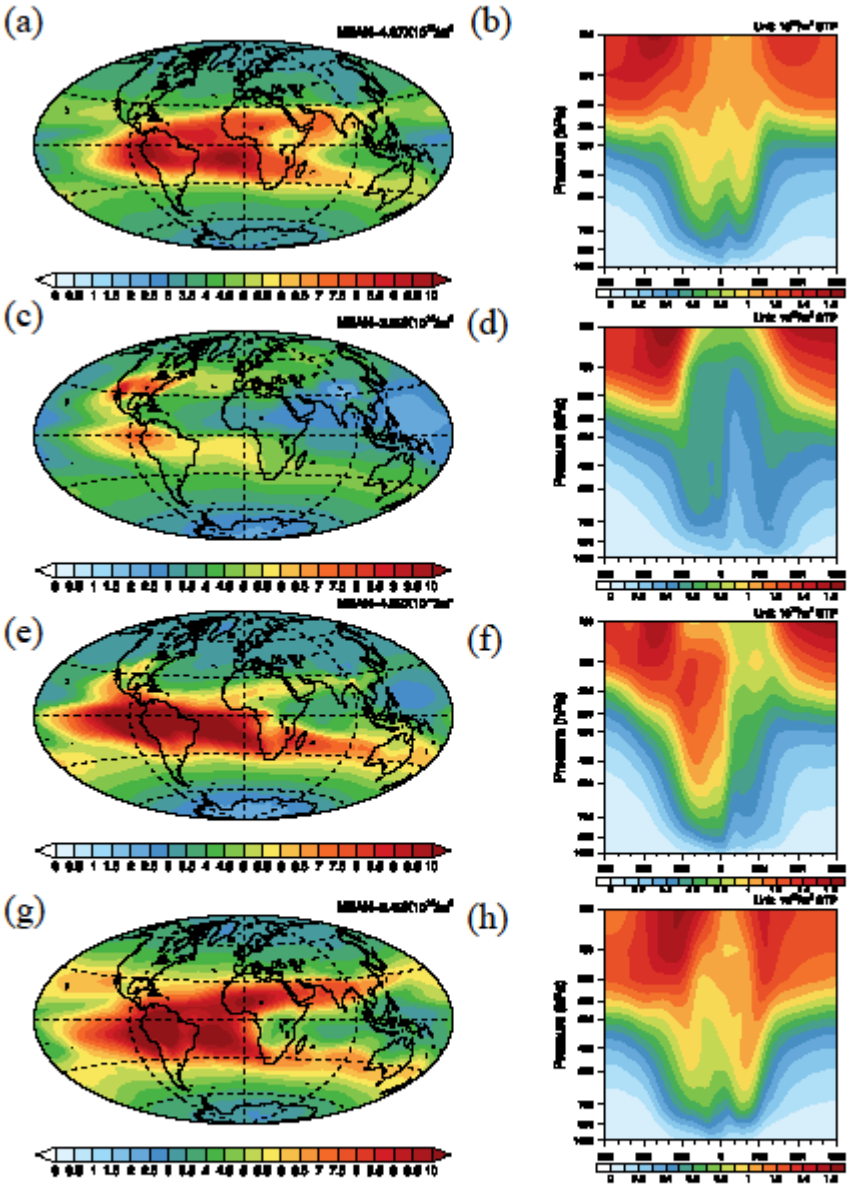


Figure 7. The column number concentration (a,c,e,g) and zonal average number concentration (b,d,f,h) of newSOA in the Aitken mode in boreal spring (a,b), summer (c,d), fall (e,f) and winter (g,h) for the BASE case.

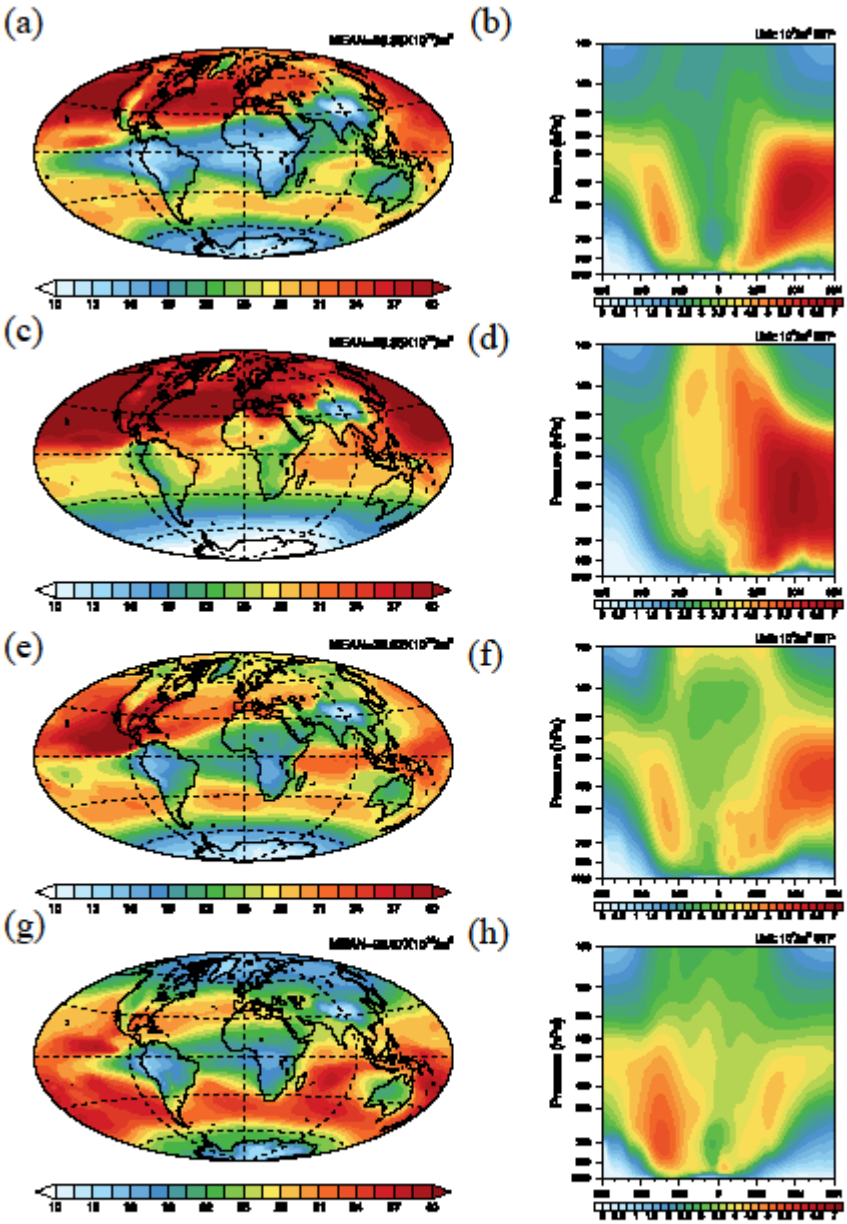


Figure 8. The column number concentration (a,c,e,g) and zonal average number concentration (b,d,f,h) of newSOA in the accumulation mode in boreal spring (a,b), summer (c,d), fall (e,f) and winter (g,h) for the BASE case.

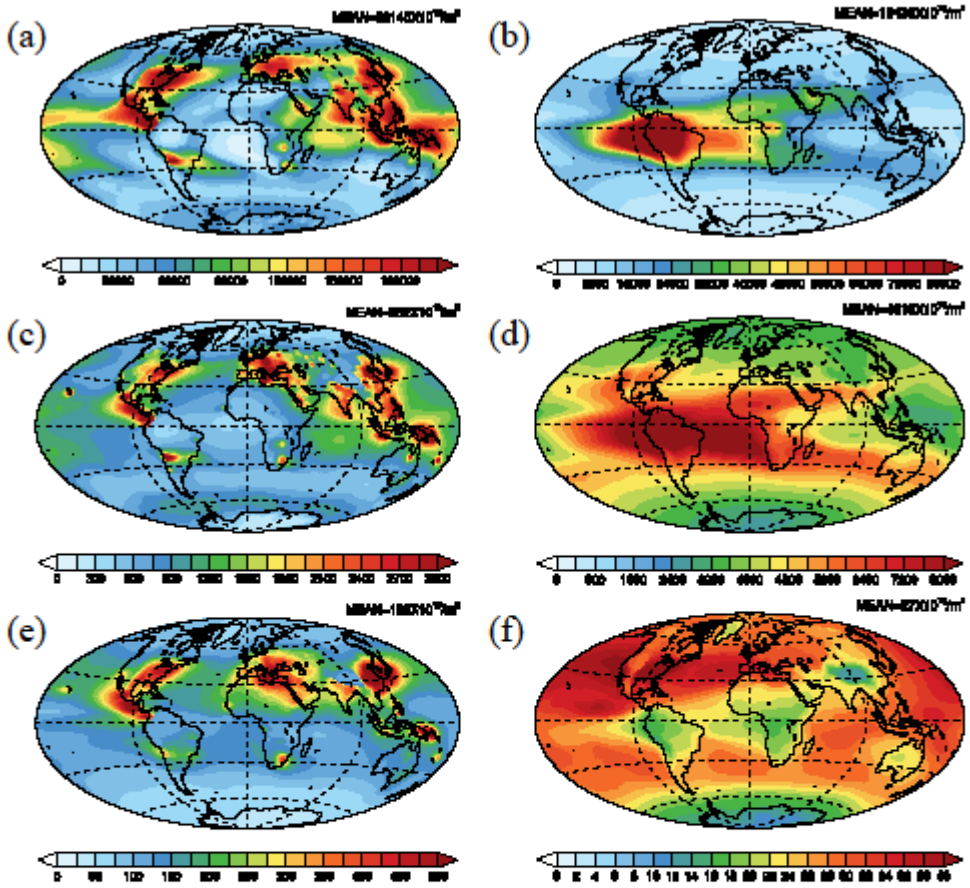


Figure 9. The annual average column number concentration of new sulfate (a, c, e) and newSOA (b, d, f) in the nucleation mode (a, b), Aitken mode (c, d) and accumulation mode (e, f) for the BASE case. Note that the color bars limits are different for sulfate and newSOA.

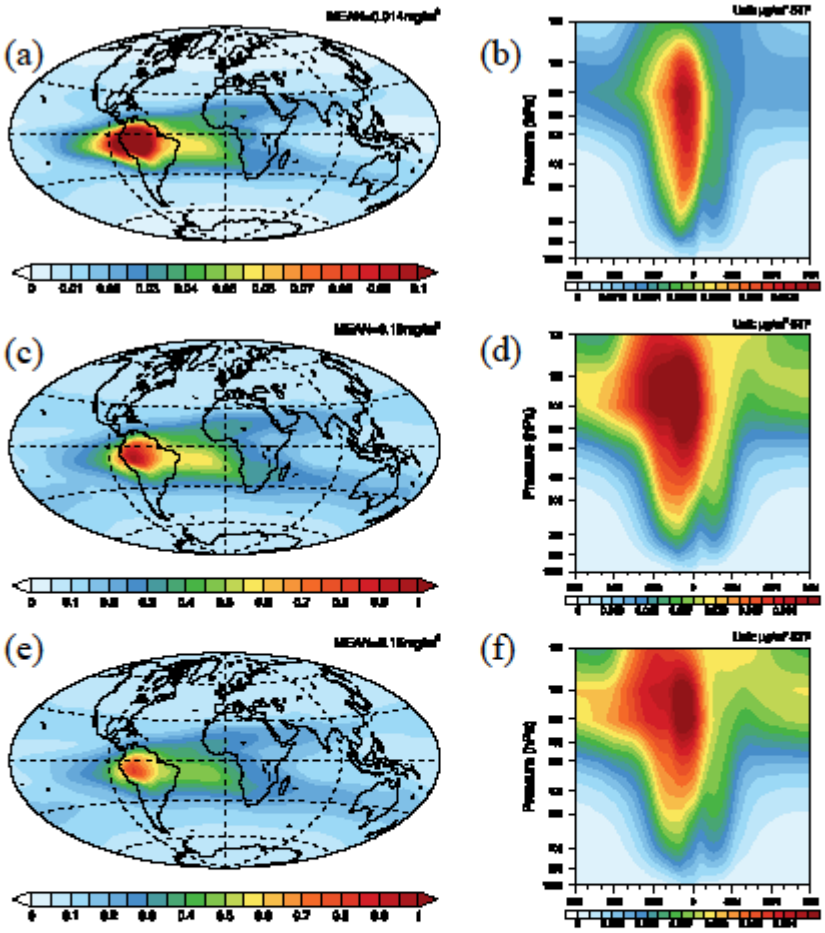
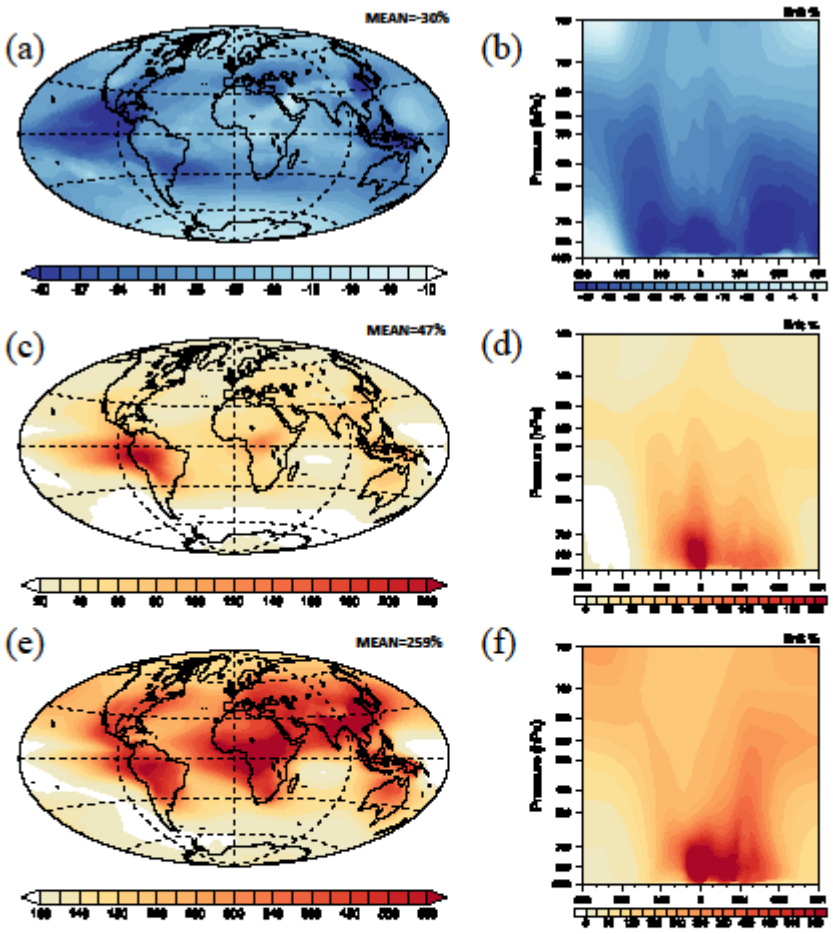


Figure 10 The annual average burden (a,c,e) and zonal average mass concentration (b,d,f) of newSOA in the nucleation mode (a,b), Aitken mode (b,d) and accumulation mode (e,f) for BASE case.



11

Figure 11 The annual average percentage difference in column number concentration (a,c,e) and zonal average number concentration (b,d,f) of newSOA in the nucleation mode (a,b), Aitken mode (b,d) and accumulation mode (e,f) between EX1 and BASE.

Tables

Table 1. The NMB as well as temporal and spatial correlation coefficient of OC concentration between simulation and observation in IMPROVE and EMEP network

Network	Observation (μg C/m^3)	Simulation (μg C/m^3)	NMB	Temporal R	Spatial R
IMPROVE	0.96	0.69	-29%	0.726	0.698
EMEP	2.80	0.919	-67%	-0.050	0.432

NMB: normalized mean bias
R: correlation coefficient of the OC concentration between the simulation and observation

Table 2. The NMB as well as the average temporal and spatial correlation coefficient of particle number concentration between simulation and observation in different schemes

Scheme	NMB	Temporal R	Spatial R
w/o organic nucleation	-4%	0.124	0.842
BASE	19%	0.134	0.846
EX1	48%	0.349	0.819
EX2	-0.1%	0.112	0.852
EX3	19%	0.163	0.843

NMB: normalized mean bias
R: correlation coefficient of the aerosol number concentration between the simulation and observation

Table 3. The seasonal and annual global average organic nucleation rate vertically integrated within the whole troposphere and PBL (Unit: 10^6 m^{-2} s^{-1})

	Scheme	Spring	Summer	Fall	Winter	Annual
Within the whole Troposphere	HET	36.12	47.14	31.53	22.61	34.35
	NON	0.72	1.17	1.22	0.71	0.96
	ION	49.98	35.90	71.00	54.60	52.87
	Total	86.82	84.21	103.75	77.92	88.18
Within the PBL	HET	7.79	10.19	4.59	3.37	6.48
	NON	0.06	0.16	0.11	0.05	0.10
	ION	0.80	1.88	0.91	0.70	1.07
	Total	8.65	12.23	5.61	4.12	7.65

Table 4. Summary of the boreal seasonal and annual global average SOA number and burden in BASE

		Spring	Summer	Fall	Winter	Annual
Aerosol Number (10^{10} m^{-2})	newSOA (nucleation)	22133	13190	20395	22032	19438
	newSOA (Aitken)	4868	3985	4921	5488	4815
	newSOA (accumulation)	26.4	29.4	26.6	26.7	27.3
Aerosol Burden (mg m^{-2})	newSOA (nucleation)	0.018	0.007	0.014	0.019	0.014
	newSOA (Aitken)	0.213	0.090	0.130	0.208	0.160
	newSOA (accumulation)	0.205	0.095	0.110	0.190	0.150
	mixSOA with sulfate	0.966	1.172	1.072	0.826	1.009
	mixSOA with soot (fossil/bio-fuel)	0.268	0.339	0.315	0.218	0.289
	mixSOA with soot (biomass burning)	0.161	0.334	0.415	0.181	0.273
	mixSOA with sea salt and dust	0.007	0.004	0.003	0.004	0.005
	Total SOA	1.84	2.04	2.06	1.65	1.90

Total SOA: newSOA and SOA internal mixed with other preexisting aerosols

Table 5. The summary of annual global average organic nucleation rate, column number concentration and burden of SOA in different schemes

		BASE	EX1	EX2	EX3
Organic nucleation Rate (10^6 m^{-2} s^{-1})	HET	34.35	22.07	11.53	33.70
	NON	0.96	0.94	0.41	1.02
	ION	52.87	50.80	17.49	55.52
	Total	88.18	73.81	29.43	90.24
Aerosol Number (10^{10} m^{-2})	newSOA(nucleation)	19438	13593	7994	19352
	newSOA(Aitken)	4816	7074	1990	5068
	newSOA(accumulation)	27.3	98.1	20.7	28.1
Aerosol Burden (mg m^{-2})	newSOA(nucleation)	0.014	0.032	0.009	0.014
	newSOA(Aitken)	0.16	0.87	0.06	0.17
	newSOA(accumulation)	0.15	0.32	0.23	0.14
	SOA mixed with sulfate	1.01	0.23	1.10	1.01
	SOA mixed with soot (fossil fuel)	0.29	0.22	0.29	0.29
	SOA mixed with soot (biomass burning)	0.27	0.22	0.27	0.27
	SOA mixed with sea salt and dust	0.005	0.005	0.004	0.005
	Total SOA	1.90	1.90	1.96	1.90

Total SOA: newSOA and SOA internal mixed with other preexisting aerosols