UNIVERSITY OF LEEDS

This is a repository copy of *Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/107267/

Version: Accepted Version

Article:
Gordon, H orcid.org/0000-0002-1822-3224, Sengupta, K, Rap, A orcid.org/0000-0002-2319-6769 et al. (78 more authors) (2016) Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation. Proceedings of the National Academy of Sciences, 113 (43). pp. 12053-12058. ISSN 0027-8424

https://doi.org/10.1073/pnas.1602360113

This is an author produced version of a paper published in Proceedings of the National Academy of Sciences. Uploaded in accordance with the publisher's self-archiving policy. In order to comply with the publisher requirements the University does not require the author to sign a non-exclusive licence for this paper.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation

Hamish Gordon, Kamalika Sengupta, Alexandru Rap, Jonathan Duplissy, Carla Frege, Christina Williamson, Martin Heinritzi, Mario Simon, Chao Yan, João Almeida, Jasmin Tröstl, Tuomo Nieminen, Ismael K. Ortega, Robert Wagner, Eimear Dunne, Alexey Adamov, Antonio Amorim, Anne-Kathrin Bernhammer, Federico Bianchi, Martin Breitenlechner, Sophia Brilke, Xuemeng Chen, Jill Craven, Antonio Dias, Sebastian Ehrhart, Lukas Fischer, Richard C. Flagan, Alessandro Franchin, Claudia Fuchs, Roberto Guida, Jani Hakala, Christopher R. Hoyle, Tuija Jokinen, Heikki Junninen, Juha Kangasluoma, Jaeseok Kim, Jasper Kirkby, Manuel Krapf, Andreas Kürten, Ari Laaksonen, Katrianne Lehtipalo, Vladimir Makhmutov, Serge Mathot, Ugo Molteni, Sarah Monks, Antti Onnela, Otso Peräkylä, Felix Piel, Tuukka Petäjä, Arnaud P. Praplan, Kirsty Pringle, Nigel A. D. Richards, Matti P. Rissanen, Linda Rondo, Nina Sarnela, Siegfried Schobesberger, Catherine E. Scott, John H. Seinfeld, Sangeeta Sharma, Mikko Sipilä, Gerhard Steiner, Yuri Stozhkov, Frank Stratmann, Antonio Tomé, Annele Virtanen, Alexander L. Vogel, Andrea C. Wagner, Paul E. Wagner, Ernest Weingartner, Daniela Wimmer, Paul M. Winkler, Penglin Ye, Xuan Zhang, Armin Hansel, Josef Dommen, Neil M. Donahue, Douglas R. Worsnop, Urs Baltensperger, Markku Kulmala, Joachim Curtius, and Kenneth S. Carslaw

November 11, 2016

Abstract

The magnitude of aerosol radiative forcing caused by anthropogenic emissions depends on the baseline state of the atmosphere under pristine preindustrial conditions. Measurements in the CERN CLOUD

chamber show that particle formation in atmospheric conditions can occur solely from biogenic vapours. Here we evaluate the potential effect of this new source of particles on pre-industrial cloud condensation nucleus (CCN) concentrations and on aerosol-cloud radiative forcing over the industrial period. Model simulations show that the pure biogenic particle formation mechanism has a much larger relative effect on CCN concentrations in the pre-industrial atmosphere than in the present atmosphere because of the lower aerosol concentrations. Consequently, pre-industrial cloud albedo is increased more than under present-day conditions, so the cooling forcing of anthropogenic aerosols is reduced. The new mechanism increases CCN concentrations by 20-100% over a large fraction of the pre-industrial lower atmosphere and the magnitude of annual global mean radiative forcing caused by changes of cloud albedo since 1750 is reduced by $0.22\,\mathrm{Wm}^{-2}$ (27%) to $-0.60\,\mathrm{Wm}^{-2}$. Model uncertainties, relatively slow formation rates and limited available ambient measurements make it difficult to establish the significance of a mechanism that has its dominant effect under pre-industrial conditions. Our simulations predict more particle formation in the Amazon than is observed. On the other hand, the first observation of pure organic nucleation has now been reported for the free troposphere. Given the potentially significant effect on anthropogenic forcing, effort should be made to better understand such naturally-driven aerosol processes.

1 Significance Text

A new mechanism for the formation of atmospheric aerosols via the gas-to-particle conversion of highly oxidised organic molecules is found to be the dominant aerosol formation process in the pre-industrial boundary layer over land. The inclusion of this process in a global aerosol model raises baseline pre-industrial aerosol concentrations, and could lead to a reduction of 27% in estimates of anthropogenic aerosol radiative forcing.

2 Article

Measurements in the CERN CLOUD chamber under atmospheric conditions show that new particles can form purely from the oxidation products of α-pinene, a compound emitted by the biosphere [1]. Nucleation of new aerosol

particles via gas-to-particle conversion has been studied for fifty years [2] and is responsible for around half of global cloud condensation nuclei (CCN) [3], which affect Earth's radiation balance via aerosol-cloud interactions. The involvement of oxidised organic molecules in the process, alongside sulphuric acid, was proposed in early studies, and has been well-established for some time [4, 5]. The new mechanism for organic particle formation without sulphuric acid presented in Ref. [1] could be important for Earth's climate because it provides a way to form particles in the pristine pre-industrial atmosphere, when the concentrations of sulphuric acid and ammonia were much lower. The pre-industrial environment forms the baseline for calculations in global models of the radiative forcing caused by anthropogenic emissions [6], and uncertainties in this baseline are the largest component of the overall uncertainty on aerosol radiative forcing [7]. This is because an incremental increase in particle concentrations when they are low has a much stronger radiative effect than when they are high. Previous model uncertainty analyses suggested that the sensitivity of radiative forcing to particle formation rates is low compared to many other factors [7]. However, these studies varied the nucleation rate assuming that sulphuric acid is required for particle production. Here we show that the inclusion of a new nucleation mechanism that does not require sulphuric acid could have a more significant effect on radiative forcing than previously thought [8, 7].

Our modelling study is inspired by and based on measurements in which α-pinene, a volatile organic compound (VOC) emitted into the atmosphere by vegetation, was oxidised by ozone and hydroxyl radicals in the CLOUD chamber under ultra-clean conditions without sulphuric acid [1]. The mass spectra of the highly oxidised multifunctional organic molecules (HOMs) produced from the VOCs closely resemble those observed in the atmosphere [9]. Therefore, while the concentrations of some reactive gases in the chamber do not perfectly match those in the troposphere, we have confidence in our assumption that the chamber results can be generalised to the atmosphere. Particle counters show that typical atmospheric concentrations of the HOMs produce particles at significant rates, even when sulphuric acid is absent from nucleating clusters. We describe this process as pure biogenic nucleation.

In this paper, we examine the implications of pure biogenic nucleation for atmospheric aerosol and Earth's radiation balance using the GLOMAP global model of aerosol microphysics [10]. A parametrisation of the pure biogenic nucleation rate that depends on the HOM concentration and the concentration of ions is provided in supplementary materials of Ref. [1]. We

assume for this study that this can be added linearly to parametrisations of the nucleation rate involving sulphuric acid only [11] and sulphuric acid with organics similar to HOMs [5]. Ref. [1] also provided the yields of HOMs from the oxidation of α-pinene by ozone (2.9%) and by the hydroxyl radical (1.2%). The yield of HOM from endocyclic monoterpenes such as α-pinene is higher than that from exocylic monoterpenes, so we separate these classes in our model and use the yields from β-pinene in Ref. [12] to produce HOM from exocylic monoterpenes. The rate of formation of 1.7 nm diameter aerosols by gas-to-particle conversion is therefore described by the sum of the following parametrisations:

1. Binary homogeneous nucleation of sulphuric acid and water [11].

2. Nucleation of organics with sulphuric acid [5], also used in Ref. [13]:

$$J_{\mathrm{sa-org}} = k_{\mathrm{sa-org}}[\mathrm{H_2SO_4}]^2[\mathrm{BioOxOrg}] \tag{1}$$

   where BioOxOrg refers to the oxidation products of monoterpenes with OH and $k_{\mathrm{sa-org}} = 3.27 \times 10^{-21}\,\mathrm{cm^6 s^{-1}}$ (see Methods).

3. Pure biogenic nucleation, a sum of neutral (J_n) and ion-induced (J_{iin}) components [1]:

$$J_{\mathrm{org}} = J_n + J_{iin} \tag{2}$$
$$J_{\mathrm{n}} = a_1[\mathrm{HOM}]^{a_2+a_5/[\mathrm{HOM}]} \tag{3}$$
$$J_{\mathrm{iin}} = 2[n_\pm]a_3[\mathrm{HOM}]^{a_4+a_5/[\mathrm{HOM}]} \tag{4}$$

   where HOMs are produced as described above but given here for convenience in units of 10^7 molecules per cubic centimetre, $n_\pm$ is the ion concentration and a are free parameters. Ions in the model are produced from radon and galactic cosmic rays (see SI Appendix).

Ammonia and amines can also contribute to nucleation by stabilising sulphuric acid clusters, but the binary homogeneous mechanism has been shown to be a reasonable representation of free tropospheric nucleation [14], and nucleation at low altitudes involving amines or ammonia is important only in polluted regions where the changes in radiative forcing calculated here are very insensitive to nucleation rates.

In our model, aerosols formed in this way, and those emitted directly from Earth's surface, grow by condensation and coagulation, are transported in

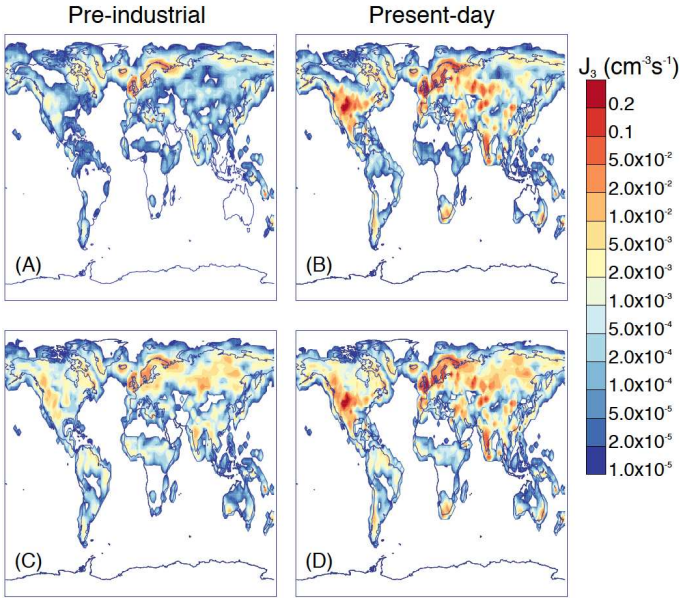

Figure 1: Nucleation rates at 3 nm diameter (J_3, $\mathrm{cm^{-3}s^{-1}}$) within approximately 500 m of the surface averaged over June without pure biogenic nucleation in (A) pre-industrial and (B) present-day conditions, and with pure biogenic nucleation in (C) pre-industrial and (D) present-day conditions.

the atmosphere, and are ultimately removed by dry or wet deposition. We consider the radiative forcing between 1750 and 2008 via the effect of these aerosols on cloud albedo, which is evaluated at the top of the atmosphere (0.03 Pa atmospheric pressure). To determine the effects of pure biogenic nucleation, particle formation rates, aerosol concentrations and radiative forcing from model runs with and without mechanism 3 are compared.

3 Biogenic nucleation rates and observational evidence

Fig. 1 shows the effect of pure biogenic nucleation on the pre-industrial and present-day atmospheres. When sulphuric acid is required for nucleation to proceed, substantially less nucleation is expected for pre-industrial times (Fig. 1A) compared with the present (Fig. 1B). However, when pure biogenic nucleation is included, the nucleation rates in pre-industrial (C) and present-day times (D) become more similar. While pure biogenic nucleation

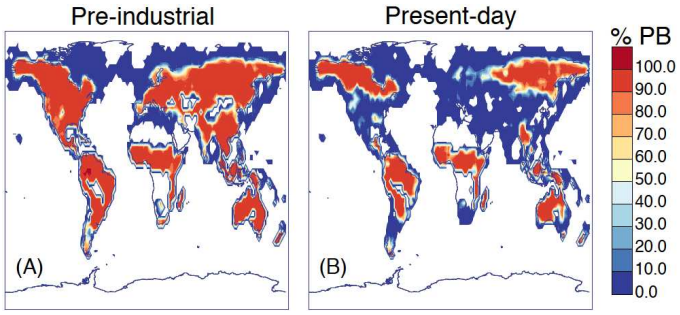

Figure 2: Percentage of particles produced via pure biogenic (PB) nucleation within approximately 500 m of the surface, averaged over June in (A) pre-industrial and (B) present-day conditions. We note that our model predicts large changes to particle formation at the surface and very little change above the boundary layer.

is much less important today (compare the change from B to D with that from A to C), it is still expected to be significant in some continental regions remote from pollution, for example boreal regions, Australia and, according to our simulations (discussed later), the Amazon. Within around 500 m of the surface pure biogenic nucleation increases total production of particles of at least 3 nm in diameter via nucleation by 2.1% globally in the present-day atmosphere, but by 90% in pre-industrial conditions.

Fig. 2 shows that pure biogenic nucleation is predicted to be the dominant mechanism for particle formation over large parts of the land surface above 50°N in summer even in the present-day. However, both pure biogenic and sulphuric acid particle formation rates are often insufficient to produce detectable nucleation events (see SI Appendix, Fig. S4). Pure biogenic nucleation has more effect in June than in January because terpene emissions are higher in June. The diurnal cycles of nucleation rates at Hyytiälä and Pallas in Finland, shown in SI Appendix Fig. S4, indicate that nucleation rates in these areas are occasionally higher than around $0.1\,\mathrm{cm}^{-3}\mathrm{s}^{-1}$. Experience from these boreal forest sites [15] suggests that nucleation rates above this value will result in detectable nucleation events. This is confirmed by the modelled size distributions shown in SI Appendix Fig. S6. As is observed, simulated nucleation rates are substantially higher during the day than at night.

To our knowledge, Hyytiälä and Jungfraujoch are the only locations with published measurements from the APi-TOF and CI-APi-TOF mass spec-

trometers needed to unambiguously detect pure biogenic nucleation [15].There is strong evidence in Ref. [16] that pure organic nucleation proceeds alongside sulphuric acid-driven nucleation at Jungfraujoch. For example, their Fig. 2 shows that, on the Nucleation Day 3, most organic clusters of masses of up to 400 amu contain no sulphuric acid, there is no inorganic nucleation, and the nucleation rate exceeds $10\,\mathrm{cm}^{-3}\mathrm{s}^{-1}$ when sulphuric acid concentrations are less than $5 \times 10^5 \mathrm{cm}^{-3}$.

There are no measurements of pure biogenic nucleation so far from Hyytiälä since almost all the nucleation rates measured in Ref. [15] are at $[\mathrm{H_2SO_4}] > 1 \times 10^6\,\mathrm{cm}^{-3}$. Observations at Hyytiälä were, however, used alongside those from Melpitz and Hohenpeissenberg to derive parameterizations of particle formation rates in Ref. [17]. The authors found that nucleation could be described well by

$$J_2 = k_1[\mathrm{H_2SO_4}]^2 + k_2[\mathrm{H_2SO_4}][\mathrm{org}] + k_3[\mathrm{org}]^2, \tag{5}$$

for constant k_{1-3}, suggesting that pure biogenic nucleation is a statistically detectable component of nucleation in these environments.

In addition to the Jungfraujoch observations, there is extensive circumstantial evidence for pure biogenic nucleation. The Amazon, where the lowest SO_2 concentrations over land are found, is an obvious place to look. While some nucleation mode particles are seen in pristine regions of the Amazon [18] (on 19% of days sampled in the study referenced), no clear nucleation events or conclusive evidence for biogenic nucleation have yet been published, and growth of nucleation mode particles to CCN size is rarely observed there. Our model does not produce Hyytiälä-like nucleation events (see SI Appendix Figs. S5-S7) but it does predict non-zero particle formation rates. It slightly overestimates CCN concentrations compared to Ref. [19] in the Amazon even without pure biogenic nucleation, and pure biogenic nucleation further increases the discrepancy, by around a factor two. This may point to a chemical suppression of HOM yields by isoprene [20] or NO_x [21], but could also be due to other sources of model error, for example, underestimation of particle size and therefore condensation sink. Overprediction of particle concentrations over the Amazon seems to be a common feature among models [22]. Comparing models with observations in this region is challenging due to large uncertainties in emissions of biogenic VOCs and a complex wet scavenging environment.

Pure biogenic nucleation is also predicted to be the dominant source of secondary particles in the cleanest high latitude boreal regions. Low SO_2

concentrations, often below 100 ppt, and nocturnal nucleation were reported in a study at Värriö, Finland (67°N) at similar temperatures to the CLOUD chamber [23]. Similar observations of nocturnal nucleation were made at Abisko, Sweden [24] and Tumbarumba, Australia [25], although SO_2 concentrations were not reported. At Pallas, Finland, H_2SO_4 concentrations are reported below $3 \times 10^5 \mathrm{cm}^{-3}$ in a large number of new particle formation events [26]. The air masses in Pallas are usually of marine origin, which leads to low condensation sinks favourable to nucleation, but may also allow halogens of marine origin to locally influence nucleation. Three instances of new particle formation with $[H_2SO_4] < 3 \times 10^5\,\mathrm{cm}^{-3}$ shown in Ref. [26] Fig. 6 are unambiguously continental. This should also allow the contribution of halogens to be excluded, making it highly likely the nucleation was pure biogenic.

With only sparse or indirect observational evidence for pure biogenic nucleation, an alternative strategy is to compare modelled particle concentrations against observations. However, this is also inconclusive because there are many compensating causes of model error [7], making attribution of biases ambiguous. Substantial changes in total particle number concentration are caused by pure biogenic nucleation (SI Appendix Fig. S2). However, when we compare the monthly mean model predictions to particle number concentrations at 37 surface sites [27, 28], and the daily mean concentrations to those measured during the ARCTAS aircraft campaign [29] in 2008 (SI Appendix Figs. S8 and S9), we find that the effect of pure biogenic mechanism, increasing summertime particle concentrations by up to a factor 2, is also comparable to, or smaller than, existing discrepancies between observations and the model.

4 Impact on CCN and radiative forcing

Fig. 3 shows the effect of pure biogenic nucleation on present-day and pre-industrial CCN concentrations, calculated at 0.2% supersaturation. When pure biogenic nucleation is included, global annual average concentrations of these particles at cloud base level (approximately 600 m altitude) increase by 4% in the present-day and 12% in the pre-industrial atmospheres. Although nucleation rates are affected mostly close to sources of biogenic gases, CCN are affected over much wider areas due to the slower removal rate of larger aerosol particles. This spread is important because it carries the particles to

cloudy marine regions where most of the anthropogenic aerosol-cloud radiative forcing occurs [30]. The change in CCN production across the pristine pre-industrial atmosphere is particularly important for global climate because cloud droplet concentrations and albedo are both more sensitive to CCN changes in pristine environments.

The change in aerosol radiative forcing from 1750 to 2008 attributable to pure biogenic nucleation was calculated by comparing simulations with and without pure biogenic nucleation. We only consider changes in the cloud albedo effect. The aerosol direct forcing is unlikely to be substantially influenced by the new nucleation mechanism as it is not strongly affected by the aerosol size distribution [31]. The change in radiative forcing when pure biogenic nucleation is included is presented in Fig. 4. We estimate that the global annual mean cloud albedo forcing since 1750, after including pure biogenic nucleation, is $-0.60\,\mathrm{Wm^{-2}}$. The *change* in calculated aerosol radiative forcing due to pure biogenic nucleation is $+0.22\,\mathrm{Wm^{-2}}$, corresponding to a 27% reduction in the negative forcing. This change is a result of the non-linear dependence of the forcing on the baseline CCN concentration [7]. We note that our simulations may underestimate the net effect since they do not account for possible increases in cloud fraction and thickness, which, in pristine regions (CCN below $100\,\mathrm{cm^{-3}}$), may be highly sensitive to small changes of CCN [32]. We also do not account for the possibility of pure biogenic nucleation involving sesquiterpenes. However, we also emphasise that including pure biogenic nucleation in our model leads to an over-prediction of CCN in the Amazon region, which may indicate that it is chemically suppressed. Inhibition of nucleation, if it happens, may be local to the tropical rainforest environment or more widespread. If we artifically set pure biogenic nucleation rates to zero within 10° latitude of the Equator, the effect on aerosol forcing when pure biogenic nucleation is included changes only slightly, to $+0.20\,\mathrm{Wm^{-2}}$.

The largest changes in radiative forcing occur over the NH, especially over oceans with high annual cloud cover (Fig. 4b) where CCN concentrations are most strongly perturbed by anthropogenic emissions. The NH is also where pure biogenic nucleation causes the largest reduction in contrast between pre-industrial and present day CCN concentrations driven by the large continental source of biogenic gases. However, the relative change in forcing in the SH is greater than the NH: pure biogenic nucleation reduces the annual southern hemispheric mean from $-0.25\,\mathrm{Wm^{-2}}$ to $-0.14\,\mathrm{Wm^{-2}}$ (compared to a change in the NH of $-1.39\,\mathrm{Wm^{-2}}$ to $-1.06\,\mathrm{Wm^{-2}}$). In some tropical and

southern regions, there are higher CCN in pre-industrial times than today, and a positive radiative forcing. In these regions and nearby, pre-industrial OH· and HOMs were higher than today and particle condensation sinks were lower, while SO_2 levels (largely marine) were comparable.

We consider the principal uncertainties in our analysis to be associated with a) VOC, SO_2 and primary particle emissions as in Ref. [7], b) how representative α-pinene and the pinanediol used in Ref. [5] are of VOCs in the atmosphere, c) yields of HOM from α-pinene oxidation in the presence of other vapours such as NO_x, and d) temperature dependence of the nucleation rates.

To investigate the effect of a plausible temperature dependence we reran the model multiplying all boundary-layer nucleation rates by $\exp(-(T-278)/10)$. The charged nucleation rate remained limited by the ion production rate and the overall rate by the kinetic limit. We find annually averaged changes to cloud albedo radiative forcing over the industrial period from pure biogenic nucleation are reduced to $+0.14\,\mathrm{Wm^{-2}}$ from $+0.22\,\mathrm{Wm^{-2}}$.

The yields of HOM have an experimental uncertainty around a factor two (and were reported to be about a factor two higher in an earlier chamber study [33]). These uncertainties are comparable to uncertainties in the VOC emissions themselves [34]. The yields could be affected by nitrogen oxides [21], and were found to differ substantially between monoterpenes [12]. To test the sensitivity to the uncertainty in yields, which is a proxy for the overall intrinsic uncertainty on the experimental measurements, we repeated our analysis with the yield of the HOMs that participate in pure biogenic nucleation perturbed by a factor 3. This gives an uncertainty range for the increase in CCN due to the pure biogenic mechanism of $4-19\%$ in the pre-industrial and $1-6\%$ in the present-day, as shown in Table S2. The lower limit still leads to a significant change to cloud albedo forcing of $0.10\,\mathrm{Wm^{-2}}$ when the corresponding parametrisation is added to the model.

We have also investigated the sensitivity of our radiative forcing estimate to other sources of uncertainty. We perturb the pre-industrial volcanic SO_2 emissions and find this does not strongly affect our reported CCN changes. When we perturb the biomass burning and sea spray emissions (see SI Appendix for details) we find larger changes both to CCN and forcing, especially when emissions are reduced. The model becomes slightly more sensitive to pure biogenic nucleation when different baseline nucleation mechanism from Ref. [17] instead of Ref. [5] is used. The percentage changes to CCN from including pure biogenic nucleation under these scenarios are given in SI Ap-

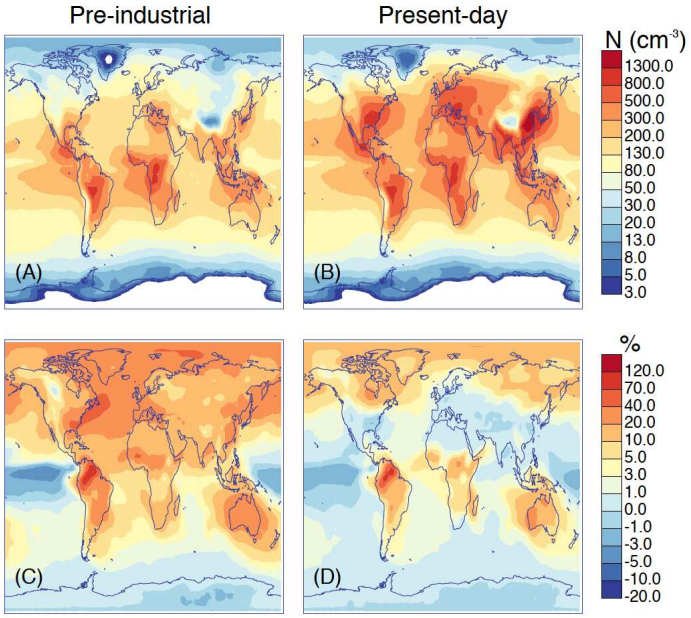

Figure 3: Concentrations of cloud condensation nuclei calculated at 0.2% supersaturation, in cm^{-3}, annually averaged at cloud base level in (A) pre-industrial and (B) present-day conditions, and (C, D) percentage changes to these concentrations when pure biogenic nucleation is introduced. In this Figure we assume HOM formation and pure biogenic nucleation proceed at the rates measured at the CLOUD chamber.

pendix Table S2, and the changes to forcing in Table S3.

5 Discussion and conclusions

Our global aerosol simulations indicate that pure biogenic nucleation [1] dominates particle formation in the pre-industrial boundary layer, producing 59% of new particles below approximately 500 m altitude and 36% below around 1.5 km. For the organic system, laboratory measurements are currently the only route to a comprehensive understanding of the processes leading to particle formation. This is particularly the case for a mechanism that is difficult to decouple from sulphuric acid-driven nucleation pathways in the polluted present-day atmosphere. This mechanistic understanding is required to perform accurate extrapolations from present-day conditions back to the pre-industrial. Improving such extrapolations is of critical importance as un-

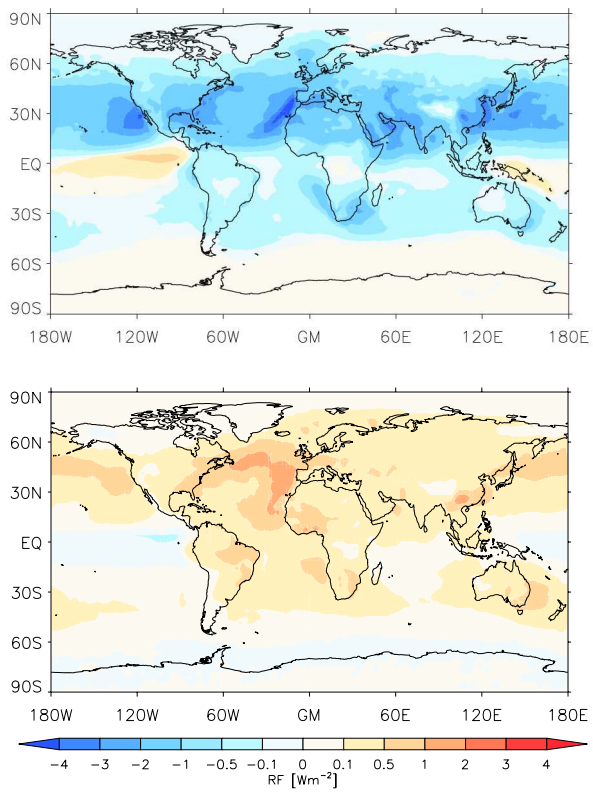

Figure 4: Distribution of (A) cloud albedo radiative forcing and (B) change to this distribution when pure biogenic nucleation is included in the model.

certainties in pre-industrial aerosol are a large component of the uncertainty in IPCC estimates of radiative forcing. While nucleation in tropical environments is relatively unimportant for global mean cloud albedo radiative forcing in our model, discrepancies between modelled and observed nucleation in these regions suggest further investigation of Amazon aerosol chemistry could significantly improve our understanding of pristine aerosol processes.

Based on the nucleation rates reported by CLOUD [1], we show here that pure biogenic nucleation may reduce the magnitude of pre-industrial to present-day aerosol cloud albedo forcing by as much as $0.22\,\mathrm{Wm}^{-2}$, or 27%. This change in forcing is greater than the combined one standard deviation uncertainty of twenty-eight parameters related to emissions and aerosol processes in this model [7], which is 19%. Other forcing mechanisms or uncertainties in the results quoted here could still lead to stronger effects. Although the calculated change in forcing is comparable to the model parametric uncertainty, it shifts the entire probability distribution of forcing, and therefore represents a significant downward revision in the likelihood of high negative aerosol-cloud forcings in this model. Similar revisions are likely to occur in other models [35] due to the same chain of processes: 1) proportionally greater increases in aerosol concentrations in the cleaner pre-industrial atmosphere than in the present day; 2) high sensitivity of cloud albedo and adjustments on the pre-industrial aerosol concentrations; 3) reduction in the magnitude of anthropogenic aerosol radiative forcing by raising the pre-industrial baseline aerosol concentration. To remain consistent with the observed temperature rise over the industrial period, reduced aerosol forcing implies reduced climate sensitivity [30, 36].

6 Materials

The modal version of the global aerosol model GLOMAP [10] is used to determine the impact of the biogenic nucleation mechanism reported in Ref. [1]. The model resolution is $2.8^\circ \times 2.8^\circ$ horizontally, and there are 31 vertical levels from ground level to 10 hPa. GLOMAP is embedded within a chemical transport model, TOMCAT [37], and simulates the formation or emission, growth, coagulation, advection, cloud processing and deposition of aerosol in seven log-normal size modes. Four modes (nucleation, Aitken, accumulation and coarse) are hydrophyllic, and there are also hydrophobic Aitken, accumulation and coarse modes. The composition of each mode is determined

by the relative fractions of the sulfate, sea-salt, black carbon, and organic carbon compounds. Dust is not included, as it was not found to contribute significantly to CCN [38]. Meteorology is forced by fields from the European Centre for Medium-range Weather Forecasting. Total monoterpene emissions are taken from Ref. [34] and the ratio of endocyclic to exocyclic monoterpenes was calculated from a run of the MEGAN model with the settings prescribed to follow Ref. [39]. Ref. [40] suggests that terpene emissions are (within uncertainties) unchanged through the industrial period.

While sulphuric acid, ammonia, amines, halogens and HOMs can all participate directly in nucleation, here we consider only sulphuric acid and HOMs. The HOMs are formed via the oxidation of monoterpenes (MT) by ozone (O_3) and hydroxyl radicals (OH·). The concentrations of these oxidants are read in every six hours from a dedicated TOMCAT simulation. Instead of modelling the full reaction mechanism, we represent the HOM concentrations by

$$[\mathrm{HOM}] = (Y_{\mathrm{AP.O_3}} k_{\mathrm{AP.O_3}} [\mathrm{AP}][\mathrm{O_3}] + Y_{\mathrm{BP.O_3}} k_{\mathrm{BP.O_3}} [\mathrm{BP}][\mathrm{O_3}] + Y_{\mathrm{AP.OH\cdot}} k_{\mathrm{AP.OH\cdot}} [\mathrm{AP}][\mathrm{OH\cdot}] + Y_{\mathrm{BP.OH\cdot}} k_{\mathrm{BP.OH\cdot}} [\mathrm{BP}][\mathrm{OH\cdot}]) / CS$$

where $Y_{\mathrm{AP.O_3}} = 2.9\%$ and $Y_{\mathrm{AP.OH\cdot}} = 1.2\%$ are the yields of HOM from α-pinene (AP) oxidation with ozone and hydroxyl radicals in the CLOUD chamber, described below, $Y_{\mathrm{BP.O_3}} = 0.12\%$ and $Y_{\mathrm{BP.OH\cdot}} = 0.58\%$ are taken from Ref. [12] and CS is the condensation sink (s^{-1}), determined assuming the diffusion characteristics of a typical α-pinene oxidation product (see Appendix A1 of Ref. [10]). The temperature-dependent reaction rate constants k for oxidation of α and β-pinene by ozone and hydroxyl radicals are taken from IUPAC [41].

The ozonolysis yield is determined with chemical ionisation time-of-flight mass spectrometers in the presence of a hydroxyl scavenger (0.1% H_2), replicating the effect of atmospheric OH· sinks such as methane and carbon monoxide. The HOM yield from reaction with hydroxyl radicals is determined from measurements in the absence of ozone, and where photolysed HONO provides the OH· source.

BioOxOrg in nucleation mechanism 2 and HOM in mechanism 3 play equivalent roles but the former refers to the parametrised oxidation products derived from pinanediol, a first-generation oxidation product of α-pinene. Its concentration, as described in Ref. [5], is

$$[\mathrm{BioOxOrg}] = k_{\mathrm{MT.OH\cdot}}[\mathrm{MT}][\mathrm{OH\cdot}]/CS$$

where CS is the condensation sink. The BioOxOrg concentration was not measured directly in a mass spectrometer, but calculated from the pinanediol concentration assuming a yield of 100%. The nucleation rate in mechanism 2 is measured as a function of this BioOxOrg, so the yield is incorporated into the rate constant for nucleation. In Ref. [5] monoterpenes are assumed to be equivalent to α-pinene, and so we assume only endocyclic monoterpenes participate in this nucleation mechanism.

Particles are formed according to the mechanisms described in the main text at a critical diameter usually around 1.7 nm. Ion concentrations are determined by balancing production from radon and galactic cosmic rays with losses to pre-existing particles and to ion-ion recombination (see SI Appendix). The formation rates are then adjusted to account for losses during the initial growth with the Kerminen-Kulmala equation [42] using growth rates taken from the parametrisation of Ref. [43].

Particles subsequently grow by kinetic condensation of organic molecules produced from oxidation of terpenes or isoprene by nitrate or hydroxyl radicals, or ozone, with a 13% assumed yield for terpenes [10] and a 3% yield for isoprene [44]. They also coagulate, and hence the overall particle number is determined by solving the coagulation-nucleation equation [10]. Finally, particles may be lost by dry or wet deposition.

Present-day simulations are run for 2008 and pre-industrial simulations are run with 2008 meteorology and 1750 emissions. For the 1750 simulation, anthropogenic sources of SO_2 and H_2SO_4 were removed from the model, OH, NO_3 and ozone concentrations were adjusted to pre-industrial levels determined from a dedicated TOMCAT simulation, and black and organic carbon primary emissions were adjusted to a representation of pre-industrial levels.

Cloud condensation nuclei (CCN) and cloud droplet number concentrations (CDNC) are calculated for each simulation from the particle size distributions using the parametrisation of Ref. [45], assuming for the CDNCs constant updraft velocities of $0.15\,\mathrm{ms}^{-1}$ over sea and $0.30\,\mathrm{ms}^{-1}$ over land. The hygroscopicity parameters assigned to each chemical component follow Ref. [44]: sulphate (0.61, assuming ammonium sulphate), sea salt (1.28), black carbon (0.0), and organics (0.1). The change in cloud droplet effective

radii corresponding to the CDNC change is calculated in accordance with Ref. [31], while the cloud albedo is estimated using the radiative transfer model of Ref. [46].

7 Acknowledgments

We would like to thank CERN for supporting CLOUD with important technical and financial resources, and for providing a particle beam from the CERN Proton Synchrotron. We also thank P. Carrie, L.-P. De Menezes, J. Dumollard, F. Josa, I. Krasin, R. Kristic, A. Laassiri, O.S. Maksumov, B. Marichy, H. Martinati, S.V. Mizin, R. Sitals, A. Wasem and M. Wilhelmsson for their important contributions to the experiment. We thank D. Veber from Environment and Climate Change Canada for maintenance and calibrations of instruments at East Trout Lake and NOAA/ESRL for collaboration with data collection and QA/QC software. We thank A. D. Clarke and C. L. S. Reddington for making available processed data from the ARCTAS campaign. The global modelling simulations were performed on ARC1 and ARC2, part of the High Performance Computing facilities at the University of Leeds, U.K. This research has received funding from the EC Seventh Framework and Horizon 2020 Programmes (Marie Curie Initial Training Network MC-ITN CLOUD-TRAIN no. 316662, Marie Sklodowska-Curie grant agreements Nos. 656994 and 600377, ERC-Consolidator grant NANODYNAMITE no. 616075 and ERC-Advanced grant ATMNUCLE no. 227463), the German Federal Ministry of Education and Research (project no. 01LK1222A), the Swiss National Science Foundation (project nos. 200020_135307, 200021_140663, 206021_144947/1 and 20FI20_149002/1), the Academy of Finland (Center of Excellence project no. 1118615), the Academy of Finland (135054, 133872, 251427, 139656, 139995, 137749, 141217, 141451), the Finnish Funding Agency for Technology and Innovation, the Väisälä Foundation, the Nessling Foundation, the Austrian Science Fund (FWF; project no. L593), the Portuguese Foundation for Science and Technology (project no. CERN/FP/116387/2010), the Swedish Research Council, Vetenskapsrådet (grant 2011-5120), the Presidium of the Russian Academy of Sciences and Russian Foundation for Basic Research (grant 12-02-91522-CERN), the U.K. Natural Environment Research Council (grant NE/K015966/1), the Royal Society (Wolfson Merit Award), the U.S. National Science Foundation (grants AGS1136479,AGS1439551, AGS1447056 and CHE1012293), Caltech ESE Grant (Davidow Foundation),

Dreyfus Award EP-11-117, the French National Research Agency (ANR), the Nord-Pas de Calais, the European Funds for Regional Economic Development (FEDER, Labex-Cappa, ANR-11-LABX-0005-01), and the French Civil Aviation Office (MERMOSE).

References

[1] Kirkby, J et al. (2016) Ion-induced nucleation of pure biogenic particles. *Nature* **533**, 521–526.

[2] Went, F. W. (1960) Blue hazes in the atmosphere. **187**, 641–643.

[3] Merikanto, J, Spracklen, D. V, Mann, G. W, Pickering, S. J, & Carslaw, K. S. (2009) Impact of nucleation on global CCN. *Atmospheric Chemistry and Physics* **9**, 8601–8616.

[4] Kavouras, I. G, Mihalopoulos, N, & Stephanou, E. G. (1998) Formation of atmospheric particles from organic acids produced by forests. *Nature* **395**, 683–686.

[5] Riccobono, F et al. (2014) Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles. *Science* **344**, 717–721.

[6] Hansen, J et al. (1981) Climate impact of increasing atmospheric carbon dioxide. *Science* **213**, 957–966.

[7] Carslaw, K. S et al. (2013) Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature* **503**, 67–71.

[8] Pierce, J. R & Adams, P. J. (2009) Uncertainty in global CCN concentrations from uncertain aerosol nucleation and primary emission rates. *Atmospheric Chemistry and Physics* **9**, 1339–1356.

[9] Schobesberger, S et al. (2013) Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *Proceedings of the National Academy of Sciences* **110**, 17223–17228.

[10] Mann, G. W et al. (2010) Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model. *Geoscientific Model Development* **3**, 519–551.

[11] Vehkamäki, H et al. (2002) An improved parameterization for sulfuric acid-water nucleation rates for tropospheric and stratospheric conditions. *Journal of Geophysical Research* **107**, 4622–4622.

[12] Jokinen, T et al. (2015) Production of extremely low volatile organic compounds from biogenic emissions: Measured yields and atmospheric implications. *Proceedings of the National Academy of Sciences* **112**, 7123–7128.

[13] Yu, F et al. (2015) Spring and summer contrast in new particle formation over nine forest areas in north america. *Atmospheric Chemistry and Physics* **15**, 13993–14003.

[14] Spracklen, D. V, Pringle, K. J, Carslaw, K. S, Chipperfield, M. P, & Mann, G. W. (2005) A global off-line model of size-resolved aerosol microphysics: Ii. identification of key uncertainties. *Atmospheric Chemistry and Physics* **5**, 3233–3250.

[15] Kulmala, M et al. (2013) Direct observations of atmospheric aerosol nucleation. *Science* **339**, 943–946.

[16] Bianchi, F et al. (2016) New particle formation in the free troposphere: A question of chemistry and timing. *Science*.

[17] 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**, 11223–11242.

[18] Rizzo, L, Artaxo, P, Karl, T, Guenther, A, & Greenberg, J. (2010) Aerosol properties, in-canopy gradients, turbulent fluxes and VOC concentrations at a pristine forest site in Amazonia. *Atmospheric Environment* **44**, 503 – 511.

[19] Martin, S et al. (2010) Sources and properties of amazonian aerosol particles. *Reviews of Geophysics* **48**.

[20] Kiendler-Scharr, A et al. (2009) New particle formation in forests inhibited by isoprene emissions. *Nature* **461**, 381–384.

[21] Wildt, J et al. (2014) Suppression of new particle formation from monoterpene oxidation by NO_x. *Atmospheric Chemistry and Physics* **14**, 2789–2804.

[22] 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**, 4679–4713.

[23] Vehkamäki, H et al. (2004) Atmospheric particle formation events at Värriö measurement station in Finnish Lapland 1998-2002. *Atmospheric Chemistry and Physics* **4**, 2015–2023.

[24] Svenningsson, B et al. (2008) Aerosol particle formation events and analysis of high growth rates observed above a subarctic wetland–forest mosaic. *Tellus B* **60**, 353–364.

[25] Suni, T et al. (2008) Formation and characteristics of ions and charged aerosol particles in a native Australian Eucalypt forest. *Atmospheric Chemistry and Physics* **8**, 129–139.

[26] Asmi, E et al. (2011) Secondary new particle formation in Northern Finland Pallas site between the years 2000 and 2010. *Atmospheric Chemistry and Physics* **11**, 12959–12972.

[27] 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**, 4775–4793.

[28] Torseth, K et al. (2012) Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009. *Atmospheric Chemistry and Physics* **12**, 5447–5481.

[29] Jacob, D. J et al. (2010) The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results. *Atmospheric Chemistry and Physics* **10**, 5191–5212.

[30] Charlson, R. J et al. (1992) Climate forcing by anthropogenic aerosols. *Science* **255**, 423–430.

[31] Rap, A et al. (2013) Natural aerosol direct and indirect radiative effects. *Geophysical Research Letters* **40**, 3297–3301.

[32] Koren, I, Dagan, G, & Altaratz, O. (2014) From aerosol-limited to invigoration of warm convective clouds. *Science* **344**, 1143–1146.

[33] Ehn, M et al. (2014) A large source of low-volatility secondary organic aerosol. *Nature* **506**, 476–479.

[34] Sindelarova, K et al. (2014) Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics* **14**, 9317–9341.

[35] Myhre, G et al. (2013) Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. *Atmospheric Chemistry and Physics* **13**, 1853–1877.

[36] Knutti, R, Stocker, T. F, Joos, F, & Plattner, G.-K. (2002) Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature* **416**, 719–723.

[37] Chipperfield, M. P. (2006) New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments. *Quarterly Journal of the Royal Meteorological Society* **132**, 1179–1203.

[38] Manktelow, P. T, Carslaw, K. S, Mann, G. W, & Spracklen, D. V. (2010) The impact of dust on sulfate aerosol, CN and CCN during an East Asian dust storm. *Atmospheric Chemistry and Physics* **10**, 365–382.

[39] Guenther, A et al. (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**, 1471–1492.

[40] Acosta Navarro, J. C et al. (2014) Global emissions of terpenoid VOCs from terrestrial vegetation in the last millennium. *Journal of Geophysical Research: Atmospheres* **119**, 6867–6885.

[41] Atkinson, R et al. (2006) Evaluated kinetic and photochemical data for atmospheric chemistry: Volume ii: gas phase reactions of organic species. *Atmospheric Chemistry and Physics* **6**, 3625–4055.

[42] Kerminen, V.-M & Kulmala, M. (2002) Analytical formulae connecting the real and the apparent nucleation rate and the nuclei number concentration for atmospheric nucleation events. *Journal of Aerosol Science* **33**, 609 – 622.

[43] Tröstl, J et al. (2016) The role of low-volatility organic compounds in initial particle growth in the atmosphere. *Nature* **533**, 527–531.

[44] Scott, C. E et al. (2014) The direct and indirect radiative effects of biogenic secondary organic aerosol. *Atmospheric Chemistry and Physics* **14**, 447–470.

[45] Morales, R, Nenes, A, Jonsson, H, Flagan, R. C, & Seinfeld, J. H. (2011) Evaluation of an entraining droplet activation parameterization using in situ cloud data. *Journal of Geophysical Research: Atmospheres* **116**, D15205.

[46] Edwards, J. M & Slingo, A. (1996) Studies with a flexible new radiation code. i: Choosing a configuration for a large-scale model. *Quarterly Journal of the Royal Meteorological Society* **122**, 689–719.

8 Supplementary: Ion concentrations in the GLOMAP aerosol model

We consider two sources of ions in the atmosphere: radon and galactic cosmic rays. Radon is dominant at the land surface, where most biogenic nucleation is likely to happen. Ion production rates from radon are read in from look-up tables [1]. Above the surface and over the ocean, cosmic ray ionisation is more important. The ionization rates from cosmic rays are calculated from lookup tables [2] which are provided for several solar cycles, so the effect of the Sun's magnetic field can be incorporated via the heliospheric modulation potential. The technique of Fraser-Smith [3] is used to calculate the geomagnetic cut-off rigidity from the International Geomagnetic Reference Field coefficients. These are available with five-yearly time resolution so are interpolated within the five-year periods, then the atmospheric depth (which determines the interaction probability of a cosmic ray) and the heliospheric modulation potential are spatially interpolated across the model grid-boxes.

The small-ion concentration of either sign, $[n_\pm] = [n_+] = [n_-]$, is calculated from the steady state solution of the ion balance equation [4]

$$\mathrm{d}[n_\pm]/\mathrm{d}t = q - \alpha[n_\pm]^2 - k_i[n_\pm] \tag{6}$$

where q is the ion pair production rate from GCRs and α is the ion-ion recombination coefficient ($\mathrm{cm^3 s^{-1}}$). The factor 2 in Eq. 4 accounts for nucleation from both positive and negative ions. The ion loss rate, k_i, is due to the condensation sink, CS, and ion-induced nucleation, so that $k_i = CS + J_{iin}/2\,[n_\pm]$ where $J_{iin}/2\,[n_\pm]$ is given by Eq. 4 and the steady state concentration of small ions is $[n_\pm] = [(k_i^2 + 4\alpha q)^{0.5} - k_i]/2\alpha$. From Eq. 6, J_{iin} saturates at $2q$ at high nucleation rates (see Ref. [5] Fig. 2).

9 Supplementary: Simulating the pre-industrial atmosphere

The concentrations of key precursor gases for particle formation are compared between present-day and pre-industrial in Fig. S1. The percentage changes between pre-industrial and present-day are compared in Table S1 for summer and winter in the two hemispheres. The sulphuric acid concentration is substantially higher in the present day atmosphere due to much higher

emissions, while the organic concentrations are higher in the pre-industrial atmosphere due to lower sinks.

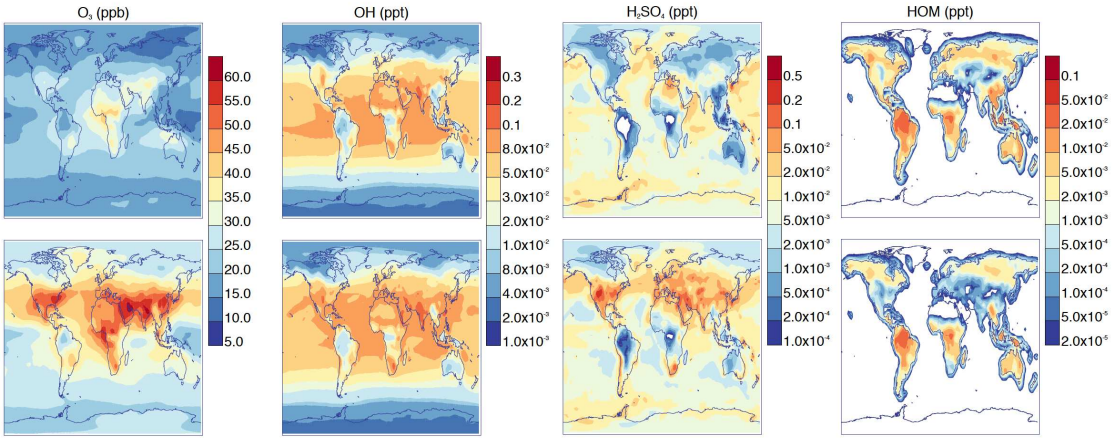

Figure S1: Concentrations of key gases: ozone, hydroxyl radicals, sulphuric acid and HOMs in pre-industrial (top row) and present-day atmospheres (bottom row) at cloud base level, annually averaged.

10 Supplementary: modelled changes in particle concentrations and further discussion of particle numbers in the Amazon region

In Fig. S2 we present the seasonal cycle in surface 3 nm particle concentrations with and without pure biogenic nucleation, and the change when pure biogenic nucleation is included. This figure shows that the strongest effects are in the present-day in summertime in boreal regions, Australia, southern Africa and the Amazon region. Fig. S3 shows the effect of pure biogenic nucleation on cloud-level CCN concentrations in months chosen to reflect the Amazon wet and dry seasons (February and August). In Sect. 12, we further show that the present-day concentrations are in good agreement with observations at a diverse range of surface sites.

As discussed in the main text, on average our model predicts greater numbers of particles in the Amazon than observations suggest. According to the review by Martin *et al* [6], the mean number concentration in the

Table S1: Relative changes between present-day and pre-industrial atmosphere at cloud level, averaged over the month for January and July, and averaged over the entire year in the "Annual" column. The value quoted is the percentage increase in the mean in the present-day compared to the pre-industrial atmosphere. Pure biogenic nucleation is included.

Quantity	Change w.r.t. pre-industrial (%)				
	Jan. NH	Jul. NH	Jan. SH	Jul. SH	Annual
O_3	76.9	70.2	32.5	37.9	59.1
OH	41.7	21.3	-0.9	-0.1	14.6
H_2SO_4	361.4	84.4	0.8	42.3	79.3
HOM	-69.4	-36.1	-21.6	-26.4	-39.8
total J	1255.0	445.3	23.0	341.5	491.6
J_R	1264.5	164.3	24.5	343.2	484.0
J_{org}	-90.3	-47.5	-33.6	-65.3	-57.8
N_3	136.4	36.8	3.0	10.6	42.9
N_{70}	117.0	36.9	6.0	7.5	41.8
CCN 1%	122.2	42.2	4.6	7.3	44.4
CCN 0.2%	99.3	72.0	7.5	9.5	54.2

Aitken mode is $239\,\mathrm{cm}^{-3}$ and that in the accumulation mode is $177\,\mathrm{cm}^{-3}$, so the total concentration of particles of at least 70 nm in diameter (N70, a reasonable proxy for CCN, usually equivalent to a supersaturation between 0.4% and 1%) is $\sim 300\,\mathrm{cm}^{-3}$ in the wet season. Our model predicts N70 of $\sim 500 - 800\,\mathrm{cm}^{-3}$ in the wet season (higher near Manaus, lower near the coast, see Fig. S3) so it is still high, but within a factor 2 or 3, which may be larger than the measurement uncertainty but is certainly within our best estimate of the model parametric uncertainty [7]. In February in Manaus, we predict CCN concentrations at 0.2% supersaturation to be $127\,\mathrm{cm}^{-3}$ without pure biogenic nucleation, and pure biogenic nucleation increases this to $299\,\mathrm{cm}^{-3}$. Observations in the wet season in the pristine forest near Manaus are considerably lower, at $30 - 80\,\mathrm{cm}^{-3}$. However, our model averages over both the pristine forest and the Manaus pollution plume in this area so would be expected to yield higher concentrations than the pristine observations. There is the additional challenge that Manaus is so close to the Equator that it is close to the boundary between the wet season and the dry season. The observed transition season concentrations of $200 - 300\,\mathrm{cm}^{-3}$ may be more appropriate, which would agree with our model. In the dry

season, our model is in reasonably good agreement with observations, with on average $400 - 800\,\mathrm{cm}^{-3}$ CCN 0.2% (Fig. S3, Ref. [6]).

Overprediction of nucleation rates in the Amazon region suggests that pure biogenic nucleation may be suppressed there. The implications of the CLOUD results for the global atmosphere thus depend on whether a mechanism that suppresses pure biogenic nucleation exists, and if it does, whether or not it has a seasonal dependence, or is localised to regions with similar characteristics to the Amazon. For example, it could be associated with high humidity which increases the effective condensation sink [8, 9], high isoprene concentrations [10], high temperatures, or high levels of peroxy radicals, or it could be present globally all year round.

We note that overprediction of CCN (even without pure biogenic nucleation) is a feature of many global aerosol models. The Aerocom assessment of 15 global aerosol models [12] shows annual mean N100 concentrations in the Amazon of over $500\,\mathrm{cm}^{-3}$ and N30 concentrations of closer to $1000\,\mathrm{cm}^{-3}$. This suggests that there are general model weaknesses in the Amazon which makes it difficult to say anything with confidence about the effects an additional aerosol source would have on this region.

In Ref. [6], it is observed that particles in the nucleation and Aitken modes in the wet season rarely grow to larger sizes (although it is certainly the case that some condensational growth does occur). This is *a priori* surprising since the terpene and condensable organic concentrations are very high in this area. The most likely explanation must be very high losses: high precipitation frequency in the wet season and high condensation sink in the dry season, or very strong vertical mixing that prevents particle growth being observed adequately from a single surface location. Any losses are also likely to be inhomogenously distributed, and therefore their non-linear effects could be easily underestimated in a low-resolution model like ours.

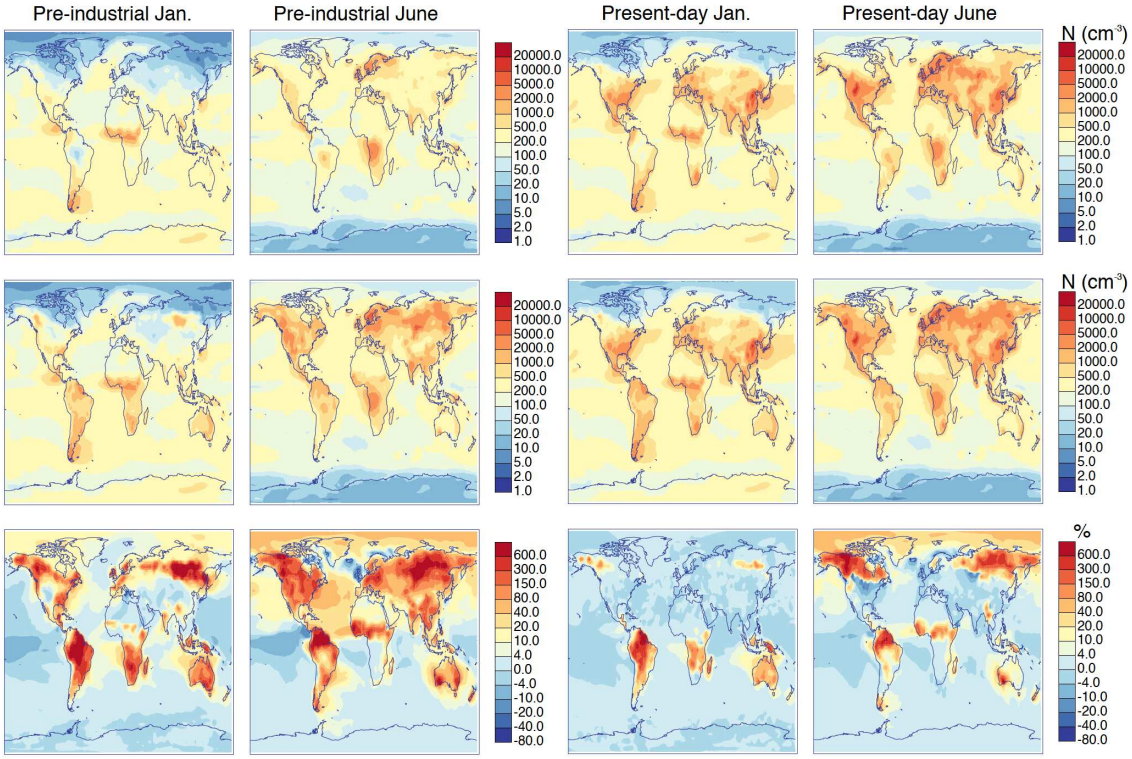

Figure S2: Monthly average surface level concentrations of particles larger than 3 nm in diameter, along each row pre-industrial January, pre-industrial June, present-day January and present-day June. Top: without pure biogenic nucleation. Middle: with pure biogenic nucleation. Bottom: percentage changes when pure biogenic nucleation is added.

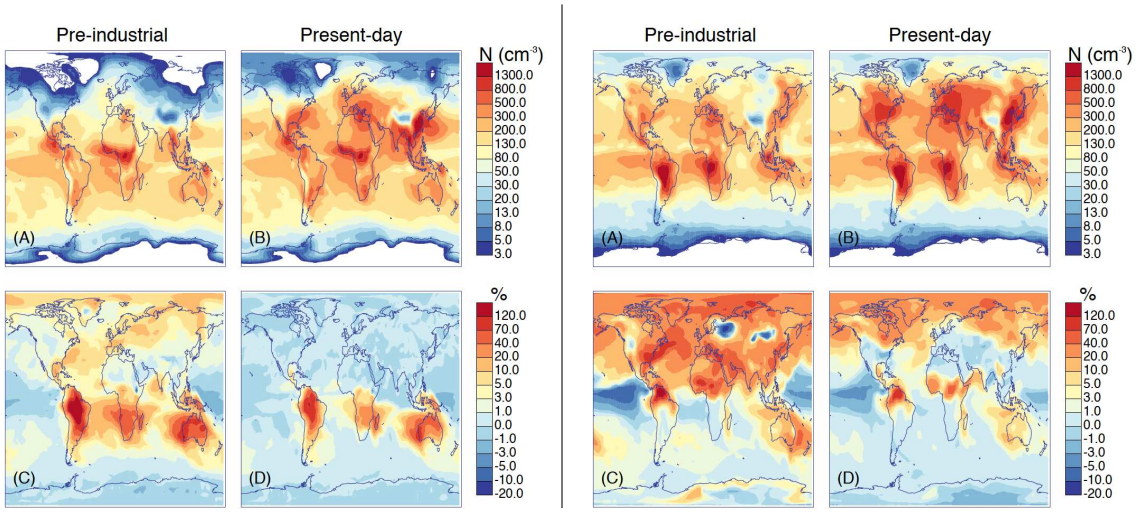

Figure S3: Concentrations of cloud condensation nuclei calculated at 0.2% supersaturation, in cm^{-3}, in February (left four panels) and August (right four panels). On either side of the line, average CCN concentrations at cloud base level over the month in (A) pre-industrial and (B) present-day conditions are shown, and, below these in subfigures (C, D), the percentage changes to these concentrations when pure biogenic nucleation is introduced.

11 Supplementary: diurnal cycles of particle formation rates

In Fig. S4, we show model predictions of the diurnal cycle of particle formation in July at Pallas and at the most studied field site, Hyytiälä, also in Finland. We predict that pure biogenic nucleation contributes significantly to the nucleation rate at both sites in July. One would expect APi-TOF data at Hyytiälä in July to show clusters of HOMs both with and without sulphuric acid. However, identifying the absence of sulphuric acid from clusters large enough to be equivalent to nucleated particles in mass spectra from field measurements at Hyytiälä has not been possible [13]. Furthermore, nucleation measurements at Hyytiälä are usually made in spring, when pure biogenic nucleation is predicted to make a much smaller contribution (see also Fig. S8, below). At Pallas, on the other hand, nucleation events will be rarer and signals are likely to be smaller, but the background particle concentrations responsible for the condensation sink are lower. Therefore, the peaks in nucleation at 200 and 520 hours into the month in Fig. S4 might well lead to observable 'banana'-type events, and our model suggests that all nucleation at this site should be dominated by pure biogenic processes. We note that while the pure biogenic nucleation mechanism can in principle operate both day and night, the model does not predict any nucleation at night, principally because terpene emissions are higher during daytime.

We also show the diurnal cycle in February (wet season) and August (dry season) at the most studied observation site in the Amazon, Manacapuru, in Fig. S5. Manacapuru is in the same model gridbox as the Amazon Tall Tower Observatory and as Manaus. At the surface level, approximately 0 to 30 m above ground level, we predict quite a substantial amount of nucleation, but rarely, perhaps never, enough to produce observable banana-type events, especially in the dry season.

Further, we show the evolution of the size distributions at the surface level at Hyytiala and Pallas in August in Fig. S6, and in Manacapuru for February and August in Fig. S7. Clear nucleation events are seen at both Finnish sites which are similar to the observations detailed in, for example, Refs. [14] or [15] for Hyytiala and [16] or [17] for Pallas. The size distributions also show that in the dry season in the Amazon essentially no nucleation is predicted, while in the wet season very weak additions to the Aitken mode are predicted. We speculate that these signals would rarely, if ever, be observable

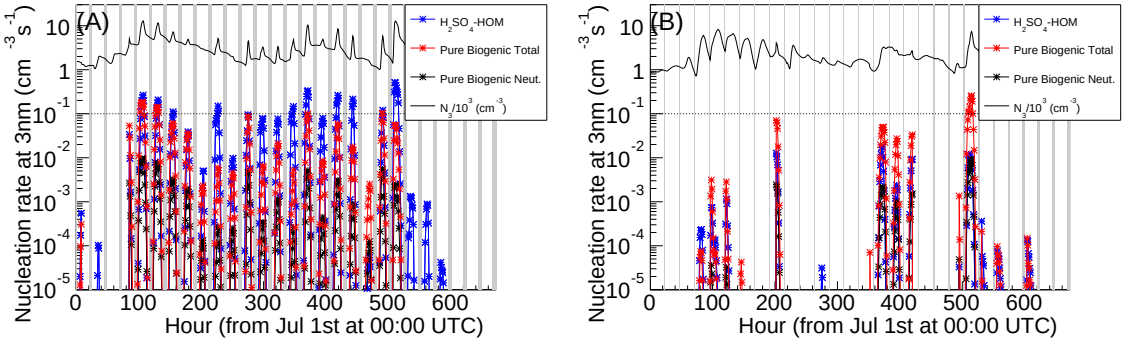

Figure S4: Modelled diurnal cycles of nucleation rates and particle concentrations at present-day (A) Hyytiälä (61.85°N, 24.28°E) and (B) Pallas (68.00°N, 24.23°E) in the first four weeks of July 2008. July is the month where the pure biogenic nucleation rate at Hyytiälä is strongest. Intervals between sunset and sunrise are marked in grey. The wind changes direction around 90 hours into the month. Observable nucleation events are likely when the nucleation rate is above around $0.1\,\mathrm{cm}^{-3}\mathrm{s}^{-1}$, indicated by the dotted line. At Pallas, Aitken mode particles are transported from nucleation happening elsewhere, which explains the daytime peaks in particle number concentration even when nucleation rates at Pallas are very low.

as nucleation events due to the complicated and inhomogenous meteorology.

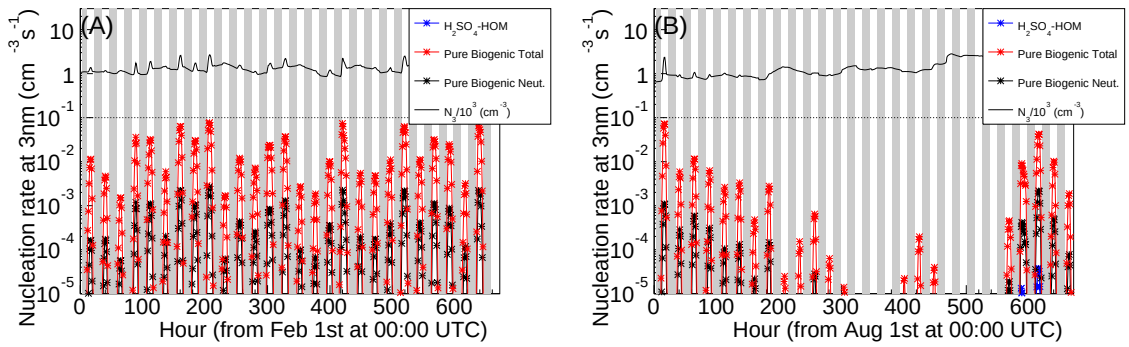

Figure S5: Modelled diurnal cycles of nucleation rates at present-day Manacapuru (3.30°S, 60.62°W) in (A) the wet season (February) and (B) the dry season (August), in 2008. The particle number concentration shows small spikes during nucleation events, often increasing from around 1000 cm^{-3} to around 2000 cm^{-3} due to nucleation. However, in these figures, our spatial model resolution smears out much larger, more local fluctuations in the particle concentrations (which can easily be a factor 10), and so in reality such small spikes would be difficult, and probably impossible, to discern in atmospheric observations.

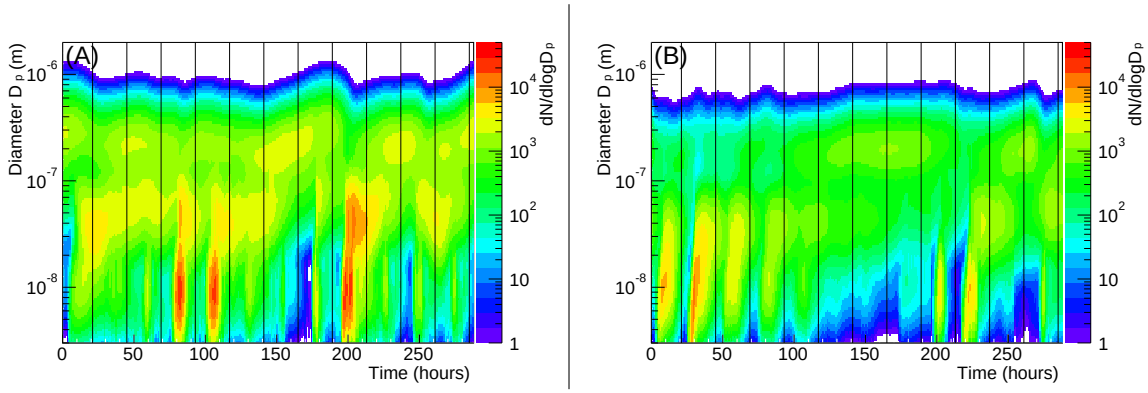

Figure S6: Modelled diurnal cycles of particle size distribution $dN/d\log D_p$ at present-day (A) Hyytiala and (B) Pallas, in the first twelve days of August (UTC time). The vertical lines mark midnight Finnish local time (UTC+3). In this figure $d\log D_p = 0.02$.

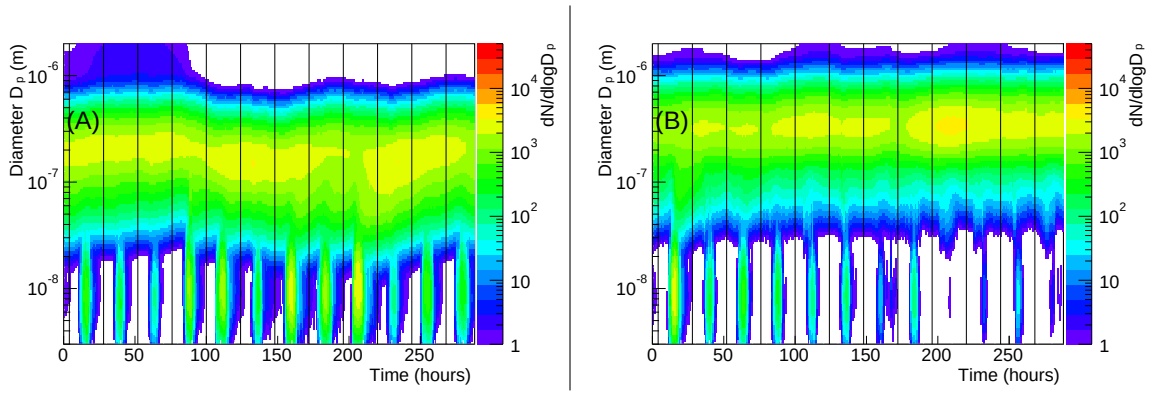

Figure S7: Modelled diurnal cycles of particle size distribution at present-day Manacapuru in (A) the wet season (the first twelve days of February, UTC time) and (B) the dry season (the first twelve days of August). Like the fluctuations in particle number concentration shown in Fig. S5, the small spikes in the nucleation mode would be difficult to see in observation data. The vertical lines mark midnight local time (UTC-4). In this figure $d \log D_p = 0.02$.

12 Supplementary: Model evaluation against measurements

The model was evaluated by comparing the particle number concentrations it predicts to those measured at 37 surface sites (Fig. S8). The first thirty-six are those used in Ref. [18]. We also added previously unpublished data recorded in 2010, 2011 and 2013 from a condensation particle counter with a 4 nm cut-off diameter at the East Trout Lake Global Atmosphere Watch station (54.35° N 104.98° W) because our model predicts a significant contribution from pure biogenic nucleation in central Canada. When measurements from multiple years are available, the data from the months in each year were averaged. Averaging over all sites and over the whole year, we find including pure biogenic nucleation leads to a modest improvement in the model bias from -42% to -41%. In summer the bias changes from -36% to -34% when pure biogenic nucleation is included and in winter it is unchanged at -53%. The overall low bias, particularly in winter, is likely to be because we do not include the effects of ammonia or anthropogenic organic molecules on nucleation in our model. Fig. S8 shows that pure biogenic nucleation strongly affects particle concentrations only at East Trout Lake, Listvyanka and Point Barrow.

We also compared the daily mean particle number concentrations from the model to observations made during the ARCTAS campaign [19] in spring and summer 2008 (Fig. S9). We interpolate these modelled particle concentrations within model grid boxes to match the locations of one-minute-averaged condensation particle counter measurements from a NASA P3-B aircraft. We note that the campaign during summer, when pure biogenic nucleation has most effect, was designed to investigate the influence of boreal forest fires. Our low model resolution and averaged fire emissions inventory mean that close agreement between the model and the measurements is not expected. However, we attempt the comparison anyway as the measurements are, unusually, in a particularly relevant region where pure biogenic nucleation is predicted to have a strong effect. Averaged over all altitudes, the model without pure biogenic nucleation is biased low (-57.9%); including pure biogenic nucleation leads to a smaller bias of -37.4%. Most of the pure biogenic contribution is in a band of latitudes from 51 to 62°N, corresponding to flights starting from Cold Lake between 26 June and 14 July 2008. While the relatively large underprediction of particle number without pure biogenic

nucleation suggests a particle formation pathway for pristine regions like ours may be needed, large local deviations of the model and measurements evident in Fig. S9 preclude any firm conclusion being drawn. These deviations could be attributed to the fire emissions or low model resolution discussed earlier, temperature or chemistry effects on new particle formation that we did not account for, or uncertainties in the observation data.

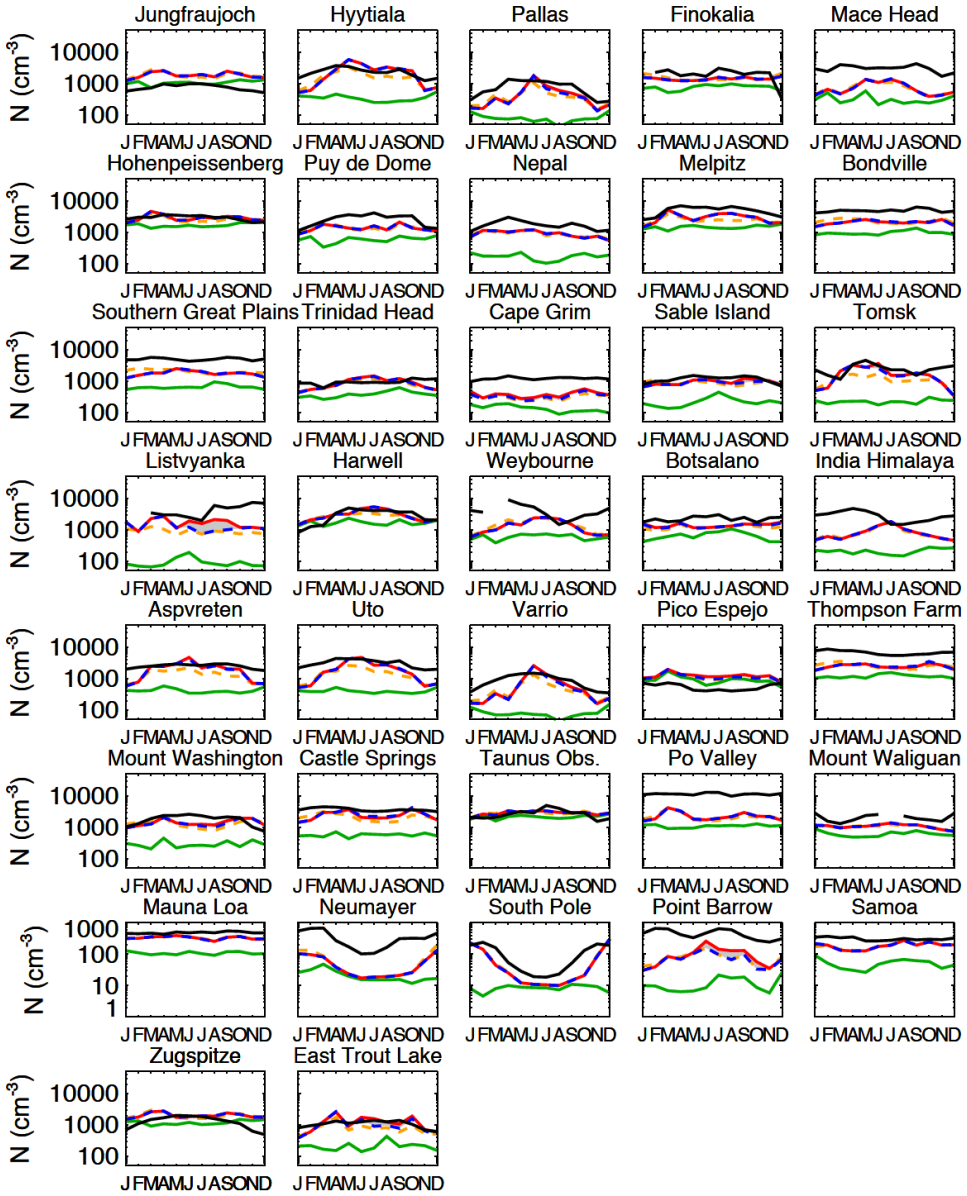

Figure S8: Particle concentrations at selected measurement sites [18, 20], in black, measured by counters with cut off sizes varying from 3 nm to 14 nm, compared to model predictions.The red curve shows the particle concentrations predicted by the baseline nucleation mechanisms, numbered 1-3 in the main text and including pure biogenic nucleation. The blue dotted curve shows the particle concentrations predicted without pure biogenic nucleation. A grey band is drawn between these two curves. The orange dotted curve shows particle concentrations predicted by the parametrisation of Paasonen et al [21], including a component of pure biogenic nucleation proportional to the square of the organic concentration. The dark green curve shows particles from primary emissions only.

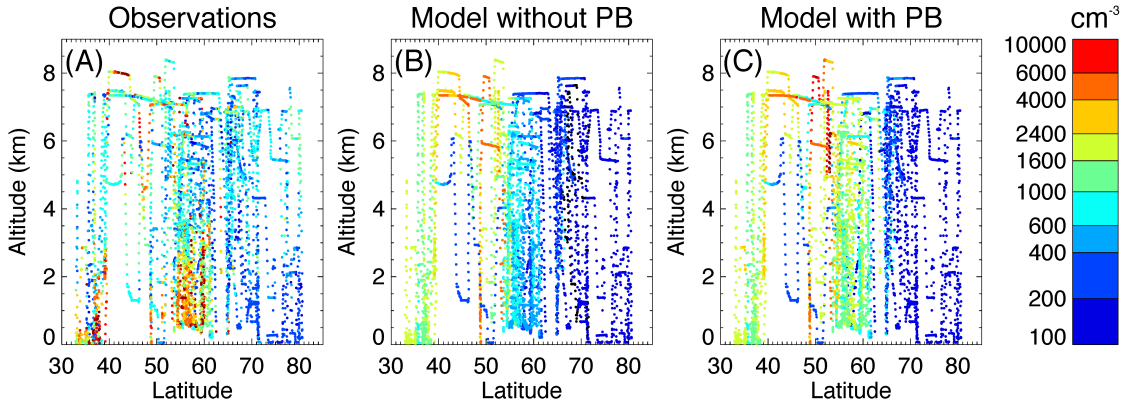

Figure S9: Vertical profiles of particle number concentrations calculated at S.T.P. (all particles of at least 3 nm in diameter), (A) measured during the ARCTAS campaign [19] in 2008, (B) modelled without pure biogenic nucleation (PB) and (C) modelled including pure biogenic nucleation.

13 Supplementary: further discussion of uncertainties and summary tables of sensitivity studies

Table S2 gives the changes to CCN concentrations due to pure biogenic nucleation in different scenarios (e.g. if organic nucleation is temperature-dependent, or if primary emissions are higher in our simulation than in reality). Following Ref. [7], the high primary emissions scenario corresponds to a doubling of biomass burning primary particulate emissions, a reduction in the mode diameter of biomass burning primary emissions from 150 nm to 100 nm, and a factor 2.5 increase in the sea spray flux. The low emissions scenario corresponds to a halving of biomass burning emissions, an increase in the mode diameter to 175 nm and a reduction in the sea spray flux by 60%. Table S3 gives the corresponding changes to radiative forcing.

In the main text, we discussed various sources of uncertainty, such as uncertainties in the CLOUD experimental measurements or HOM yields, in the possible temperature dependence of pure biogenic nucleation, and in primary emissions. The experiment in which we determine the sensitivity to

Table S2: Annual average global mean changes to the concentrations of particles larger than 70 nm in diameter, a proxy for CCN, when pure biogenic nucleation is introduced. The differences between the contents of the first row and the subsequent rows give an indication of the sensitivity of the analysis to different perturbations. The temperature dependence is added for both the baseline organic and the pure biogenic nucleation mechanisms, so changes both the control and perturbed simulations.

Variation	Change to CCN 0.2% (PI) (%)	Change to CCN 0.2% (PD) (%)
Add pure biogenic nucleation (PBN)	12	4
Add PBN with trebled yield	19	6
Add PBN with one-third yield	4	1
Add T dependence	7	2
Double pre-industrial volcanic SO2	11	-
Halve pre-industrial volcanic SO2	13	-
High primary emissions	7	2
Low primary emissions	14	5
Use baseline J from Ref. [21]	14	5

Table S3: Annual average global mean radiative forcings (including pure biogenic nucleation) and absolute changes to forcing when pure biogenic nucleation is introduced.

Variation	Forcing (PD-PI) Wm^{-2}	Change to forcing with pure biogenic nucleation (Wm^{-2})
Add pure biogenic nucleation (PBN)	-0.60	+0.22
Add PBN with trebled yield	-0.52	+0.30
Add PBN with one-third yield	-0.72	+0.10
Add T dependence	-0.64	+0.14
High primary emissions	-0.63	+0.17
Low primary emissions	-0.54	+0.33
Use baseline J from Ref. [21]	-0.63	+0.24
Exclude area close to Equator	-0.62	+0.20

primary emissions also tests our sensitivity to uncertainty in the condensation sink.

In the particularly interesting Amazon region, this uncertainty will be strongly influenced by the quantity and mode diameter of biomass burning emissions [7]. The 150 nm diameter used, following Ref. [22], is already larger than that of fresh smoke particles (averaged over all vegetation types) of 117 nm [23], and in theory condensation of secondary organic vapours in our model should increase this diameter to the aged diameter of 235 nm in a few days. However, the aged diameter in reality depends on cloud processing of organics as well as condensation, and this is not included in our model. To compensate for this, the larger fresh diameter of 150 nm is used by default in GLOMAP, but it has a large uncertainty [7]. We increase it to 175 nm in our sensitivity study with high primary emissions, which is still well within the uncertainty and the ranges for different phases of typical Amazon burning found in Ref. [24].

Another possible source of uncertainty in the Amazon region, which we are not able to quantify directly, is the condensation sink during periods of high humidity [8, 9] or aerosol-cloud interaction [25]. Clouds in the Amazon region are likely to be strongly affected by both of these. For the latter, impaction scavenging in GLOMAP is due to raindrops but the only loss mechanism to cloud droplets is nucleation scavenging. In a cloud, the coagulation sink should surely dramatically increase. This is not modelled. The effective condensation sink in a monodisperse cloud with 100 droplets cm^{-3} of diameter 15.6 μm (corresponds to LWC 0.2 gm^{-3}) is 0.16 s^{-1}. We investigated this further by including a crude treatment of this effect in our model. The perturbation to the model results in this test was small, mostly because clouds rarely cover more than half a model gridbox in areas where pure biogenic nucleation is important, and therefore the average effect is smeared out. With higher model resolution, however, the effect would likely be stronger due to the nonlinear nature of nucleation and survival probability.

References

[1] Zhang, K et al. (2011) Radon activity in the lower troposphere and its impact on ionization rate: a global estimate using different radon emissions. *Atmospheric Chemistry and Physics* **11**, 7817–7838.

[2] Usoskin, I. G, Kovaltsov, G. A, & Mironova, I. A. (2010) Cosmic ray induced ionization model CRAC:CRII: An extension to the upper atmosphere. *Journal of Geophysical Research* **115**, 6.

[3] Fraser-Smith, A. C. (1987) Centered and eccentric geomagnetic dipoles and their poles, 16001985. *Reviews of Geophysics* **25**, 1–16.

[4] Franchin, A et al. (2015) Experimental investigation of ion-ion recombination under atmospheric conditions. *Atmospheric Chemistry and Physics* **15**, 7203–7216.

[5] Kirkby, J et al. (2016) Ion-induced nucleation of pure biogenic particles. *Nature* **533**, 521–526.

[6] Martin, S et al. (2010) Sources and properties of Amazonian aerosol particles. *Reviews of Geophysics* **48**.

[7] Lee, L. A et al. (2013) The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei. *Atmospheric Chemistry and Physics* **13**, 8879–8914.

[8] Rose, C et al. (2015) Frequent nucleation events at the high altitude station of Chacaltaya (5240 m a.s.l.), Bolivia. *Atmospheric Environment* **102**, 18 – 29.

[9] Falvey, M & Garreaud, R. D. (2005) Moisture variability over the South American Altiplano during the South American low level jet experiment (SALLJEX) observing season. *Journal of Geophysical Research: Atmospheres* **110**.

[10] Kiendler-Scharr, A et al. (2009) New particle formation in forests inhibited by isoprene emissions. *Nature* **461**, 381–384.

[11] Crounse, J. D, Nielsen, L. B, Jørgensen, S, Kjaergaard, H. G, & Wennberg, P. O. (2013) Autoxidation of organic compounds in the atmosphere. *The Journal of Physical Chemistry Letters* **4**, 3513–3520.

[12] 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**, 4679–4713.

[13] Schobesberger, S et al. (2013) Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *Proceedings of the National Academy of Sciences* **110**, 17223–17228.

[14] Dal Maso, M et al. (2005) Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiala, Finland. *Boreal Environment Research* **10**, 323.

[15] Kulmala, M et al. (2012) Measurement of the nucleation of atmospheric aerosol particles. *Nature Protocols* **7**, 1651–1667.

[16] Väänänen, R et al. (2013) Analysis of particle size distribution changes between three measurement sites in northern Scandinavia. *Atmos. Chem. Phys* **13**, 11887–11903.

[17] Komppula, M, Sihto, S.-L, Korhonen, H, Lihavainen, H, Kerminen, V.-M, Kulmala, M, & Viisanen, Y. (2006) New particle formation in air mass transported between two measurement sites in northern finland. *Atmospheric Chemistry and Physics* **6**, 2811–2824.

[18] 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**, 4775–4793.

[19] Jacob, D. J et al. (2010) The Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) mission: design, execution, and first results. *Atmospheric Chemistry and Physics* **10**, 5191–5212.

[20] Torseth, K et al. (2012) Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972-2009. *Atmospheric Chemistry and Physics* **12**, 5447–5481.

[21] 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**, 11223–11242.

[22] Stier, P et al. (2005) The aerosol-climate model ECHAM5-HAM. *Atmospheric Chemistry and Physics* **5**, 1125–1156.

[23] Janhäll, S, Andreae, M. O, & Pöschl, U. (2010) Biomass burning aerosol emissions from vegetation fires: particle number and mass emission factors and size distributions. *Atmospheric Chemistry and Physics* **10**, 1427–1439.

[24] Costa, M. A. M et al. (2012) Real-time sampling of particulate matter smaller than 2.5 μm from Amazon forest biomass combustion. *Atmospheric environment* **54**, 480–489.

[25] Pierce, J. R, Croft, B, Kodros, J. K, D'Andrea, S. D, & Martin, R. V. (2015) The importance of interstitial particle scavenging by cloud droplets in shaping the remote aerosol size distribution and global aerosol-climate effects. *Atmospheric Chemistry and Physics* **15**, 6147–6158.