

SCIENTIFIC REPORTS

OPEN

Novel insights on new particle formation derived from a pan-european observing system

Received: 4 May 2017

Accepted: 20 November 2017

Published online: 24 January 2018

M. Dall'Osto^{1,2,3,4}, D. C. S. Beddows², A. Asmi⁵, L. Poulain⁶, L. Hao⁷, E. Freney⁸, J. D. Allan⁹, M. Canagaratna⁴, M. Crippa^{10,24}, F. Bianchi^{5,10}, G. de Leeuw^{11,12}, A. Eriksson¹³, E. Swietlicki¹⁴, H. C. Hansson¹⁵, J. S. Henzing¹², C. Granier^{16,17}, K. Zemankova¹⁸, P. Laj^{5,19,20}, T. Onasch⁴, A. Prevot¹⁰, J. P. Putaud²¹, K. Sellegri⁸, M. Vidal²², A. Virtanen⁷, R. Simo¹, D. Worsnop^{4,5}, C. O'Dowd³, M. Kulmala⁵ & Roy M. Harrison^{2,23}

The formation of new atmospheric particles involves an initial step forming stable clusters less than a nanometre in size ($< \sim 1$ nm), followed by growth into quasi-stable aerosol particles a few nanometres (~ 1 – 10 nm) and larger ($> \sim 10$ nm). Although at times, the same species can be responsible for both processes, it is thought that more generally each step comprises differing chemical contributors. Here, we present a novel analysis of measurements from a unique multi-station ground-based observing system which reveals new insights into continental-scale patterns associated with new particle formation. Statistical cluster analysis of this unique 2-year multi-station dataset comprising size distribution and chemical composition reveals that across Europe, there are different major seasonal trends depending on geographical location, concomitant with diversity in nucleating species while it seems that the growth phase is dominated by organic aerosol formation. The diversity and seasonality of these events requires an advanced observing system to elucidate the key processes and species driving particle formation, along with detecting continental scale changes in aerosol formation into the future.

¹Institute of Marine Science, Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain. ²National Centre for Atmospheric Science Division of Environmental Health & Risk Management School of Geography, Earth & Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom. ³School of Physics, Centre for Climate & Air Pollution Studies, National University of Ireland Galway, University Road Galway, Galway, Ireland. ⁴Aerodyne Research, Inc., Billerica, MA, USA. ⁵Department of Physics, University of Helsinki, P.O. Box 64, 00014, Helsinki, Finland. ⁶Leibniz Institute for Tropospheric Research, Permoserstr. 15, 04318, Leipzig, Germany. ⁷University of Eastern Finland, Department of Applied Physics, P.O. Box 1627, FIN-70211, Kuopio, Finland. ⁸Laboratoire de Météorologie Physique, CNRS-Université Blaise Pascal, UMR6016, 63117, Clermont, Ferrand, France. ⁹School of Earth, Atmospheric and Environmental Sciences, The University of Manchester, Manchester, UK. ¹⁰Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, 5232, PSI, Villigen, Switzerland. ¹¹Finnish Meteorological Institute, Climate Change Unit, P.O. Box 503, 00101, Helsinki, Finland. ¹²Netherlands Organisation for Applied Scientific Research TNO, Princetonlaan 6, 3508 TA, Utrecht, The Netherlands. ¹³Division of Ergonomics and Aerosol Technology, Lund University, Box 118, SE-22100, Lund, Sweden. ¹⁴Division of Nuclear Physics, Lund University, Box 118, SE-22100, Lund, Sweden. ¹⁵Department of Environmental Science and Analytical Chemistry, Stockholm University, 10691, Stockholm, Sweden. ¹⁶Laboratoire d'Aérodologie, Toulouse, France. ¹⁷NOAA Earth System Laboratory and CIRES, University of Colorado, Boulder, USA. ¹⁸Charles University, Faculty of Mathematics and Physics, Dept. of Atmospheric Physics, Prague, Czechia. ¹⁹Univ. Grenoble-Alpes, CNRS, IRD, INPG, Institut des Géosciences de l'Environnement, Grenoble, France. ²⁰Univ. Grenoble-Alpes, CNRS, IRD, Observatoire des Sciences de l'Univers, Grenoble, France. ²¹European Commission, Joint Research Centre, Institute for Environment and Sustainability, 21027, (VA), Italy. ²²Department of Evolutionary Biology, Ecology and Environmental Sciences, Universitat de Barcelona, Av. Diagonal 643, 08028, Barcelona, Catalonia, Spain. ²³Department of Environmental Sciences / Center of Excellence in Environmental Studies, King Abdulaziz University, PO Box 80203, 21589, Jeddah, Saudi Arabia. ²⁴Present address: European Commission, Joint Research Centre (JRC), Directorate for Energy, Transport and Climate, Air and Climate Unit, Via E. Fermi 2749, I-21027, Ispra, (VA), Italy. Correspondence and requests for materials should be addressed to M.D'O. (email: dallosto@icm.csic.es)

Atmospheric aerosol formed by nucleation is hypothesized to provide an important source of global cloud condensation nuclei (CCN)^{1,2}. Such processes have a major influence on the microphysical properties of clouds and the radiative balance of the global climate system³. However - despite its importance - atmospheric nucleation is still poorly understood; it is not clearly known whether it is dominated by a single nucleation pathway, or whether multiple different mechanisms are competing with each other.

There is strong evidence that atmospheric new particle formation (NPF) involves clusters with at least one molecule of sulphuric acid^{4,5}; but nucleation theories have failed to predict correctly either the observed nucleation rates or their functional dependence on sulphuric acid concentrations. Recently, the complexity and variability of the atmosphere has hindered elucidation of the fundamental mechanism of NPF from gaseous precursors in experiments performed with the CLOUD (Cosmics Leaving Outdoor Droplets) chamber at CERN. The CLOUD chamber enables precise control of experimental parameters and provides the exceptionally clean experimental conditions that are essential when performing experiments with extremely low concentrations of participating vapours. Overall, molecular analysis of the mechanisms of the initial stage ($\sim 1\text{--}3\text{ nm}$) reveals that a number of primary vapours may be responsible. Atmospherically relevant ammonia concentrations increase the nucleation rate of sulphuric acid particles by more than a factor of 100 to 1,000⁶. Ions are also expected to enhance nucleation; nevertheless - even with the large enhancements in rate caused by ammonia and ions - the remarkable study of Kirby *et al.* (2011)⁶ concluded that atmospheric concentrations of ammonia and sulphuric acid are insufficient to account for observed boundary layer nucleation. Better agreement with predicted nucleation rates is achieved when including efficient stabilization of the acids by bases such as amines⁷. Indeed, amines strongly enhance nucleation rates already in the low pptv range⁸. The involvement of oxidized organic molecules in the process, alongside sulphuric acid, was proposed in early studies and has been now well-established^{9–11}. Additionally, the existence and formation mechanisms of essentially nonvolatile highly oxidised molecules deriving from biogenic organic vapours in the atmosphere were elucidated, both in laboratory studies and in the ambient atmosphere^{12,13}. Lately, evidence for the formation of biogenic aerosol particles from highly oxygenated molecules (HOMs) in the absence of sulphuric acid in a large chamber under atmospheric conditions was presented¹⁴. On the basis of the combined modelling results and experimental data, such low-volatility organic vapours are the key to particle growth at the initial sizes¹⁵. There is also strong experimental evidence that pure organic nucleation proceeds alongside sulphuric acid-driven nucleation in the free troposphere¹⁶. Indeed, simulations and a comparison with atmospheric observations show that nearly all nucleation throughout the present-day atmosphere involves ammonia or biogenic organic compounds, in addition to sulphuric acid^{17,18}. The new mechanism for organic particle formation without sulphuric acid provides a way to form particles in the pristine preindustrial atmosphere, when the concentrations of sulphuric acid and ammonia were much lower¹⁹.

In marine and coastal environments - biogenic iodocarbons emitted from marine algae may control the formation of marine aerosols and cloud condensation nuclei^{20,21}. Additional marine biogenic components - namely amines and methanesulfonic acid - may also play a role in the process²².

Overall, the clustering of vapours is a major source of new particles of $\sim 1\text{--}10\text{ nm}$ in diameter, but these small particles must grow in size to act as CCN and influence clouds.

It is well established that oxidation products of volatile organic compounds (VOCs) are important for particle growth^{3,23}. As the particles increase in size ($\sim >10\text{ nm}$), condensation of organic compounds probably becomes increasingly important, although the exact identities of the organic molecules driving the growth of atmospheric particles are largely unresolved. Our understanding on the role of organics and other chemical species in nanoparticle growth is advancing quickly; but there is a considerable gap between modelling and laboratory studies on the one hand, and direct ambient experimental evidence on both new particle formation occurrence and relative chemical composition on the other. So far, there is very limited experimental evidence upon the chemical processes driving that growth, or the chemical composition of the particles comprising the nucleation mode in the atmosphere³.

Coordinated field measurement studies of the atmospheric composition and size distribution of aerosols are essential to bridge the gap between these two extremes. This can help to elucidate the processes responsible for the occurrence and the growth of atmospheric nanoparticles. However, the atmospheric real time detection of the frequency and the chemical composition of NPF events is difficult because it requires the expensive, long term deployment of multiple state-of-the-art instruments to make field measurements in different environments. Here - for the first time - we use novel methodologies applied to the most comprehensive dataset for aerosol size distributions and chemical composition available to date (EUCAARI - European Aerosol Cloud Climate and Air Quality Interactions; EMEP - European Monitoring and Evaluation Programme and EUSAAR - European Supersites for Atmospheric Aerosol Research^{24–26}). Specifically - by means of real time, NPF ambient measurements - we aim to throw new light upon the nature of particle nucleation processes across Europe. The main objectives of this work were (a) a categorization and quantification of ambient NPF events by means of aerosol size distributions (mobility diameter $\sim 17\text{--}30\text{ nm}$ ^{27,28}; Methods), (b) elucidating the observed spatial regional NPF variability; and (c) identify the real time chemical composition of maturing/grown nucleation mode particles (vacuum aerodynamic diameter, $\sim 30\text{--}60\text{ nm}$ ^{29,30}; Methods).

Results and Discussion

Detecting New Particle Formation Events Across Europe. A unique dataset, spanning 24 supersites across Europe was analysed for the frequency of occurrence and seasonal patterns associated with the formation of new particles (Table S1, Methods). The vast dataset, comprising 117,000 hourly size distributions over the years 2008–2009, was exposed to K-means cluster analysis of a maturing nucleation mode size range extending from 17 nm to 30 nm , which in turn, revealed the presence of four clusters describing the entire aerosol population, one of which represents the size distributions associated with a maturing nucleation mode (Methods, Figure S1a–e). An example of how the cluster analysis identifies the occurrence of the nucleation mode is shown in Fig. 1.

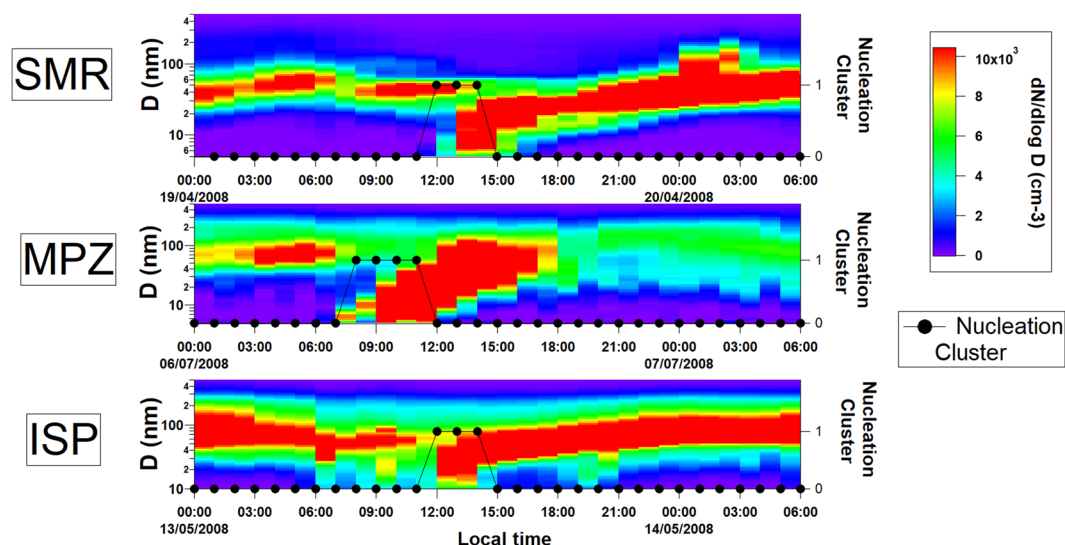


Figure 1. New particle formation events at three sites in Europe. The right hand scale and black dots indicate on a binary scale those hours for which our method attributes the size distribution to the nucleation cluster.

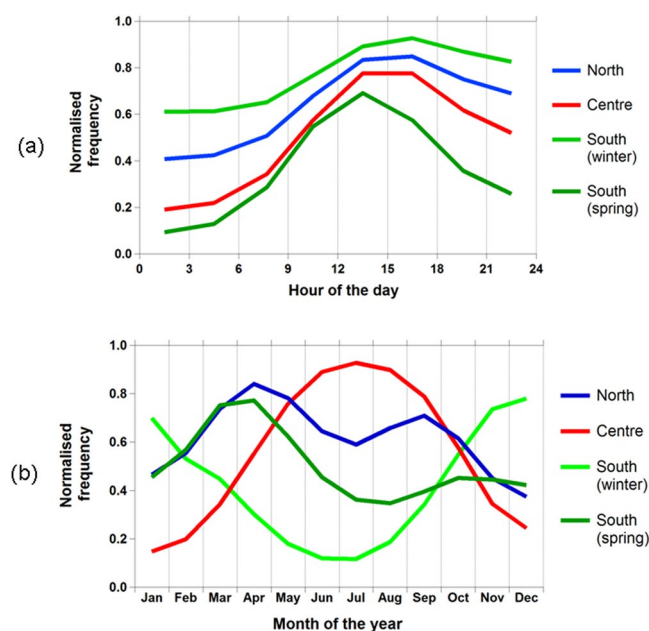


Figure 2. Seasonal prevalence of new particle formation events across Europe in the period 2008–2009 and relative diurnal profile (**a** and **b**, respectively).

Previous analyses of the EUCAARI dataset have presented a detailed overview of the sampling sites and seasonally disaggregated size distributions excluding these sizes²⁷, or have discussed mainly larger aerosol size modes²⁸. Our results are compared with previous continuous atmospheric cluster and particle measurements taken with different types of air ion and cluster mobility spectrometers³¹, and found to be comparable in the overlap particle size region (17–42 nm), as shown in Figure S1f. Our new methodology allows detection in real time of NPF events across the monitoring stations, accounting for $7 \pm 4\%$ of the sampled time at each site.

When NPF events for each of the 24 stations were analysed independently, strikingly different seasonal cycles are seen. These are summarised in four main categories, shown in Fig. 2a. The northern group of monitoring stations (ZEP, PAL, SMR, ASP, MHD) show a peak frequency mainly in spring and in autumn. By contrast, central European monitoring stations (VHL, WAL, MPZ, OBK, BIR) mainly show a clear seasonality trend peaking during the summer months. Southern group stations present more complex scenarios, with some having a maximum occurrence in winter (SHC, ZSF, PDD, BEO, CMN) and others in the spring (KPO, FKL, HPB, ISP, JFJ). The average diurnal occurrence of NPF K-mean clusters (Fig. 2b) shows - for all different seasonal categories - a similar profile peaking during daytime³, although less sharp for North and South (winter) categories relative to South (spring) and Centre ones. The overall results of our K-means SMPS analysis are summarised and plotted in

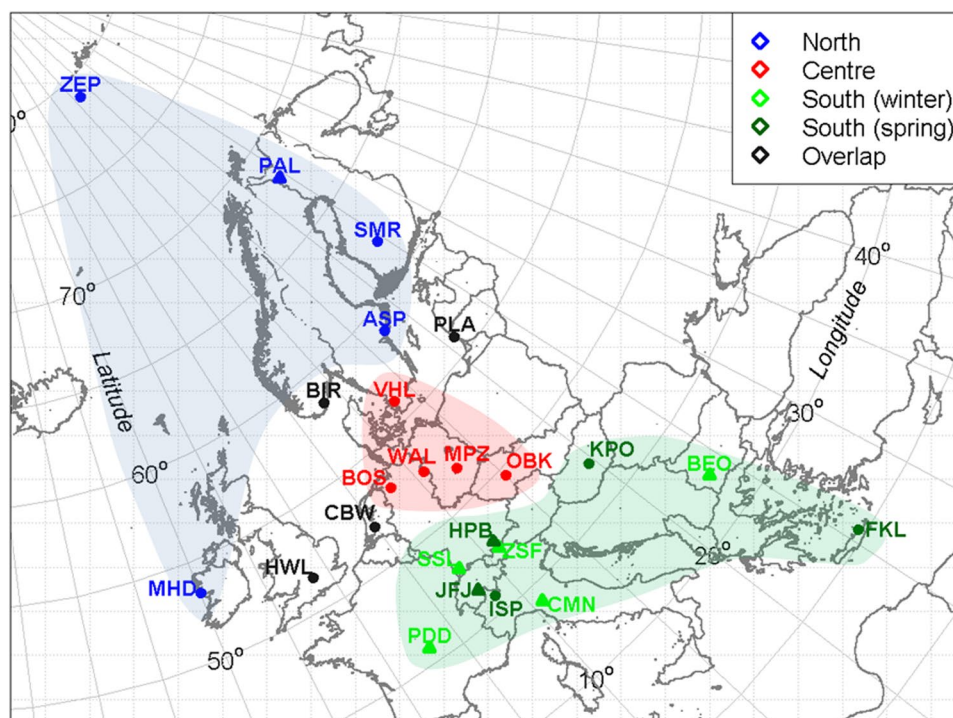


Figure 3. Spatial distribution of the different patterns of new particle formation.

the map shown in Fig. 3, which is the current best representation of the largest real time aerosol size distribution dataset available today and analysed at a continental level. Overall, we successfully allocated 20 of the 24 studied stations to three broad NPF European regions (North, Centre, South). Four monitoring stations (Black dots; HWL, CBW, BIR, PLA) do not well fit the regional pattern, likely due to an overlap of the geographical regions; or in the case of HWL and CBW experience substantial influence of local anthropogenic sources which may affect their seasonality. For this reason these four stations are categorised as “overlap” and are not discussed in the present work.

Elucidating the Observed Spatial Regional Variability. Any interpretation of the regionality of the formation of nucleation mode particles must consider both the nucleation process itself and early growth (~ 1 – 10 nm), and the subsequent particle growth ($\sim > 10$ nm); as both processes are an essential pre-requisite to particles being recognised as within the nucleation mode in this study. Thus, factors which need to be considered include the following:

- The condensation sink is a very important factor in influencing the nucleation process. Homogeneous nucleation is unlikely to occur in environments with a high condensation sink as under such circumstances, condensable molecules and clusters are likely to attach to existing surfaces rather than self-nucleating to form new particles.
- Chemical substances which are critical to the nucleation process itself and initial growth (~ 1 – 10 nm). The strongest evidence relates to sulphur dioxide as a precursor of sulphuric acid together with basic species (ammonia and amines) but there is also strong evidence for active participation of other species such as oxidised VOC or iodine compounds co-nucleating or acting alone in the nucleation process.
- Chemical substances involved in the growth ($\sim > 10$ nm) of newly nucleated particles to tens of nanometres in diameter. This requires condensable vapours and the strongest evidence in continental environments is for sulphates and oxidised organic compounds, although other species of low vapour pressure may have the capacity to contribute.

Consequently, the spatial and temporal evolution of the NPF events depends primarily upon two parameters: a high vapour formation rate and a low condensation sink (CS) due to pre-existing particles³². In other words, the lifetimes of nucleating particles depend on the competition between their condensational growth and cluster scavenging³³. The calculated CS is highly variable across the European monitoring stations, although lower in general in north European stations (Figure S2b). The observed meteorological conditions alone (temperature, relative humidity, solar radiation) cannot explain our spatial differences in seasonal behaviour as, although different in amplitude, the same seasonal pattern in meteorology can be seen across Europe (Figure S2a–c). Previous analysis³⁴ revealed that while the increased uptake of water by particles does affect the concentration of nucleating vapors and survival of nucleating clusters to some extent, these effects are typically minor in comparison to the reduced OH effect and sulphuric acid limitation.

Nucleation occurred most frequently on sunny days with below-average CS. To test this further, NPF events were explored by correlating nucleation activity and H_2SO_4 production - based on an existing concept^{35–37} - Figure S2d shows the relationship between the hourly-average product of [UV intensity * SO_2 concentration] (a proxy for sulphuric acid production) versus the condensation sink; helping us to view the data in terms of condensation and new particle formation. A total of 3691 hours of specific NPF events from seven sites with available data are plotted. The conclusions that can be drawn from comparing the different NPF events presented in Figure S2d are as follows: for a given level of condensation sink (CS), NPF events in the north of Europe frequently occur at a lower level of SO_2 concentration than events occurring in central and southern Europe (assuming similar levels of UV, as indicated in Figure S2a). Specifically, for the interval 0.007–0.013 (s^{-1}) of CS, average x-values in Figure S2d range by a factor of seven (521, 115, 75 W m^{-2} ppb, for Northern, Central and Southern groups, respectively). The scatter in Figure S2d suggests that the CS, solar radiation and SO_2 concentrations alone cannot fully explain the variability of the detected ambient NPF events, as pointed out recently¹⁷.

The variability of vapour phase precursor distributions across the European continent is large. Unfortunately, the only inorganic chemical species continuously monitored in some air quality stations - and thought to be involved in the nucleation process - is SO_2 . In recent decades much lower SO_2 levels have been recorded following strict implementation of controls upon power generation, manufacturing industry and road transport fuels³⁸, although in central and eastern Europe coal-fired power generation plants still play a major role (Figure S3a). The road vehicle fleet is responsible for substantially higher NO_2 concentrations in Southern Europe than in Western and Northern Europe³⁹, as reflected in our data (Figure S3a–c). Ammonia emissions in Europe - 94% originating from agriculture - have fallen since 1990, but by not as much as emissions of other air pollutants⁴⁰.

None of the volatile organic compounds were sampled continuously for the 2008–2009 studied period at the European aerosol monitoring stations, therefore Figure S2d cannot be generated for chemical species other than SO_2 . Nevertheless, we estimated European volatile organic compound emissions calculated by the MEGAN model^{41,42} (Methods) at monthly resolution. Here, we aim to briefly present and discuss the spatial variability of VOC across the three main geographical European regions reported. Biogenic VOC emissions are greater in warmer southern Europe (Figure S3d), although ratios of VOC and Oxidised VOC (O-VOC) are different. When considering the different stations around Europe, the main difference is found between the North and the South regions, whereas the centre shows intermediate values most of the time (S3d–h). Overall, when looking at the ratio of different biogenic VOC between south and north regions (Figure S3e), the South has about two times higher concentrations of VOC relative to the North. Anthropogenic VOC (A-VOC) are also distributed differently (Figure S3g), and also present different seasonal variations^{43–45}. Ad-hoc coordinated field studies monitoring both NPF events and organic and inorganic gaseous precursors are needed at continental level in order to address this current large knowledge gap.

Real Time Chemical Composition of Nucleation Mode Particles. There is limited, but nonetheless useful information upon the chemistry of ultrafine particles during the growth phase of the frequently observed nucleation events across Europe through direct atmospheric ambient measurements. Recently developed research instruments have helped immensely in understanding the processes occurring in the 1–5 nm size range^{46–48}. CLOUD laboratory experiments have dramatically advanced knowledge of chemical substances which are critical to the nucleation process itself and initial growth (~ 1 –10 nm), although field observations are limited^{16,21}. Further insights into the processes driving particle growth (> 10 nm) can be gained from knowledge of the chemical composition of the nucleation mode particles detected in real time conditions. However, this cannot be discerned from normal air quality observations. In fact, it is difficult to measure it directly with state of the art instrumentation - as usually deployed in research field studies. The best available evidence derives from the Aerodyne Aerosol Mass Spectrometer (AMS)⁴⁹. The time-of-flight (ToF) AMS allows the quantitative measurement of size resolved chemistry of submicrometer non-refractory aerosol with high time resolution and high sensitivity^{50,51}. Indeed, the AMS has been successfully applied in an environmental chamber study on new particle formation and growth^{20,52}. Additionally - using an AMS - Zhang *et al.* (2004)²⁹ were able to show that the composition of the growing particles (33–60 nm vacuum aerodynamic diameter or about 18–33 nm in physical diameter) was predominantly sulphuric acid during the earliest observable stages of formation events in the urban area of Pittsburgh. Also using an AMS in Hyytiälä, a forested site in southern Finland, Allan *et al.* (2006)⁵³ were also able to characterise by AMS the particles in the < 50 nm regime several hours after a nucleation event, demonstrating that the particles were principally organic in composition. However, despite convincing evidence of the presence of organics in growing particles in individual locations, no harmonised analysis of a large region composed of many monitoring stations has been conducted to date.

Therefore, here we use another unique dataset collected over three intensive field measurement campaigns to investigate the aerosol chemical composition over Europe by means of AMS⁵⁴. Such AMS campaigns were carried out also within the framework of EUCAARI/EUSAAR/EMEP during 2008 (May–June and September–October) and 2009 (February–March). Our unique dataset derives from combining the size resolved chemical composition data from the EMEP-EUCAARI-EUSAAR ToF-AMS, temporally overlapped with the presented aerosol size distribution data from SMPS instruments collected over the 24 European monitoring sites. Only 9 of those stations were equipped to collect AMS data and only 5 had size resolved aerosol ToF-AMS data suitable for this study (Table S1). Hourly data were available for both size and composition at KPO (686 hours), MPZ (2,289 hours), PDD (852 hours), and VHL (859 hours). The AMS at SMR did not work during 2008–2009 but a research campaign from May–July 2014 was included (SMR; 2,160 hours). ToF-AMS data were summarised in five logarithmically equally spaced size bins between the range 20 and 500 nm (D_{vac} , Vacuum aerodynamic diameter), reporting quantitative hourly concentrations of sulphate, nitrate, ammonium and organic matter (Methods). Collection efficiency of the particles at the detector approaches 100% for particles of aerodynamic diameters in

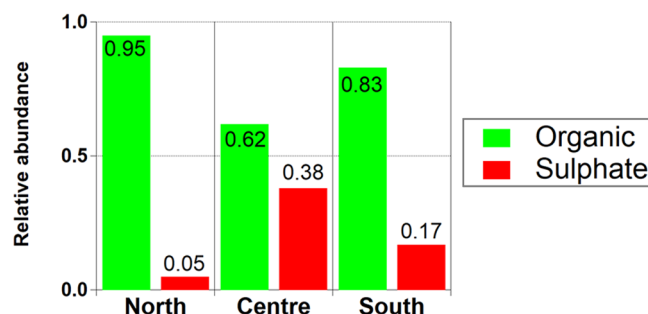


Figure 4. Relative abundance of organic and sulphate in nucleation mode particles (<50 nm) for the three different European regions.

the range of 70 to 500 nm, but dramatically decreases for the smaller diameter particles - reaching 0% at about 20–30 nm⁴⁹ (Methods).

Overall, the new combined SMPS-AMS dataset was composed of a matrix of 6,846 hours. Whilst they represent only about 5% of the SMPS dataset, they sum up to about 280 days of combined SMPS-AMS measurements. We applied Positive Matrix Factorization (PMF) to a combination of aerosol size distribution (SMPS) and size resolved chemical composition data (AMS) in order to identify associations between characteristic modes in the size distribution, and the size-resolved chemical composition of particles. This allows us to link directly the chemical composition of particles to size association. Results of the PMF analysis are shown in Figure S4a–c; AMS size bins have been converted from vacuum aerodynamic diameters to mobility diameters^{29,30} using a density of 1.4 g cm^{-3} .

Figure S4a–c shows the F-matrix for each of the solutions when we progressively include an additional factor into the model. Starting with a three factor solution, Figure S4aF1–3 shows three number size distribution modes at 25 nm (with a strong organic component), 60 nm (with a strong organic, as well as inorganic sulphate, nitrate and ammonium component) and 150 nm (mainly with an inorganic component composed of ammonium nitrate and sulphate). Adding a fourth factor (Figure S4b), the modal diameter of the factor describing the smallest number size distribution maintains its features. An extra accumulation mode is observed, mainly composed of organic components, but not adding additional information. The additional factor (S4bF3) along with factor S4bF2 likely represent the broad Aitken organic mode represented in the 3 factor solution (S4aF2). The five factor solution also does not add additional information, but starts to present similar factors (S4cF2 and S4cF3) as well as not presenting clearly defined particle number size distributions. In summary, the 3 factor solution gives a clear nucleation mode and a clear Aitken mode, as well as an accumulation mode. The nucleation mode shows the strongest association with organic matter in the AMS 30 nm and 55 nm size bins (D_{va}).

We also looked directly at the ToF-AMS data during new particle formation events. As case studies, we examined 16 events occurring at three stations (SMR (North), PDD (South), MPZ (Centre)), and provide size resolved chemical abundances in the lowest two size bins (vacuum aerodynamic diameter < 70 nm, corresponding to a physical diameter ~50 nm). Figure 4 shows the relative contributions of organics and sulphate to particles in the lowest size range at the three stations. Results point to a major role of organics in particle growth in northern Europe, with the lowest in the Central Europe - where sulphate plays a bigger role, presumably due to higher average concentrations of SO_2 (Fig. S3a–c). The growth of newly formed particles from sizes of ~1–5 nm up to the sizes of cloud condensation nuclei (~100 nm) in many continental environments requires abundant very low volatility organic vapours^{55–57}. Mentel *et al.*, (2009)⁵⁸ found that plant-emitted OVOC may play a role in the process of new particle formation consistent with the importance of oxygenated organics in nucleation events over boreal forests⁵⁹. The variation between AMS systems⁶⁰ and the absence of standardized data processing, calibration practices and data treatment in the low diameter AMS size bins creates difficulty in deriving harmonised information at the level of research based field studies^{29,30}.

Inferences. The information presented, and other data presented in the Supporting Information allow some tentative inferences to be drawn regarding the processes likely to determine the seasonal effects upon nucleation/growth events seen clearly in this dataset. These are as below:

- The central group of sites show a peak in nucleation mode particles between June and September (Fig. 2), occurring at midday. These sites show the highest mean concentration of biogenic VOC and similar concentrations of anthropogenic VOC to southern Europe. The condensation sink shows no clear pattern (Figure S2c). It seems likely that at these sites, the nucleation process, favoured by high concentrations of SO_2 and high insolation favouring its oxidation is the determinant of the seasonal profile.
- The northern group of sites show peak occurrence of nucleation mode particles in April and September and greatest occurrence at midday (Fig. 2). These sites have the lowest sulphur dioxide (Figure S3a) and the lowest concentrations of both anthropogenic and biogenic VOC (Figures S3e and S3g). The condensation sink is low in August and September, favouring nucleation at this time of year, but not in April and May (Figure S2c). As the release of biogenic VOC is low in April following the winter minimum, the explanation for a peak occurrence in Spring is unclear, unless lower temperatures at this time of year are influential. It is notable that several of these sites, most notably MHD, ZEP and ASP are heavily exposed

to maritime air masses, and in this context the seasonality of biogenic marine emissions may be important. The contribution of iodine oxides to the nucleation process is known to be important, especially at MHD^{20,21}. MSA (methanesulphonic acid) and DMA (dimethylammonium salts) were already identified as a possible connection between marine air masses and particle formation events in North European boreal forests (SMR station⁶¹).

- (c) The southern group of sites show either a winter or spring maximum in the frequency of nucleation events (Fig. 2a). Those showing a spring maximum (KPO, SIP, HBP, FNK and JFJ) on average show a peak frequency in occurrence at midday in March, while the winter sites (SSL, ZSF, PDD, BEO, CMN) have peaks at midday in December and February (Figure S2a). The southern European sites show intermediate concentrations of sulphur dioxide and high concentrations of nitrogen dioxide (Figure S3a). They show a pronounced summer maximum in biogenic VOC (Figure S3f) and higher levels than other regions. Concentrations of anthropogenic VOC are similar to central Europe. For these sites, the existence of a minimum in the condensation sink at the relevant time of year (Figure S2c) appears to be the probable determinant of frequent nucleation events.

This study shows that new particle formation and growth is a frequent occurrence, but that the processes responsible differ substantially across the European continent. This is the first analysis of the field measurement evidence showing regional differences in nucleation events across a continent, and the first direct evidence that organic compounds dominate the growth of new particles at continental level. The results highlight the importance of a minimum in the condensation sink correlating with a maximum frequency of nucleation at many of the sites. However, this does not explain the higher frequency of nucleation in summer at the central European sites, where high sulphur dioxide and insolation appear to be more important. At many northern European sites, and especially those located close to the coast, it may be marine biogenic precursor concentrations which determine the production of nucleation mode particles. These factors all point to the nucleation, rather than the particle growth process as being the driver of the events when particles reach sizes in the tens of nanometre range from which they can begin to show activity as CCN.

This work highlights some of the benefits which can be gained from coordinated networks of observations. Not only do these yield novel insights into fundamental processes, they also provide the data essential to develop and constrain numerical models of atmospheric new particle formation. The significant costs associated with coordinated European multiplatform atmospheric observational strategies return vastly more information than each of the platforms operating independently. Our work shows that major multiplatform field campaigns and a long term monitoring network are essential to address important research questions.

Methods

The monitoring sites used to collect data are presented in Table S1 and the reader is referred to ref.²⁷ for a full account and description. It is important to note that all these sites are considered as remote or rural regional monitoring sites.

Measurements of aerosol size distributions. Data were collected using either Differential Mobility Particle Sizer (DMPS) or Scanning Mobility Particle Sizer (SMPS) instruments and then harmonised into a single data set. Further information can also be found elsewhere²⁸. Although the instruments within the 24-site network of SMPS/DMPS devices used several different size ranges, all the data collected were harmonised into one large matrix by interpolating the data onto a common size bin scale; 121 size bins spanning 1 to 1000 nm with 40 channels per decade were used. In all, the reduced matrix had 54 size bins (17.8–375.8 nm) and 117 000 hourly particle size spectra (given the capture rates of the instruments across the EUSAAR network in 2008/2009). The data are extracted from the EBAS database (<http://ebas.nilu.no>), located at NILU (the Norwegian Institute for Air Research), which is a database hosting data for projects and programmes such as the European Monitoring and Evaluation Programme (EMEP), the Global Atmospheric Watch – World Data Centre for Aerosols (GAW-WDCA) and the European Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS) network^{27,28}.

Calculation of condensation sink. The condensation sink (CS) describes how rapidly condensable vapour molecules will condense on the existing aerosol. Specifically this quantity describes the loss rate of molecules with diameter d_p , diffusion coefficient D , and mean free path λ_v onto a distribution $n(d_p)$ (or N_i in the discrete case) of existing particles and as such, can be obtained from integrating over the particle size spectrum⁶². Calculations are described elsewhere⁶³.

SMPS K-Means cluster analysis. Although the instruments within the 24-site network of SMPS/DMPS devices used several different size ranges, all the data collected were harmonised into one large matrix by interpolating the data onto a common size bin scale; 121 size bins spanning 1 to 1000 nm with 40 channels per decade were used. More information can be found in ref.²⁷. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, i.e. given a set of observations $(x_1, x_2, \dots, x_n)$, where each observation is a d -dimensional real vector, k -means clustering aims to partition the n observations into k ($\leq n$) sets $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the inter-cluster sum of squares (equation 1):

$$\arg \min_s \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (1)$$

where μ_i is the mean of points in S_i ^{28,64}. The analysis works given a predefined number of clusters and an optimum needs to be decided upon. In this work, the optimum cluster number was derived using the total residual sum-of-squares (RSS Index) calculated for the clustered normalised data. This was calculated for 2 to 30 clusters and plotted on a log-log graph Figure S1a. The RRS is a measure of the discrepancy between the data and the fitted clusters. A large number of clusters (20–30) will in no doubt produced a small RSS due to very tight fit but will not identify the main clusters in the data – the data will be over fitted. Too few clusters (2–3) will indeed under fit the data and as additional clusters are added the RSS will be reduced. There will however be an optimum number of clusters between these two extreme cases where additional clusters will not produce the same reduction in RSS and this is argued to be the point where the main clusters have been fitted in the data and identified as being the point in the RSS vs cluster number point where a ‘knee’ in the data can be seen. To help identify this point, tangents are drawn on the plot which run along the curve at the low and middle points of the data and where these two meet the ‘knee’ in the data is defined. The first point was located at 5 clusters and the second point was located. To verify this finding, the Calinski and Harabasz (CH) Index was also calculated for the 2–30 clusters and this was observed to be a maximum for 5 clusters when applied to non-normalised data.

Measurement of Gaseous Pollutants During EUCAARI-ACTRIS. *Inorganic gases.* Standard continuous gas measurements were taken for SO₂ (UV fluorescence) and NO₂ (chemiluminescence) as described elsewhere^{24–26}.

Volatile organic compounds (VOC). VOC are not measured directly at the 24 EUCAARI stations. Hence, data were obtained by the Model of Emissions of Gases and Aerosols from Nature (MEGANv2.1) together with the Modern-Era Retrospective Analysis for Research and Applications (MERRA) meteorological fields. This dataset is called MEGAN-MACC and allows creation of a global emission dataset of biogenic volatile organic compounds (BVOC) available on a monthly basis for the time period of 1980–2010⁴². Four model grids closest to the EUCAARI stations were chosen, allowing a resolution of 4x (0.5 deg latitude × 0.666 deg longitude).

Measurements of particle composition. Mass concentrations and size distributions of submicron nonrefractory sulfate, nitrate, ammonium, and organics were measured with the AMS. Detailed descriptions of the Aerodyne AMS can be found elsewhere^{50,51}. The AMS has nearly 100% transmission efficiency for the considered particle size range of 70–500 nm vacuum aerodynamic diameters (D_{va}), whereas it is lower for the smallest particle sizes (~20–70 nm). No attempt was made to correct the measured size distributions for partial transmission of larger and smaller particles. While this may lead to an underestimation of the growth rate of ultrafine species, it does not affect our ability of identifying the species that are responsible for the growth (Zhang *et al.*, 2004; 2005). Although techniques for improving the determination of the NH₄⁺ size distributions for ultrafine particles exists^{29,30}, in this study the two smallest bins (D_{va} 20–38 nm and D_{va} 38–72 nm) were not considered for this chemical species, because low signal-to-noise conditions were experienced for some of the ToF-AMS instruments deployed. Similar to ammonium^{29,30}, the removal of gaseous interference from the size distributions was also investigated for organics (m/z 28 CO⁺ and m/z 44 CO₂⁺). Other m/z generally chosen in particle ToF mode contain negligible amounts of gas signals compared to the particle signals because the aerodynamic lens and skimmers of the AMS reduce the concentration of gas phase species by a factor of 10⁷ relative to aerosol species⁶⁵.

PMF Analysis of SMPS-AMS Combined Dataset. SMPS data were obtained from previous studies^{27,28}. Although the instruments within the 24-site network of SMPS/DMPS devices used several different size ranges, all the data collected were harmonised into one large matrix by interpolating the data onto a common size bin scale; 121 size bins spanning 1 to 1000 nm with 40 channels per decade were used. AMS data were used also from previous studies⁵⁴. Aerosol mass spectrometer (AMS) measurements were carried out during 26 field campaigns at 17 different sites. Only five monitoring stations were overlapping with AMS and SMPS data. Particle time of flight (PToF) AMS data were obtained for nitrate, ammonium, sulphate and organics. Five equally spaced bins were obtained, 20–38 nm, 38–72 nm, 72–137 nm, 137–262 nm, 262–500 nm. More information can be found elsewhere⁵¹.

PMF analysis was applied to the AMS-SMPS dataset, following the same approach recently described in ref.⁶⁶. Compared to cluster analysis, which groups similar data together, Positive Matrix Factorisation is used to identify the common ‘building blocks’ within the data. PMF solves the general receptor modelling problem using constrained, weighted, least-squares applied to the input data x which represent a matrix of concentrations, albeit particle or PM, measured at specific intervals during the study⁶⁷. The general model assumes there are p factors F which are interpreted as fixed emission source profiles and impact the receptor site by various amounts - represented by the scores G - during the measurement. PMF determines the profiles of these factors and calculates their contribution G such that the sum of linear combinations $G \times F$ of closely matches the measured concentration. Mathematically, the observation x_{ij} , at the receptor is represented in the matrix equation $X = G \times F + E$ whose elements are,

$$x_{ij} = \sum_{h=1}^p g_{ij} \cdot f_{hj} + e_{ij} \quad (2)$$

The measurements (AMS or NSD concentrations) are indexed by the integer j for the i^{th} time step (hour or day). The term g_{ik} is the contribution of the k^{th} factor to the receptor site on the i^{th} hour/day, f_{kj} is the fraction of the k^{th} factor (AMS or NSD concentrations) that contributes to measurement j . Matrix E , comprises of elements e_{ij} which are the residual values for the j^{th} measurement on the i^{th} hour.

In PMF, only x_{ij} are known and the goal is to estimate the contributions (g_{ik}) and the fractions (f_{ij}). It is assumed that the contributions and number fractions are all non-negative, hence the “constrained” part of the least-squares. Furthermore, PMF uses uncertainties measured for each of the x_{ij} size-bin. measurements with high uncertainty are not allowed to influence the estimation of the contribution and fractions as much as those with small uncertainty, thus giving the “weighted” part of the least squares.

Given the above, it is task of PMF to minimise the sum of the squares Q calculated using equation 3.

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{e_{ij}}{s_{ij}} \right)^2 \quad (3)$$

where s_{ij} is the uncertainty in the j^{th} measurement for hour/day i and PMF can be operated in a robust mode, meaning that “outliers” are also not allowed to influence the fitting of the contributions and profiles^{68,69}.

References

- Spracklen, D. V. *et al.* Contribution of particle formation to global cloud condensation nuclei concentrations. *Geophys. Res. Lett.* **35**, L06808, <https://doi.org/10.1029/2007GL033038> (2008).
- Merikanto, J. *et al.* Impact of nucleation on global CCN. *Atmos. Chem. Phys.* **9**, 8601–8616 (2009).
- Kulmala, M. *et al.* Chemistry of atmospheric nucleation: on the recent advances on precursor characterization and atmospheric cluster composition in connection with atmospheric new particle formation. *Annual Rev. Phys. Chem.* **65**, 21–37 (2014).
- Weber, R. J. *et al.* Measured atmospheric new particle formation rates: Implications for nucleation mechanisms. *Chem. Eng. Commun.* **151**, 53–64 (1996).
- Sipilä, M. *et al.* The role of sulfuric acid in atmospheric nucleation. *Science* **32**, 1243–1246 (2010).
- Kirkby, J. *et al.* Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation. *Nature* **476**, 429–433 (2011).
- Chen, M. *et al.* Acid–base chemical reaction model for nucleation rates in the polluted atmospheric boundary layer. *Proc. Natl Acad. Sci. USA* **109**, 18713–18718 (2012).
- Almeida, J. *et al.* Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere. *Nature* **502**, 359–363 (2013).
- Kavouras, I. G., Mihalopoulos, N. & Stephanou, E. G. Formation of atmospheric particles from organic acids produced by forests. *Nature* **395**, 683–686 (1998).
- Schobesberger, S. *et al.* Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *Proc. Nat. Acad. Sci. USA* **110**, 17223–17228 (2013).
- Riccobono, F. *et al.* Oxidation products of biogenic emissions contribute to nucleation of atmospheric particles. *Science* **344**, 717–721 (2014).
- Ehn, M. *et al.* A large source of low-volatility secondary organic aerosol. *Nature* **506**, 476–479 (2014).
- Jokinen, T. *et al.* Production of extremely low volatile organic compounds from biogenic emissions: Measured yields and atmospheric implications. *Proc. Nat. Acad. Sci. USA* **112**, 7123–7128 (2015).
- Kirkby, J. *et al.* Ion-induced nucleation of pure biogenic particles. *Nature* **533**, <https://doi.org/10.1038/nature17953> (2016).
- Tröstl, J. *et al.* The role of low-volatility organic compounds in initial particle growth in the atmosphere. *Nature* **533**, <https://doi.org/10.1038/nature18271> (2016).
- Bianchi, F. *et al.* New particle formation in the free troposphere: a question of chemistry and timing. *Science* **352**, <https://doi.org/10.1126/science.aad5456> (2016).
- Dunne, E. M. *et al.* Global particle formation from CERN CLOUD measurements. *Science*, <https://doi.org/10.1126/science.aaf264> (2016).
- Kürten, A. *et al.* Experimental particle formation rates spanning tropospheric sulfuric acid and ammonia abundances, ion production rates, and temperatures. *J. Geophys. Res. Atmos.* **121**(12), 377–12,400 (2016).
- Gordon, H. *et al.* Reduced anthropogenic aerosol radiative forcing caused by biogenic new particle formation. *Proc. Nat. Acad. Sci.* **113**, 12053–12058 (2016).
- O’Dowd, C. *et al.* Marine aerosol formation from biogenic iodine emissions. *Nature* **417**, 632–636 (2002).
- Sipilä, M. *et al.* Molecular-scale evidence of aerosolparticle formation via sequential addition of HIO₃. *Nature* **1–3**, <https://doi.org/10.1038/nature19314> (2016).
- Dawson, M. L. *et al.* Simplified mechanism for new particle formation from methanesulfonic acid, amines, and water via experiments and ab initio calculations. *Proc. Nat. Acad. Sci. USA* **109**, 18719–18724 (2012).
- Riipinen, I. *et al.* The contribution of organics to atmospheric nanoparticle growth. *Nature Geosci.* **5**, 453–458 (2012).
- Kulmala, M. *et al.* Introduction: European Integrated Project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI) – integrating aerosol research from nano to global scales. *Atmos. Chem. Phys.* **9**, 2825–2841 (2009).
- Tørseth, K. *et al.* Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972–2009. *Atmos. Chem. Phys.* **12**, 5447–5481 (2012).
- Henne, S. *et al.* Assessment of parameters describing representativeness of air quality *in-situ* measurement sites. *Atmos. Chem. Phys.* **10**, 3561–3581 (2010).
- Asmi, A. *et al.* Number size distributions and seasonality of submicron particles in Europe 2008–2009. *Atmos. Chem. Phys.* **11**, 5505–5538, <https://doi.org/10.5194/acp-11-5505-2011> (2011).
- Beddows, D. C. S. *et al.* Variations in tropospheric submicron particle size distributions across the European continent 2008–2009. *Atmos. Chem. Phys.* **14**, 4327–4348 (2014).
- Zhang, Q. *et al.* Insights into nucleation burst and particle growth in Pittsburgh based on aerosol mass spectrometry. *Environ. Sci. Technol.* **38**, 4797–4809 (2004).
- Zhang, Q. *et al.* Time- and size-resolved chemical composition of submicron particles in Pittsburgh: Implications for aerosol sources and processes. *J. Geophys. Res.* **110**, D07S09, <https://doi.org/10.1029/2004JD004649> (2005).
- Manninen, H. E. *et al.* EUCAARI ion spectrometer measurements at 12 European sites – analysis of new particle formation events. *Atmos. Chem. Phys.* **10**, 7907–7927 (2010).
- Kulmala, M. & Kerminen, V.-M. On the growth of atmospheric nanoparticles. *Atmos. Res.* **90**, 132–150 (2008).
- Kerminen, V.-M. & Kulmala, M. Analytical formulae connecting the “real” and the “apparent” nucleation rate and the nuclei number concentration for atmospheric nucleation events. *J. Aerosol Sci.* **33**, 609–622 (2002).
- Hamed, A. *et al.* The role of relative humidity in continental new particle formation. *J. Geophys. Res.* **116**, D03202, <https://doi.org/10.1029/2010JD014186> (2011).
- Pirjola, L. *et al.* Formation of sulphuric acid aerosols and cloud condensation nuclei: An expression for significant nucleation and model comparison. *J. Aerosol Sci.* **30**, 1079–1094 (1999).

36. Wexler, A. S., Lurmann, F. W. & Seinfeld, J. H. Modelling Urban and Regional Aerosols—I. Model Development. *Atmos. Environ.* **28**, 531–546 (1994).
37. Stanier, C. O., Khlystov, A. Y. & Pandis, S. N. Nucleation events during the Pittsburgh air quality study: Description and relation to key meteorological, gas phase and aerosol parameters. *Aerosol Sci. Tech.* **38**, 253–264 (2004).
38. Crippa, M. *et al.* Forty years of improvements in European air quality: regional policy-industry interactions with global impacts. *Atmos. Chem. Phys.* **16**, 3825–3841 (2016).
39. Eeftens, M. *et al.* Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM_{coarse} in 20 European study areas; results of the ESCAPE project. *Environ. Sci. Technol.* **46**, 11195–11205 (2012).
40. EAA. Air pollution from agriculture: EU Exceeds International Limit in 2014. ISSN 1977–8449 (2016).
41. Guenther, A. B. *et al.* The Model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **5**, 1471–1492 (2012).
42. Sindelarova, K. *et al.* Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmos. Chem. Phys.* **14**, 9317–9341 (2014).
43. Steinbrecher, R. *et al.* Intra- and inter-annual variability of VOC emissions from natural and seminatural vegetation in Europe and neighbouring countries. *Atmos. Environ.* **43**, 1380–1391 (2009).
44. Oderbolz, D. C. *et al.* A comprehensive emission inventory of biogenic volatile organic compounds in Europe: improved seasonality and land-cover. *Atmos. Chem. Phys.* **13**, 1689–1712 (2013).
45. Koohkan, M. R., Bocquet, M., Roustan, Y., Kim, Y. & Seigneur, C. Estimation of volatile organic compound emissions for Europe using data assimilation. *Atmos. Chem. Phys.* **13**, 5887–5905 (2013).
46. Junninen, H. *et al.* A high-resolution mass spectrometer to measure atmospheric ion composition. *Atmos. Meas. Tech.* **3**, 1039–1053 (2010).
47. Kurten, T. *et al.* The effect of H₂SO₄-amine clustering on chemical ionization mass spectrometry (CIMS) measurements of gas-phase sulfuric acid. *Atmos. Chem. Phys.* **11**, 3007–3019 (2011).
48. Jokinen, T. *et al.* Atmospheric sulphuric acid and neutral cluster measurements using CI-API-TOF. *Atmos. Chem. Phys.* **12**, 4117–4125 (2012).
49. Jayne, J. T. *et al.* Development of an aerosol mass spectrometer for size and composition analysis of submicron particles. *Aerosol Sci. Technol.* **33**, 49–70 (2000).
50. Drewnick, F. *et al.* A new Time-of-Flight Aerosol Mass Spectrometer (ToF-AMS): Instrument Description and First Field Deployment. *Aerosol Sci. Technol.* **39**, 637–658 (2005).
51. De Carlo, P. *et al.* Field-Deployable, High-Resolution, Time-of-Flight Aerosol Mass Spectrometer. *Anal. Chem.* **78**, 8281–8289 (2006).
52. Jimenez, J. L. *et al.* New particle formation from photooxidation of diiodomethane. *J. Geophys. Res.* **108**(D10), 4318 (2003).
53. Allan, J. D. *et al.* Size and composition measurements of background aerosol and new particle growth in a Finnish forest during QUEST 2 using an aerodyne aerosol mass spectrometer. *Atmos. Chem. Phys.* **6**, 315–327 (2006).
54. Crippa, M. *et al.* Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach. *Atmos. Chem. Phys.* **14**, 6159–6176, <https://doi.org/10.5194/acp-14-6159-2014> (2014).
55. Kulmala, M. *et al.* Direct observations of atmospheric aerosol nucleation. *Science* **339**, 943–946 (2013).
56. Riipinen, I. *et al.* Organic condensation: a vital link connecting aerosol formation to cloud condensation nuclei (CCN) concentrations. *Atmos. Chem. Phys.* **11**, 3865–3878 (2011).
57. Paasonen, P. *et al.* Warming-induced increase in aerosol number concentration likely to moderate climate change. *Nature Geosci.* **6**, 438–442 (2013).
58. Mentel, T. F. *et al.* Photochemical production of aerosols from real plant emissions. *Atmos. Chem. Phys.* **9**, 4387–4406 (2009).
59. Laaksonen, A. *et al.* The role of VOC oxidation products in continental new particle formation. *Atmos. Chem. Phys.* **8**, 2657–2665 (2008).
60. Pieber, S. M. *et al.* Inorganic Salt Interference on CO₂ in Aerodyne AMS and ACSM Organic Aerosol Composition Studies. *Environ. Sci. Technol.* **50**, 10494–10503 (2016).
61. Makela, J. M. *et al.* Chemical composition of aerosol during particle formation events in boreal forest. *Tellus* **53B**, 380–393 (2001).
62. Dal Maso, M. *et al.* Condensation and coagulation sinks and formation of nucleation mode particles in coastal and boreal forest boundary layers. *J. Geophys. Res.* **107**(D19), 8097, <https://doi.org/10.1029/2001JD001053> (2002).
63. Dall'Osto, M. *et al.* On the spatial distribution and evolution of ultrafine particles in Barcelona. *Atmos. Chem. Phys.* **13**, 741–59 (2013).
64. Beddows, D. C. S., Dall'Osto, M. & Harrison, R. M. Cluster Analysis of Rural, Urban, and Curbside Atmospheric Particle Size Data. *Environ. Sci. Technol.* **43**, 4694–4700 (2009).
65. Allan, J. D. *et al.* A generalised method for the extraction of chemically resolved mass spectra from Aerodyne aerosol mass spectrometer data. *J. Aerosol Sci.* **35**, 909–922 (2004).
66. Beddows, D. C. S., Harrison, R. M., Green, D. C. & Fuller, G. W. Receptor modelling of both particle composition and size distribution from a background site in London, UK. *Atmos. Chem. Phys.* **15**, 10107–10125 (2015).
67. Refr, A., Eberly, S. I. & Bhav, P. V. Receptor modeling of ambient particulate matter data using positive matrix factorization: Review of existing methods. *JAWMA* **57**, 146–154 (2007).
68. Ogulei, D., Hopke, P. K. & Wallace, L. A. Analysis of indoor particle size distributions in an occupied townhouse using positive matrix factorization. *Indoor Air* **16**, 204–215 (2006).
69. Ogulei, D. *et al.* Source apportionment of Baltimore aerosol from combined size distribution and chemical composition data. *Atmos. Environ.* **40**, S396–S410 (2006).

Acknowledgements

We gratefully acknowledge support by the German Federal Environment Agency (Umweltbundesamt), grants no. 351 03 031 and 351 01 038, and UFOPLAN grant 3703 43 200 (title “Erfassung der Zahl feiner und ultrafeiner Partikel in der Außenluft”). We acknowledge support by the European Integrated Project on Aerosol Cloud Climate and Air Quality Interaction (EUCAARI). We further appreciate extensive instrumental support by Joachim Grüner and Thomas Tuch at the TROPOS Melpitz field site. The study was funded by the Spanish Ministry of Economy through projects BIO-NUC (CGL2013-49020-R), and by the EU through the FP7-PEOPLE-2013-IOF programme (Project number 624680, MANU – Marine Aerosol NUCleations) and the Ramon y Cajal fellowship (RYC-2012-11922). The National Centre for Atmospheric Science NCAS Birmingham group is funded by the UK Natural Environment Research Council. Thanks to the British Atmospheric Data Centre, which is part of the NERC National Centre for Atmospheric Science (NCAS).

Author Contributions

M.D., D.C.S.B. and R.M.H. designed the study and equally contributed to the study. A.A., L.P., L.H., E.F., J.D.A., M.C.A., G.de L., A.E., H.C.H., J.S.H., P.L., A.P., J.P.P., K.S., E.S., A.V. collected the data and contributed to the analysis. K.Z. and C.G. provided VOC emissions with the MEGAN model. T.O., M.C., P.L., A.P., C.O.D., M.K. provided useful discussions. M.D., D.C.S.B., A.A., L.P., L.H., E.F., J.D.A., M.C.R., M.C.A., F.B., G.L., A.E., H.C.H., J.S.H., H.H., C.G., P.L., N.M., T.O., A.P., J.P.P., K.S., E.S., M.V., A.V., K.Z., R.S., A.W., D.W., C.O.D., M.K., R.M.H. commented on the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-17343-9>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018