Supplementary Materials for:

Is reducing new particle formation a plausible solution to mitigate particulate air pollution in Beijing and other Chinese megacities?

1 Materials and Methods

1.1 Measurement site location

The measurement site is at the campus building of the Beijing University of Chemical Technology (BUCT) in the western part of Beijing (39°56'31"N,116°17'50"E). BUCT is located near the 3rd Ring Road of Beijing; and therefore, close to fresh traffic emissions and also surrounded by residential areas. However, the distance to the nearest forest area in the northwest is only around 10 km. The city itself lies in the western part of the North China Plain with mountains to the north, northwest and west. The location of the site on the Beijing map is shown in figure S1. For more details on the measurement location see Zhou et al. [1].

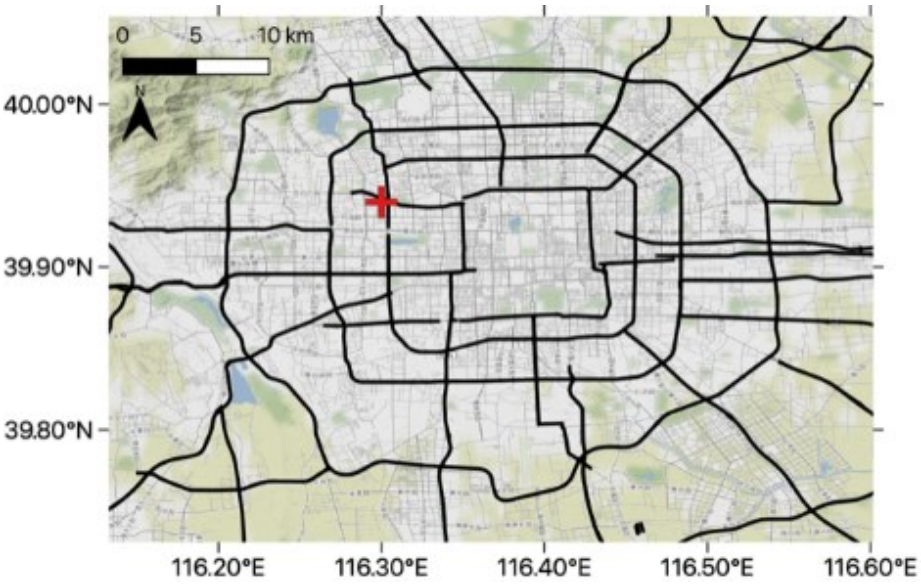


Figure S1 Map of Beijing City. The measurement site at the west campus of the Beijing University of Chemical Technology is marked with the red cross.

1.2 Measurement of particle number size distribution

The particle number size distribution of 1 nm–10 μm was measured by with a homemade Diethylene Glycol Scanning Mobility Particle Sizer (DEG-SMPS[2]) equipped with core sampling [3] and a homemade Particle Size Distribution (PSD) system[4]. In parallel, the number concentration of particles larger than 1.2 nm and the size distribution of clusters and sub-2.5 nm particles was measured with a Particle Sizer Magnifier (PSM; Airmodus A11[5]), with a total sample flow rate of 5.5 liters per minute, of which a sample flow rate of 2.5 liters per minute is extracted to the A11 with a core sampling system[3]. Additionally, ions in the mobility diameter range of 0.8–42 nm were measured with the ion mode of

Neutral cluster and Air Ion Spectrometer (NAIS; Airel Ltd)[6]. The particle number size distribution in the diameter range 2.5–42 nm was measured with the total mode of NAIS.

1.3 Meteorological variables

Meteorological variables were measured by Vaisala automatic weather station AWS310 at the BUCT site. They included: ambient air temperature T (K), relative humidity RH (%), pressure p (hPa), visibility (m), and horizontal wind speed U (m s^{-1}) and direction WD (∘). Additionally, the atmospheric boundary layer height BLH (m) was retrieved from ceilometer measurements (CL-51; Vaisala) of the optical backscattering by applying a three-step idealized-profile as proposed by Eresmaa, et al. [7].

1.4 Trace gases

Trace gases including carbon monoxide (CO), sulfur dioxide (SO2), nitrogen oxides (NOx) and ozone (O3) concentrations were sampled through a common inlet through the roof of the building and measured using Thermo Environmental Instruments models 48i, 43i-TLE, 42i, and 49i, respectively. The length of the sampling tube was around 3 m and the time resolution was 5 min.

1.5 Composition of PM2.5

Online Time-of-Flight Aerosol Chemical Speciation Monitor (ToF-ACSM[8], equipped with a standard vaporizer) observations were conducted at the BUCT station from February 21 to April 7, 2018, equipped with a PM2.5 lens. The mass concentrations of non-refractory fine particles (NR-PM2.5, including organics, sulfate, nitrate, ammonium, and chloride) were obtained by the ToF-ACSM standard data analysis software (Tofware) within Igor Pro (Wavemetrics). The ToF-ACSM was regularly calibrated (ionization efficiency) and relative ionization efficiency (RIE) of NH_4 (3.76) and SO_4 (0.88) were determined experimentally by using nebulizing aqueous solutions of pure NH_4NO_3 and pure $(NH_4)_2SO_4$ into the ToF-ACSM. Default relative ionization efficiency (RIE) values were used for organics (1.4), nitrate (1.1), and chloride (1.3). To account for the incomplete detection of aerosol, a composition dependent collection efficiency (CDCE)[9] was computed and resulted in more than 70% of the data points in a CDCE of 0.45. Since the agreement with PM2.5 analyses from surrounding stations was not improved by applying the CDCE, we applied a CE of 0.45 to the entire dataset.

Black Carbon (BC) mass concentration and Absorption Coefficient were measured by a dual spot Aethalometer[10,11] at BUCT, which aims to explicitly correct for filter loading effects using two simultaneous absorption measurements (from two separate spots exposed to different flow rates and thus different filter loadings). However, this state-of-the-art correction algorithm did not provide satisfactory results (related to the use of the discontinued filter tape type 8050). In fact, whenever the filter tape, on which PM was collected for the measurement, advanced, a jump in the black carbon (BC) mass concentration occurred. This indicates that even after applying this type of correction, the filter loading induced an artefact on the BC concentration. Therefore, we also computed BC using the approach for the single-spot aethalometer (AE31) presented in Virkkula, et al. [12] and used the average of both approaches as a best estimate of BC. This correction uses the same basic equations as Drinovec, et al. [11]. While the underlying principle of the correction for the dual-spot Aethalometer to correct for the filter-loading artifact uses the difference of two simultaneously collected filter spots with different loadings, the single-

spot correction relies on the measured difference right before and after the filter tape advancement[11,12]. We use the average BC from both approaches as best estimate.

2 General characteristics of NPF and haze days

2.1 New particle formation (NPF) events classification

We classified new particle formation (NPF) events following the method introduced by Dal Maso, et al. [13]. All the data (January 15, 2018 – March 31, 2019) measured with NAIS and PSD system were analysed day by day; i.e. each 24 hours starting at 00:00 and ending by 24:00. We identified 189 NPF event days, which were ~40% of all measurement days. The classification is documented in Figures S2-S4.

2.2 Haze events classification

We performed a classification for haze occurrence based on the visibility and relative humidity (RH) following the standards provided by the Chinese Meteorology Bureau. Haze events were identified when the visibility was gradually decreasing from 20 km (maximum) and reached less than 10 km, and was accompanied with RH lower than 85%, which is a condition considered for haze not to be mixed with fog.

The visibility sensor was operated starting at the beginning of February. We were able to identify 158 haze days during February 1, 2018 – March 31, 2019 (Figure S2-S4). NPF and haze can occur on the same day, but the two phenomena did not occur simultaneously. An NPF and haze event calendar is displayed in Figure S3, showing that all the haze events were preceded by an NPF event. However, not all NPF events translated into haze events.

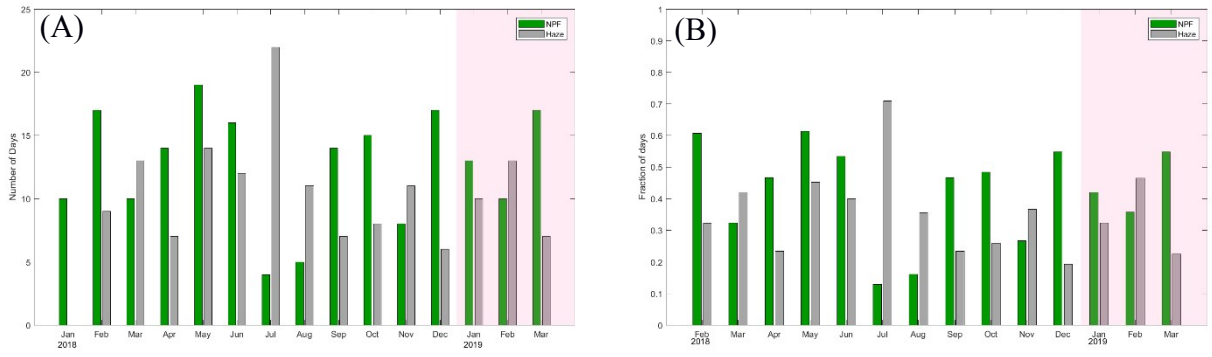


Figure S2 (A) Number of NPF events and haze days at the BUCT site between January 12th, 2018 and March 31st, 2019. (B) Fraction of NPF events and haze days at the BUCT site between February 1st, 2018 and March 31st, 2019.

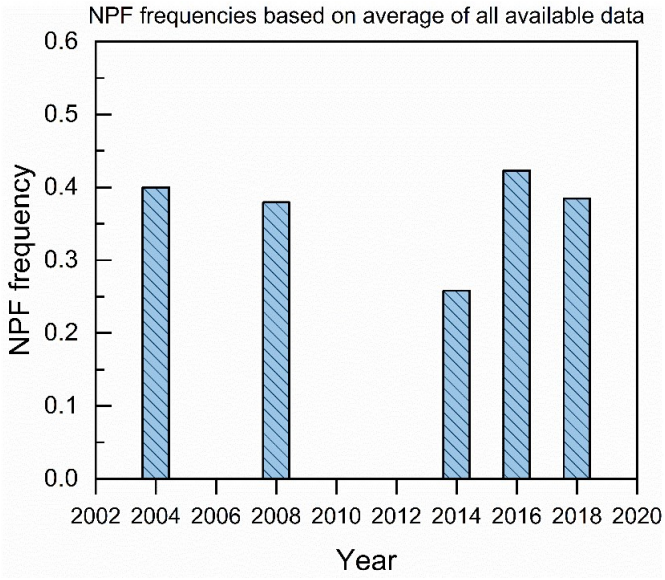


Figure S3 NPF frequency (fraction of days having NPF events) in Beijing on different years. The bars represent averages of all available data, including our own measurements and those reported in the literature[64,65]. Years with no bar did not have reported data. Our results show a constant NPF frequency throughout the years when we have measurements regardless of the decrease in SO2 concentration which follows from the regulations implemented in Beijing.

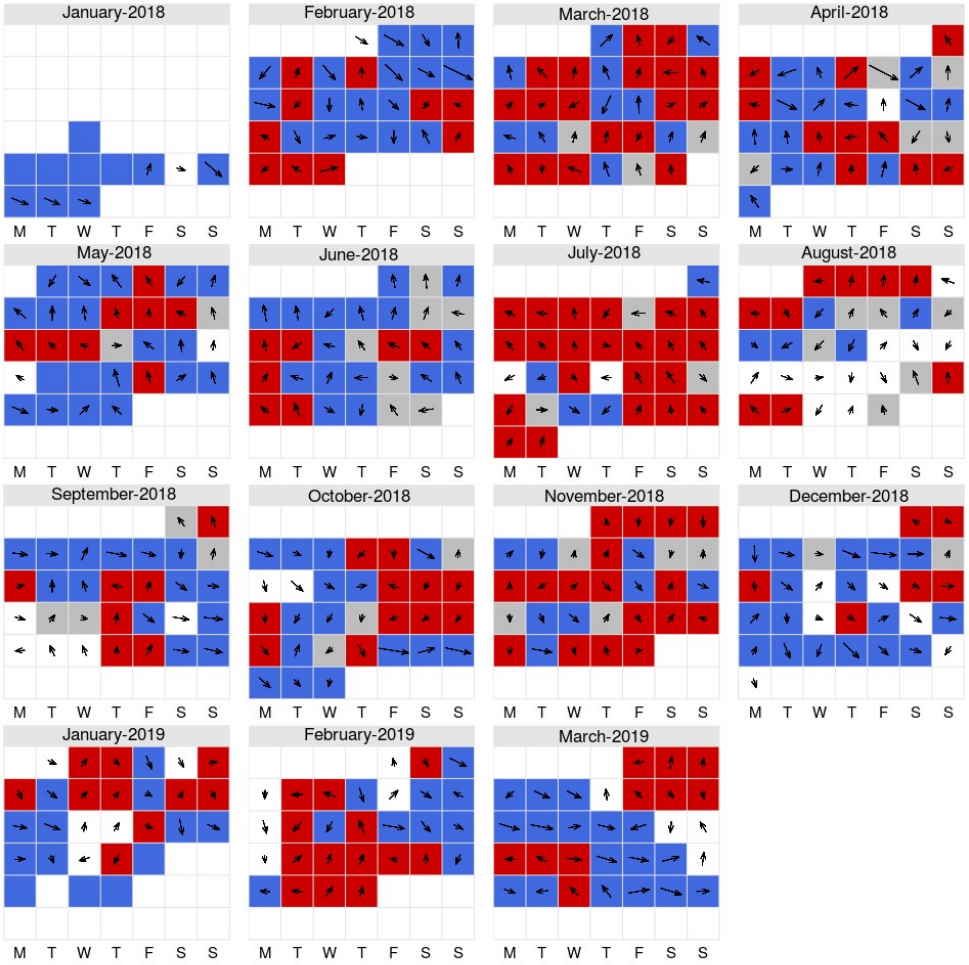


Figure S4 Calendar for haze event days (red), NPF event days (blue), wind speed (length of arrow) and direction between January 12, 2018 –March 31, 2019 (meteorological data from January 26). White slots indicate no data and grey ones are days with neither NPF nor haze event. Sometimes it takes more than one day to develop haze after NPF. If NPF and haze are observed on the same day, the day is classified as NPF. Usually haze is observed in the morning before NPF starts, or develops in the evening after the particles from NPF have grown past 20 nm. A day is counted as haze day in case haze extends for at least 8 hours.

2.3 Case studies of NPF to haze

An example of an NPF day which is subsequently followed by haze within the next day is shown in Fig. S5. After the NPF event happens, we observe a gradual decrease in visibility reaching below 10 km, accompanied by an increase of PM2.5 reaching more than 100 µg/m^3. An example of two days on which an NPF event was interrupted and was not followed by haze event is shown in Figure S6. Here the particles originating from NPF cannot be distinguished from local traffic emissions.

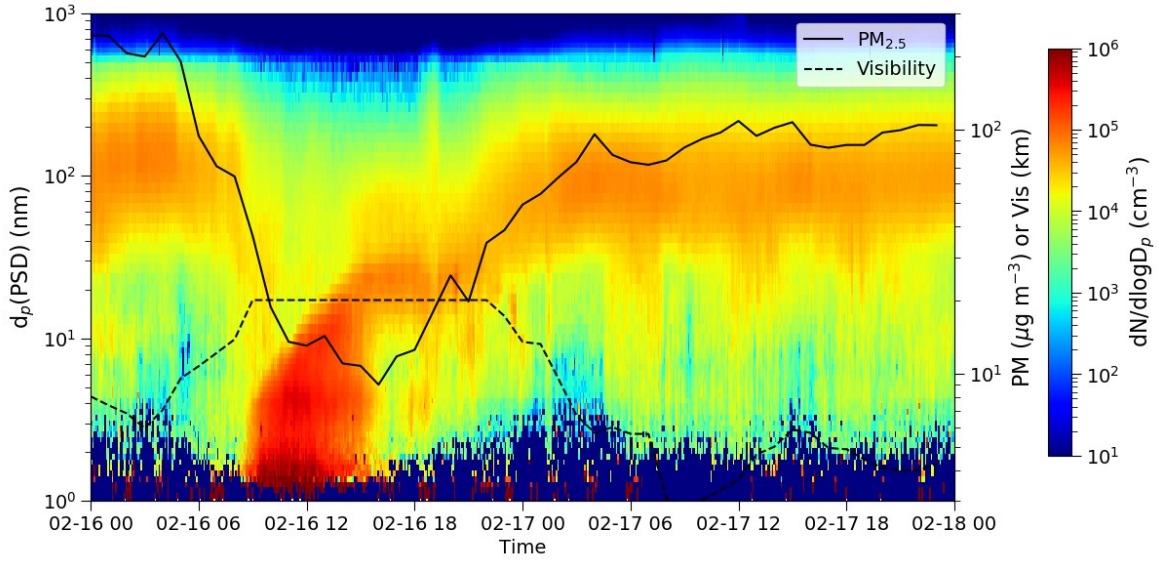


Figure S5 Particle size distribution on a 2-day period starting with an NPF day that is followed by haze formation (size distributions measured with the PSD). Dashed black line is visibility (km) on the right axis. Black solid line is PM2.5 (µg/m3) on the right axis.

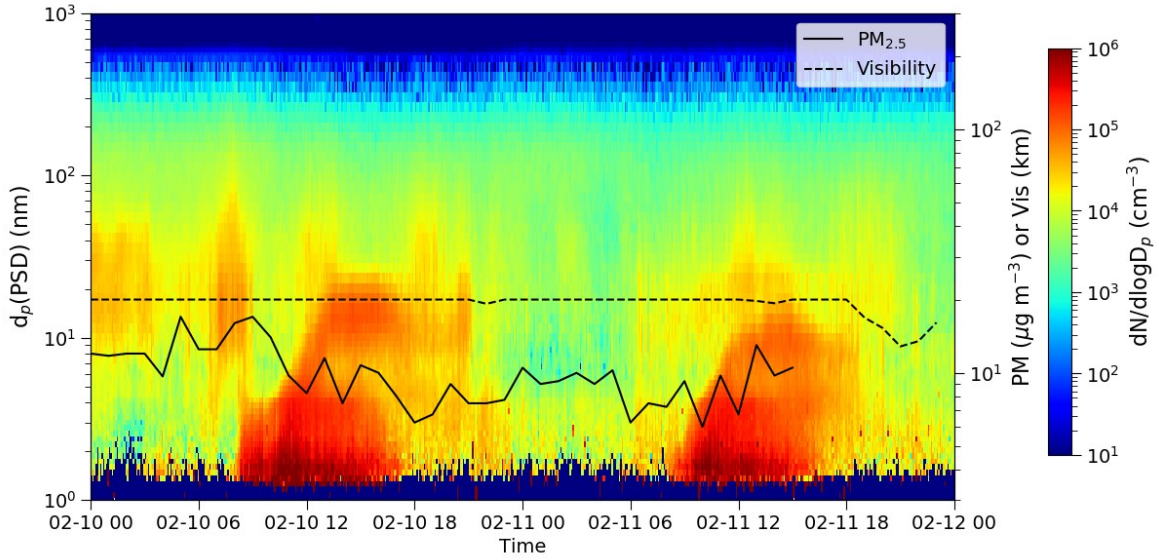


Figure S6 Particle size distribution for two example days when NPF does not grow further (size distribution measured with the PSD). Dashed black line is visibility (m) on the right axis. Solid black line is $PM_{2.5}$, both on the right axis.

2.4 Air mass origin

Air mass history was studied by calculating particle retroplumes using a Lagrangian particle dispersion model FLEXPART (FLEXible PARTicle dispersion model) version 9.02[14]. ECMWF (European Centre for Medium-Range Weather Forecast) operational forecast with 0.15° horizontal and 1 h temporal resolution was used as the meteorological input into the model. The number of height levels in the meteorological data is 137.

The model was run for a time period from January 15 2018 to April 11 2018. During this period, a new release of 50 000 model particles, distributed evenly between 0–100 m above the measurement site, occurred every 1 hour. The released particles were traced backwards in time for 72 h, unless they exceeded the model grid (20–60°N, 95–135°E, resolution: 0.05°). The model time step was 10 min, but in the calculation of turbulent wind components, the use of a shorter time step, determined internally by the model, was allowed. The parametrization of moist convection was also set on to improve the quality of the model run. Particle wet and dry deposition were not considered, since in this work, the model is only used to study the movements of the air masses.

In Figure S7 we show examples of air mass trajectories during NPF events (NPF + 0h) and observe how the trajectories change in time as we progress towards haze formation. From this figure, we can see that NPF events typically take place in clean air masses that arrive from the sparsely populated areas located north of Beijing. This is most likely because the condensation sink is low during these periods. As the development towards haze begins, we start to see a southern loop, over more polluted areas, in the trajectories. This loop allows for continuous particle growth in the area of high precursor vapor emissions, which seems essential for haze formation (Figure S4). In the case where no such loop develops, we often observe interrupted NPF events (Figure S5), where the size of the observed particles in Beijing is limited by the time they have spent in the polluted region before their arrival. It is important to note that if we were to observe these same particles at a measurement station located further south, we would likely see their growth to larger sizes, but since Beijing is located so close to the clean regions in the north, the travel time in the polluted region can easily become a limiting factor.

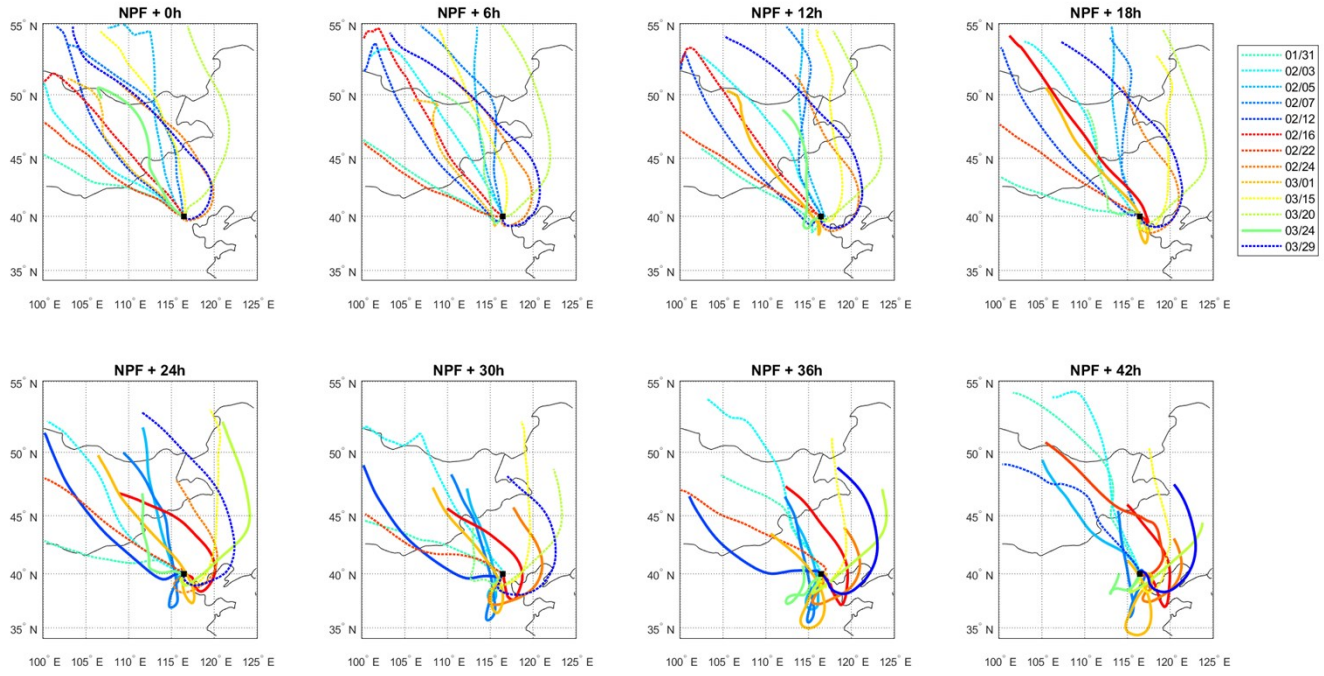


Figure S7 60-h Back trajectories for 13 selected cases where a NPF event is followed by haze (Table S2). The text above the panels indicates the trajectory release time in time after NPF on the days shown in the legend. The trajectory lines turn from dotted to solid if the trajectory release occurs during haze (i.e. hourly median visibility

has been <=10km.) In most of the cases, the haze development seems to be related to a southern loop over more polluted areas. However, even in the last panel (NPF+42 h), the air mass source regions (trajectory end points) at NPF - 18 h are clearly pointing towards the clean regions in the north. Therefore, it is likely that NPF has also occurred in these air masses.

2.5 Growth rates, formation rates and condensation sink

Growth rates at different particle diameters were calculated at the onset of the NPF events by using the 50% appearance time method[15,16] observed in the NAIS (in ion positive and negative mode). The growth rates for Beijing are shown in Table S2. Additionally, we separately applied the new mode fitting/peak finding method[17] to our BUCT PSD data. Figure S8 shows that we were able to follow the growing mode of particles up to accumulation mode during 4 days where after an NPF event followed by haze. Our results show that at sizes larger than 25 nm, the growth rate is independent of particle size.

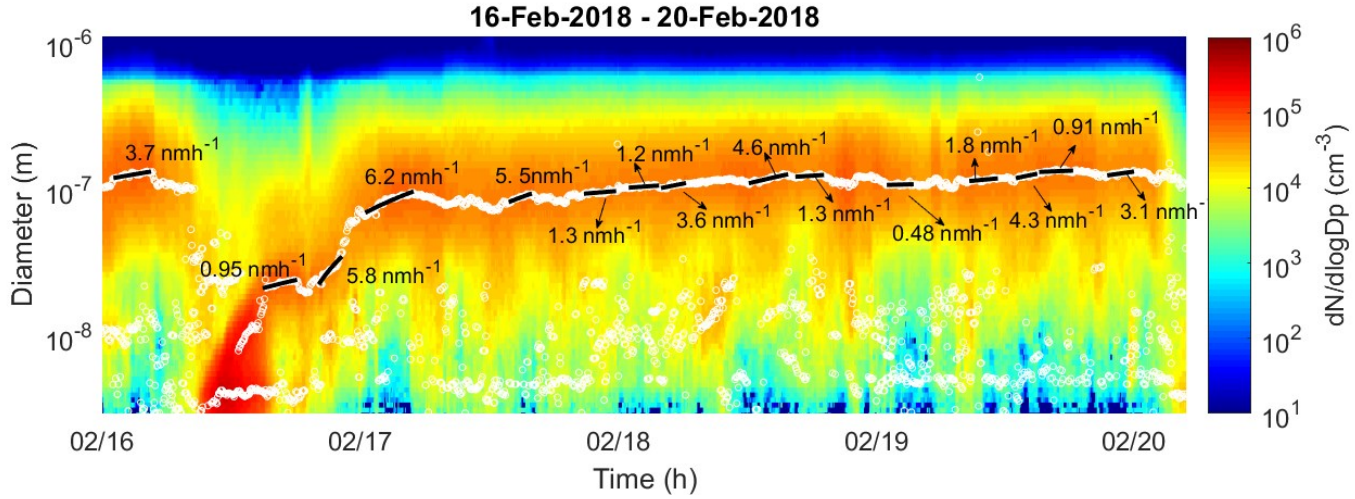


Figure S8 Growth rate by mode fitting/peak finding on Feb 16- 20, 2018. White circles are the mode maxima found by the automatic code while black lines indicate monotonous growth period.

Formation rates were calculated by using the balance equation proposed by Kulmala et al.[15] and more details on the calculation of formation rates can be found therein.

The formation rate of 1.5 nm particles ($J_{1.5}$) was calculated by using PSM data in the size range of 1.5 to 3 nm. J_3, J_5 and J_{10} were calculated using NAIS data (total mode - negative polarity) of size ranges 3 -7 nm, 5 - 9 nm, and 10 -14 nm; respectively. During events when the GR cannot be calculated, a median growth rate of all the events in the same month was used to calculate the formation rate. Daily median formation rates during peak times of NPF events (11:00–12:00) are presented in Table S1.

The condensation sink (CS, s^{-1}), which is the rate at which vapors are lost to pre-existing particles, is calculated following the method described in Kulmala, et al. [15] and the values during NPF days are also give in Table S1. NPF events are characterized by a median CS being up to one order of magnitude lower in comparison to haze days.

Table S1 Characteristic growth rates (± error on the fit), formation rates and condensation sink during peak NPF events in Beijing. Dates marked in green are chosen for further analyses.

	Growth Rates (nm h^{-1})			**Formation Rates (cm^{-3} s^{-1})**				**CS (s^{-1})**
Size range	***< 3 nm***	***3 - 7 nm***	***7 - 25 nm***	***J1.5***	***J3***	***J5***	***J10***	***CS***
17/01/2018	1.3±0.4	3.4±1.2	9.1±1.6	44	26	15	4	0.05
22/01/2018	1.4±0.1	3.3±0.2	11.8±5.6	138	89	42	14	0.01
23/01/2018	0.7±0.04	2.6±0.1	10.3±4.6	95	56	18	7	0.01
24/01/2018				100	41	8	1	0.01
25/01/2018				127	34	12	5	0.02
28/01/2018	1.5±0.2	2.3±0.4	5.8±3.8	123	67	21	2	0.01
29/01/2018	1.3±0.2	2.7±0.1	8.9±2.2	141	88	35	7	0.01
30/01/2018	1.3±0.1	1.3±0.2	4.5±3.4	182	72	26	2	0.02
31/01/2018	2.2±0.1	2.0±0.3	5.7±2.4	214	87	37	5	0.02
02/02/2018	0.7±0.1	4.6±0.3	4.7±1.2		73	21	2	0.01
03/02/2018				224	73	24	4	0.01
04/02/2018	0.3±0.04	3.0±0.1	5.1±0.9	1356	257	47	7	0.06
05/02/2018	0.7±0.1	2.8±0.3	5.7±1.1	258	122	36	5	0.01
07/02/2018				255	113	34	6	0.01
09/02/2018	0.6±0.1	8.1±0.1	5.4±1.0	75	49	29	6	0.01
10/02/2018	0.7±0.3	2.5±0.3	6.7±1.8	125	82	28	6	0.01
11/02/2018	1.0±0.1	2.6±0.2	6.3±0.6	98	58	16	3	0.01
12/02/2018	1.5±0.1	2.7±0.3	5.4±1.8	84	71	28	7	0.01
14/02/2018	1.2±0.1	1.8±0.2	4.8±0.7	37	28	15	8	0.02
15/02/2018				149	27	9	2	0.02
16/02/2018	0.8±0.3	4.4±0.5	4.8±1.2	161	114	52	10	0.01
20/02/2018	3.4±0.1	2.6±0.1	4.6±1.0	190	116	48	13	0.01
21/02/2018	1.2±0.1		5.0±1.2	28	18	10	2	0.04
22/02/2018	0.6±0.05		4.9±1.6	56	36	18	5	0.02
24/02/2018	0.7±0	2.9±0.6	6.1±0.6	147	108	31	6	0.02
01/03/2018	0.5±0.2	2.7±0.1	5.5±0.7	53	67	36	8	0.01
08/03/2018				9	5	4	2	0.03
15/03/2018	3.0±0.9	10.0±0.3	3.2±0.9		61	27	6	0.01
19/03/2018		3.5±1.3	7.5±2.7		122	37	10	0.04
20/03/2018	0.6	6.1±0.2	12.8±1.5	15	36	14	5	0.03
24/03/2018		3.0±1.3	3.1±2.1	22	23	10	6	0.08
29/03/2018				18	111	34	7	0.05

3 Contribution of NPF to Accumulation mode haze particles

3.1 NOx Tracer method (Number concentration)

We explored the relative contribution of primary and secondary sources to the accumulation mode particles by estimating the number concentration of primary particles using NOx as the main tracer of traffic and other primary emissions. Detailed description of the method and its limitations is presented by Kulmala, et al. [18] and Rodríguez and Cuevas [19]. The number concentration of primary particles is estimated using the relation:

$$N_1 = S_1 \times \mathrm{NOx}, \quad (3)$$

where N_1 is the estimated number of primary accumulation mode particles emitted by traffic and other sources, while S_1 is amount of primary accumulation mode particles emitted per ppb of NOx. S_1 is derived semi-empirically from simultaneous observations of the accumulation mode particles and NOx. We computed the ratio of the total accumulation mode number concentration to NOx (N_{tot}(acc)/NOx) for all measurements and used the 0.1^{th}, 0.5^{th}, 1^{th}, 1.5^{th}, 2^{th} and 5^{th} percentiles as S_1 at different thresholds as shown in Figure S9. We estimate the fraction of secondary particles in Fig. S10. The latter is calculated as follows:

$$N_2\,(\mathrm{acc}) = N_{tot}(\mathrm{acc}) - N_1(\mathrm{acc}) \quad (4)$$

Here N_2 is the number of accumulation mode particles from NPF, N_{tot} is the total measured concentration of particles in the accumulation mode and N_1 is the number of accumulation mode particles from traffic, calculated as described above. Our results in Fig. S10 (see also Fig. 2A) show that the primary fraction of accumulation mode particles is less than 20% at higher PM2.5 levels, corresponding to a secondary fraction larger than 80%.

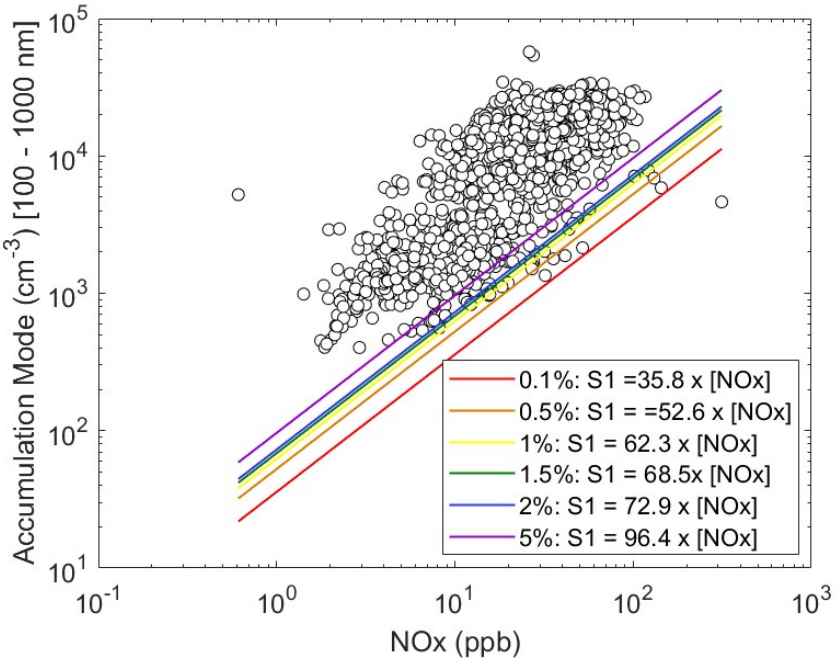


Figure S9 Accumulation mode number concentration at the Beijing station as a function of NOx concentration illustrates the way, how the value of S1 is determined following the method by Kulmala et al.[28]. In the constrained fittings, 0.1% to 5% of the data points are located below the lines determining the values of S1. This plot includes data from January 22 until March 18 2018.

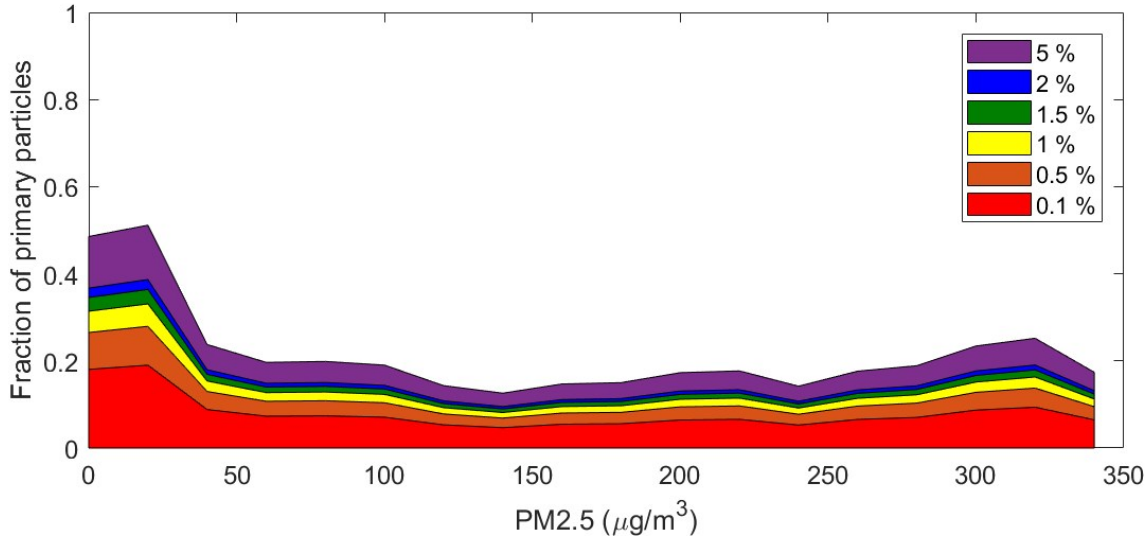


Figure S10 Constrained fits from NOx method at different sensitivity values showing the fraction of primary (colored area) and secondary particles within the accumulation mode. This plot includes data from January 22 until March 18 2018.

3.2 Secondary mass fraction from source apportionment of organic aerosol

The organic mass spectra from the ToF-ACSM were analyzed by positive matrix factorization (PMF)[20] to identify and quantify the potential sources of organic aerosol (OA). We solved the PMF by the multi-linear engine (ME-2) algorithm implemented within the toolkit SoFi, Source Finder[20-23]. The PMF is a statistical unmixing model that explains the variability in of a given dataset ($x_{i,j,}$ i being a point in time and j being one single m/z) with linear combinations of p constant factor profiles ($f_{k,j,}$, chemical composition of OA source/category, k being a factor), and their concentration varying in time ($g_{i,k,}$) and a residual term ($e_{i,j,}$):

$$x_{i,j} = \sum_{k=1}^{p} g_{i,k} f_{k,j} + e_{i,j} \tag{1}$$

Using ME-2, it is possible to use a priori information such as the chemical composition of specific OA sources by constraining their spectral fingerprint within a range defined by the scalar a (a-value, $0 \leq a \leq 1$).

$$f_{k,j} = f'_{k,j} \pm a * f'_{k,j} \tag{2}$$

Exploratory unconstrained PMF runs separated primary OA (POA) from traffic (HOA), and cooking (COA), a component representing residential heating (mixture of biomass burning, BBOA, and coal combustion, CCOA), and a component representing secondary OA (SOA). However, the unconstrained PMF runs exhibited mixing between HOA and COA, on one hand, and residential heating and OOA, on the other hand. Thus, the mass spectral fingerprints of HOA[24], COA[24] and CCOA[25] were constrained. When using five factors, a fraction of the solutions separated BBOA mixed with other components. Therefore, as a final step, we constrained HOA, COA, CCOA, and BBOA[25].

Q/Q_{exp} is used as a metric for assessing the mathematical quality of the PMF solution $\left(Q/Q_{exp} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} \frac{e_{i,j}}{\sigma_{i,j}}}{n * m - p * (m+n)} \right)$ with $\sigma_{i,j}$ being the measurement uncertainty, m the number of time points and n the number of m/z-values[20]. While Q/Q_{exp} decreased when increasing to five or more factors, the 5-factor solution explains the entire dataset well (Figures S11). Further, allowing for additional factors (more than five factors), leads to separations of SOA that could not be related to environmental processes. Therefore, we opted for a 5-factor PMF solution. To account for variability of the composition of aerosol components, we performed sensitivity tests by running PMF 4500 times using a-values randomly picked for each factor independently in a range between 0.0 and 1.0 (increments 0.2).

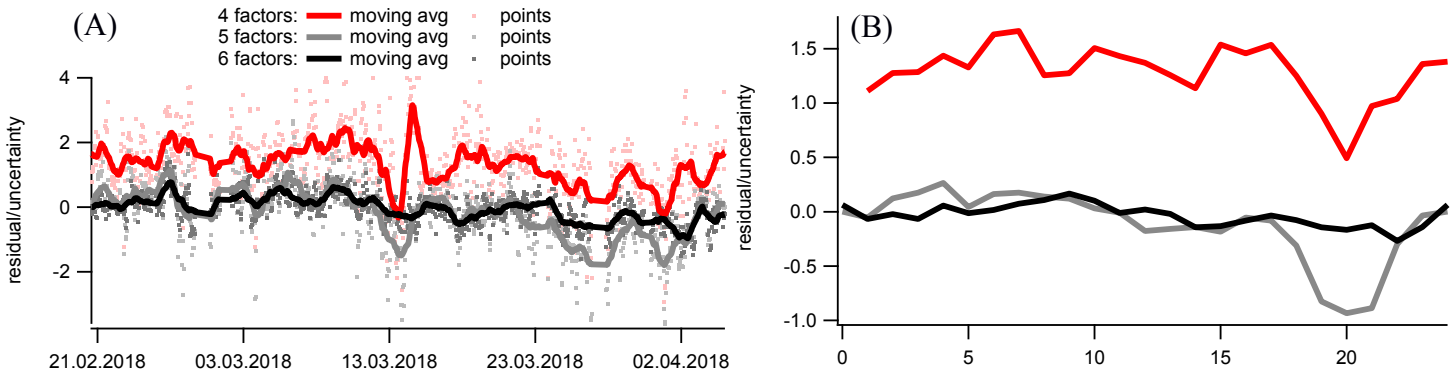


Figure S11 (A) Time series of PMF residuals normalized to measurement uncertainty as a function of the number of factors (HOA, COA, BBOA, CCOA constrained) (moving average is computed using 24 points). (B) Median diurnal of PMF residuals normalized to measurement uncertainty as a function of the number of factors (HOA, COA, BBOA, CCOA constrained).

We evaluated the environmental interpretability of every single PMF run and accepted only those fulfilling the following criteria:

1. Result from multiple linear regression (CO = *a* * HOA + *b* * BBOA + *c* * CCOA) correlates better with CO than OOA (significance level 0.05)
2. Result from multiple linear regression (BC = *a* * HOA + *b* * BBOA + *c* * CCOA) correlates better with BC than OOA (significance level 0.05)
3. Result from multiple linear regression (NO_x = *a* * HOA + *b* * CCOA) correlates better with NO_x than OOA (significance level 0.05)
4. Result from multiple linear regression (NO_x = *a* * HOA + *b* * CCOA) correlates better with NO_x than BBOA (significance level 0.05)
5. Result from multiple linear regression (NO_x = *a* * HOA + *b* * CCOA) correlates better with NO_x than OOA (significance level 0.05)
6. Lunchtime peak of COA is larger than the one of HOA
7. Mass spectral signature of CCOA is more similar to CCOA reference profile than to COA, than to HOA, than to BBOA reference profiles (assessed by Kendall-Tau)
8. Mass spectral signature of BBOA is more similar to BBOA reference profile than to COA, than to HOA, than to CCOA reference profiles (assessed by Kendall-Tau)
9. f44(HOA) < 0.04 and f44(COA) < 0.04 based on assessing f44 content of reference profiles[26].

Thereby, we accepted 96 PMF solutions. Mass spectral fingerprints, time series and diurnal patterns of the resolved OA factors are presented in Figures S12 and S13.

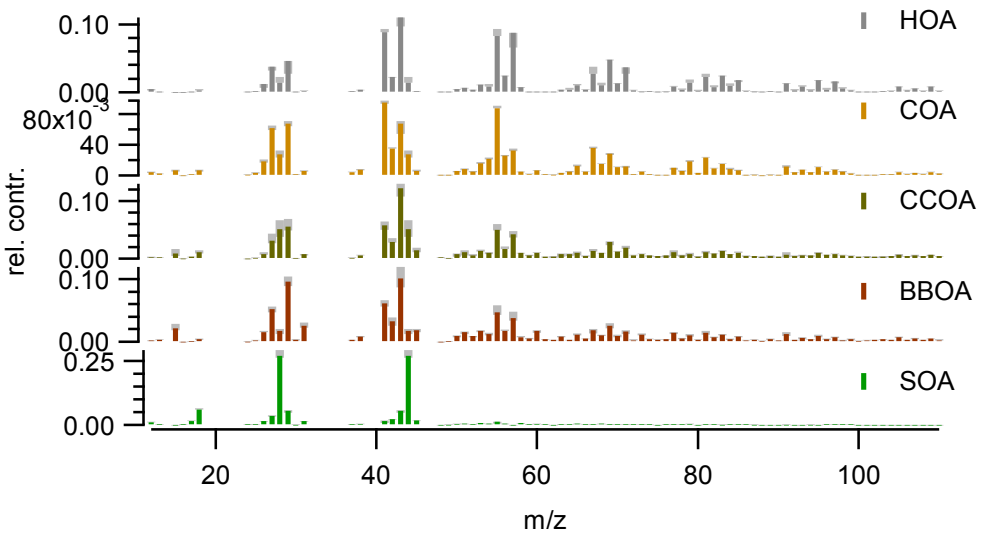


Figure S12 Mass spectral fingerprints of factors separated by PMF (median: sticks, first to third quartile: grey shaded area): primary traffic OA (HOA), primary cooking OA (COA), primary coal combustion OA (CCOA), primary biomass burning OA (BBOA), secondary OA (SOA).

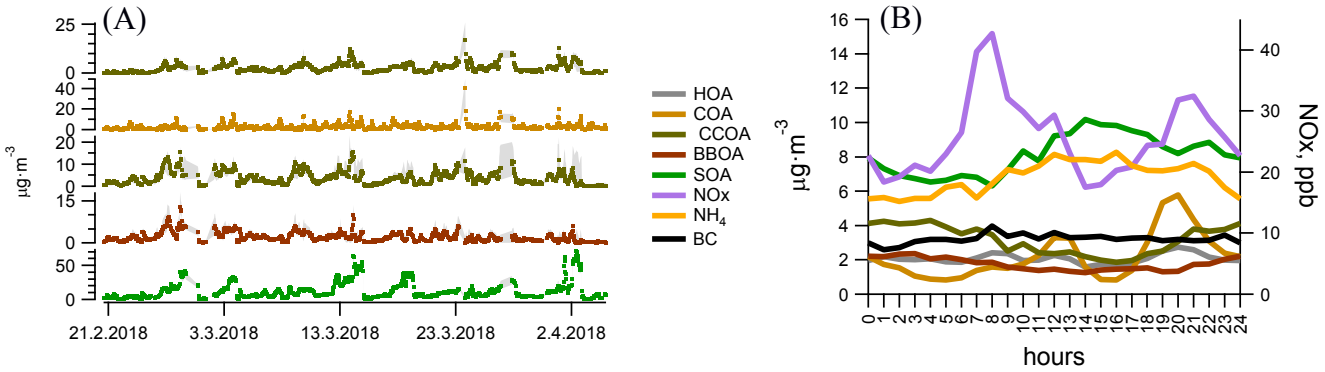


Figure S13 (A) Time series of factors separated by PMF (dots: median, grey shaded area: first to third quartile): primary traffic OA (HOA), primary cooking OA (COA), primary coal combustion OA (CCOA), primary biomass burning OA (BBOA), secondary OA (SOA). (B) Diurnal pattern of primary traffic OA (HOA), primary cooking OA (COA), primary coal combustion OA (CCOA), primary biomass burning OA (BBOA), and secondary OA (SOA) from PMF. For comparison the diurnal pattern of ammonium (NH_4, measured by ToF-ACSM), black carbon (BC, measured by aethalometer), and NO_x are additionally displayed.

To fill gaps in the BC data, we made use of the OA PMF results. We created predictive multilinear regression model explaining BC with HOA, CCOA, and BBOA as predictors which recreated the measured BC concentrations well (Figure S14, validation of gap-filling BC time series predictive model).

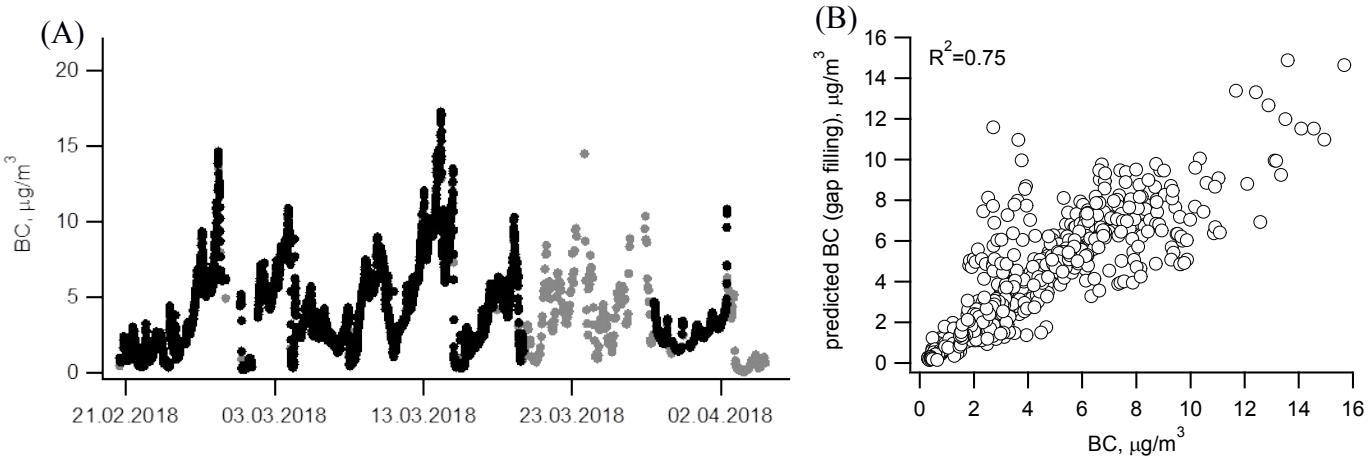


Figure S14 (A) Time series of black carbon concentration (BC, 950 nm). Measured points are shown in black and gap-filled points in grey. (B) Validation of gap-filling predictive model for BC: the predicted BC concentrations using gap-filling predictive model in comparison to the measured BC concentrations.

3.3 Estimating CS_{eff} based on heterogeneous nucleation theory

To participate in the particle growth, vapors need to b-e able to either condense onto the surface of pre-existing aerosol particles or react chemically with it. In practice this means that existing particles need to be activated by this vapor (e.g. Kulmala, et al. [27] and Kulmala, et al. [28]). This activation process can be described by using heterogeneous nucleation theory (see e.g. Wagner, et al. [29] and Winkler, et al. [30]). Recently, Tuovinen et al. [31] showed that if the activation is described by heterogeneous nucleation, inefficient heterogenous nucleation can lead to efficient condensation sink (CS_{eff}) being clearly lower than CS obtained with the traditional assumption that clusters are scavenged by larger particles upon every collision with them. In this study, we used the same approach as Tuovinen et al. [37] and determined CS_{eff}/CS for sulfuric acid–DMA clusters as a function of condensing cluster mass assuming a constant contact angle of 45° (Fig. 4A). For calculating CS, we used median particle size distribution measured during NPF events observed at our measurement site in Beijing between January 17 and March 31 2018. One should note that heterogeneous nucleation probability, and thus CS_{eff}/CS, is sensitive to the assumptions about contact angle of heterogeneous nucleation and the properties of nucleating vapor, such as molecular mass, surface tension and saturation vapor pressure. As the exact values of these variables are unknown for the vapors participating in nucleation in Beijing, the actual effect of inefficient heterogeneous nucleation on CS_{eff} remains to be solved. A more detailed description of the calculation of CS_{eff} based on heterogenous nucleation theory and the analysis of the effects of different assumptions on CS_{eff} are presented in Tuovinen et al.[31].

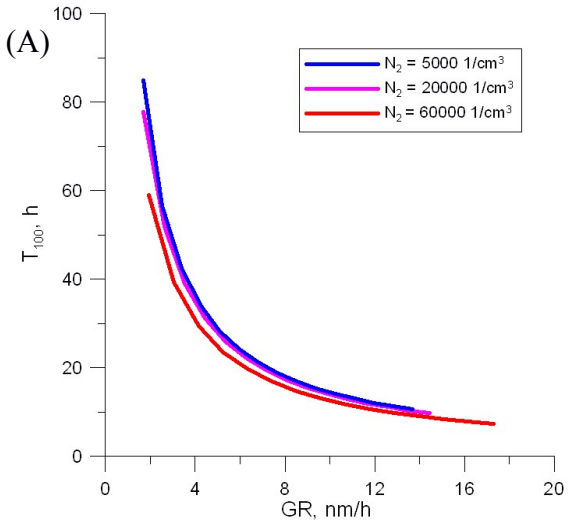
3.4 NPF ON/OFF Model

The model describes time evolution of the monodisperse aerosol population due to vapor condensation onto the particles in the system a constant vapor supply. The model is a generalization of the equations for the kinetic regime[32] to the transition regime using a modified Fuchs-Sutugin coefficient[33]. There are two modes (growing and accumulation modes) and two different vapors in the system. The first vapor is allowed to condense onto both modes, while the second vapor is able to condense onto the growing mode only. Both vapors are nonvolatile (equilibrium vapor concentrations are equal to zero). However, the second vapor is not able to condense on the accumulation mode because before starting to condense onto it, this vapor needs to go over the free energy barrier for heterogeneous nucleation (see section 3.4 for

more details). In practice, the surface of the accumulation mode is phobic for this vapor. On the other hand, the second vapor can condense onto the growing mode because this mode consists largely of the same material as the vapor.

The set of ordinary differential equations for the vapor condensation into the two modes was solved numerically using Runge-Kutta method of the 4th order. Model calculations are initiated with the measurements data corresponding approximately to the end of NPF event and represent cases with small initial PM2.5 concentrations. The basic setup parameters for the simulations were 10 nm initial diameter for the growing mode and 100 nm initial diameter for the accumulation mode and sources rates of 10^6 cm^{-3} s^{-1} for the vapor condensing on both modes and source rates between 10^6 and 1.5 10^7 cm^{-3} s^{-1} for the vapour condensing only on the accumulation mode. The simulations were terminated when the PM2.5 concentration reached 100 µg/m3. The growth rate GR was estimated from the simulations as the diameter of the growing mode divided by time ($GR = d_{100}/T_{100}$), both parameters corresponding to the end of simulations.

Figure S15 shows the sensitivity of the model to different concentrations of the accumulation and growing mode, demonstrating that the relation between growth rate GR and haze appearance time (T_{100}) is very sensitive to growing mode number concentrations not to the initial accumulation mode number concentration.



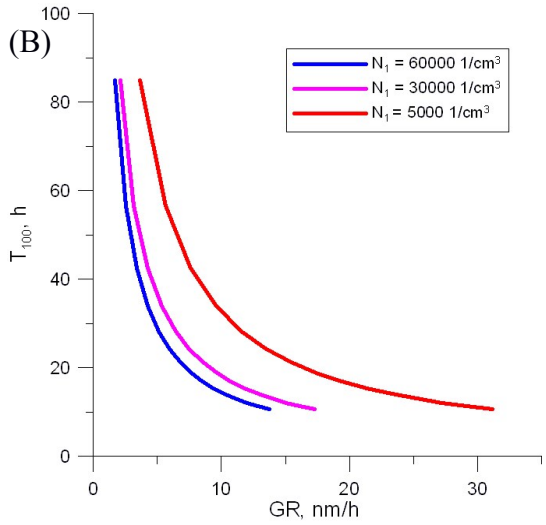


Figure S15 Sensitivity analysis. (A): sensitivity of T100 (the time difference between end of NPF and time when the PM2.5 concentration reaches 100 µg/m³) to the number concentration of particles in the accumulation mode. Note that we assume that initial PM2.5 is the mass concentration of submicron modes. In this case, initial mass concentration of aerosol population is determined by the accumulation mode. N_2 is the number concentration of the accumulation mode: N_2 = 5000 #/cm³ corresponds to the initial PM2.5 equal to 2.5 µg/m³, N_2 = 20000 #/cm³ corresponds to the initial PM2.5 equal to 10 µg/m³, N_2 = 60000 #/cm³ corresponds to the initial PM2.5 equal to 30 µg/m³. The difference in the initial PM2.5 leads to the noticeable difference in T100 only at small growth rates. (B): sensitivity of T100 to the number concentration of the particles in the growing mode.

3.5 Estimating current SO_2 and NH_3 emissions

Estimates of present-day SO_2 and NH_3 emissions are made with GAINS model (Greenhouse gas Air pollution Interactions and Synergies[34]). The calculations are made with ECLIPSE-V5a-CLE-base

scenario[35] for year 2020 and include Beijing, Tianjin and Hebei provinces. Emissions from GAINS model have been previously shown to represent well the real emissions in the same region[36].

References

1 Zhou, Y. *et al.* Variation of size-segregated particle number concentrations in wintertime Beijing. *Atmos. Chem. Phys.* **20**, 1201-1216, doi:10.5194/acp-20-1201-2020 (2020).

2 Jiang, J., Chen, M., Kuang, C., Attoui, M. & McMurry, P. H. Electrical Mobility Spectrometer Using a Diethylene Glycol Condensation Particle Counter for Measurement of Aerosol Size Distributions Down to 1 nm. *Aerosol Science and Technology* **45**, 510-521, doi:10.1080/02786826.2010.547538 (2011).

3 Kangasluoma, J. *et al.* Operation of the Airmodus A11 nano Condensation Nucleus Counter at various inlet pressures and various operation temperatures, and design of a new inlet system. *Atmospheric Measurement Techniques* **9**, 2977-2988, doi:10.5194/amt-9-2977-2016 (2016).

4 Liu, J. *et al.* A spectrometer for measuring particle size distributions in the range of 3 nm to 10 µm. **10**, 63-72, doi:10.1007/s11783-014-0754-x (2016).

5 Vanhanen, J. *et al.* Particle Size Magnifier for Nano-CN Detection. *Aerosol Science and Technology* **45**, 533-542, doi:10.1080/02786826.2010.547889 (2011).

6 Mirme, S. & Mirme, A. The mathematical principles and design of the NAIS – a spectrometer for the measurement of cluster ion and nanometer aerosol size distributions. *Atmospheric Measurement Techniques* **6**, 1061-1071, doi:10.5194/amt-6-1061-2013 (2013).

7 Eresmaa, N. *et al.* A Three-Step Method for Estimating the Mixing Height Using Ceilometer Data from the Helsinki Testbed. *Journal of Applied Meteorology and Climatology* **51**, 2172-2187, doi:10.1175/jamc-d-12-058.1 (2012).

8 Fröhlich, R. *et al.* The ToF-ACSM: a portable aerosol chemical speciation monitor with TOFMS detection. *Atmospheric Measurement Techniques* **6**, 3225-3241, doi:10.5194/amt-6-3225-2013 (2013).

9 Middlebrook, A. M., Bahreini, R., Jimenez, J. L. & Canagaratna, M. R. Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer using Field Data. *Aerosol Science and Technology* **46**, 258-271, doi:10.1080/02786826.2011.620041 (2012).

10 Hansen, A. D. A., Rosen, H. & Novakov, T. The aethalometer — An instrument for the real-time measurement of optical absorption by aerosol particles. *Science of The Total Environment* **36**, 191-196, doi:10.1016/0048-9697(84)90265-1 (1984).

11 Drinovec, L. *et al.* The "dual-spot" Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation. *Atmospheric Measurement Techniques* **8**, 1965-1979, doi:10.5194/amt-8-1965-2015 (2015).

12 Virkkula, A. *et al.* On the interpretation of the loading correction of the aethalometer. *Atmospheric Measurement Techniques* **8**, 4415-4427, doi:10.5194/amt-8-4415-2015 (2015).

13 Dal Maso, M. *et al.* Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiala, Finland. *Boreal Environment Research* **10**, 323-336 (2005).

14 Stohl, A., Forster, C., Frank, A., Seibert, P. & Wotawa, G. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmospheric Chemistry and Physics* **5**, 2461-2474, doi:10.5194/acp-5-2461-2005 (2005).

15 Kulmala, M. *et al.* Measurement of the nucleation of atmospheric aerosol particles. *Nature Protocols* **7**, 1651–1667, doi:10.1038/nprot.2012.091 (2012).

16 Lehtipalo, K. *et al.* Methods for determining particle size distribution and growth rates between 1 and 3 nm using the Particle Size Magnifier. *Boreal Environment Research* **10**, 323-336 (2014).

17 Paasonen, P. *et al.* Comprehensive analysis of particle growth rates from nucleation mode to cloud condensation nuclei in boreal forest. *Atmospheric Chemistry and Physics* **18**, 12085-12103, doi:10.5194/acp-18-12085-2018 (2018).
18 Kulmala, M. *et al.* On the mode-segregated aerosol particle number concentration load: contributions of primary and secondary particles in Hyytiala and Nanjing. *Boreal Environ Res* **21**, 319-331 (2016).
19 Rodríguez, S. & Cuevas, E. The contributions of "minimum primary emissions" and "new particle formation enhancements" to the particle number concentration in urban air. *Journal of Aerosol Science* **38**, 1207-1219, doi:10.1016/j.jaerosci.2007.09.001 (2007).
20 Paatero, P. & Tapper, U. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **5**, 111-126, doi:10.1002/env.3170050203 (1994).
21 Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U. & Prévôt, A. S. H. SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data. *Atmospheric Measurement Techniques* **6**, 3649-3661, doi:10.5194/amt-6-3649-2013 (2013).
22 Paatero, P. Least squares formulation of robust non-negative factor analysis. *Chemometrics and Intelligent Laboratory Systems* **37**, 23-35, doi:10.1016/S0169-7439(96)00044-5 (1997).
23 Paatero, P. The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n-Way Parallel Factor Analysis Model. *Journal of Computational and Graphical Statistics* **8**, 854-888, doi:10.1080/10618600.1999.10474853 (1999).
24 Crippa, M. *et al.* Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris. *Atmospheric Chemistry and Physics* **13**, 961-981, doi:10.5194/acp-13-961-2013 (2013).
25 Elser, M. *et al.* New insights into PM2.5 chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry. *Atmospheric Chemistry and Physics* **16**, 3207-3225, doi:10.5194/acp-16-3207-2016 (2016).
26 Daellenbach, K. R. *et al.* Long-term chemical analysis and organic aerosol source apportionment at nine sites in central Europe: source identification and uncertainty assessment. *Atmospheric Chemistry and Physics* **17**, 13265-13282, doi:10.5194/acp-17-13265-2017 (2017).
27 Kulmala, M. *et al.* Initial steps of aerosol growth. *Atmospheric Chemistry and Physics* **4**, 2553-2560, doi:10.5194/acp-4-2553-2004 (2004).
28 Kulmala, M., Kerminen, V.-M., Anttila, T., Laaksonen, A. & O'Dowd, C. D. Organic aerosol formation via sulphate cluster activation. *Journal of Geophysical Research* **109**, Article ID D04205, doi:10.1029/2003jd003961 (2004).
29 Wagner, P. E. *et al.* Nucleation probability in binary heterogeneous nucleation of water-n-propanol vapor mixtures on insoluble and soluble nanoparticles. *Physical Review E* **67**, Article ID 021605, doi:10.1103/PhysRevE.67.021605 (2003).
30 Winkler, P. M. *et al.* Heterogeneous Nucleation Experiments Bridging the Scale from Molecular Ion Clusters to Nanoparticles. *Science* **319**, 1374-1377, doi:10.1126/science.1149034 (2008).
31 Tuovinen, S., Kontkanen, J., Jiang, J. & Kulmala, M. Investigating the effectiveness of condensation sink based on heterogeneous nucleation theory, *J. Aerosol Sci*., in press (2020).
32 Ezhova, E., Kerminen, V. M., Lehtinen, K. E. J. & Kulmala, M. A simple model for the time evolution of the condensation sink in the atmosphere for intermediate Knudsen numbers. *Atmospheric Chemistry and Physics* **18**, 2431-2442, doi:10.5194/acp-18-2431-2018 (2018).

33 Wu, C.-Y. & Biswas, P. Particle Growth by Condensation in a System with Limited Vapor. *Aerosol Science and Technology* **28**, 1-20, doi:10.1080/02786829808965508 (1998).
34 Amann, M. et al. Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, *Environ. Modelling & Software*, **26**, 12, 1489-1501 (2011)..
35 Klimont, Z., et al. Global anthropogenic emissions of particulate matter including black carbon, *Atmos. Chem. Phys.*, **17**, 8681–8723, https://doi.org/10.5194/acp-17-8681-2017, (2017).
36 Wang, S., et al. Verification of anthropogenic emissions of China by satellite and ground observations, *Atmos. Environ.*, **45**, 35, 6347-6358 (2011).