Contributions of traffic emissions and new particle formation to the ultrafine particle size distribution in the megacity of Beijing

Hao Wu, Zhanqing Li, Mengjiao Jiang, Chunsheng Liang, Dongmei Zhang, Tong Wu, Yuying Wang, Maureen Cribb

A Key Laboratory of China Meteorological Administration Atmospheric Sounding, School of Electrical Engineering, Chengdu University of Information Technology, Chengdu, 610225, China
b Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, 20742, USA
c Sichuan Provincial Key Laboratory of Plateau Atmosphere and Environment, School of Atmospheric Sciences, Chengdu University of Information Technology, Chengdu, 610225, China
d Collaborative Innovation Center for West Ecological Safety, Lanzhou University, Lanzhou, 730000, China
e College of Global Change and Earth System Science, Beijing Normal University, Beijing, 100875, China
f Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Nanjing University of Information Science & Technology, Nanjing, 210044, China

HIGHLIGHTS

- Source contributions were determined by analyzing NPF characteristics and using k-means clustering and PMF.
- Two-year particle evolution and NPF features reveal that traffic emissions contribute significantly to the Aitken mode.
- NPF is a major contributor to the PNSD and identified key influential factors on the ultrafine particles.

ARTICLE INFO

Keywords:
Traffic emissions
New particle formation
Source characteristics

ABSTRACT

The important role of traffic emissions in air pollution has been widely recognized. However, the corresponding contributions to the ultrafine particle distribution and new particle formation (NPF) are poorly understood due to the mixed influences of various sources. Measurements of particle evolution and NPF were made at a heavy traffic site near the 5th Ring Road of Beijing from August 2017 to October 2019. A positive matrix factorization receptor model (version 5.0) consisting of the log-normal fit method and the k-means clustering method was used to identify and quantify the major sources of particle number concentrations. Based on the characteristics of NPF, additional analyses were performed to identify formation sources, including back-trajectory clustering and correlations between auxiliary variables and the relative contributions of all resolved factors. Traffic emissions mainly influenced Aitken mode (25–100 nm) particles (>60%), and NPF significantly contributed to the ultrafine particle distribution (47.2%). Additionally, regional transport notably affected the occurrence of NPF at the observation site, with northwest airmass transport influencing the nucleation mode the most and southeast airmass transport influencing the Aitken mode the most.

1. Introduction

Traffic emissions are a major source of aerosol particles in the urban environment (Ronkkö et al., 2017), significantly altering the particle number size distribution (PNSD) in the urban atmosphere (Vu et al., 2015; Harrison et al., 2019; Saha et al., 2019). Urban traffic emissions mainly contribute to ultrafine particles (UFPs, < 100 nm) (Guo et al., 2020), thus play a key role in particle size distributions. Aitken-mode particles (25–100 nm) appear to be closely associated with road traffic (Weber, 2009; Padró-Martínez et al., 2012; Pikridas et al., 2015) because they reach a peak in concentration during rush hours (Masiol et al., 2017). Beddows et al. (2009) used the k-means clustering method to analyze particle clusters in urban and rural environments. A study of PNSDs from sites across Europe indicated that two

* Corresponding author.
E-mail address: zhanqing@umd.edu (Z. Li).

https://doi.org/10.1016/j.atmosenv.2021.118652
Received 26 April 2021; Received in revised form 28 July 2021; Accepted 28 July 2021
Available online 30 July 2021
1352-2310/© 2021 Published by Elsevier Ltd.
of the nine clusters identified were influenced by air masses (Beddows et al., 2014). Harrison (2011) noted a peak in PNSD at 20 nm, attributed to primary road traffic emissions near a major highway in the U.K. Traffic was found to be a determining factor for particles with a modal diameter of ~30 nm at an urban background site in west London (Beddows et al., 2015; Vu et al., 2016) and for the volume concentration peak from 20 to 40 nm in central Los Angeles (Sowlat et al., 2016). Street canyons also limit pollutant dispersion due to low wind speed and poor mixing conditions (Kumar et al., 2010; Zhu et al., 2017).

New particle formation (NPF) also increases UFP at different time scales in the urban atmosphere (Ma and Birmili, 2015). Newly produced nanoparticles are involved in the initial step of the nucleation of gaseous precursor clusters, which are transformed into particles, generating all kinds of particles in the atmosphere. Furthermore, they may grow via hygroscopic or condensational processes to sizes large enough to serve as cloud condensation nuclei (CCN) (Svensmark et al., 2017; Sanchez et al., 2018; Kalkavouras et al., 2019; J. Kim et al., 2019), affecting clouds, Earth’s radiation budget, and the climate (Zhang et al., 2016; Li et al., 2017). However, the contributions of traffic emissions to NPF and CCN growth remain unclear. NPF events can contribute to CCN during particle growth (Yue et al., 2016; Shen et al., 2018; S.-H. Lee et al., 2019). The concentration of CCN also increases significantly during NPF events, and the nucleation-induced concentration enhancement of CCN can vary to different degrees (Lihaavainen et al., 2003; Matsui et al., 2011; Leng et al., 2014; Zhang et al., 2014). The role of traffic emissions in promoting nucleation has rarely been considered, and the key processes involved should be quantified (Guo et al., 2020). The overall driving factor has yet to be identified, remaining a major challenge (Wang et al., 2017; Yang et al., 2021).

Freshly nucleated particles with ~3-nm diameters coagulate very efficiently with larger particles within a short period through Browninan motion and are consumed in collision processes (Sarangi et al., 2015; Olenius et al., 2018; Pichelstorfer et al., 2018). Newly formed aerosols become climatically important when they reach a size at which they scatter light efficiently, thereby having a direct cooling effect on the climate (Gordon et al., 2016; Wang et al., 2017; Nieminen et al., 2018). The concentration of CCN also increases significantly during NPF events, and the nucleation-induced concentration enhancement of CCN can vary to different degrees (Lihavainen et al., 2003; Matsui et al., 2011; Leng et al., 2014; Zhang et al., 2014). The role of traffic emissions in promoting nucleation has rarely been considered, and the key processes involved should be quantified (Guo et al., 2020). The overall driving factor has yet to be identified, remaining a major challenge (Wang et al., 2017; Yang et al., 2021).

2. Site, data, and methodology

2.1. The observation site

Measurements used in this study were acquired at the Beijing Meteorological Observatory (39.81° N, 116.82° E) surrounded by the 5th-Ring Beltway of Beijing to the south with generally heavy traffic (mean traffic flow: ~156,300 cars/day) and two smaller city roads to the west (Santaishan Road, traffic flow: 17,200 cars/day) and to the west (Santaishan Road, traffic flow: ~2,000 cars/day). There are no significant industrial emissions in the area, which includes a large grassy park to the southeast (Fig. 1a).

The site is influenced by mixed-emission sources such as traffic and burning, biogenic, and residential activities. It thus provides a unique opportunity to estimate the contributions of traffic to the occurrence of NPF and to understand their relationships, which may improve numerical simulations because the impact of exhaust emissions is often underestimated (Rönkkö et al., 2017).

Ample measurements were made during three intensive observation periods (IOPs) from July 2017 to October 2019 (Fig. 1b). The first IOP took place from July 2017 to December 2017 using instruments located in a container near the roadside. In December 2017, the instruments were moved to the roof of a building 13 m above the ground and ~30 m away from the roadside. On 7 November 2018, they were moved back to the ground level in the Integrated Atmospheric Observatory (BNU-MOIAS) container (details in Fig. S1).

![Fig. 1. Map of (a) the sampling location (black and yellow stars in each panel) and (b) three observation periods and location/instrument-configuration changes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)](image-url)
2.2. Instruments and datasets

The observation system was placed in a container on a grass surface. The aerosol samples were collected on the rooftop (~4 m above the ground) through a cyclonic inlet (2.5-μm cutoff size) connected by stainless steel tubes. The internal diameter was 6 mm, and the flow length was less than 4 m. All sample inlets were placed near the road base to ensure that the dust elevated by vehicles and other gas exhaust particles were directly transported to the sample inlet tubes. The whole system airflow was dried through a silica gel tube, and the infill material was changed approximately twice a month to ensure that the RH was less than 30%. The refill frequency depended on environmental variations. Table 1 conveys more information about the instruments used in this study.

2.2.1. PNSD measurements

This study utilized a scanning mobility particle sizer spectrometer (SMPS; model 3938) system, a condensation particle counter (CPC; model 3772), a nano-SMPS, and multi-instrument manager software (TSI, v3.0). The SMPS was calibrated with polystyrene latex particles (40/80/100/200/400 nm), which had a 5% bias. The condensation particle counter (CPC; model 3772) was calibrated with the TSI second standard CPC curve, with a bias within the 10% range. The system sample flow rate was set to 1.0 L min\(^{-1}\), and the sheath flow rate was maintained at 4.5 L min\(^{-1}\). The particle size ranged from 11.3 nm to 552.3 nm, with 64 channels/decade. A nano-SMPS was added to the instrument suite in January 2019 to measure particle sizes in the range of 4 nm–64.8 nm. Data quality was assured by following the steps outlined by Liang et al. (2020) and Wu et al. (2020). Monthly and diurnal variations were examined. Particles were classified into four modes: nano mode (or Sub0 mode, 4–10 nm), nucleation mode (10–25 nm), Aitken mode (25–100 nm), and accumulation mode (>100 nm) (Vu et al., 2015). The geometric mean diameter (GMD) and CS were then estimated based on the particle size distribution. NPF, non-NPF, and undefined days were selected to investigate NPF frequencies and characteristics in this urban environment (Dada et al., 2018). Log-normal fits and k-means clustering were used to illustrate which source dominated in each period. The positive matrix factorization (PMF) approach was finally applied to evaluate the source contribution factors.

2.2.2. Precursor gas measurements

Quality-controlled gas measurements were retrieved from the Beijing Municipal Environmental Monitoring Center measurement network (http://www.bjmemc.com.cn/) at the nearest site, Yizhuang. Data for six species in total were collected at this site, including SO\(_2\), NO\(_2\), O\(_3\), CO, particulate matter with a diameter less than 2.5 μm (PM\(_{2.5}\)), and PM\(_{10}\).

2.3. Description of NPF parameters and methodology

2.3.1. GMD and growth rate (GR)

The observed particle GR was calculated based on the log-normal distribution function method (Kulmala et al., 2012). The PNSD is typically fitted with the least-squares log-normal fitting method, yielding parameters in one log-normal mode (Hussein et al., 2005). First, the temporal variation in the GMD is obtained, then the particle GR is estimated (given in nm h\(^{-1}\)). The first step is to fit the GMD. Particles in different modal ranges during period \(i\) are then used to calculate the GR:

\[
GMD(d_i) = \exp \left[ \frac{\sum N_i \log(d_i)}{\sum N_i} \right]
\]

where \(d_i\) is the GMD of the particles, \(d_p\) is the particle diameter in size bin \(i\), and \(N_i\) is the particle number concentration in size bin \(i\). The observed particle GR is defined as the rate of change in the diameter, \(d_p\), representing the growth of the particle population.

2.3.2. CS

The CS (condensation sink) variable reflects the ability of the size distribution of a preexisting particle population to act as a sink for condensable vapors in the atmosphere and freshly formed particles. This result describes the loss rate (in molecules s\(^{-1}\)) of vapors to the aerosol phase in the atmosphere. When particle concentrations are high, CS can be a limiting factor in the formation of new particles (Dal Maso et al., 2002). The CS is calculated as follows:

\[
CS = 2\pi D \beta_{\text{dp}} \times N_i (D_p) \Delta D = 2\pi D \sum \beta_{\text{dp}} D_i N_i
\]

where \(D\) is the diffusion coefficient of the condensing vapor, usually assumed to be sulfuric acid, for which a value of 0.80*10\(^{-5}\) was used in this study; \(D_p\) represents the particle size distribution; \(N_i\) is the particle number concentration; and \(\beta_m\) is defined by the following equation:

\[
\beta_m = \frac{1 + Kn}{1 + 1.677Kn + 1.333K_n^2}
\]

In Eqn. (4), \(K_n = 2\beta/d_p\), and if \(K_n < 1\), the droplet is in the continuum regime, where macroscopic laws such as Fick’s law of diffusion or Fourier’s law of thermal conduction can be applied. In the kinetic regime, i.e., \(K_n > 1\), the kinetic gas theory is used to calculate collisions and the resulting heat or matter exchange between particles and gas-phase molecules.

2.3.3. K-means cluster classification

Variables in the particle size distribution and prevailing meteo-

Table 1

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Parameter</th>
<th>Manufacturer</th>
<th>Model</th>
<th>Time Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerosol Properties</td>
<td>SMPS</td>
<td>TSI</td>
<td>3938</td>
<td>5 min</td>
</tr>
<tr>
<td></td>
<td>Particle number size distribution (10.7–532.8 nm)</td>
<td>TSI</td>
<td>3938N56</td>
<td>5 min</td>
</tr>
<tr>
<td></td>
<td>Nano-SMPS</td>
<td>TSI</td>
<td>3321</td>
<td>5 min</td>
</tr>
<tr>
<td></td>
<td>Particle number size distribution (4–64.8 nm)</td>
<td>TSI</td>
<td>100</td>
<td>1 s</td>
</tr>
<tr>
<td></td>
<td>APS</td>
<td>DMT</td>
<td>Aerodynamic</td>
<td>Q-ACSM</td>
</tr>
<tr>
<td></td>
<td>Aerosol composition</td>
<td>Kipp &amp; Zonen</td>
<td>CNR4</td>
<td>1 min</td>
</tr>
<tr>
<td>Solar Radiation</td>
<td>Pyranometer</td>
<td>Li-COR</td>
<td>LI-190R</td>
<td>1 min</td>
</tr>
<tr>
<td></td>
<td>Net irradiance</td>
<td>Li-COR</td>
<td>Li-7500A</td>
<td>5 min</td>
</tr>
<tr>
<td>Gas Analyzers</td>
<td>Nondispersive infrared</td>
<td>Thermo Fisher</td>
<td>43L/43i</td>
<td>5 min</td>
</tr>
<tr>
<td></td>
<td>CO(_2) concentration</td>
<td>Gill</td>
<td>Wind Master Pro</td>
<td>1 s</td>
</tr>
<tr>
<td></td>
<td>Gaseous precursors</td>
<td>Vaisala</td>
<td>HMP 155A</td>
<td>1 s</td>
</tr>
<tr>
<td></td>
<td>SO(_2)/NO(_x) concentration</td>
<td>LI-COR</td>
<td>SmartFlux2</td>
<td>1 s</td>
</tr>
</tbody>
</table>

ACSM: aerosol chemical speciation monitor; APS: aerodynamic particle sizer; SMPS: scanning mobility particle sizer spectrometer.
logical condition observations were taken into consideration. The set \( x_i \) for \( i = 1, \ldots, n \) is a set of \( n \) points in the three modes to be clustered into \( k \) clusters, namely, \( C_k \) (Beddows et al., 2009; Carslaw and Beevers, 2013; Masiol et al., 2017). The basic \( k \)-means algorithm for \( k \) clusters is obtained by minimizing the following equation:

\[
\sum_{k=1}^{k} \sum_{x_i \in C_k} \left| x_i - \mu_k \right|^2
\]

(7)

\[d_{xy} = (x_j - y_j)^2 (x_j - y_j)^2)^{1/2}
\]

(8)

where \( x_i \) and \( y_j \) are two \( j \)-dimensional vectors that have been standardized by subtracting the mean and dividing by the standard deviation. Here, \( j \) has a length of three and includes the wind components \( u \) and \( v \) and concentration \( C \), each of which is standardized as follows:

\[
x_j = \left( \frac{x_j - \bar{x}}{s_x} \right)
\]

(9)

Standardization is necessary because the wind components \( u \) and \( v \) have different scales. The \( k \)-means clustering approach helps clarify potential source characteristics that, combined with bivariate polar-plot features, can be identified and grouped.

2.3.4. PMF model principles

The uncertainty in the PMF model used here was computed using the method described by Thimmiah et al. (2009). The units of PNSD and mass concentration are \( \# \text{ cm}^{-3} \) and \( \mu g \text{ m}^{-3} \), respectively. Specifically, the uncertainty of the PNSD (with \( x \) representing concentration) is calculated as \( 1 + x^{0.5} + 0.1x \). The uncertainties of the mass concentrations of \( \text{O}_3, \text{SO}_2, \text{NO}_2, \) and \( \text{PM}_{2.5} \) are calculated as \( 0.5 \times \text{min}(x) + 0.1x \). An object function, \( Q \), for which PMF provides a solution that minimizes it, is defined as

\[
Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[ \frac{E_i}{U_i} \right]^2
\]

(10)

or

\[
Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[ \frac{x_i - \sum_{k=1}^{k} P_{ik} f_{kj}}{u_{ij}} \right]^2
\]

(11)

where \( U_i \), i.e., \( u_{ij} \) in Eqn. (11), is the uncertainty estimate in particle size bin \( j \) measured in sample \( i \) (E. Kim et al., 2004), \( E \) is the residual error (Krecl et al., 2015), \( n \) is the total number of samples, and \( m \) is the total number of particle size bins (Krecl et al., 2008). The quantities \( f_{ij} \) and \( g_{ik} \) are the particle volume fraction in the \( j \)th particle size bin from the \( k \)th source and the particle concentration from the \( k \)th source contributing to the \( j \)th sample, respectively. For a given number of factors, PMF (Paatero and Tapper, 1994) adjusts the values of \( f_{ij} \) and \( g_{ik} \) by the method of least squares (with the constraint that \( f_{ij} \) and \( g_{ik} \) values are non-negative) until the minimum \( Q \) value is found (Krecl et al., 2008; Norris et al., 2014).

3. Results and discussion

3.1. Classification of NPF days and characteristic analysis

Data were collected from two instruments: all channels of the nano-SMPS, ranging from 4 to 63.8 nm, and the SMPS, ranging from 11.3 to 532.8 nm. Aerodynamic particle sizes were first converted to mobile sizes, assuming that the relative air density equals \( 1.4 \text{ kg m}^{-3} \) (Shen et al., 2011; DeCarlo et al., 2004). Particle size distributions covering the size range of 4 nm–20,000 nm were then generated. Note that a Gaussian function was applied to smooth the number concentration at the conjunction between the different instruments, i.e., 64 nm. Although there was some noise below 50 nm, NPF events were identified (Fig. 2b). An NPF day is a day where particle growth starts at 4 nm, and the number concentration exceeds \( 3 \times 10^6 \# \text{ cm}^{-3} \) for at least 3–4 h. If the growth process is very short or no significant new mode occurs, the day is classified as an undefined day, i.e., there is no increase in the nucleation mode, or the number concentration remains at a low level. From 18 to 20 September 2019, NPF and subsequent growth to above 100 nm occurred, contributing significantly to \( \text{PM}_{2.5} \) concentrations. This trend suggests a relation between NPF and haze pollution.

3.1.1. Geometric mean diameter (GMD) and CS variations

The GMD and CS during each of the three IOPFSs were estimated based on the particle size distribution. NPF, non-NPF, and undefined days (c.f. Fig. S2) were selected according to the criteria established by Kulmala et al. (2012). NPF was identified when the maximum concentration of particles smaller than 25 nm in diameter was higher than \( 3 \times 10^6 \# \text{ cm}^{-3} \) (excluding the background concentration), and particle growth was sustained for several hours. Undefined days were selected according to the method proposed by Dal Maso et al. (2005), and the duration and NPF window (12:00 a.m. ± 3:00) were defined according to Dada et al. (2018). The average GMD ranged from 9.252 to 1435.67 nm, and the median GMD was 55.920 ± 5.36 nm. The mean CS ranged from 0.0015 to 0.1757, and the median CS was 0.021 ± 0.009. This GMD size range is very close to motor vehicle emission results (Morawska et al., 2008), indicating that the particle number concentration was easily impacted by regional transport processes, e.g., dust transport, regional nucleation, and haze pollution events.

The average GMD in periods 1 and 2 remained nearly at the same level, i.e., ~50 nm (Fig. 3b). After the nano-SMPS was added to the suite of instruments, the average GMD decreased to ~22 nm, and more nucleation processes were observed in period 3. Pollution and haze events appear to have had a significant influence on the GMD time series. The CS was also much higher on non-NPF days than on other days (Fig. 3a). Note that there is considerable uncertainty in the non-NPF day classification because pre-existing particle concentrations can increase the CS. Frequent periodic checks, instrument maintenance, and inlet cleaning were performed throughout the field campaign, and all data were averaged in 5-min intervals.

3.2. NPF observation and results

3.2.1. NPF characteristics and frequency

Eighty-four NPF cases were identified. The GR exhibits no clear seasonal variation (Fig. 4), but its relation to the duration of growth is clearly seen. Before January 2019, the GR (11–25 nm) was determined by the SMPS only, with durations much longer than those for the GR (11–25 nm) retrieved from the nano-SMPS after January 2019. Additionally, more NPF cases were found after the nano-SMPS went into operation. The higher GRs in 2019 (green-shaded area in Fig. 4) were associated with shorter durations. The GR (4–25 nm) occurs close to the initial step of cluster nucleation for critical-sized particles, which is a rapid process, explaining why the GR was higher. The NPF frequency is influenced by the following factors: solar radiation, CS, the air mass in place, and the concentrations of sulfuric acid or highly oxygenated molecules in the atmosphere (c.f. Fig. S3).

3.2.2. NPF diurnal evolution and CS features

Fig. 5a shows a clear peak in the nano mode at ~10:00 local time (LT) on NPF days compared to the trends on non-NPF and undefined days. Dual peaks in the morning and evening occurred every day in the three categories. A significantly higher peak is seen at noon on NPF days than on undefined days, a unique feature of NPF. Particle formation is initiated at the nano size and grows to the nucleation and Aitken modes.
Fig. 2. (a) Time series of different mass concentrations measured by the ACSM and (b) the total particle size distribution from 17 to 26 September 2019. Particle sizes smaller (bigger) than 65 nm were measured by the nano-SMPS (SMPS). The two shaded areas (EP1 and EP2) show typical NPF and mass concentration variation, respectively.

Fig. 3. (a) CS and (b) GMD time series over the two studied years, with period 1 (roadside) shaded in pink, period 2 (roof) shaded in purple, and period 3 (roadside) shaded in green. Blue triangles, red circles, and green squares represent NPF days, non-NPF days, and undefined days, respectively, curves show the average value variation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
as NPF progresses. There is a clear peak at night before 06:00 LT, which is even higher on undefined days than on other days. This finding indicates that small particles gathering at night can influence nucleation the next day by decreasing the CS or coagulation sink (CoagS). This process is also linked to the diurnal evolution of the planetary boundary layer (PBL), increasing the particle concentration at night. Fig. 5b shows the mean CS in the NPF window (09:00 to 15:00 local time) for NPF, non-NPF, and undefined days. The CS was much lower on NPF days than on the other days. The CS on all NPF days was less than 0.02 s⁻¹, as was the CS on non-NPF days, especially near the roadside in period 3. This is likely because of the daylong traffic. The CS describes the ability of the aerosol population to remove condensable vapors from the atmosphere. The low CS in the NPF window indicates that conditions were favorable for promoting NPF. In general, NPF occurred on low CS days and when the GMD was small (cf. Fig. S2). These conditions can thus be used as reliable indicators of NPF. The highest CS concentrations were observed from 00:00 to 02:00 LT in the northeast on NPF days due to heavy trucks driving along the highway located northeast of the observation site, generating large particles that increased the CS at midnight (Fig. 5c). The CS was much lower around noon, which is when the most suitable
conditions for NPF occur. Plotting the data as an annulus rather than a circle avoids the difficulty associated with interpreting values close to the origin to a certain extent. A low CS occurred in the morning at ~08:00 LT, with northern winds decreasing the CS level. Compared to the trends on undefined days, the high CS late in the evening in the southeast led to no NPF. The non-NPF-day annulus reveals CS reductions under northwesterly wind conditions and near-constant CS under the other wind conditions (Fig. S4). The CS evolution estimated by the PNSD reveals that the trends on undefined days, the high CS late in the evening in the northeast at all wind speeds. The results for mode 3 suggest the transport of larger particles from the north to the south and smaller particles from the northeast (see Fig. S5 for information from the other months).

### 3.2.3. Regional transport contributions to the PNSD

The concentration-weighted trajectory analysis reveals sources from different directions. Air masses from Mongolia carrying aged particles to the observation site had the greatest influence on the accumulation mode (Fig. 6a). The bulk of the transport originated from the north polar region, with easterly directions becoming significant for the nucleation mode (Fig. 6c). Pollution from the North China Plain was the major source contributing to local nucleation processes. The highest percentages of the nucleation mode were also associated with northwest air mass transport, but the dominant sources of the Aitken mode came from the southeast. Furthermore, there was a significant source from the north and northeast directions for the gaseous precursors (Figs. S6–8), closely linked to long-distance transport from Hebei Province and the air mass transport from highly polluted regions.

### 3.3. Source contribution evaluation based on k-means clustering

Particle size distributions were divided into three modes, i.e., nucleation, Aitken, and accumulation modes. The k-means procedure first randomly chooses k points from the space represented by the objects that are being clustered into k groups.

The Dunn index (DI; Masiol et al., 2016; Agudelo-Castañeda et al., 2019) is a metric for evaluating clustering algorithms using an internal evaluation scheme based on the clustering results. The higher the Dunn index, the better is the clustering. If the dataset contains compact and well-separated clusters, the diameter of the clusters is expected to be small, and the distance between the clusters is expected to be large. The DI is expressed as

\[ DI = \min_{1 \leq i < j \leq k} \left\{ \min_{m} \left( \frac{d_{\text{min}}(c_i, c_j)}{\max_{1 \leq k \leq j} \text{diam}(C_k)} \right) \right\}, \tag{12} \]

where \( c_i \) and \( c_j \) are the intercluster distance, respectively. Here, the Dunn index is 0.31, and the average silhouette is 0.38. Based on this, six clusters were chosen as the optimal classification scheme, with clusters comprised of points separated by a small distance relative to the distance between clusters (Liang et al., 2020).

**Nucleation mode**: The diurnal evolution shows that the nucleation-mode particle concentration decreased to <5000 # cm\(^{-3}\) around noon, then increased to 10,000 # cm\(^{-3}\) in the evening and overnight (top panel in Fig. 7a). Cluster 1 (dark blue) was the dominant cluster among the six clusters in the southern sector (top panel in Fig. 7b), indicating that local and long-distance transport from the south were the most significant contributors to the PNSD, regardless of the wind speed (0–6 m s\(^{-1}\)). This process was attributed to the heavy traffic during rush hours in these two periods. Cluster 6 (maroon) from the north with high wind speeds (>2 m s\(^{-1}\)) is linked to NPF, with the nucleation-mode particle concentration increasing from sunset onward, reaching a peak (10,000 # cm\(^{-3}\)) in the afternoon, then decreasing into the night (<1000 # cm\(^{-3}\)).

**Aitken mode**: There was an abundant source throughout the day (>4000 # cm\(^{-3}\)), distinguishable from the other clusters (middle panel in Fig. 7a), reflecting the sole influence of this cluster on the Aitken mode. Cluster 4 (yellow) covered almost every direction around the study area at wind speeds ranging from 0 to 3 m s\(^{-1}\), with sources possibly originating from nearby areas (middle panel in Fig. 7b). The Aitken mode may also be easily impacted by local emissions, especially those nearby. Cluster 3 (aqua), observed in the east at higher wind speeds (>3 m s\(^{-1}\)), was a significant contributor at night, suggesting that heavy truck traffic after midnight could be a potential source of Aitken-mode particles.

**Accumulation mode**: Cluster 3, with a source from the northeast, had the highest concentration throughout the day (~3000 # cm\(^{-3}\), bottom panel of Fig. 7a). Clusters 3, 4, and 5 (aqua, yellow, and red, respectively) contributed significantly to the accumulation mode, covering all wind speeds in the range of 0–6 m s\(^{-1}\), suggesting that the accumulation mode had no clear persistent source (bottom panel of Fig. 7b). This cluster showed no clear diurnal variation. Cluster 4 had two peaks: one at approximately noon (3000 # cm\(^{-3}\)) due to NPF growth to 100 nm and the other at approximately midnight (2000 # cm\(^{-3}\)) due to the relatively low PBL height.

---

**Fig. 6.** Concentration-weighted trajectories using the back-trajectory approach for the (a) accumulation, (b) Aitken, and (c) nucleation modes. The blue triangle in each panel shows the location of Beijing (BJ). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
3.4. PMF results

The PMF method has been used for the source apportionment of traffic emissions influencing UFPs (Harrison et al., 2019; Saha et al., 2019; Weber et al., 2019). The PMF case study uses data collected during the summer months (June, July, and August), a time of the year with the highest number of NPF events observed. Data exceeding the 99.5th percentile were removed before performing the analysis to eliminate extreme peak values. Here, $\alpha = 0.01$ and $C_3 = 0.105$ were selected to estimate the uncertainty in the PMF model. Table 2 lists the uncertainties. Separating the mean influence of traffic emissions from the influences of other sources of particle growth is preferred. and two stages of nucleation and growth in NPF can result in particle source apportionment (Masiol et al., 2017; Rönkkö et al., 2017).

The factors resolved by the PMF model include particle modes, gas concentrations, auxiliary atmospheric variables, and turbulence features. The factor that contributed the most was normalized as 1, and the other factors were normalized accordingly (Beddows and Harrison et al., 2019). Factors has divided into regional transport (16.3%), secondary aerosols (19.7%), new particle formation (22.7%), fresh traffic emissions (23.2%), and the urban background (18.1%). Traffic emissions accounted for 72% of the overall influence on the particle distribution, and new particle formation occupied 64% of the UFPs.

**Regional Transport (Fig. 8a):** Dominated by the accumulation mode of particles (assigned a value of 1 and referred to as factor 1), the high CS (0.83) related to pre-existing particles suggests that factor 1 likely reflected a large particle concentration. This factor is also sensitive to RH (0.9), mostly attributed to the hygroscopic growth of particles in the ambient humid environment. Wind from the south also indicates that local transport from the south may be a major contributor to the observed trend, with a large GMD (0.7) and high CoagS (0.7) suppressing the nucleation process by coagulating small particles. NO$_2$ (0.8)

![Fig. 7. (a) K-means clustering of the three modes of particles, and (b) polar cluster plots of the six clusters. From top to bottom: the nucleation (10-25 nm), Aitken (25-100 nm), and accumulation (>100 nm) modes.](image)

**Table 2**

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Species</th>
<th>Units</th>
<th>Description</th>
<th>$\alpha$ coefficient</th>
<th>$C_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPF parameters</td>
<td>GMD</td>
<td>nm</td>
<td>Geometric mean diameter</td>
<td>0.01</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>CS</td>
<td>s$^{-1}$</td>
<td>Condensation sink</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CoagS</td>
<td>cm$^{-3}$s$^{-1}$</td>
<td>Coagulation sink</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNSD modes</td>
<td>Nucleation</td>
<td>dN/(dlogDp) cm$^{-3}$</td>
<td>-10 nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aitken</td>
<td>-</td>
<td>25-100 nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accumulation</td>
<td>-</td>
<td>&gt;100 nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas concentrations</td>
<td>CO</td>
<td>$\mu$g m$^{-3}$</td>
<td>Carbon monoxide</td>
<td>0.03</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>NO$_2$</td>
<td>-</td>
<td>Nitrogen dioxide</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SO$_2$</td>
<td>-</td>
<td>Sulfur dioxide</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O$_3$</td>
<td>-</td>
<td>Surface ozone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical compounds</td>
<td>SO$_4$</td>
<td>-</td>
<td>Sulfate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NO$_3$</td>
<td>-</td>
<td>Nitrate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NH$_4$</td>
<td>-</td>
<td>Ammonium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Org</td>
<td>-</td>
<td>Organics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chl</td>
<td>-</td>
<td>Chloride</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meteorological &amp; turbulence conditions</td>
<td>T</td>
<td>°C</td>
<td>Temperature</td>
<td>0.05</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>RH</td>
<td>%</td>
<td>Relative humidity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ws</td>
<td>m s$^{-1}$</td>
<td>Wind speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TKE</td>
<td>m$^2$s$^{-2}$</td>
<td>Turbulent kinetic energy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>m</td>
<td>Monin–Obukhov length</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
suggesting that local emissions could increase the levels of tracer gases (0.38) and NO (0.7) gas concentrations had significant impacts on factor 1, suggesting that local emissions could increase the levels of tracer gases at the observation site.

**Secondary Aerosols** (Fig. 8b): Factor 2 had a significant relation with the concentration of secondary inorganic aerosols (NO3, assigned a value of 1) due to the rapid generation of mass concentrations involving CHl (0.99) and NH4+ (0.99) compounds. The secondary formation of inorganic aerosols was closely related to winds from the east and south, where the 5th Ring beltway is located. All secondary aerosol compounds display modal structures with a dominant peak above 0.85. The windrose diagram also reveals that the source was from nearby, with the greatest concentrations associated with the southern wind regime.

**NPF 1 (nucleation)** (Fig. 8c): Solar radiation was a strong contributor, causing photochemical oxidation and the formation of O3 (1) in the atmosphere, generating high total kinetic energy (TKE) levels (0.97) along with high wind speeds (0.58). This factor suggests that the vertical mixing process was more predominant than emission-related factors. The nano (0.31), accumulation (0.32), and GMD (0.18) factors were much higher than others, highlighting the transport of urban background particles from the north. The sink effect was reduced, and nucleation was initiated, favoring NPF.

**Fresh Traffic** (Fig. 8d): The Aitken mode (25–100 nm) contributed the most in this case (1), related to the high CoagS (0.78) and large GMD (0.76). Thus, traffic emissions with particles ranging from 40 to 70 nm were most abundant, based on previous studies (E. Lee et al., 1999; Ogulei et al., 2006; Pey et al., 2009; Vu et al., 2015). SO4 (0.25) and the nucleation mode (0.2) formed in the northwest at low wind speeds. The low RH (0.03) did not facilitate hygroscopic growth in this case.

**Urban Background** (Fig. 8e): A solo peak in the Monin-Obukhov length (1) instead of peaks in the wind speed and other parameters was observed under stable atmospheric conditions, reflecting the partial influence of traffic sources at low wind speeds (2–3 m s⁻¹). Gaseous pollution was accompanied by high CS and temperature values. SO2 (0.38) and NO2 (0.16) may be tracers of local emissions from the urban background. The high contribution from the east may reflect the effect of human activities on the urban environment.

**NPF 2 (growth)** (Fig. 8f): This factor indicates the classical contribution to NPF growth. Nano-mode particles (<10 nm) dominated this trend, implying rich cluster transfer from gas-phase particles. A strong increasing signal was observed in the growth process from the nano to nucleation (0.86) and Aitken modes (0.59). A high SO2 (0.5) concentration may favor growth into sulfuric acid. The above characteristics are all typical of a nucleation factor. New particles were formed via photochemical events under high temperature, high wind speed (0.43), and low RH (0.36) conditions. Particle contributions from nearby directions suggest that NPF was not influenced by regional emissions, i.e., traffic or other sources.

**4. Summary**

Urban pollution stems from a variety of sources, in which traffic emissions are prominent. To provide detailed insight into the effect of traffic emissions on urban air quality, this study used comprehensive atmospheric and environmental measurements made from August 2017 and October 2019 near the 5th Ring beltway in Beijing to characterize new particle formation. Also performed was a source apportionment of the contributions from traffic emissions. Integrated parameters were used to identify formation sources, the evolution of particle size distributions, and new particle formation characteristics. The back-trajectory results indicate that regional transport had a significant influence on the nucleation mode. The contributions of different factors to diurnal variations in the evolution of particle size distributions were determined. By coupling with the k-means clustering method and positive matrix factorization, we identified key mechanisms and influential factors on the ultrafine particles observed. They include regional transport (16.3%), secondary aerosols (19.7%), new particle formation (22.7%), fresh traffic emissions (23.2%), and the urban background (18.1%). Traffic emissions are the most significant contributor, accounting for 72% of the overall influence on the particle distribution, especially in the Aitken mode. New particle formation (accounting for 64%) was the major factor significantly affecting ultrafine particles.

**Author contributions**

Zhanqing Li proposed and guided the study; Hao Wu conducted the bulk of the investigation and drafted the original manuscript; Mengjiao
Jiang and Maureen Cribb revised the manuscript; Chunsheng Liang and Tong Wu helped calculate key parameters; Yuying Wang and Dongmei Zhang operated and maintained the instruments used in the study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by National Key R&D Program of China (Grant No. 2017YFC1501702), the Natural Science Foundation of China (NSFC) research project (Grant No. 42030606) and the National Public Welfare Industry (Meteorology) special scientific research project (Grant No. GHY2021406038).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2021.118652.

References


