Fingerprints of the urban particle number size distribution in Helsinki, Finland: Local versus regional characteristics

Tareq Hussein1(2), Bjarke Mølgaard1), Hanna Hannuniemi3), Jyrki Martikainen1), Leena Järvi1), Tobias Wegner4(5), Giovanna Ripamonti6), Stephan Weber4), Timo Vesala1) and Kaarle Hämeri1)

1) Department of Physics, P.O. Box 48, FI-00014 University of Helsinki, Finland
2) The University of Jordan, Department of Physics, Amman 11942, Jordan
3) Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland
4) Technische Universität Braunschweig, Institute of Geocology, Climatology and Environmental Meteorology, Langer Kamp 19c, D-38106 Braunschweig, Germany
5) University of Duisburg-Essen, Faculty of Biology, Department of Applied Climatology and Landscape Ecology, Campus Essen, D-45141 Essen, Germany
6) Politecnico di Milano, DIIAR Environmental Engineering Section, P.za L. da Vinci n°32, IT-20133 Milano, Italy

Received 17 Sep. 2012, final version received 9 Apr. 2013, accepted 27 Feb. 2013


Understanding the fingerprints of urban aerosols is very important in urban model development. Cluster analysis combined with visual classification, air mass back-trajectories, and local meteorology form a comprehensive analysis tool to understand the fingerprints of urban aerosol particles and relate them to their source origin as local or regional. Here we identified seven fingerprints of urban aerosols in Helsinki during 2006. The fingerprints of fresh emissions (Clusters 1–2) from local sources including traffic are characterized by a dominant nucleation mode (GMD < 25 nm and 62%–82% of the submicron particle number concentration). Cluster 3 is characterized by aged ultrafine particle modes with a dominant Aitken mode (diameter 25–100 nm). The fingerprint (Cluster 0) of New Particle Formation (NPF) events is characterized by a second nucleation mode (GMD < 10 nm and a fraction more than 65% of the submicron particle number concentration); the inclusion of particles with \(D_p < 7\) nm in the analysis is important to identify this unique fingerprint. The fingerprints (Clusters 4–5) of aerosols originated via Short-Range or Long-Range Transport (SRT/LRT) from Russia; middle Europe and the Baltic Sea are characterized by dominant Aitken and accumulation modes (as high as 70% of the submicron particle number concentration). Cluster 6 emerged from a mixture between locally emitted aerosols and those originated via SRT/LRT with roughly 50% contribution of the nucleation mode in the submicron particle number concentration. While the data used in this analysis were for the year 2006 only, we foresee the fingerprints are generally valid for the Helsinki Metropolitan Area.
Introduction

Urban areas are of great interest regarding human exposure assessment as well as air quality analysis and modeling because they are hotspots for population and anthropogenic air pollution (e.g. Fenger 1999, Jones 1999). Besides local anthropogenic air pollutants, the urban atmosphere is influenced by regional sources, both natural and anthropogenic.

In the urban atmosphere, the local anthropogenic emissions originate from a wide range of sources: traffic, power plants, and industrial emissions. Traffic emissions are mainly of two major types: tailpipe (combustion) and non-tailpipe (non-combustion). Tailpipe emissions include both primary aerosols, which are emitted at any particle size, and secondary aerosols that are formed with relatively small particle diameters (e.g. Hussein et al. 2007, Maricq 2007, Wehner et al. 2009, Yoon and Lee 2011, Giechaskiel et al. 2012). In subarctic cities, re-suspended road dust as a non-tailpipe source of urban aerosols is a common problem in spring (e.g. Omstedt et al. 2005, Ketzel et al. 2007, Amato et al. 2008). Road dust is accumulated on streets during winter as a result of street sanding and the use of studded tires that wear the road surface. When the road surface becomes dry in spring, the accumulated road dust becomes available for re-suspension behind vehicles. Road dust can also be a problem at any time of the year nearby quarries, industrial areas, and construction sites (e.g. Hussein et al. 2008, DeLuca et al. 2012). While traffic emissions and small-scale combustion emissions are dominant at the surface level, power plants and industrial emissions are typically injected into the urban atmosphere at high altitudes, where they can be further transported across long distances (e.g. Janhäll et al. 2005, Pirjola et al. 2006, Pey et al. 2009, Karppinen et al. 2000, Roldin et al. 2011).

New particle formation (NPF) is commonly observed in the urban atmosphere (e.g. Wehner and Wiedensohler 2003, Wu et al. 2007, Hussein et al. 2008, Salma et al. 2011, Cheung et al. 2012). NPF events might occur locally as a result of secondary particle formation related to local anthropogenic emissions or can occur naturally over a large-spatial scale. A new aerosol particle mode with a geometric mean diameter (GMD) smaller than 30 nm is observed in the particle number size distribution. The newly formed mode is often associated with an observable growth during several hours up to several days depending on the spatial-scale of its occurrence (e.g. Hussein et al. 2009).

Aerosol particles can be transported across several hundreds to couple of thousands of kilometres, where they influence local characteristics of aerosols far away from the original source (e.g. Petäjä et al. 2007, Tunved et al. 2008, Niemi et al. 2009, He et al. 2012). For example, biomass burning can emit large amounts of both secondary and primary aerosols in addition to many gaseous species that affect large areas (e.g. Li et al. 2010, Riche et al. 2012). Saharan dust episodes are also very serious problems during certain seasons (e.g. Wehner et al. 2004, Cheng et al. 2005, Zhang et al. 2006, Hussein et al. 2011). During such episodes the urban particle number size distribution and total particle number concentration are dominated by large particles. Consequently, ultrafine particles (UFP, diameter < 0.1 µm) are scavenged by the large dust particles; therefore, their relative contribution to the total particle number concentration decreases.

Good understanding of the characteristics of local urban aerosols as well as regional aerosols and their sources, dynamics, and interaction with the surrounding is essential in developing better urban air quality models (e.g. Wagstrom and Pandis 2011, Clifford et al. 2011, Thurston et al. 2011, Molgaard et al. 2012), and thus, better human exposure assessment. So far, measurement campaigns as well as long-term measurement initiatives have provided valuable information about the influence of the above-mentioned processes and factors on the urban particle number size distributions. However, only few of them linked the fingerprints (or in other words, signatures) of the urban particle number size distribution to their sources and origin (e.g. Charon et al. 2008, Beddows et al. 2009, Gu et al. 2011). For instance, we previously investigated the spatial and temporal variations of the particle number size distributions within the Helsinki Metropolitan Area (e.g. Hussein et al. 2006, 2005, 2004, Aarnio et al. 2008, Järvi et al. 2009).
We also utilized cluster analysis for a long-term aerosol data set (Wegner et al. 2012), but so far, we have not made an extensive investigation of the clusters of the particle number size distributions to reveal and distinguish the local characteristics from the regional ones.

In this study, we aimed to investigate fingerprints of particle number size distributions observed in the urban background atmosphere of Helsinki, and relate them to their sources of origin, local or regional. For that purpose, we utilized aerosol data collected during the year 2006 at an urban background site in Helsinki. During that year, southern Finland experienced a variety of atmospheric events such as long-range transport (LRT) episodes, new particle formation events, marine air masses, and unique meteorological conditions. Our analysis consisted of two main parts: (1) development and application of a modified classification scheme for urban aerosol characteristics, and (2) a cluster analysis of the particle number size distribution. The cluster analysis provided fingerprints of particle number size distributions, whereas the classification was used to determine the connection between each fingerprint (or cluster) and atmospheric conditions as well as geographic origin. Another important objective of this study was to investigate the improvement in the fingerprint analysis by including particles smaller than 7 nm in diameter in the cluster analysis.

Material and methods

Measurement sites

We utilized particle number size distributions measured at the SMEAR II and SMEAR III stations; the term SMEAR stands for Station for Measuring Ecosystem–Atmosphere Relations. SMEAR III (60°12′N, 24°57′E) is an urban background station that was established in Helsinki (Hussein et al. 2008, 2009), whereas SMEAR II (61°51′N, 24°17′E) represents a remote boreal forest location in Hyytiälä (Hari and Kulmala 2005). SMEAR II is located about 220 km north of SMEAR III (Fig. 1). In addition, meteorological parameters, traffic counts, and air mass back-trajectories were available at SMEAR III. The aerosol data measured at the SMEAR III station was the main focus of this study, whereas
the aerosol data from the SMEAR II station were used to identify the occurrence of regional NPF events as well as the arrival of LRT in southern Finland.

The urban background station SMEAR III was officially started in the year 2004 on the Kumpula campus of the University of Helsinki. The campus itself is located about 5 km north of the Helsinki downtown. The measurement site is located within the campus on the top of a 20-m high hill. The distance between the nearest road and the measurement site is about 200 meters. The surroundings of the campus are heterogeneous with a mixture of residential buildings, urban forest, and small roads. The Helsinki Metropolitan Area is located in southern Finland on the shore of the Gulf of Finland (Baltic Sea). The area comprises four cities (Helsinki, Vantaa, Espoo, and Kauniainen). Helsinki is the capital of Finland with about 600,000 inhabitants, and an area of about 214 km². The Gulf Stream and the prevailing atmospheric circulation provide relatively mild climate as compared with that in most cities at the same latitude.

According to air quality reports by the Helsinki Region Environmental Services Authority (HSY, www.hsy.fi), the particulate matter emissions for the year 2011 were 45% from traffic, 27% from energy production, 5% from the harbors, and 4% from the largest point sources (HSY 2011). In addition to these local sources, Long-Range Transport (LRT) and suspension/re-suspension of aerosol particles is a significant fraction in the PM$_{2.5}$ concentrations (e.g. Kärppinen et al. 2004, Saarikoski et al. 2007, Kauhamiemi et al. 2011).

Particle number size distributions

We utilized the particle number size distributions (dry diameter 3–950 nm, time resolution 5–10 minutes, and size bins 24–48) measured during the year 2006 at both SMEAR II and SMEAR III. At both stations, the measurements were made with a Twin Differential Mobility Particle Sizers (Twin DMPS) consisting of a bipolar charger, a differential mobility analyzer, and a condensation particle counter (e.g. Wiedensohler et al. 2012). The particle number size distributions were extracted from the DMPS by data inversion, which takes into account the particle charging probability, particle losses inside the different parts of the DMPS, and the transfer function of the DMA.

The particle number size distributions were subject to quality assurance individually by the operator of the measurement. Even though the DMPS setups were slightly different at both sites, the laboratory calibration and data quality assurance provide reliable data for further processing and analysis. According to the quality assurance of the SMEAR III data, 11 days were excluded from the analysis because of missing or bad measurements for all size fractions during more than half a day. Another 10 days were also excluded from the analysis because they were reported to have bad measurement for particles smaller than 30 nm or larger than 80 nm during either the whole day or a part of it. In total, we had 344 days valid for this study.

A particle number size distribution can be described with the multi-lognormal distribution function (e.g. Whitby 1978)

$$\sum_{i=1}^{n} \frac{N_i}{\sqrt{2\pi} \log(\sigma_{g,i})} \exp \left[ -\frac{(\log D_p - \log D_{p,i})^2}{2 \log^2(\sigma_{g,i})} \right],$$

where $n$ is the number of individual log-normal modes and $D_p$ (m) is the aerosol particle diameter, and the three log-normal parameters that characterize an individual log-normal mode $i$ are the mode number concentration $N_i$ (cm$^{-3}$), the geometric variance $\sigma_{g,i}^2$, and the geometric mean diameter $D_{p,i}$ (m).

Local weather conditions

Local weather conditions were monitored from the rooftop (about 29 meters from the ground) of the Department of Physics on the Kumpula campus. The Department of Physics is located about 100 meters from the site of the aerosol measurement site. Air temperature and relative humidity were measured with a platinum resistance thermometer and a thin film polymer sensor (Vaisala HMP243, Vaisala Ltd., Vantaa, Finland), wind speed and direction with a cup.
In 2006, the annual variations in the daily air temperature were large with values ranging from –20.9 °C in January to 24.5 °C in July (Fig. 2a). February was exceptionally cold (monthly average –8.2 °C) when compared with the 30-year monthly average air temperature (–4.9 °C; Drebs et al. 2002). July and August in Helsinki (monthly average 19.0 and 18.8 °C, respectively) were warmer than the 30-year monthly average. The daily average wind speed varied between 0.9 and 7.9 m s⁻¹ (Fig. 2b), which is typical for Helsinki, with prevailing wind direction from SW. Due to the high latitude, solar radiation had a strong annual cycle with daytime (10:00–14:00) averages ranging from 10 W m⁻² in winter to 600 W m⁻² in summer (Fig. 2c). The atmosphere was relatively dry in late April and early May with daily relative humidities below 50% (Fig. 2d). The highest rainfall was measured in autumn when the daily values reached 34.5 mm (Fig. 2e).

Air mass back-trajectories

Air mass back-trajectories were obtained from the HYSPLIT 4 model developed by NOAA/ARL. HYSPLIT 4 is a single particle Lagrangian trajectory dispersion model (Stohl 1998). We made model runs by using the Global FNL meteorological archive with 190 x 190 km² spatial resolution, 96 h back in time, and 100 m arrival height every hour. Current literature suggests that the error in a trajectory is within 15%–30% of the travel distance.

We considered the 3-day back-trajectories to distinguish three types of air masses: marine, continental, and mixed. Because of the geographical location of Finland, and specifically Helsinki, about 73% of the air masses had a mixed history; i.e. continental and marine. Less than 9% of the daily air masses were of pure marine history with two main possibilities: (1) originated and passed over the Baltic Sea, or (2) originated over the Atlantic Ocean and found their way towards Helsinki by crossing over Denmark and the Baltic Sea. Therefore, pure continental air masses (about 18%) could (a) originate over the northern parts of Scandinavian countries and crossed over Finland from north to south, or (b) originate over Russia. On some occasions, air...
masses rapidly changed origin and path from one region to another.

Traffic information

Traffic count data from the main highways within the Helsinki Metropolitan Area were provided by the Finnish Traffic Agency. Regardless of the absolute value of the traffic counts, the daily pattern was similar on all main highways. Therefore, traffic count data from one of the closest highways were sufficient as an indicator for the Helsinki Metropolitan Area (Fig. 3).

Cluster analysis

Clustering is an unsupervised classification of patterns into groups, i.e. clusters (e.g. Jain et al. 1999). Although clustering is a difficult combinatorial problem due to differences in assumptions, it has proven its usefulness as one of the steps in exploratory data analysis; and thus, cluster analysis of particle size distributions has been utilized (e.g. Tunved et al. 2004, Dall’Osto et al. 2011, Wegner et al. 2012).

We applied the cluster analysis according to Beddows et al. (2009). The particle number size distributions can be considered a \( n \times m \) matrix, where \( n \) is the number of size bins and \( m \) is the number of size distributions. We normalized the particle number size distribution by using the Euclidian norm and applied the \( k \)-means algorithm from the MATLAB statistics toolbox. For a given number of clusters \( (k) \), the \( k \)-means algorithm minimizes the total sum of the square Euclidian distance from each point to the center of the cluster that it belongs to. The center is the mean of all the points in a cluster. The \( k \)-means algorithm starts with \( k \) random clusters and then modifies these until a local minimum is found. For each \( k \in \{2, 3, ..., 20\} \) of clusters, we ran the algorithm 15 times to improve the chances of finding the global minimum. Finally, the Dunn index as suggested by Beddows et al. (2009) was used as an indicator for the optimum number of clusters that best describes the particle number size distributions. The Dunn index is the minimum distance between points in separate clusters divided by the maximum distance between points belonging to the same cluster.

We utilized the aerosol data in two forms: (1) omitting particle size sections below 7 nm in diameter from the particle number size distributions, and (2) using the whole measured particle size range 3–950 nm performing careful data cleaning for particle size sections below 7 nm. Size sections below 7 nm might include negative values after the inversion process of the DMPS data. After that, we calculated the half-hourly medians for the particle number size distributions and applied the cluster analysis.

Classification scheme for urban aerosols

We developed a classification scheme for the
particle number size distributions based on the classification suggested previously by Hussein et al. (2008) for regional NPF events observed in the urban atmosphere of Helsinki. As compared with our previous classification, the current classification was adopted for urban conditions, local sources, and long-range transport based on the findings by Hussein et al. (2004, 2009). Therefore, the current classification consisted of three main classes: (1) regional new particle formation events (NPF), (2) short-range and long-range transport episodes (SRT/LRT), and (3) clear influence of traffic emissions (TRAFFIC). Additional flags were attached to each main class in order to further distinguish individual subclasses.

The classification was made on a daily basis. A simple routine was created in MATLAB to plot the particle number size distribution spectra at both SMEAR II and SMEAR III stations for several days along with local weather conditions and air mass back-trajectories. A researcher examined these plots visually and classified each day to at least one class and assigned an additional flag. The classification was performed by two additional researchers independently, after which the three classification results were revised by an expert to match similarities and approve differences.

Influence of traffic emissions (TRAFFIC)

In this class, we looked for daily patterns in the particle number concentration of both submicron and ultrafine fractions that are similar to those found in the traffic count daily patterns (Fig. 3). The daily pattern of traffic counts in the Helsinki Metropolitan Area is characterized by two peaks on workdays and one peak on weekends (e.g. Hussein et al. 2004); those peaks coincide with the traffic rush hours. This behavior of traffic counts is very common in other cities as well. Based on this fact, this class has two flags: a “clear” traffic influence whenever the particle number concentration of submicron and ultrafine fractions showed a similar daily pattern to that of the traffic counts. Otherwise, the day was classified as “unclear”.

Regional new particle formation events (NPF)

An NPF event is identified whenever a distinctly new mode of aerosol particles emerges in the nucleation mode size range (diameter < 25 nm) and shows a growth pattern for several hours. If both conditions are met, then we identify the day as an “event”. If only one of these conditions is observed, then we identify the day as an “undetermined”. Otherwise, it is a “non-event” day.

The event is flagged as “large-scale” if it is observed at both sites (SMEAR II and SMEAR III). Otherwise it is flagged as “small-scale”. The spatial-scale of an event can be determined from its time-span and the spatial scale of the back-trajectories during the event (e.g. Hussein et al. 2009). The longer the time-span of the growth pattern is, the larger the spatial scale for the regional new particle formation event becomes. It should be noted that regional NPF are rarely observed over a spatial scale as large as that observed during transport of aerosols (SRT/LRT). In fact, the spatial scale of a SRT episode can be equivalent to the spatial scale of a NPF event.

Short-range transport (SRT) and long-range transport (LRT) episodes

We classified a transport episode of aerosols whenever the following conditions were met:

1. The size fraction with diameter > 75 nm has concentrations significantly higher than the seasonal average. This size fraction includes the upper part of the Aitken mode and the whole size range of the accumulation mode.
2. That size fraction does not show rapid variations in either the particle number size distribution or the number concentration.
3. Similar temporal variation in the concentration of that particle size fraction is observed at both SMEAR II and SMEAR III stations.

Depending on the direction and speed of the air mass movement, the difference in temporal variation between the SMEAR stations varies. SMEAR stations lie on a north–south
line and an episode arriving from central Europe would appear first at SMEAR III and later on at SMEAR II after crossing a distance of about 220 km to the north. On the other hand, an episode crossing over southern Finland from either east or west would appear almost at the same time at both SMEAR stations.

We introduced a flag for the air mass back-trajectories during a transport episode as “short-range” or “long-range”. Trajectories observed within a radius smaller than 500 km from the station were remarked as “short-range”, otherwise, the trajectories were recognized as “long-range”. Then for each episode, we calculated the fraction of “short-range” or “long-range” trajectories and reported this fraction as the additional flag for the classified episode.

Here, the “long-range” flag is analogous to the approach of Wagstrom and Pandis (2011), who previously identified trajectories longer than 550 km as long-range. On the other hand, they identified short-range and mid-range transport for trajectories < 100 km and 100–550 km, respectively. In Finnish conditions, we did not find significant differences between episodes within circles with radii between 100 and 550 km, therefore, we only considered SRT and LRT as described above.

Results and discussion

Overview about the average concentrations

The annual average particle number concentration (± SD) of the submicron fraction was 10 700 ± 9400 cm⁻³. The monthly maximum was registered in February (18 600 ± 13 300 cm⁻³); June–December values varied between 5800 and 9500 cm⁻³ (Fig. 4a). The annual average value was similar to the one for the preceding three years (2003–2005). During the following three years (2007–2009), the annual average value decreased below 9000 cm⁻³. However, the seasonal variation was similar throughout 2003–2009.

The ultrafine particle (UFP, diameter < 100 nm) fraction showed a similar trend as that of the submicron fraction (Fig. 4b) and it accounted for more than 90% of the submicron fraction during much of the winter (Fig. 4d). The contribution of ultrafine fraction ranged between 70% and 90% during almost the whole summer and early autumn. Occasionally, the contribution dropped below 70% when Helsinki experienced severe LRT episodes (such as forest fires smoke during May and August), and during the road dust re-suspension episodes in late April. During the 2007 New Year’s Eve, the ultrafine fraction also dropped down to ~65% as a result of extensive fireworks. This was a result of increased accumulation mode (diameter 0.1–1 µm) particle number concentration that enhanced the coagulation sink (Zhang et al. 2010).

The accumulation mode fraction had different seasonal variation than the UFP fraction (Fig. 4c). Its concentration was well below 2000 cm⁻³ during 80% of the year with an overall average concentration 1400 ± 1200 cm⁻³. During the April–May air pollution episodes, the daily mean particle number concentration of the accumulation mode increased to above 4000 cm⁻³, and it was well above 2000 cm⁻³ during the August air pollution episode.

Clusters of the particle number size distribution

We applied the cluster analysis to the particle diameter range 7–950 nm by assuming 1 to 20 clusters. The highest Dunn index value of 0.0336, 0.0328, 0.0348, and 0.0348 was obtained for 3, 6, 9, or 20 clusters, respectively. In practice, it is not wise to describe the size distributions with either too few or too many clusters. For instance, using too many clusters (say 9 or 20) would make the analysis more complicated and it would be difficult to relate each cluster to its origin. Besides that, the differences between the clusters would become very small. On the other hand, too few clusters (say 3) are not enough to explain variations and detailed differences in the particle number size distributions observed in the urban atmosphere. Therefore, six is the optimal number of clusters to interpret the fingerprints of urban background particle number size distributions within the diameter range 7–950 nm in Helsinki (Fig. 5). Repeating the cluster analysis for the particle diameter
range 3–950 nm revealed a 7th cluster (denoted by Cluster 0). The inclusion/exclusion of Cluster 0 did not affect the modal structure or the frequency of occurrence of the other six clusters because its frequency of occurrence was very small (cf. Figs. 5–6 and Tables 1–2).

It should be noted that in our previous study (Wegner et al. 2012), we applied the cluster analysis to three years (2006–2008) and particle diameter range 3–950 nm. Here, we repeated the cluster analysis in a similar way but for two particle diameter ranges in order to evidently show that inclusion of particles smaller than 7 nm is important to reveal the unique fingerprint (Cluster 0) of regional NPF as will be verified later on. Otherwise, it is evident that Clusters 1–6 are characteristic for Helsinki. Charron et al. (2008) suggested seven clusters to characterize the fingerprints of particle number size distributions at a rural background site (Harwell) in the UK. Six of those clusters had very similar modal structure as our Clusters 1–6. In another study,
Fig. 5. Normalized fingerprints of the particle number size distributions: (a) Cluster 1, (b) Cluster 2, (c) Cluster 3, (e) Cluster 4, (f) Cluster 5, (g) Cluster 6, and (i) Cluster 0. The corresponding daily patterns of occurrences of these fingerprints is shown in d and h. Legend for a–c, e–g, and i: solid thick lines: 7-cluster approach by using particle diameter range 3–950 nm, cross (x): 6-cluster approach by using particle diameter range 7–950 nm, and thin gray lines: individual lognormal modes.
Fig. 6. Frequency of occurrence of fingerprints throughout the year 2006 (7-cluster approach by using particle diameter range 3–950 nm, and 6-cluster approach by using particle diameter range 7–950 nm): (a) Cluster 1, (b) Cluster 2, (c) Cluster 3, (d) Cluster 4, (e) Cluster 5, (f) Cluster 6, (g) Cluster 0.
Beddows et al. (2009) suggested 10 clusters for the rural site and 15 clusters for the urban sites.

**Fingerprints of the urban particle number size distributions**

**Traffic influence**

We classified 132 days as clear TRAFFIC; this is about 38% of the valid days. The majority of these days were during January–March and October–December, when low radiation and low temperature led to a shallow boundary layer where surface emissions are trapped within a small air volume. The low occurrence of clear TRAFFIC days does not mean that traffic has a minor influence at an urban background site; but instead, other factors such as local meteorology influence the urban aerosol characteristics besides traffic emissions (e.g. Kassomenos et al. 2012, Pey et al. 2008, Weber et al. 2008).

In order to clearly understand the effect of traffic on urban aerosols, we omitted days that were simultaneously classified as NPF events and/or SRT/LRT episodes. This yielded 8 weekend days and 32 workdays. The analysis of these days revealed that Clusters 1, 2, and 3 were the dominant ones with a combined frequency of more than 80% (Fig. 7). As clearly seen in Fig. 7f, Clusters 1 and 2 occurred mainly during the daytime, and thus, they are most likely representing fingerprints of fresh traffic emissions. According to measurements in real-life conditions as well as dynamometer tests, the modal structure of fresh aerosols from tailpipe traffic emissions is characterized by dominant nucleation and Aitken modes as a result of secondary/primary particles (e.g. Harris and Maricq 2001, Lonati et al. 2001, Charron and Harrison 2003, De Filippo and Maricq 2008, Klose et al. 2009, Keskinen and Rönkkö 2010, He et al. 2011). In the afternoon, when the traffic activity declines, the nucleation and Aitken mode particles that were freshly emitted into the urban atmosphere undergo condensation, coagulation, and dilution processes that in turn shift the diameter of the particle number size distribution towards greater

**Table 1.** Modal structure of the particle number size distribution fingerprints. $D_{pg}$ is the mode geometric mean diameter, $\sigma_{pg}$ is the mode variance, and $N$ is the mode number concentration.

<table>
<thead>
<tr>
<th>Mode 0</th>
<th>Mode 1</th>
<th>Mode 2</th>
<th>Mode 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{pg}$ (nm)</td>
<td>$\sigma_{pg}$</td>
<td>$N$ (%)</td>
<td>$D_{pg}$ (nm)</td>
</tr>
<tr>
<td>Cluster 0</td>
<td>5.1</td>
<td>2.0</td>
<td>64.6</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Table 2.** Percentile occurrence of clusters based on the half-hourly average aerosol data, their interpretation in brief, spatial range, and origin of air mass back-trajectories.

<table>
<thead>
<tr>
<th>6-cluster</th>
<th>7-cluster</th>
<th>Fingerprint</th>
<th>Spatial scale/transport range</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 0</td>
<td>–</td>
<td>4.02</td>
<td>New particle formation</td>
<td>Large-scale</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>15.78</td>
<td>14.98</td>
<td>Fresh emissions from transport</td>
<td>Local</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>19.58</td>
<td>17.67</td>
<td>Fresh emissions from transport</td>
<td>Local</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>16.36</td>
<td>15.44</td>
<td>Aged aerosols/soot</td>
<td>Local, small-scale</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>16.89</td>
<td>15.01</td>
<td>Soot/marine</td>
<td>Short-range</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>12.63</td>
<td>14.05</td>
<td>Forest fire smoke</td>
<td>Long-range</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>18.75</td>
<td>18.83</td>
<td>Mixed/urban background</td>
<td>–</td>
</tr>
</tbody>
</table>
values. This results in Cluster 3 as a fingerprint of aged urban aerosols. The wind sector analysis for workdays and especially during 1 January–15 April, 1 October–15 November, and 15–31 December supports that Clusters 1–3 originated from local sources (Fig. 8). The nearby potential sources around the measurement site are: downtown area, main railroad (Helsinki), and ships harbours located to the south (150°–210°); the second main railroad (Pasila) and cargo center located to the northwest (270°–330°); main road towards the downtown via Kalasatama (30°–180°); and small roads and residential areas located to the north of the measurement site.

**Regional new particle formation events**

We classified 30 days as NPF event. Seven of these events were observed at SMEAR III only, which indicates a small spatial-scale. Typical for Scandinavian region, the origin of air masses during NPF events was from the North Atlantic Ocean passing over continental and
marine areas. The fingerprint of NPF events is Cluster 0, which is evident from its occurrence being as high as 60% around noon on the days classified as NPF events. The modal structure of the NPF fingerprint here is different than that suggested by Charron et al. (2008), which had a larger GMD for the nucleation mode. Note that the cluster analysis by Charron et al. (2008) was applied for particle number size distributions within the diameter range 11–450 nm. According to our analysis, it was not possible to identify Cluster 0 by applying the cluster analysis for the particle size range 7–950 nm. Therefore, it is important to include particles smaller than 7 nm in the cluster analysis in order to correctly identify the unique fingerprint of NPF.

Short-range and long-range transport episodes

We found six transport episodes longer than five days (Table 3): E1 (26 February–12 March), E2 (12–20 April), E3 (24 April–7 May), E4 (15–27 June), E5 (1–27 August), and E6 (13–28 September). We verified these episodes by retrieving reports from the local newspapers and scientific reports about LRT issued by the Helsinki Region Environmental Services Authority (HSY). During 2006, HSY reported three poor air-quality episodes in the local newspaper “Helsingin Sanomat”. These episodes were consistent with our classification for E1, E3 and E5. The HSY-reported episodes involved forest-fire smoke or days with extremely high PM concentrations. However, our classification revealed three more episodes that were of marine origin.

E1 was an SRT episode and originated from east and southeast crossing over marine and continental regions. The particle number size distributions during E1 were a mixture between local emissions and aerosols from SRT. The fraction of the aerosols originated from local fresh emissions (i.e. Clusters 1 and 2) was around 54% and they were dominant during the daytime (Table 3), whereas Cluster 6 had a fraction around 20% and it appeared during the nighttime. Clusters 3–5 had a fraction around 26% and they appeared as a transition state from aged local aerosols. Cluster 6 is a fingerprint of mixed aerosols between local emissions, which are dominated by UFP, and a foreign mode (diameter > 100 nm) via SRT.
Table 3. Summary of the episodes' characteristics.

<table>
<thead>
<tr>
<th>Short description</th>
<th>E1 Local mixed with continental short-range transport</th>
<th>E2 Local mixed with marine short-range transport</th>
<th>E3 Forest fire smoke via long-range transport</th>
<th>E4 Marine via short-range transport</th>
<th>E5 Forest fire smoke via short-range transport</th>
<th>E6 Marine via long-range transport</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-day back-trajectory history</td>
<td>Short-range fraction</td>
<td>90%</td>
<td>55%</td>
<td>46%</td>
<td>71%</td>
<td>85%</td>
</tr>
<tr>
<td>Source origin</td>
<td>E ↔ SE</td>
<td>SW ↔ W</td>
<td>E ↔ SE</td>
<td>S ↔ W</td>
<td>N ↔ E ↔ S</td>
<td>SW ↔ S</td>
</tr>
<tr>
<td>History of air masses</td>
<td>Continental</td>
<td>40%</td>
<td>–</td>
<td>54%</td>
<td>–</td>
<td>59%</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>60%</td>
<td>67%</td>
<td>46%</td>
<td>57%</td>
<td>41%</td>
</tr>
<tr>
<td>Marine</td>
<td>–</td>
<td>33%</td>
<td>–</td>
<td>43%</td>
<td>–</td>
<td>22%</td>
</tr>
<tr>
<td>Clusters fraction (%)</td>
<td>Cluster 0</td>
<td>5.8</td>
<td>8.8</td>
<td>8.7</td>
<td>0.8</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Cluster 1</td>
<td>17.5</td>
<td>5.4</td>
<td>3.4</td>
<td>1.8</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>Cluster 2</td>
<td>36.3</td>
<td>6.1</td>
<td>4.0</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Cluster 3</td>
<td>11.1</td>
<td>20.5</td>
<td>8.1</td>
<td>10.3</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>Cluster 4</td>
<td>8.9</td>
<td>26.3</td>
<td>3.9</td>
<td>34.1</td>
<td>19.2</td>
</tr>
<tr>
<td></td>
<td>Cluster 5</td>
<td>0.8</td>
<td>9.5</td>
<td>49.0</td>
<td>35.6</td>
<td>38.2</td>
</tr>
<tr>
<td></td>
<td>Cluster 6</td>
<td>19.6</td>
<td>23.5</td>
<td>23.0</td>
<td>14.6</td>
<td>24.8</td>
</tr>
<tr>
<td>Accumulation-to-Aitken mode concentration ratio</td>
<td>Mean</td>
<td>1.1 ± 0.7</td>
<td>2.0 ± 1.7</td>
<td>1.8 ± 1.7</td>
<td>3.4 ± 3.2</td>
<td>2.1 ± 1.7</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Number concentration (cm⁻³)</td>
<td>Nucleation</td>
<td>9000 ± 6850</td>
<td>4600 ± 4500</td>
<td>4950 ± 4850</td>
<td>2100 ± 2200</td>
<td>2900 ± 3400</td>
</tr>
<tr>
<td></td>
<td>Aitken</td>
<td>8100 ± 6850</td>
<td>5300 ± 3700</td>
<td>5700 ± 4100</td>
<td>4150 ± 2700</td>
<td>4000 ± 2400</td>
</tr>
<tr>
<td></td>
<td>Accumulation</td>
<td>1800 ± 1200</td>
<td>1460 ± 770</td>
<td>3900 ± 2300</td>
<td>1800 ± 870</td>
<td>2450 ± 1500</td>
</tr>
<tr>
<td></td>
<td>Total submicron</td>
<td>18700 ± 13150</td>
<td>11200 ± 7500</td>
<td>14400 ± 9050</td>
<td>7950 ± 4750</td>
<td>9300 ± 5250</td>
</tr>
</tbody>
</table>
The accumulation-mode/Aitken-mode number concentration ratio exceeded 1 with a maximum value around 6 during the nighttime.

Episodes E2, E4 and E6 represent marine transport episodes during three seasons: early spring, summer and autumn, respectively. During these episodes, the air masses originated from the south and west and the fractions of short-range trajectories were 55%, 71%, and 69%, respectively (Table 3). The particle number size distributions were characterized by a dominant Aitken mode with a GMD between 50 and 100 nm; the longer the range of a trajectory, the larger the GMD. The occurrence of Cluster 4 was about 26.3%, 34.1% and 24.1% respectively, during the marine episodes E2, E4 and E6. Therefore, it can be considered a fingerprint of SRT marine aerosols (Table 3). Note that E6 had about 69% of the air masses trajectories classified as long-range and this explains the low occurrence of Cluster 4 (24.1%) and higher occurrence of Cluster 5 (42.8%) when compared with E2 and E4. During E2, the influence of local aerosols and NPF events was more pronounced than during E4 and E6, and that explains the low fraction of Cluster 4 during that episode.

Episodes E3 (~50% LRT) and E5 (~85% SRT) brought forest fire smoke transported from Russia and eastern Europe (Table 3). To a certain extent, E5 was also characterized by northern air masses, which are usually clean. E3 was stronger than E5 with overall average particle number concentrations around 9500 cm\(^{-3}\) for particles larger than 25 nm. The particle number size distributions clearly showed that forest-fire smoke had rather stable particle mode with GMD close to 100 nm, which can be recognized by Cluster 5 as a fingerprint of forest fire.

Although we claim that Cluster 4 is a fingerprint of marine aerosols, it occurred with a fraction of 19.2% during E5. On the other hand, we claimed that Cluster 5 is a fingerprint of forest-fire smoke but it occurred with a fraction of 42.8% during episode E6. This should not be a contradiction because E5 had a large fraction (about 85%) of short-range trajectories whereas E6 had a large fraction (about 69%) of long-range trajectories. As we mentioned before, the longer the range of a trajectory, the larger the GMD. Anyway, it should be noted that in episodes E2–E6, the accumulation to the Aitken mode concentration ratio was seldom below 1 (Table 3).

Charron et al. (2008) obtained two regional fingerprints (clusters M4 and M6 in their paper) and traced them back to the European continent. These fingerprints are, in fact, very similar to the fingerprints obtained here (denoted by Clusters 4 and 5) that were traced back to the Baltic Sea region, middle and eastern Europe, and Russia. Charron et al. (2008) also suggested a fingerprint (M5) traced back to maritime air masses and this corresponds to our Cluster 4 that was traced back to the Baltic Sea region, where we expect a significant amount of ship emissions. Ships emit large amounts of particles larger than 20 nm, most of which consist of volatile material and soot (Jonsson et al. 2011, Li et al. 2010, Fridell et al. 2008, Kasper et al. 2007). Marine air masses are also loaded with salt particles that usually have a GMD in the Aitken mode size range.

Furthermore, aerosols observed during SRT and LRT episodes are distinguishable from those grown after intense formation rates from several successive NPF events as reported by Hussein et al. (2009). During such events, the newly formed mode of aerosols steadily grows during several days reaching the accumulation mode where it further continues its growth, and the particles in it may act as cloud condensation nuclei or simply be washed out from the atmosphere. In 2006, we observed several events as that described by Hussein et al. (2009) but with a shorter time-span; the most pronounced ones were during 16–19 May and 10–14 June 2006.

**Summary and conclusions**

Urban particle number size distributions have a complex structure in time and space. Spatial differences are due to differences in nearby sources, air mass transport patterns, and topography, whereas the temporal variations are related to the activities of nearby aerosol sources, ambient conditions, and regional characteristics of aerosols. In this study, we identified the fingerprints of particle number size distributions observed in the urban background atmosphere of Helsinki and related them to their sources, local or
regional. Our study consisted of cluster analysis of the particle number size distribution, and a classification scheme designed especially for urban aerosols to distinguish new particle formation (NPF) events, short-range and long-range transport (SRT/LRT) episodes, and the influence of traffic emissions.

We identified seven fingerprints of urban aerosols in Helsinki (Table 1). Clusters 1–3 are the fingerprints of aerosols originated from local sources including transport sector and other anthropogenic activities within the city. Clusters 1 and 2 are characterized by a dominant fresh nucleation mode (GMD < 25 nm and 62%–82% of the submicron particle number concentration). Cluster 3 is characterized by aged ultrafine particle modes with a dominant Aitken mode. In general, the ultrafine modes of these fingerprints are more than 90% of the submicron particle number concentration.

The unique fingerprint (Cluster 0) of NPF events is characterized by a fourth log-normal mode (GMD < 10 nm and a fraction more than 65% of the submicron particle number concentration). According to the cluster analysis, the inclusion of particles smaller than 7 nm in diameter is important to identify this unique fingerprint.

Clusters 4 and 5 are the fingerprints of aerosols originated via SRT/LRT from Russia, middle and eastern Europe, and the Baltic Sea. These fingerprints are characterized by dominant Aitken and accumulation modes with a fraction as high as 70% of the submicron particle number concentration. Cluster 6 emerged from a mixture between locally emitted aerosols and those originated via SRT/LRT with roughly 50% contribution of the nucleation mode in the submicron particle number concentration.

Finally, the cluster analysis combined with visual classification, air mass back-trajectories, and local meteorology forms a comprehensive analysis tool to understand the fingerprints of urban aerosol particles and relate them to their sources of origin, local or regional. Although we applied the analysis here to the year 2006 only, the fingerprints should generally be valid for the Helsinki Metropolitan Area.

Acknowledgement: This study was supported by the Finnish Center of Excellence funded by the Academy of Finland (project 1118615).

References


