Past, present, and future of ultrafine particle exposures in North America

Albert A. Presto*, Provat K. Saha, Allen L. Robinson

Department of Mechanical Engineering and Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA

ARTICLE INFO

Keywords: Ultrafine particles PM$_{2.5}$ Exposure

ABSTRACT

There is growing concern that exposure to ultrafine particles (UFP, particles with diameter < 100 nm) may have health effects distinct from exposure to fine particulate matter mass (PM$_{2.5}$). This investigative review examines spatial and temporal trends in UFP concentrations in North America. We analyze (i) multiyear (2006–2016) datasets from 11 stationary sampling sites and (ii) shorter duration, but highly spatially resolved mobile/stationary sampling data (2017–2019) in three cities (Baltimore, Oakland, and Pittsburgh). UFP concentrations have fallen by an average 30% over the past decade, similar to the reduction of PM$_{2.5}$ mass concentration (35%). UFP reductions are likely a co-benefit of PM$_{2.5}$ and other pollutant regulations. PM$_{2.5}$ have a factor of two to three spatial variation both within and between cities. Traffic is a major factor influencing intra-urban spatial variations. New particle formation (nucleation) is also an important source of UFP in many places. Regulations to reduce SO$_2$ emissions from coal combustion have reduced nucleation events in the Eastern U.S., but some coastal areas with Mediterranean climates still have consistent new particle formation events. Highly spatially resolved UFP exposures in urban areas can be estimated using spatial models such as land use regression (LUR) fit to high spatial resolution data. Data collection for these models often uses mobile monitoring or other short-term sampling strategies because there is not a national-scale monitoring network for UFP. Short-term sampling produces LURs with modest ($R^2 < 0.5$) performance; model performance can be improved with additional sampling. The current ability to estimate exposure at high spatial resolution over larger (e.g., national) scales is limited by a lack of data. We discuss strategies to improve UFP quantification and therefore exposure estimates.

1. Introduction

Exposure to airborne fine particulate matter mass (PM$_{2.5}$, particles smaller than 2.5 μm) dominates the human health burden from environmental pollutants both in the U.S. and globally. (Brauer et al., 2012, 2016; Burnett et al., 2014) PM$_{2.5}$ exposure is associated with cardiovascular disease (Pope et al., 2002; Grame and Schlesinger, 2010), lung cancer (Raaschou-Nielsen et al., 2013; Pope et al., 2002), asthma (Anenberg et al., 2018), and decreased life expectancy (Correia et al., 2013; Pope et al., 2009). These exposures are estimated to contribute around 100,000 deaths per year in the U.S. (Goodkind et al., 2019; Fann et al., 2012, 2013) As a result, ambient PM$_{2.5}$ concentrations have been regulated in the U.S. since 1997. While the health effects of PM$_{2.5}$ exposure are well documented (Pope et al., 2019), there is uncertainty about causation; e.g., whether specific components or PM from specific sources are more or less harmful (Health Effects Institute, 2013; Wyzga and Rohr, 2015; Krall and Strickland, 2017; Krall et al., 2017; Krall et al., 2013).

Ultrafine particles (UFP) are a subset of PM$_{2.5}$. UFP is commonly defined as particles smaller than 100 nm, but other definitions exist (e.g., particles smaller than 200 nm) (Ostro et al., 2015). Because of their small size, UFP often constitute a minority of PM$_{2.5}$ mass, though these particles dominate total number concentrations. Total particle number concentration (PNC) is commonly used as a proxy for UFP.

There are concerns that exposures to UFP have distinct health effects from PM$_{2.5}$ mass. Toxicological studies suggest that health effects of UFP may differ from PM$_{2.5}$ (Schmid et al., 2009; Baldauf et al., 2016). For example, UFPs can penetrate more deeply into the respiratory system and also have a potential for translocation to the brain (Health Effects Institute, 2013; Li et al., 2003). A 2013 review by the Health Effects Institute (HEI) found that there was insufficient epidemiological evidence to conclude that UFPs had distinct health effects from larger particles for either short- or long-term exposures (Health Effects Institute, 2013). The report noted that the existing epidemiology provided suggestive but inconsistent evidence of adverse health effects associated with short-term exposures. Further, it noted a lack of both large and...
long-term epidemiological studies aimed at understanding impacts of chronic exposures. Today, despite significant interest in UFP exposures, the situation remains much the same. Compared to PM$_{2.5}$, relatively few long-term and/or large-scale epidemiologic studies have been conducted to examine the potential health effects of UFP; for example, a recent review by Ohlwein et al. identified only ten epidemiological studies that considered long-term UFP exposures (Ohlwein et al., 2019). Existing studies have not revealed definitive evidence for independent health effects of UFP from PM$_{2.5}$ (Ohlwein et al., 2019; Heinzerling et al., 2016).

There is a large body of literature on the climatology of UFP from the perspective of global climate. This includes the contribution of new particle formation (nucleation) versus primary emissions to cloud condensation nuclei (CCN) concentrations. Globally, nucleation is the dominant source of new particles to the atmosphere and contributes up to half of CCN (Merikanto et al., 2009; Wang and Penner, 2009; Kulmala et al., 2009). Significantly less is known about UFP from a human exposure perspective, especially for chronic exposures in urban areas where the majority of the population lives.

There are several reasons why we know less about UFP exposures than PM$_{2.5}$ exposures. One major reason is the lack of large-scale monitoring networks. The U.S. EPA regulates seven criteria pollutants under the Clean Air Act (PM$_{2.5}$, PM$_{10}$, O$_3$, CO, NO$_2$, SO$_2$, Pb) and there is a nationwide network of ~4000 sites that measure ambient concentrations of these pollutants. In many locations, 20+ years of continuous (e.g., hourly) measurements of criteria pollutant concentrations are available. These data have been used to derive national exposure estimates at high spatial resolution (Kim et al., 2020). They also enable quantification of long-term reductions in pollutant concentrations in both rural and urban areas. UFP is not a criteria pollutant, so there is not a large-scale monitoring network and limited monitoring data exist.

The lack of a national long-term monitoring network means that the available data on UFPs are sparse in both space and time. Some states, including California, have instituted UFP monitoring at a small number of sampling sites, though most states have not. Much of the available UFP exposure data comes from research studies that are limited in scope (e.g., one city) (Masiol et al., 2018) and/or time (e.g., one year or less) (de Jesus et al., 2019; Wolf et al., 2017; Montagne et al., 2015; Weichenthal et al., 2016a; Minet et al., 2018; Saraswat et al., 2013; Hankey and Marshall, 2015). The net result is that there is not a comprehensive, nationwide, long-term picture of UFP concentrations and exposures the way there is for criteria pollutants.

The sparse data are further complicated by spatial heterogeneity of UFP. Like most pollutants, UFP exhibit intra- and inter-city variations. UFP have significant intra-urban variation (Saha et al., 2019a), which means that concentrations measured at one location can be a strong function of the specifics of the sampling site (e.g., proximity to nearby sources). The combination of large intra-city variation and sparse sampling complicates comparisons both between and within cities; in this paper we explore this issue in more detail.

Another complication for characterizing UFP exposures arises from uncertainty in the proper metric or indicator. Ultimately, we want a metric for UFP that can be easily measured and is a robust indicator of health effects. There are problems on both of these fronts. As noted above, UFP are typically defined as the number concentration of particles smaller than 100 nm diameter, though some studies use PM$_{2.5}$ mass (Venecek et al., 2019; Xue et al., 2020). Since measuring “true” UFP often requires expensive instrumentation, many studies use total particle number concentration (PNC, e.g., as measured with a condensation particle counter) as an alternative metric. There is not an established reference method for measuring PNC and there can be substantial differences between instruments using different operating principles. For example differences in the lower cut point can lead to higher UFP concentrations when measuring total particle counts with a condensation particle counter (CPC, lower cut point <10 nm) than with differential or scanning mobility particle sizers (DMPS or SMPS, cut point generally 10–20 nm) (Morawska et al., 2008). There is also insufficient epidemiological data to determine whether “true” UFP, PNC, PM$_{2.5}$, or some other metric are most connected to health.

In this manuscript, we use measured PNC, either total particle counts from a condensational particle counter (CPC) or from integrated particle size distributions, as the metric for UFP concentrations. This is the most common UFP metric in the literature. We use the terms UFP and PNC interchangeably throughout.

A final issue is that the analyses presented here, as with most air pollution exposure estimates, focus on outdoor concentrations as the indicator for exposure. However, people in the U.S. spend ~90% of their time indoors (Klepis et al., 2001). There are important indoor sources of UFP, including cooking, biomass burning for home heating, new particle formation, and gas-burning appliances (Patel et al., 2020; Singer et al., 2006; Bhangar et al., 2011). Outdoor UFP penetrate indoor environments, but the penetration efficiency depends on a number of factors related to both the building (Hodas et al., 2016) and the particle size and composition (Lundgren et al., 2003; Riley et al., 2002). While we recognize the importance of indoor sources and outdoor-to-indoor UFP penetration to human exposure, these issues are beyond the scope of this manuscript.

The goal of this investigative review is to synthesize and evaluate what is known about long-term UFP exposures and trends in North America, and to suggest ways to improve UFP exposure estimates. This includes analysis and synthesis of several existing datasets. We structure the discussion around the following key questions relevant to UFP exposures. These questions are grouped into three categories: current concentrations and spatial/temporal trends, major sources, and exposure assessment:

1. UFP concentrations and spatial/temporal trends
   - What are typical outdoor UFP concentrations?
   - What are the changes over time and how do they compare to PM$_{2.5}$ trends?
   - What are the inter- and intra-city spatial patterns?

2. UFP sources
   - What are the major sources, and how do they influence temporal and spatial patterns?
   - What are contributions of primary and secondary particles?

3. UFP exposures
   - What is the current ability to estimate exposures at the urban and national scales?
   - What is needed to improve exposure estimates?

Each of these categories is a section of the Results and Discussion. Lastly, we provide a future outlook on urban UFP concentrations and exposure assessment. To examine these issues, we analyze two different types of datasets. Long-term trends come from multiyear datasets from 11 stationary sampling sites. Intra-city spatial trends are based on highly spatially resolved mobile/stationary sampling data collected in three cities: Baltimore, MD; Oakland, CA; and Pittsburgh, PA.

2. Methods

2.1. Decadal stationary datasets to assess long-term temporal trends

Fig. 1 shows the locations of the 11 sites used to quantify long-term trends in PNC. A full description of the sites is available in Section S1 and Table S1 of the SI. These sites were selected for three reasons: (1) they provide data on multi-year trends, (2) these sites all have either co-located or nearby PM$_{2.5}$ measurements that can be used to compare trends in PNC and PM$_{2.5}$, (3) these sites span a range of land uses from near road to urban background to rural, so enable us to examine trends across different land use types. Many of the sites are in cities or nearby suburban areas. The rural site is the U.S. Department of Energy sampling site at Southern Great Plains (SGP) in Oklahoma.
The analyses presented here use PNC as the UFP concentration metric. All of the PNC and corresponding PM$_{2.5}$ datasets were cleaned for null or out-of-bounds readings. Hourly average concentrations were used to compute annual average and hourly diurnal average concentrations. The exception was Toronto, where data were reported as monthly averages, so only annual average concentrations are computed.

2.2. Spatially distributed measurements to assess intra-city variability

Mobile measurements of PNC have been conducted to quantify intra-urban spatial patterns. We examined data from three cities: Pittsburgh (2016–2017), Oakland (summer 2017), and Baltimore (summer 2019). Mobile measurements in Pittsburgh (Ye et al., 2020; Li et al., 2019; Gu et al., 2018) and Oakland (Shah et al., 2018) have been described previously. Figs. S2 and S3 in the SI show the sampling domain for all three cities. Briefly, in each city we drove a mobile laboratory in specified neighborhoods on multiple days. PNC concentrations were measured with a water-based CPC (Aerosol Devices MAGIC CPC) mounted inside of a mobile laboratory.

Raw mobile sampling data are aggregated into 200 m grids for analysis. The mean concentration in each grid is computed following Shah et al. (2018). We first calculate the median concentration measured in each grid cell on each sampling day, and then the mean of medians across all sampling days. In order to ensure robust average concentrations that are representative of long-term (e.g., annual) trends, only grid cells with sufficient data collected on multiple days are included for analysis. Previous mobile sampling studies have shown that grid cells with data from 7 or more days provide a reasonable estimate of long-term concentrations (Li et al., 2019; Apte et al., 2017).

For Pittsburgh, we also collected multi-week continuous PNC data at 30 sites with a rotating network of CPCs in winter 2017 and 2018 (Saha et al., 2019a). Section S.1.3 of the SI compares the mobile and stationary data from Pittsburgh. There are not systematic biases between the on-road and fixed-site sampling.

3. Results and discussion

3.1. UFP concentrations and spatial/temporal trends

Fig. 2(A) shows annual average PNC from 2006 to 2016. Of the cities shown in Fig. 1, only four have data that cover this entire period: Boston, Rochester, SGP, and Toronto. The annual mean PNC in those four locations, indicated by the thick black line in Fig. 2(A), fell from ~12,500 to ~8800 cm$^{-3}$, a decrease of 30% between 2006 and 2016. In three of these four locations (Rochester, SGP, and Toronto), the declining trend of annual average PNC is statistically significant ($p < 0.05$).

The reductions in PNC mirror changes in PM$_{2.5}$ across the U.S. (Fig. 2(C)). Over the same period from 2006 to 2016, PM$_{2.5}$ fell by 35%. Two major drivers for the well-documented changes in PM$_{2.5}$ are reduced emissions from motor vehicles and electric utilities (e.g., Fig. S5). (Ridley et al., 2018; Lawal et al., 2018) The similar fractional reductions in PM$_{2.5}$ and PNC suggest that the PNC reductions are a co-benefit of PM$_{2.5}$ controls. For example, diesel particulate filters that are effectively required on new diesel vehicles can reduce direct PM$_{2.5}$ mass and UFP emissions (Preble et al., 2015), though in some cases diesel particulate filters can enhance UFP emissions because of nucleation in the cooling exhaust (Vaaraslahl et al., 2004). Sulfur controls on coal-fired power plants reduce both particulate sulfate mass and the frequency of new particle formation (nucleation) events (Saha et al., 2018a). The impacts of both of these sources on UFP are discussed in further detail below.

Concentrations of source specific PM$_{2.5}$ components that can be co-emitted with UFP had similar decreases from 2006 to 2016. Several transition metals, including those associated with fossil fuel combustion (e.g., Fe, Ni, V) had similar decreasing trends over this period (Hennigan et al., 2019). Additionally, organic PM$_{2.5}$ attributable to vehicle emissions fell by about 50% at the Central LA site between 2005 and 2015 (Altuwayjiri et al., 2021).

While the broad trend is a decrease over time, there are city-specific variations for both PNC and PM$_{2.5}$ mass, some of which can be large. For example, PNC increased by ~50% in Boston from 2010 to 2012, and Redwood City had a ~40% decline from 2015 to 2016. The causes of the site-specific trends are not known and are beyond the scope of this manuscript. Nonetheless, the overall trend is that PNC concentrations...
have decreased. This is consistent with data collected in other locations (Sun et al., 2020).

Figs. 1 and 3 highlight the variability in PNC. Data are for 2016–2017. SGP, a rural, continental-background site, has the lowest annual average PNC (3884 cm⁻³). Many of the urban locations (e.g., Toronto, Boston, Pittsburgh) have PNC ~14,000 cm⁻³ (factor of 3–4 higher than SGP, right axis of Fig. 3). The variation among these urban sites is less than a factor of two. The highest concentrations are at a near-road site in Oakland with PNC >25,000 cm⁻³ (factor of 7 higher than SGP).

One challenge in interpreting these data is separating inter- and intra-city variability (e.g., high concentrations at a specific site because of proximity to UFP sources), because PNC is known to have large intra-urban gradients. For example, are PNC in Central LA intrinsically higher than other cities or does that reflect the nature of the location of the monitoring site? We previously showed for Pittsburgh (Saha et al., 2019a) that there is a factor of 2–3 spatial variation in PM2.5 at the urban scale. Other studies report similar spatial variability in other cities (Wolf et al., 2017; Klompmaker et al., 2015; Rivera et al., 2012). Fig. 3 qualitatively labels the sites based on nearby land use. Concentrations broadly scale with activity level; for example, the near-highway Oakland site has the highest PNC concentrations, and the more suburban and rural sites have lower concentrations. This variability in PNC is significantly larger than the accompanying variation in PM2.5; while Oakland has 7 times higher PNC than SGP, it only has 1.25 times higher PM2.5.

We investigated intra-city variations in PNC with a combination of mobile sampling and distributed stationary sampling. Results for three cities (Baltimore, Pittsburgh, and Oakland) are shown in Fig. 4. In each city, sampling was conducted across a variety of land uses, e.g., near-highway, near local roads, urban background, etc.

Fig. 4 shows the intra-city variability of PNC. The spatial variation, defined here as the ratio of 95th to 5th percentiles, is a factor of 2–3 in each city. Much of this variation is due to traffic, and the highest PNC during mobile sampling in each city was measured near busy roads and in the city center (“downtown”). Other sources, such as industrial emissions, can also contribute to the intra-urban variation. The stationary dataset in Pittsburgh has a similar mean and median PNC as the mobile data, but larger variation (e.g., a larger 95th percentile). This is because the stationary sampling campaign included high concentration sites downwind of large industrial sources. These areas were not included in the mobile sampling domain. Similarly, the stationary dataset included 2 sites that were more suburban/rural than any locations in the mobile sampling domain.

Fig. 4(D) reveals that there are clear inter-city variations. There is a factor of two difference in mean concentration between the three cities. By sampling across a representative portion of each city, we demonstrate that PNC in Oakland are systematically higher than in Baltimore and Pittsburgh. The implication for Fig. 3 is that Oakland does seem to have higher mean PNC than Pittsburgh (and potentially other cities such as Boston and Toronto). However, Fig. 3 overstates the difference between Oakland and other cities because PNC at the near-highway site shown in Fig. 3 is elevated by ~30% relative to the citywide mean (Fig. 4). This underscores the importance of controlling for the location of sampling sites when comparing PNC data between cities.

Both the inter- and intra-city spatial patterns are less prominent for PM2.5. The bottom panel of Fig. 3 shows that there is significantly less inter-city variability for PM2.5 than PNC. All of the sites, including the...
near-highway site in Oakland, are within ~50% (~3 μg m\(^{-3}\)) of SGP for PM\(_{2.5}\). The full range of PM\(_{2.5}\) is a factor of 2, from ~6 μg m\(^{-3}\) in Boston and Rochester to ~12 μg m\(^{-3}\) in Central LA. Additionally, multiple studies have shown that since PM\(_{2.5}\) mass is regional and dominated by secondary components (Fine et al., 2008; Robinson et al., 2007), there is less intra-city variation in PM\(_{2.5}\) than PNC (Li et al., 2019; Wu et al., 2015).

This section addressed three questions pertaining to PNC and spatial/temporal trends. We summarize the findings here.

**What are typical outdoor concentrations?** PNC range from <5000 cm\(^{-3}\) in rural areas to >20,000 cm\(^{-3}\) near roadways. Urban areas have ~10,000–15,000 cm\(^{-3}\).

**What are the trends over time?** PNC have decreased by ~30% from 2006 to 2016. This mirrors reductions in PM\(_{2.5}\) and is likely a co-benefit of PM\(_{2.5}\) regulations.

**What are the inter- and intra-city spatial patterns?** There is roughly a factor of 2–3 intra-city variation of long-term mean concentrations. There is also a factor of 2 variation in the long-term (e.g., annual) mean concentration between cities. A key consequence of intra-city variation is that long-term average PNC can be a strong function of the details of a specific sampling site. This is much less the case for PM\(_{2.5}\), which is a more regional pollutant.

### 3.2. UFP sources: traffic and nucleation

Traffic is a well-known, major source of direct, or primary, UFP emissions. Vehicles primarily emit carbonaceous UFPs consisting of organic carbon (OC) and black carbon (BC) soot (Maricq, 2007). BC is formed during combustion under fuel-rich conditions in diesel engines or in fuel-rich zones in gasoline engines. OC components include lubricating oil and products of incomplete combustion (Li et al., 2016; Worton et al., 2014). OC in vehicular UFP can occur as coatings on BC cores and as pure OC particles that nucleate in the cooling vehicle exhaust (Li et al., 2018). Vehicular UFP also include contributions from sulfate (due to sulfur in fuel) and metals (e.g., Zn, Ca) that are used as fuel and oil additives (Ronkkö et al., 2007, 2013). Historically, sulfate aerosols have been an important contributor to PNC emissions from diesels, but that has dramatically reduced due to the introduction of low-sulfur fuels (Ruehl et al., 2015).
controlled laboratory experiments (i.e., dynamometer) (Saliba et al., 2017; Herner et al., 2011) and real-world settings such as near roadways (Wang et al., 2017; Saha et al., 2018b; Zimmerman et al., 2016) and inside of traffic tunnels (Li et al., 2018; Ban-Weiss et al., 2009, 2010; Geller et al., 2005). As we demonstrate below, traffic emissions are a major source of urban UFPs and regulations on vehicle emissions are likely a major contributor to the decreasing trend in PNC. Fig. 5 synthesizes near-road and on-road PNC measurements from a number of studies across the U.S. Details of the specific studies, including measurement distance from the road edge and traffic volume on each roadway, are provided in Table S2 in the SI. While there is variability due to differences in location, size of highway, and other study-specific factors, there is a clear decreasing trend over time. Near-road PNC fell by ~20,000 cm$^{-3}$ by 2017; Herner et al., 2011) and real-world settings such as near roadways roadway, are provided in Table S2 in the SI. While there is variability due to differences in location, size of highway, and other study-specific factors, there is a clear decreasing trend over time. Near-road PNC fell by about an order of magnitude, from ~150,000 to <20,000 cm$^{-3}$, between 2001 and 2007. Since 2007, measured near-road PNC have been nearly constant, ~20,000 cm$^{-3}$.

While Fig. 5 combines data collected downwind of roadways of different size, in different parts of the country, and during different seasons, it is unlikely that these factors are responsible for the apparent trend in near-road PNC. For example, measured PNC decrease by about a factor of 3 within the first 100–200 m away from the road edge. (Saha et al., 2018b; Karner et al., 2010) Differences in the distance to the roadway are therefore unlikely to explain the factor of nearly ten change in observed near-road PNC concentrations. UFP emissions are higher in the winter than the summer (Wang et al., 2017); data from Los Angeles in Fig. 5 shows that this impacted near-road PNC by about 20%. Data collected in traffic tunnels over the same time frame also show a significant reduction in UFP emissions from vehicles (Li et al., 2018).

Fig. 5 suggests that the major decrease in near-road PNC occurred before the implementation of the 2007 regulations that effectively required diesel vehicles to use diesel particulate filters (DPF) to control PM$_{2.5}$ emissions. This may be due to manufacturers installing DPFs in new vehicles prior to the start of enforcement. DPFs can control UFP emissions (Herner et al., 2011), but their real-world performance is often less effective than laboratory data (Kozawa et al., 2014). Real-world emissions of BC from diesel vehicles have fallen in recent years; this suggests that DPFs are effective at reducing PM$_{2.5}$ mass emissions from diesel vehicles and may also reduce real-world UFP emissions (Kozawa et al., 2014). The nearly constant near-road PNC from ~2007 - 2018 may therefore indicate that UFP emissions from gasoline vehicles are increasing. The use of gasoline direct injection (GDI) vehicles has increased in recent years, and now represents about half of all new passenger vehicles sold. GDI vehicles have higher UFP emissions than port fuel injected vehicles (Saliba et al., 2017; Zimmerman et al., 2016). The increasing population of GDI vehicles may therefore be counteracting reduced UFP emissions from diesel vehicles.

The impact of traffic emissions is also evident in PNC diurnal patterns in urban areas. Fig. 6 shows average diurnal patterns for Boston, Rochester, and Central LA. In each city there is a clear morning rush hour peak in PNC around 7 a.m. The morning rush hour peak is due to a combination of low atmospheric mixing height in the morning hours along with a spike in emissions from on-road vehicles and other sources at the start of the workday. Thus, cities with higher traffic volumes or densities should have a larger rush hour peak, and reductions in traffic emissions should reduce the size of this peak over time. For example, the rush hour peak is always larger in Boston (Fig. 6(A)), which is a more populous city with larger traffic volumes than Rochester (Fig. 6(B)).

The diurnal patterns also offer insight into long-term changes in vehicular UFP emissions. One way to quantify the direct impact of vehicle emissions on ambient UFP is to calculate the intra-day difference associated with the morning rush hour. The rush hour enhancement is defined as the concentration difference between the morning rush hour peak (7 a.m.) and the overnight minimum (2 a.m.). For Boston, this rush hour enhancement was 12,000 cm$^{-3}$ in 2003 and it fell to 7000 cm$^{-3}$ in 2016, a 42% decline. In Rochester the rush hour enhancement fell from 5000 to 1200 cm$^{-3}$ (76% decline) over the same period. The decrease in the morning maximum is a strong indication that vehicle emissions have decreased.

Figure S6 in the SI shows the time series of the annual average rush-hour enhancement in Boston and Rochester. The trends follow the general trend in PNC in each of the cities shown in Fig. 2 and S4. This suggests that changes in vehicular emissions are a major reason for the overall reduction in PNC shown in Fig. 2.

The importance of vehicle traffic to urban UFP means that this source is a major contributor to the observed spatial variations in PNC. This is an obvious consequence of traffic not being evenly distributed across an urban area. The impacts of traffic on UFP spatial patterns are discussed in more detail in the subsequent section.

Nucleation, which is the photochemically driven formation of secondary UFP from gaseous precursors, is another source of UFPs. Globally it is thought to be the dominant source (Dunne et al., 2016). Nucleation is an important source of UFP in rural and remote areas (Dunne et al., 2016; Kulmala et al., 2013; Dal Maso et al., 2005) and has also been observed in urban areas worldwide (Saha et al., 2018a; Stanier et al., 2004; Brines et al., 2015; Wu et al., 2007). Nucleation creates very little PM$_{2.5}$ mass but creates seed particles for condensation of secondary species such as sulfate and oxidized organics.

In eastern North America, nucleation is typically dominated by conversion of SO$_2$ to sulfate (Stanier et al., 2004). The major source of SO$_2$ is coal combustion for electricity generation. SO$_2$ emissions have fallen dramatically over the last two decades as power plants have either switched to low-sulfur coal, installed flue gas desulfurization, or been replaced by natural gas or renewables (Fig. 7). Nucleation events have become less frequent and less intense as a result. Fig. 7(B) shows data for Pittsburgh, where the frequency of nucleation events (fraction of days per year with observable new particle formation) fell by about half between 2002 and 2017 (Saha et al., 2018a). Similarly, Masiol et al. (2018) observed long-term reductions in the frequency of nucleation events in Rochester. Presumably, the frequency of nucleation events has decreased similarly across the northeastern U.S. as SO$_2$ emissions have fallen.

Nucleation events appear to be a small contributor to total PNC in urban areas of the eastern U.S. (Squizzato et al., 2019) While an individual nucleation event can increase PNC by 10,000–20,000 cm$^{-3}$, each event only lasts for a few hours and, with the reduction in SO$_2$ emissions, nucleation occurs on a relatively small number of days. For Pittsburgh in 2017, Saha et al. (2018a) estimated that 6% of annual average PNC measured at an urban background location could be attributed to local nucleation.

The majority of the nucleation events observed in Pittsburgh in both 2002 and 2017 were regional nucleation events that occurred over

![Fig. 5. Mean PNC measured near roadways across the U.S. since 2001. Full details of each study are given in Table S2 in the SI. The horizontal grey line shows PNC = 20,000 cm$^{-3}$](image-url)
hour enhancement (relative to overnight conditions. The midday PNC peak is nearly as large around 21,400 cm$^{-3}$ in Wood City. For Central LA, PNC peaks during the morning rush hour show the average diurnal trend at Anaheim and Central LA in 2016. At Anaheim, Central LA, Oakland, and Redwood City. Fig. 6(C) and S7 of 2015). This is the case in the four coastal California sites in our dataset: station occurs in coastal areas with Mediterranean climates (Brines et al., 2015). This indicates that most nucleation events in the eastern U.S. do not contribute to intra-urban PNC spatial variations, but could contribute to inter-urban variations.

The decreasing trend of nucleation observed in Pittsburgh and Rochester is not universal across the U.S. Nearly daily midday nucleation occurs in coastal areas with Mediterranean climates (Brines et al., 2015). This is the case in the four coastal California sites in our dataset: Anaheim, Central LA, Oakland, and Redwood City. Fig. 6(C) and S7 show the average diurnal trend at Anaheim and Central LA in 2016. At both locations there is a clear morning traffic peak from 5 to 7 am and a midday nucleation peak around 12-1 pm. A similar trend of morning traffic and midday nucleation peaks is observed at Oakland and Redwood City. For Central LA, PNC peaks during the morning rush hour around 21,400 cm$^{-3}$, which is a rush hour enhancement of $\sim 7000$ cm$^{-3}$ relative to overnight conditions. The midday PNC peak is nearly as large (20,550 cm$^{-3}$) and represents an increase of $\sim 4500$ cm$^{-3}$ over a mid-morning trough in PNC. The regularity of these nucleation events mean that nucleation contributes a significant fraction of urban PNC, approximately 10–15% (Brines et al., 2015; Sowlat et al., 2016), in areas like Central LA.

The consistent nucleation events at the coastal California sites seem to be limited to a narrow band near the coast. These nucleation events are not observed at the two inland sites in southern California (Fontana and Rubidoux; Fig. S7) that are approximately 100 km from the coast. Thus, in these coastal areas, nucleation can contribute to intraurban spatial variations in PNC in large urban areas such as Los Angeles.

Understanding the chemistry underlying the consistent nucleation events in coastal areas with Mediterranean climates is an area of active research. In Barcelona, which has similar nucleation events as the California cities in our dataset, the nucleation events occur under conditions of high ozone and low NO$_x$. Highly oxygenated organic species, perhaps oxidation products of traffic emissions, seem to be important for the growth of nucleated particles to sizes larger than 10 nm (Brean et al., 2020). A key distinction between these coastal areas and inland areas experiencing less frequent nucleation (e.g., Pittsburgh in Fig. 7) seems to be the lack of large anthropogenic SO$_2$ sources upwind of the coastal areas.

Other urban sources of UFP include cooking, biomass burning (e.g., for home heating), natural gas combustion, and industrial emissions (Wolf et al., 2017; Venecek et al., 2019; Ye et al., 2018, 2020; Vu et al., 2015; Liu et al., 2014; Xue et al., 2018; Yu et al., 2019). Overall, there is insufficient data to draw robust conclusions on the contribution of these sources to the long-term trends or spatial patterns of PNC in urban environments. Saha et al. (2019a) showed that industrial emissions can be important PNC sources in areas downwind of major sources. The highest and most variable PNC around Pittsburgh were observed in areas immediately downwind of the largest metallurgical coke facility in North America and the adjacent steel mill. Both our mobile measurements in Oakland and Apte et al. (2017) showed elevated concentrations of UFP near a metals recycling facility. Clearly, industrial emissions can be important sources of UFP, though there is significant heterogeneity in the specific emitting sources and industries between cities. Similarly, contributions of biomass burning to UFP depend on the prevalence of biomass combustion for home heating and the specific stove types used.

Cooking, especially restaurant cooking, is an important source of urban spatial variations in PM$_{2.5}$ (Gu et al., 2018; Robinson et al., 2018; Mohr et al., 2015). These cooking sources also emit UFP, and can contribute to urban UFP enhancements. However, cooking emissions are primarily in the accumulation mode (Kleeman et al., 1999) and are therefore likely more important as sources of PM$_{2.5}$ mass than UFP (Ye et al., 2020).

Emissions from airports can be an important source of urban UFP (Pirhadi et al., 2020; Hudda et al., 2014, 2020). Hudda et al. (2014) used
mobile sampling to demonstrate the impacts of airport emissions on UFPs up to 10 km downwind of Los Angeles International Airport. The datasets analyzed here are not well-suited for capturing airport emission plumes. Among the stationary sites, only Boston (8 km west) and Oakland (9 km north) are within 10 km of a major airport, and in those cases the sites are generally not downwind of the airports. Similarly, the mobile sampling domains were generally far from airports. One of the stationary sites in Pittsburgh from Saha et al. (2019a) was ~5 km downwind of an airport and showed occasional plume impacts. Nonetheless, while the datasets used here may not capture airport plumes, airport emissions may contribute to the urban background in these locations.

Shipping can also be a major source of UFP, especially in coastal areas near large seaports (Kuitinen et al., 2020). Our mobile sampling domains in Oakland and Baltimore included roadways near the major seaports, but in each case UFP concentrations in those areas were not significantly elevated relative to the urban background.

Traffic and nucleation are examples of primary and secondary UFP sources, respectively. Understanding the relative contributions of primary versus secondary sources to UFP is important for designing control strategies that can reduce exposures. If primary UFP dominate, direct emissions controls are more effective, whereas controlling secondary UFP requires reductions in the precursors (e.g., less nucleation in the eastern U.S. by reducing SO₂ emissions in Fig. 7).

Quantifying the primary/secondary UFP split is challenging without data on UFP composition. This is because most source apportionment techniques, such as the chemical mass balance (CMB) method (Schauer et al., 1996) and factor analysis (e.g., positive matrix factorization; PMF) (Paatero and Tapper, 1994; Paatero, 1997), link sources to observed PM mass via composition. Data that only quantifies PNC, even when coupled with size distributions, is therefore insufficient to completely assign UFPs to specific sources.

We used the available data to make a first-order estimation of the primary/secondary split in a typical urban environment. The mean PNC in Pittsburgh is ~14,000 cm⁻³. As noted above, Saha et al. (2018a) previously showed that about 6% (840 cm⁻³) of ambient PNC in Pittsburgh can be attributed to locally nucleated particles. We also assume that concentrations measured at SGP (~4000 cm⁻³) are representative of continental background air masses impacting Pittsburgh; the particles in these background air masses are likely a mix of primary and secondary UFP. The remaining ~9000 cm⁻³ (~66%) are the urban enhancement, a large majority of which are primary UFP emitted by local sources. Since Pittsburgh is the only city where we have detailed measurements of nucleation frequency and intensity, including particle size distributions, we cannot make a similar assessment of primary versus secondary UFPs in the other cities.

The large contribution of local, primary sources to urban PNC is in contrast to PM₂.₅, which is dominated by secondary components in both urban and rural environments. As shown in Fig. 3, urban PM₂.₅ mass enhancement is modest relative to rural areas (<50%) but there is a correspondingly large factor of >2 increase in PNC. Since mass is dominated by aged secondary particles in both urban and rural areas, the additional PNC must be dominated by fresh emissions.

This section addressed questions related to UFP sources. These questions are summarized here:

What are the major sources, and how do they influence temporal and spatial patterns? Traffic is a major UFP source in urban areas. Near road measurements suggest that traffic emissions of UFP fell dramatically from ~2001 to 2007 but appear to have stabilized since then. Traffic is also a dominant factor in diurnal temporal trends and a major contributor to intra-urban spatial patterns. Nucleation is a source of secondary UFP. It has been decreasing in the eastern U.S. because of reduced SO₂ emissions, though there is still persistent nucleation in coastal areas with Mediterranean climates. Sulfate-driven regional nucleation in the eastern U.S. is not a major contributor to intra-urban spatial patterns but can contribute to inter-city spatial patterns.

What are contributions of primary and secondary particles? It is difficult to quantitatively assign UFP as primary or secondary without composition data. However, most of the urban PNC enhancement relative to rural areas is from primary particles. In Pittsburgh, we estimate that ~60–70% of PNC are from local emissions, with traffic being a major source and restaurant cooking an important source.

### 3.3. UFP exposures

Epidemiologic studies require exposure estimates for individual participants. For recent PM₂.₅ studies in the U.S., this has been done at the zip code or census block scale (Pope et al., 2019; Lefter et al., 2019; Di et al., 2017). Some studies in Europe assign exposure based on residential addresses (Raaschou-Nielsen et al., 2013). This typically requires the development of a spatially resolved exposure model. The larger spatial variation of PNC (Fig. 3) makes these models even more important than for PM₂.₅. Exposure models are also used to quantify environmental injustice (i.e., disparities in exposure based on socioeconomic factors) (Clark et al., 2017).

High resolution city-scale UFP exposure models exist for multiple locations in the U.S., (Hankey and Marshall, 2015; Shirmohammadi et al., 2016; Delfino et al., 2010), Canada (Weichenthal et al., 2014, 2016a, 2016b; Minet et al., 2018), and Europe (Wolf et al., 2017; Montagne et al., 2015; Downward et al., 2018). The most common strategy to derive these exposure estimates is to fit a land use regression (LUR) model to measured data. The LUR model is then applied to predict concentrations across the entire domain. UFP concentration fields can also be predicted using chemical transport models and dispersion models (Kukkonen et al., 2016). Since most exposure and epidemiology studies rely on LUR or similar geospatial modeling approaches, we focus on LUR models here.

An example PNC LUR model prediction at 200 m resolution for Pittsburgh is shown in Fig. 8(A) and Table S3. Model building selects the most statistically significant covariates. For Pittsburgh these are the density of urban arterial roads (A3 roads) within a 1500 m circular buffer, the density of road intersections (1000 m buffer), and commercial land use zoning (150 m buffer). All these covariates demonstrate the importance of traffic on PNC spatial patterns. The PNC map in Fig. 8(A) shows a clear hotspot in the downtown city center. This is the area with the highest density of traffic and major roadways, and also had the highest concentrations in the measurements used to build the LUR (Fig. 4).

Many of the published PNC LURs are based on short-term sampling that was specifically conducted in order to build exposure models. The sampling typically uses either mobile monitoring or a rotating network of fixed sites. Short-term fixed-site studies collect data for short periods (e.g., 15–60 min) over a relatively small number of visits (e.g., 1–3) to a set of sampling sites (Montagne et al., 2015; Saraswat et al., 2013; Abernathy et al., 2013). Mobile studies typically sample the same roadways on repeated days, though even multiple passes on the same roadways still yield a small amount of data (e.g., 1–300 s) per road segment (Minet et al., 2018; Hankey and Marshall, 2015).

These sampling strategies may not reproduce long-term average concentrations because they do not fully separate spatial and temporal variability. As a result, LUR models built from these short-term datasets generally have modest or low explanatory power. For example, the LUR in Fig. 8(A) explains slightly more than half of the observed variability in measured PNC (10-fold cross validation R² = 0.58). Saha et al. (2019b) showed that across 16 studies using short-term sampling strategies the average R² for LUR model building was ~0.4. However, these approaches capture the spatial patterns (Paatero et al., 2010), Canada (Wold et al., e.g., higher concentration near major roadways and other sources of primary UFP).

Improved LUR performance can be achieved with longer term sampling (weeks to month), which averages out temporal variability. Fig. 8(B) shows that model performance (R²) improves with increased sampling (Saha et al., 2019b). LUR models built with between one (Eeftens
and six (Saha et al., 2019b) weeks of continuous stationary sampling yielded $R^2$ of $\sim 0.8$. Additionally, Saha et al. (2019b) showed that extending short-term sampling to 1 h of data collection on 10–15 days (rather than 1–3) should produce LURs with $R^2$ of $0.6–0.7$, thereby realizing most of the improvement captured in the multi-week stationary sampling.

The performance of PNC LURs stands in contrast to PM$_{2.5}$ (Shaddick et al., 2018; Chen et al., 2020) and NO$_2$ (Bechle et al., 2015; Novotny et al., 2011) models, which generally have higher predictive skill. In addition, the current generation of PM$_{2.5}$ and NO$_2$ spatial models combines data from national monitoring networks with satellite information to provide continental-scale exposure estimates (Novotny et al., 2011; van Donkelaar et al., 2015). Neither of these resources currently exist for building large-scale PNC exposure models. As a result, PNC exposure estimates only exist at the city scale.

The large spatial variation in PNC further complicates national exposure estimates. The significant (~factor of 2–3) intra-city spatial variation may require a larger number of sampling sites to quantify urban UFP exposures than for other pollutants (e.g., PM$_{2.5}$) (Li et al., 2019). This makes PNC data collection with a stationary network of fixed sites expensive from both an equipment standpoint – condensation particle counters cost ~$10,000 - $25,000 each – and an operations and maintenance standpoint, as each sampling site requires a temperature-controlled enclosure and frequent maintenance. The high cost of data collection is a major motivator for the adoption of mobile and short-term sampling strategies, which enable spatially distributed data collection with a relatively small number of more expensive instruments.

This section addressed questions related to UFP exposures. These questions are summarized here:

**What is the current ability to estimate exposures at the urban and national scales?** High spatial resolution UFP exposures have been estimated in a small number of cities based on targeted, often short duration monitoring campaigns. The resulting spatial models (e.g., LUR models) generally have modest performance ($R^2 < ~0.5$). Increased sampling duration can improve model performance by providing more robust estimates of long-term PNC.

**What is needed to improve exposure estimates?** Improving the performance of city specific LUR models built using short duration sampling data can be achieved by longer sampling times. For example, performing mobile sampling with 10–15 days of data collection per road segment should yield LUR models with $R^2 > 0.6$. Expanding UFP exposure estimates to nationwide may require a large investment in a stationary monitoring network and/or development of remote sensing tools.

**4. Implications and future outlook**

UFP concentrations in the U.S. appear to have decreased by 20–30% over the past 1-2 decades. Long-term reductions in PNC seem to be a co-benefit of PM$_{2.5}$ regulations, specifically reduction in emissions of primary UFP from traffic and reduced frequency of SO$_2$-driven nucleation events associated with the reduction in sulfur emissions from coal-fired power plants.

Nationwide, PM$_{2.5}$ concentrations appear to have stabilized (Fig. S4) or may even be slightly increasing (Clay and Muller, 2019). This may suggest that UFP concentrations will also stop decreasing, which is supported by the data presented in Fig. 2. For several of the sites considered here, PNC have been nearly constant since 2012. Without policy changes aimed directly at PNC, such as the recent adoption of gasoline particle filters in Europe, it is unlikely that there will be large reductions in PNC in North America in the near future. There may be additional reductions in UFP if nucleation frequency continues to decline, but that would seem to be a minor reduction in UFP.

Both inter- and intra-city variations in PNC are larger than corresponding variations in PM$_{2.5}$. Traffic has a major influence on intra-city variations. Barring major changes to traffic patterns or emission control technologies, these gradients will remain. The magnitude of intra-city spatial variations (currently a factor of $\sim 2$) may decrease if vehicular PNC emissions decrease faster than other sources.

Improving estimates for UFP exposures will require an investment in more systematic and spatially representative data collection to characterize both long-term temporal and/or intra-city spatial patterns in PNC across a larger number of cities. New and creative ways to combine different types of data (e.g., mobile sampling, low-cost sensors, or hybrid networks that combine multiple modalities) could also improve UFP exposure estimates. Below, we discuss two strategies to better characterize of UFPs: a national-scale network to characterize long-term and inter-city trends, and a more targeted approach aimed at high resolution exposure assessment in individual cities.

**National UFP trend network:** Some countries (e.g., Germany) (Sun et al., 2020) have implemented a long-term UFP network to quantify trends in inter-city UFP concentrations. In the United States, this could be accomplished by outfitting existing monitoring sites with UFP measurements. This approach is analogous to the existing Chemical Speciation Network (CSN) that tracks PM$_{2.5}$ composition at ~150 sites (a small subset of the total monitoring sites) in the U.S. Some states (e.g.,
California) have already begun deploying UFP instrumentation at select routine monitoring sites.

To better characterize inter-city differences, sampling sites, especially those in urban areas, would need to be placed in locations with similar land uses or source mixes. For example, one could envision sampling at a near-park (e.g., a second at a near-roadway site (e.g., co-located with a near-road monitoring station), and a third at a suburban or peri-urban site. This strategy would provide data on urban enhancements relative to the regional background, long-term inter-city differences based on the urban background and suburban sites, and peak (or near peak) intra-city spatial enhancements due to traffic at the highway sites.

Spatial patterns in targeted cities: A second, complementary need is to systematically quantify intra-urban spatial variations in multiple cities. This can be achieved with either mobile sampling or distributed networks of stationary monitors. Apte et al. (2017) recently demonstrated that routine measurements with vehicle fleets can be used to map urban air pollution, and Messier et al. (2018) showed that these data can be used for building citywide LUR models. Alternatively, Saha et al. (2019b) (n = 30 sites) and studies in Europe (up to 80 sites per city) (Montagne et al., 2015) used rotating networks of fixed sites to collect data for LUR model building. It is also possible to combine data from mobile and stationary sampling (Simon et al., 2018). In this hybrid format, the stationary monitors provide measurements of longer-term variation and the mobile sampling captures spatial variations imposed on top of temporal trends.

It is critical that PNC data be collected with consistent protocols and instrumentation in order to maintain consistency across cities. A sufficient number of samples or length of time is needed to deconvolve spatial and temporal patterns (e.g., Fig. 8(B)). There can be significant differences in UFP estimates obtained from integrated particle size distributions and total particle counts from CPCs (condensation particle counters). Since CPCs are less expensive than particle sizing in filtration systems, differences in UFP estimates obtained from integrated particle size distributions may lead to inaccurate estimates of UFP abundance.

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

We thank Greg Evans (Toronto), Phillip Hopke (Rochester), and Petros Koutrakis (Boston) for sharing data used in this analysis. This work is part of the Center for Air, Climate and Energy Solution (CACES, www.caces.us). This publication was developed under Assistance Agreement no. RD83587301 awarded by the U.S. Environmental Protection Agency (EPA). Mobile sampling in Oakland was partially funded by the Environmental Defense Fund, and mobile sampling in Pittsburgh was partially funded by NSF grant AGS1543786. Boston data was funded by U.S. EPA grants US RD-835872 and RD-834796 to Harvard University. This publication has not been reviewed by the EPA. The views expressed in this manuscript do not necessarily represent those of the funding agency.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aeaoa.2021.100109.

References


resolution exposure to source-resolved atmospheric particle number and chemical mixing state. Environ. Health Perspect. 128 (1).


Further reading