Where Did This Particle Come From? Sources of Particle Number and Mass for Human Exposure Estimates

NEIL M. DONAHUE,* LAURA N. POSNER, DANIEL M. WESTERVELT, ZHONGJU LI, MANISH SHRIVASTAVA, ALBERT A. PRESTO, RYAN C. SULLIVAN, PETER J. ADAMS, SPYROS N. PANDIS AND ALLEN L. ROBINSON

ABSTRACT

Atmospheric chemistry dominates the size distribution and composition of most fine particles inhaled by humans. However, it is important to distinguish between secondary particles—new particles formed in the atmosphere—and secondary mass—molecules formed in the atmosphere that condense to existing particles. In many ways the life stories of particles viewed from the perspectives of particle number concentrations and particle mass concentrations are distinct. Individual particle cores can often be said to have an individual source, while the mass on individual particles comes from myriad sources. This, plus the aforementioned chemical processing in the atmosphere, must be kept in mind when considering the health effects of fine particles.

1 Introduction

Fine particles (PM_{2.5}, or particles smaller than 2.5 micrometers in diameter) induce negative human health effects, including mortality.\textsuperscript{1,2} The current

*Corresponding author.
best estimates are that inhalation of PM$_{2.5}$ causes between 2 and 3 million deaths per year worldwide$^{3,4}$ and roughly 75 million Disability Adjusted Life Years (DALYs).$^3$ The health effects appear worst in vulnerable populations, including children and the elderly, and measurable effects, including mortality, extend to very low ambient concentrations.$^1,5$

The evidence for these effects is strong, yet we lack mechanistic understanding connecting particle exposures to health effects. Toxicologists and epidemiologists agree that both particle size and composition influence the PM$_{2.5}$ dose response.$^6,7$ However, the strongest epidemiological association is with the PM$_{2.5}$ mass concentration, and so the health standard remains a simple mass standard. The common statutory size cutoff at 2.5 micrometers is also a complication; from the perspective of aerosol science alone, a cutoff at 1 micrometer (PM$_1$) would be more sensible, as the sources and composition of the submicron fraction differ in many ways from the supermicron fraction, even for PM$_{2.5}$, where particles between 1 and 2.5 micrometers are most often the small tail of the coarse particle mode.$^8,9$

Here we shall for the most part address only PM$_1$.

There has also been a growing interest in connecting PM health effects directly with sources. However, the relationship between individual sources and PM concentrations faces two significant complications. First, “the source” of an individual particle can often be meaningfully defined when considering PM number concentrations, though particle coagulation causes some ambiguity; however, “the” source of the mass of most particles is not meaningful. This is because a large majority of the mass of most individual particles comes from myriad sources and arrives via condensation. To the extent that health effects are driven by PM mass, “a diesel particle” may not exist very often. Second, much of the PM$_1$ mass is secondary, having been chemically transformed in the atmosphere; and most submicrometer particles are secondary (from new-particle formation). However, most of the mass on submicrometer particles inhaled by people away from urban centers probably resides on primary particles (though the mass is largely secondary). The situation in urban centers and very near roadways is uncertain.

Chemical processing is especially extensive for organics,$^{10}$ which may turn out to be potent drivers of negative health effects; even in urban centers, oxidized organic mass typically exceeds reduced organic mass by 2:1 during the summer.$^{11,12}$ Because of this, individual particles being inhaled, even in urban areas, often contain much more mass from condensation, representing the complete mix of local and regional sources, than within any core that can be associated with a single source. An additional consequence of having vapor condensation dominate fine-particle mass is that atmospheric chemistry can and does play a major role. Thus, while many specific sources certainly provide the raw material for fine-particle mass, chemistry in the atmosphere may be the dominant source of the specific compounds responsible for human health effects.

Here we shall explore what we now know (and do not know) about the different stories of particle-number and particle-mass life cycles, and discuss
how recent developments have transformed our understanding and may influence future health-effects research. For most of the case studies we shall draw on our own work in the eastern United States, but the conclusions are generalizable after considering regional differences in emissions sources; the central point here is not which sources predominate but rather the ubiquity of oxidation chemistry in the process. Recognition of the crucial role played by atmospheric chemistry is essential for designing both control strategies and health-effects studies. Nowhere is this more significant than with organic PM, where the aged organic compounds may have dramatically different effects than the more reduced precursors (including primary organic aerosols) directly emitted from sources.

2 Background

Particulate matter is a uniquely complicated pollutant. This is because “particles” comprise a huge ensemble of objects defined solely by being a condensed phase suspended in the atmosphere. Even if we restrict ourselves to PM$_1$, particles exist over a size range from approximately 1 nm diameter to 1 μm diameter—from just larger than clusters containing a handful of molecules to objects containing more than 10$^9$ molecules—with an equally enormous range of composition and physical properties, such as hygroscopicity and viscosity. Particles and constituents can range from soluble to insoluble, solid to liquid, and spherical to highly structured.

A reasonable picture for most fine particles follows a core–shell model (Figure 1), where a source-specific core is covered with a coating dominating the mass that has a similar composition on most particles. Many cores are directly emitted from primary sources. High-temperature sources dominate urban settings (combustion, cooking, etc.), but sea-spray and biological particles can also contribute to sub-micrometer primary particle emissions.

Different measurement techniques are more or less sensitive to the shell (which is typically non-refractory, relatively uniform, and dominant in mass) and the core (which may be refractory, heterogeneous, and of modest mass). We must emphasize that just the “core” and “shell” may sometimes mix freely and that the shell may include more than one separate phase (for example, an inorganic and an organic phase). In addition, in some cases the “core” is an almost abstract concept, as with new-particle formation, where it simply consists of the very lowest vapor pressure condensates and instead refers to the source (in situ formation) more than to the content of the core.

Because condensation is central to both new-particle formation and condensational growth, it deserves special attention. Particle growth rates are directly proportional to the gas-phase concentration of condensable vapors (the “condensation driving force”). However, condensation to particles, defined by the condensation sink (the suspended surface area concentration multiplied by a deposition velocity) is also generally the major sink of those vapors, and so the balance between production of condensable vapors and the condensation sink will determine the steady-state
Figure 1 A core–shell model of fine-particle composition. Particle mass is dominated by a relatively uniform shell of constituents that arrive via condensation (indicated on the left in the outer ring). These include sulfate, nitrate, ammonia, and organics (some reduced hydrocarbon like organic aerosol, HOA; most oxidized organic aerosol, OOA), and they are measured with various bulk mass measurements including filter assays and online techniques such as Aerosol Mass Spectrometry (AMS). Particle number is dominated by a diverse array of cores with numerous origins and unique signatures (indicated on the right). Cores have source-specific origins and composition, as indicated. Some cores arise from new-particle formation (NPF) and are distinguishable from the shells only by composition; these particles are completely secondary. Cores can often be measured with single-particle mass-spectrometer (SP-MS) techniques.
concentration of those vapors. That in turn will govern both the new-particle formation rate and the particle growth rate.

While it is easy to imagine that new-particle formation might be higher in relatively pristine areas because of the low condensation sink and conversely that growth rates might be higher in polluted areas because of the high production rates, this is not necessarily the case. New-particle formation rates and growth rates are positively correlated and both are buffered by the positive correlation between the production rate of condensable vapors and the total particle mass (and thus surface area) loading. Figure 2 shows the cumulative probability distributions of the condensation sink for a remote site (Hytiälä, Finland) and a polluted site (San Pietro Capofiume, Italy), based on observations originally presented by Westervelt et al. Aside from a clear bump above the 90th percentile associated with extremely hazy events in the Po Valley, the two distributions are simply shifted by roughly 1 order of magnitude, with a median value near 6 h⁻¹ in Hytiälä and near 60 h⁻¹ in San Pietro Capofiume. The condensation sink is the uptake frequency of vapors with particles (here for sulfuric acid, assuming that each collision results in uptake); vapors thus typically collide with particles once every 10 minutes in Hytiälä and once per minute in San Pietro Capofiume. This also defines the equilibration timescale for the aerosol suspension, if mass transfer within particles is sufficiently rapid. In addition, in each location,

![Figure 2](image-url) Cumulative probability distributions of the condensation sink (in h⁻¹) from a remote location (the boreal-forest site of Hytiälä, Finland, “HYY”, left) and a frequently polluted industrial region (the Po Valley regional background site, San Pietro Capofiume, Italy, “SPC”, right).
the enormous majority of cases fall within about one order of magnitude (the slope of the probability distributions is \( \approx 1.0 \)).

3 Particle Mass Concentrations

Particle mass is frequently decoupled from particle number because most mass on every particle in the atmosphere arrives either via condensation of low vapor-pressure compounds formed by gas-phase chemical reactions or via condensation of more volatile compounds followed by a condensed-phase chemical reaction that fixes the material in the condensed phase. This applies to sulfate, nitrate, ammonium, and most organics—even a substantial fraction of organics traditionally regarded as primary organic aerosol (POA) are semi-volatile.22 In total, these constituents usually comprise the large majority of the bulk particle mass11 and even most of the mass on most individual particles.23 What is left is an often relatively refractory core consisting of elemental carbon, metals, perhaps sodium chloride, some crustal material, and other constituents,24 which is typically unique and source specific, in contrast to the relative homogeneity of the particle mass.

The evidence that condensation dominates fine PM mass comes from field measurements, lab studies, and theory. The simplest generic piece of evidence is that the large majority of fine-particulate mass is found on accumulation-mode particles between 100 and 1000 nanometers in diameter, where condensation is by far the dominant growth mechanism, and where particles appear to be generally internally mixed.8 We shall go through more specific evidence for the major constituents in turn.

Sulfate is an almost trivial example as oxidation of SO\(_2\) has long been recognized as the dominant source.8 Both gas-phase oxidation by OH radicals and aqueous oxidation within cloud droplets by HOOH play a major role, but each is “condensation” under a broad definition as each consists of a secondary chemical process that adds to the particle mass by bringing vapors to the particle (in the case of aqueous sulfate, the sulfate condenses on the cloud condensation nucleus if and when the droplet evaporates). Furthermore, on a regional scale sulfate tends to be dominated by coal combustion in areas where PM mass is a health concern,25 and thus control of SO\(_2\) emissions at intense point sources remains the clear control strategy. Near roadways the sulfur content of (diesel) fuel can be a substantial contributor to ultrafine particles.26 On the other hand, because of emissions controls driven by concerns over acid deposition (and because the respiratory tract contains abundant ammonia), sulfate is often substantially neutralized by ammonia; this may limit negative health effects.

Nitrate is broadly recognized as a semivolatile constituent of fine PM, typically found as ammonium nitrate and only present when air masses have excess ammonia. As a semivolatile, it is by definition a condensation product. Nitrate forms via condensation of nitric acid followed by reaction with a base (typically ammonia), and nitric acid in turn is formed principally by gas-phase oxidation of NO\(_2\). NO\(_2\) has a more diverse set of sources
than SO$_2$, though NO$_x$ sources are mostly anthropogenic. Typically, intense emissions from power plants and area emissions from mobile sources are the largest sources, with the mobile sources dominating in urban centers. As with ammonium sulfate, there is little toxicological evidence for negative effects from modest ammonium nitrate exposure. However, for both there is epidemiological evidence that adverse health effects may indeed exist.$^6,^7,^27$

Ammonia is the most common base reacting with both sulfuric and nitric acid—highly volatile on its own it will remain in particles as part of a salt. Sources are mostly agricultural, and once again condensation is the dominant source of PM.$^28$ Because ammonia is very light, ammonium is usually a relatively small constituent of fine-particle mass, though ammonia availability can play a disproportionate role in PM composition.$^29$

Depending on location, organics will comprise from one-third to more than 90% of the fine-particle mass.$^{30}$ Where the story of inorganic mass is fairly simple, the story of the organics is complex. However, a significant finding in the past decade is that the large majority of organic mass is found in an oxidized form, which collects in a mixture that has been named “oxidized organic aerosol” (OOA) in the community of researchers employing the Aerodyne aerosol mass spectrometer (AMS) for in situ measurements.$^{11,12,31}$ In spite of the complexity of the organic fraction, various forms of factor analysis have shown that ambient AMS data can be described by a relatively small set of factors associated with different sources and processes. The OOA factors appear to be associated with atmospheric oxidation chemistry, broadly secondary organic aerosol (SOA), while various factors appear to be associated with different primary organic aerosol (POA) sources.

Little is known about the chemical composition of OOA, only that in aggregate OOA tends to produce similar, highly fragmented mass spectra upon electron ionization and that OOA appears to evolve from a “fresh” form known as semi-volatile OOA (SV-OOA) to an older, more oxidized form called low-volatility OOA (LV-OOA).$^{12,30,32}$ Both forms of OOA are highly oxidized: SV-OOA has an oxygen-to-carbon ratio (O : C) between 0.4 : 1 and 0.6 : 1, while LV-OOA has 0.6 $<$ O : C $<$ 1$^{32-35}$ In general all forms of OOA are more oxidized in the summer than in the winter.$^{32,36,37}$ In addition, thermodenuder measurements suggest that SV-OOA is relatively volatile while LV-OOA is significantly less volatile (thus the names), but in each case the quite broad temperature range of thermal vaporization suggests a relatively broad array of constituents spanning several orders of magnitude in vapor pressure.$^{38-40}$

In remote settings OOA dominates OA mass,$^{11}$ but even in urban centers OOA often makes up more than two-thirds of the OA mass, especially during the summer.$^{11,12}$ The rest of the OA mass consists of several fractions associated with primary emissions. These factors include “hydrocarbon-like organic aerosol” (HOA), which strongly resembles lubricating oil and other emissions associated with vehicles, “cooking organic aerosol” (COA), which shows a temporal pattern strongly correlated with mealtimes,$^{41,42}$ and “biomass burning organic aerosol” (BBOA), which appears to come from biomass burning.$^{12,43,44}$ It is now accepted that OOA largely derives from
condensation. In remote settings, the ratio of HOA to elemental carbon is somewhat less than that observed in urban centers,\(^{45}\) while OOA : EC grows substantially. Even very near urban centers, first SV-OOA then LV-OOA grow rapidly onto fine particles.\(^{35}\)

Laboratory evidence strongly supports this conclusion. All major primary OC emission sources investigated so far are significantly semi volatile,\(^{46–48}\) with 75–90% of the organic mass evaporating from particles as they are diluted from concentrations typical of source-testing environments to concentrations typical of the urban atmosphere. For diesel engines,\(^{48–51}\) gasoline automobile engines,\(^{52–55}\) scooters,\(^{56}\) gas turbine and jet engines,\(^{57}\) and wood burning,\(^{58,59}\) the vapors evolved from these emissions can be rapidly oxidized in the gas phase by OH radicals to produce a sharp increase in OA mass consisting of material that strongly resembles ambient OOA according to AMS analysis.\(^{59–61}\) A quick jump in OOA soon after the maximum of primary emissions from mobile sources is exactly what is observed in Mexico City, for example,\(^ {30,35}\) and this “non-traditional” SOA plays a major role in explaining that jump in model simulations.\(^ {49,62}\) Thus, far from being mostly non-volatile and non-reactive, the vast majority of organic emissions undergo a complex cycle of evaporation, oxidation, and condensation, as depicted in Figure 3.\(^ {63,64}\)

The OOA factors are specific to AMS analysis, but they appear to be highly correlated with another bulk measure—“water soluble organic carbon” (WSOC), which is also measured continuously and online.\(^ {65}\) Once again, WSOC is operationally defined and often dominates the total OC mass. Downwind of urban centers, WSOC mass rises with respect to CO (a roughly conserved tracer of mobile sources),\(^ {45,66}\) again suggesting strong gas-phase formation of condensable organic material.

Another tool for OA source attribution is Fourier Transform Infrared (FTIR) analysis of filter samples.\(^ {67}\) Because FTIR is sensitive to specific functional groups on organic compounds (e.g. aliphatic vs. aromatic carbon), FTIR analysis can provide constraints where the high degree of fragmentation in AMS mass spectra loses information. Residual functional groups can retain source-specific features during atmospheric transport and aging and thus help to constrain OA sources longer than bulk mass spectra. Key molecular markers can also be critical to source attribution, especially in conjunction with secondary molecular markers and bulk measurements, such as AMS spectra.\(^ {68}\) Complementary measurements can be made with NMR, which also provides quantitative measurements of total functional group abundance from OA filter samples.\(^ {69,70}\) Factor analysis of the NMR spectra can, for example, quantify the fraction of biogenic SOA in filter samples\(^ {71}\) as well as the primary to secondary ratio in BBOA.\(^ {72}\)

The only refractory constituent that consistently appears in accumulation mode particles with a significant mass fraction is elemental carbon (EC). However, EC on average comprises only about 10% of the PM\(_{2.5}\) mass—in the Eastern US an OC : EC of 5 : 1 is typical, and it is typical for organics to make up about half of the PM\(_{2.5}\) mass and inorganics the other half. Smaller
Figure 3 The life cycle of organics associated with organic aerosol emissions. Primary emissions typically contain a very large fraction of semi-volatile constituents (labeled “BBOA” and “HOA” in this bulk representation) on a core including black carbon and the very lowest vapor pressure organics. The semi-volatile constituents evaporate and either get oxidized in the gas phase (becoming OOA) or simply re-condense on accumulation mode particles (becoming “condensation” HOA). At the same time, inorganic components condense on the same accumulation-mode particles to form internally mixed aggregate fine particles.
contributions from, for example, crustal material\textsuperscript{73} also occur mostly at the high end of the PM\textsubscript{2.5} size distribution, so this conclusion is even stronger for PM\textsubscript{1}. Consequently, the effective (spherical) core diameter is typically about half of the particle diameter, but most of the mass is in the condensational shell.

Here we must emphasize a crucial point. Any source attribution based on some unique attribute of a single particle—for example clear biological structure or a unique combination of elemental carbon and trace metals—applies only to that core. It is easy to misattribute all of the mass on that particle to the core source—for example to a biological or a diesel particle—and this can lead to a significant over estimation of the significance of particular sources with respect to mass. However, source attribution based on mass composition—either with an AMS or with chemical mass balance (CMB) methods\textsuperscript{74,75}—is intrinsically less vulnerable to “core bias” but may suffer when CMB tracers (especially organics) are oxidized.\textsuperscript{76}

4 Particle Number Concentrations

The fine-particle number budget is far less well understood than the fine-particle mass budget, both locally and globally.\textsuperscript{19,77–79} As with mass, the major questions are the split between anthropogenic and biogenic sources, and the split between primary emissions (sea spray, fires, fine dust) and secondary formation (new-particle formation). Furthermore, ultrafine particles are extremely short lived (they either grow or are lost to coagulation) and so questions of scale in both measurements and models can strongly influence conclusions.

Now that we have established that particle mass arises largely via condensation, it follows that particle number is to first order divorced from particle mass: each particle must come from somewhere. There are two significant linkages in spite of this: new-particle formation is a specialized form of condensation, and particle surface area (and thus particle number lifetime) is also strongly influenced by condensation. In some sense particle mass devours particle number via coagulation; while this has a second-order effect on individual particle mass (and none on bulk mass), it has a first-order effect on number. This means that the sources of particle number need not have much to do with particle mass, though they might. Even within a class, such as organics, organic aerosol mass and the organic contributors (if any) to new-particle formation need not be the same, nor even correlated.

Because most of the mass on most particles arrives via condensation, it follows that most particles grow substantially during their residence in the atmosphere. The largest source of particles on a global scale is new-particle formation, at the smallest possible particle size (\(\approx 1\) nm), though as we shall discuss urban-scale exposure may paint a different picture.

The next largest source is ultrafine emissions, mostly from combustion, over a size range peaking between 10 and 50 nm. Most emissions measurements,
however, constrain the mass emissions far more accurately than the number emissions, and a shift in the modal size by 50%, conserving mass, implies more than a factor of three change in number. The largest sink of particle number is coagulation. Here by convention larger particles devour smaller ones, and most particle number losses involve the smallest particles (freshly nucleated or emitted) colliding with particles near the surface-area modal size, which is typically 100–300 nm diameter. However, those larger particles themselves will for the most part have grown up via condensation, and so it is far less obvious whether any core in the larger particle will contain more mass than the subsumed smaller particle.

Individual particle measurements can be very informative about sources of particle number, especially because the methods used to analyze single particle data often involve clustering algorithms, and these particle clusters can then be compared to measurements of known sources for ambient particle source identification. Most single-particle analysis techniques—including laser-ablation mass spectrometry, and electron and x-ray microscopies—lack the sensitivity to analyze ultrafine particles ($d_p < 100$ nm). This is unfortunately where the majority of particle number exists, and this hampers our ability to determine the sources of most particles by number. Growing mobility-size selected ultrafine particles by water condensation can facilitate their analysis by single-particle mass spectrometry. This revealed the composition of individual particles measured in La Jolla, CA, USA to be complex mixtures of combustion-derived “cores” (from cars and trucks, biomass burning, and also ships) and (presumably secondary) organic and inorganic material, and a small contribution from metallic cores. Smaller particles, down to 60 nm, contained relatively less secondary material such as ammonium, nitrate, and amines, compared to larger particles.

Other advanced mass spectrometry techniques that determine the composition of small collections of size-selected ultrafine particles have also found these to be complex mixtures of ammonium sulfate, nitrate/nitrogen, organics, and a small contribution from metals. Particles containing oxidized carbonaceous matter were generally neutralized, whereas particles containing unoxidized carbonaceous matter or no carbon at all were acidic. Measurements of nucleation-mode particles with $d_p = 8–10$ nm found these contained mixtures of amines and aminium salts, oxidized organics, and sulfate. The much larger amount of organics compared to sulfate typically measured in these nucleation-mode particles suggests that the organics play an important and even dominant role in the initial rapid growth of newly nucleated particles that is required for them to survive against death by coagulation.

The measurements of ultrafine and nucleation-mode particle composition we do have indicate that even these young and very small particles quickly become mixtures of primary and secondary inorganic and organic components. Any particle in the upper end of the ultrafine size mode will have already coagulated with several other particles and swallowed their cores in
the process, while also accumulating secondary material via condensation and heterogeneous reactions. Thus thinking about a particle as having one “core” that indicates the original source of that particle may be too simplistic as one moves away from the source. Focusing on determining a particle's refractory core is useful for distinguishing between primary and secondary contributions to particle number and mass, as all secondary components are non-refractory (they evaporate at ~500 °C when analyzed) but many primary components (combustion-derived soot, coal fly ash, and some inorganics; mechanically generated crustal, sea spray, and metallic particles) are refractory. As the sources, production mechanisms, and likely the toxicity of primary refractory particles are quite different from secondary particle components, this distinction is valuable for designing effective air pollution control strategies, and motivating PM health effects studies that target the most hazardous sources and components.

Individual particle sources and their compositions can also control the rate and type of secondary components a particle accumulates. Vanadium emitted along with combustion aerosol from ships may catalyze the oxidation of SO₂ to sulfate.⁹⁸ Oxalate, a major component of secondary organic aerosol, was found enriched in vanadium-rich particles measured in Mexico City.⁹⁹ Chelation of particulate metals by organic acids can alter the solubility and toxicity of transition metals, and also enhance the accumulation of secondary organics. The vanadium and other complex metal mixtures such as zinc chlorides found in PM were attributed to refuse burning in Mexico City.²⁴ Refuse burning and industrial and construction activities can emit significant levels of metallic particles, exposing residents and workers in urban areas to unusually high levels of particulate metals. Oxalic acid was also found enriched in Asian mineral dust particles, attributed to acid–base neutralization by alkaline carbonate minerals in the natural dust particles.¹⁰⁰ This demonstrates the potential feedbacks between “natural” and anthropogenic pollutants, where natural particle cores provide surfaces for condensation, and also for heterogeneous reactions that can increase the amount of secondary aerosol produced beyond that formed by condensation alone.

4.1 New-particle Formation

New-particle formation occurs whenever the gas-phase concentration of condensable vapors $C_i^\nu$ exceeds a critical value. Specifically, the saturation ratio $S_i^\nu = C_i^\nu / C_i^o$ must be sufficiently large. There is strong evidence that new-particle formation in the boundary layer almost always involves multiple species,¹⁰¹,¹⁰² usually including sulfuric acid,¹⁰³ but also some mixture of bases (ammonia and amines)¹⁰⁴ and oxidized organics.¹⁰⁵–¹⁰⁷

New-particle formation is surprisingly common in urban areas,¹⁰⁸–¹¹¹ even when they are heavily polluted. Even though the condensation sink of vapors to particles is usually high in polluted areas, the production rate of condensable vapors can also be high, and the end result can be high steady-state
saturation ratios of condensable (nucleating) vapors. The interplay between production and condensational loss is evident in Figure 2. If at the same time the production rate of condensable species (sulfuric acid and highly oxidized organics, specifically) is 10 times higher in the Po Valley than in the boreal forest of Finland (which is roughly true), then the steady-state concentration (and excess saturation) of the potential nucleating species will be the same in the two locations.

Many identified new-particle formation events involve well-defined formation and growth patterns known as “bananas” based on their characteristic appearance in particle size distribution plots. An example is shown in Figure 4a for an urban background site in Pittsburgh, USA. Particle size distributions are typically measured with a scanning mobility particle sizer (SMPS), in this case with a lower size limit near 3 nm. Size distributions are plotted as a surface plot over the course of a day, typically with a logarithmic scale. The new-particle formation and growth event thus appears as a curving banana-like feature, typically starting when photochemical activity is high (either near noon or some hours after daybreak). The growth rate of the feature is defined as the rate of growth of the feature mode, \( \frac{d N_{\text{max}}}{dt} \). The banana is curved because of the logarithmic y (diameter) axis—the actual growth rate of the feature is generally linear.

These banana-type events often last for 6–12 hours, during which air can travel 100 km or more. This suggests that many new-particle formation events are regional, and coordinated observations of new-particle formation events have confirmed this. However, there may be some sampling bias, as bananas are easy to identify and are strongly associated with new-particle formation. New-particle formation events in urban areas are not always be associated with “banana” events.

Under most circumstances, growth rates vary between 1 and 10 nm h\(^{-1}\). New-particle formation and growth are not easily separated, and so measured growth rates are a key diagnostic. This is partly because the same species responsible for new-particle formation may be important for growth as well, and the excess saturation of vapors thus drives both processes.

Sulfuric acid vapor measurements play an important role, both because sulfuric acid is thought to participate in new-particle formation in almost all cases and because sulfuric acid is thought to condense onto particles irreversibly. Consequently, the measured growth rate can be compared to the growth rate predicted by condensation of sulfuric acid, \( \Gamma = \frac{G_{\text{obs}}}{G_{\text{SA}}} \), with or without small corrections for co-condensation of ammonia. Sometimes sulfuric acid is sufficient to explain the observed growth, so \( \Gamma \simeq 1 \).

The events in Pittsburgh typified by Figure 4 probably fall into this category. Though the Pittsburgh Air Quality Study did not include sulfuric acid vapor measurements, larger particles (\( d_p > 50 \) nm) measured with an Aerosol Mass Spectrometer were almost pure ammonium sulfate, and \( \text{SO}_2 \) measurements could be used to model new-particle formation events. Those models reproduced the observed growth rates well using only sulfuric
acid condensation.$^{117}$ However, more often $\Gamma > 10^{10,118}$ and this excess growth has been associated with condensation of organics, partly by the process of elimination.$^{16,118}$

Three-dimensional chemical transport models can simulate new-particle formation with reasonable success. The Eastern United States has relatively high sulfuric acid vapor concentrations because of heavy reliance on coal combustion for electrical power generation.$^{74}$ In this case, it is likely that ternary nucleation involving sulfuric acid, ammonia, and water vapor is sufficient to explain new-particle formation.$^{119,120}$ In Figure 4b we show a simulation of the data from Figure 4a taken from a 3D-CTM PMCAMx-UF simulation described by Jung et al.$^{77}$ using the ternary nucleation scheme of Napari et al.$^{119}$ tempered by an empirical scaling factor of $10^{-5}$ to reduce the predicted new-particle formation rate. This factor is now understood to arise because the smallest clusters of ammonia and sulfuric acid do not chemically dissociate into ammonium and bisulfate, rendering them much less stable and thus less likely to nucleate.$^{121,122}$ The model simulations had high skill predicting new-particle formation over a 17 day period (82% predictive success) and generally predicted growth rates in addition to the presence or absence of new-particle formation with good fidelity while somewhat over predicting the number concentrations (by a factor of 2.5).$^{77}$ The overall spatial pattern for the smallest particles in the simulation ($N_{<10}$) is shown in Figure 4c.

Recent field observations suggest that a stronger base such as dimethyl amine may be required to quantitatively explain observed new-particle formation.
formation in areas dominated by inorganic (sulfuric acid + base) nucleation,\textsuperscript{101} while laboratory measurements also suggest that nucleation rates for the ammonia + sulfuric acid system are too slow at a given sulfuric acid concentration to explain the observed new-particle formation.\textsuperscript{123,124} It is thus possible that the empirical scaling factor used to obtain Figure 4a and b is itself a proxy for organic bases that either supplement or replace ammonia. Elsewhere in more remote regions there are strong indications that new-particle formation may involve stabilization of sulfuric acid by highly oxidized organic vapors,\textsuperscript{105,106} and there are at least indications from laboratory experiments that oxidation products of aromatic hydrocarbons can contribute to ternary new-particle formation as well.\textsuperscript{125}

### 4.2 Primary Particle Number Emissions

Primary emissions in the eastern US are dominated by gasoline vehicles and industrial sources, with broad emissions size distributions that peak in the 10–50 nm size range.\textsuperscript{126} Together they constitute slightly more than 70% of the total primary particle number emissions. We show spatial maps in Figure 5a. In each case we derived the number emissions estimates by

![Figure 5](image_url)

**Figure 5** (a) Main primary particle emissions over the eastern United States (in July 2001), in particles km\(^{-2}\) day\(^{-1}\) (red is $8 \times 10^{17}$). Gasoline vehicles and industrial emissions are the major primary sources, reflecting major urban centers and the industrial midwest. (b) Size distributions assumed for gasoline vehicle and industrial emissions to convert data from mass flux inventories to number flux emissions estimates.
converting relatively well constrained mass flux inventories into number fluxes by using literature data on primary particle emissions size distributions.\textsuperscript{127} Those constraints are uncertain, and thus the primary number flux estimates are also highly uncertain. However, it is clear that various combustion sources dominate the primary particle emissions (both number and mass) for sub-micrometer particles, and it is also clear that the mode for those primary emissions generally falls between 10 and 50 nm.

In most urban areas, primary particle emissions are dominated by vehicles. Typically, vehicular emissions have a bimodal distribution. A “soot mode” rich in elemental carbon has a peak in the number distribution between 50 and 100 nm, consisting of particles that form from incomplete combustion in the engine itself.\textsuperscript{128,129} An additional “nucleation” mode, composed primarily of organics, is formed as exhaust vapors are cooled with a number mode around 15–30 nm.\textsuperscript{128,129} Both gasoline and diesel vehicles emit both modes, but the soot mode is more pronounced in diesel emissions. The smaller mode dominates the number of emitted particles. Overall, gasoline vehicles probably contribute more than diesel to the total number of emissions, but the diesel contribution is non-negligible for the size range larger than 50 nm.

Other sources of primary particles that may be important in some regions or episodically include cooking, industrial combustion, biomass burning, and sea spray. Biomass-burning particles tend to be larger, with a number mode around 100 nm\textsuperscript{130,131} and a composition that is richer in organics or elemental carbon for smoldering and flaming combustion, respectively. Emissions of ultrafine sea spray are attested in several measurements,\textsuperscript{132,133} but the fluxes and sizes are poorly characterized.

In general, the relative magnitudes of these fluxes are not well constrained. As mentioned earlier, most number emissions are estimated from mass emissions inventories combined with assumptions about size distributions. The size distribution measurements are more sparse and sometimes potentially inconsistent with the mass emissions, e.g. if the number size distribution measurements do not cover the full size range contributing to particle mass or if a lognormal fit to the number distribution is made that does not represent well the “tail” of the distribution that dominates the mass emissions. The emerging emphasis on ultrafine particles and number concentrations demands more systematic and careful characterization of the number and sizes of primary particles from various sources. Nevertheless, the importance of vehicle emissions is well established. The organic-rich vehicle nucleation mode dominates number size distributions in curbside measurements and also in the urban background, where it is seen to correlate with weekday rush hour traffic patterns.\textsuperscript{134} Tunnel studies have provided very valuable information about vehicle emissions including number fluxes.\textsuperscript{127,135,136} Source apportionment studies for PM\textsubscript{0.1} have pointed to gasoline and diesel vehicles, cooking, residential wood burning, and rail.\textsuperscript{137,138}

The behavior of vehicle emissions is highly dynamic with the potential for semi-volatile organics to both condense and evaporate, shifting the particle
size distribution in the first few minutes after emission. Recent vehicle emissions controls mostly target emissions of the larger “soot mode” particles that dominate the mass emissions even in gasoline vehicles. This has caused some to worry that a mass-focused emissions approach may inadvertently increase the number of emitted particles overall by increasing the number of organic-rich particles forming in the vehicle “nucleation mode”. This is physically plausible and observed to occur in some cases, but the actual effects depend upon the detailed technology deployed and its operation. Tunnel studies that resolve individual vehicle emission factors indicate that there is a class of trucks that are super-emitters for mass emissions and a separate class that were super-emitters for number emissions. Exposure to ultrafine particles from vehicle emissions is one hypothesized explanation for epidemiological data showing adverse health effects in the near-roadway (100–200 m) environment.

4.3 Primary Emissions vs. New-particle Formation

For simulations of the eastern United States during July 2001 using PMCAMx-UF, new-particle formation dominated the total particle number source term and surface concentration as shown in Figure 6a. The $N_{<100}$ concentration is almost identical to the $N_{<10}$ concentration plotted in Figure 4c. However, Figure 6b shows that in the accumulation mode ($N_{100}$), where essentially all of the PM$_1$ mass resides, there is almost no spatial correlation between the number distribution and either $N_{<100}$ or $N_{<10}$. Figure 6c and d show the fraction of particles formed via new-particle formation for both $N_{<100}$ and $N_{100}$.

Even though new-particle formation dominates the ultrafine budget, and even though most particles larger than 100 nm “grew up” from particles in the $N_{<100}$ range, primary emissions for the most part contribute more than half of the total number concentration above 100 nm. This seeming paradox is resolved by the survival probability, which is an extremely strong function of size. The particles that actually survive to 100 nm size are heavily weighted toward the high end of the 1–100 nm size distribution, which is much more robustly represented by primary emissions.

Concentration or composition maps such as the new-particle fractions shown in Figure 6 convey the geographic distribution of a complex and variable field, but they can be misleading when considering human exposure. Many of the grid cells represented by the maps are sparsely populated, while a few are densely populated; urban centers are typically a single grid cell in these maps. One can even see in the lower panel showing the $N_{<100}$ nucleated fraction small lighter spots amid the broad field. Those are major urban centers such as Chicago, Detroit, and New York City. To overcome this bias, we integrated the US census data for 2000 onto the PMCAMx grid used in these simulations. This allows us to calculate population-weighted concentration statistics relevant to human exposure. In Figure 7 we show the cumulative distribution of the nucleated fraction for both $N_{<100}$
and $N_{100}$ in July 2001 based on census data that has been mapped onto the same 36 x 36 km grid used in the PMCAMx-UF model. The result is qualitatively similar to the maps—a large fraction of ultrafine particles being inhaled by residents of the eastern US (in July) are secondary particles, but the fraction is not as high as the maps might suggest. Furthermore, very few of the accumulation mode particles ($N_{100}$) inhaled by humans are secondary.

4.4 Issues of Scale

Another issue important to considering primary vs. secondary ultrafine particles is the question of scale. The simulation results we have shown so far are for a relatively coarse-grid model, with 36 x 36 km grid spacing. A recent empirical source-attribution study in Leipzig, Germany reported the estimated fraction of secondary particles at three sites: the town center, an urban background site (the IFT labs) and a regional background site (Melpitz, well known for new-particle formation studies). Not surprisingly, the authors found that over half of the ultrafine particles in the regional background site were local secondary particles, especially in summer.
Figure 7  Cumulative distributions of exposure to secondary particles (formed via new-particle formation) vs. primary particles, based on July 2001 population in the eastern US and the nucleation fractions shown in Figure 6. (a) Ultrafine particles ($N_{<100}$) and (b) accumulation-mode particles ($N_{100}$).
However, in the urban background, and especially the metropolitan center sites, the primary fraction rose steadily until it dominated the ultrafine number concentration even during midsummer. Likewise, in urban areas with “Mediterranean” climates, a cluster analysis of ultrafine particle observations suggests that new-particle formation can be a significant thought not usually dominant source for ultrafine particle number in urban centers.

We can turn to urban-scale EC observations as a tracer for primary particle distributions. One example of EC concentrations and OC:EC values on an urban scale is shown in Figure 8 for Pittsburgh and Allegheny County, USA. Observations at a succession of 70 sampling sites reached with an instrumented van during 2011–2014 provided the input to a land-use regression model for extrapolation over the county at a nominal spatial resolution of 100 m. Three of the explanatory variables in the land-use regression model were proximity to major highways, proximity to major industrial sources (especially two metallurgical coke works) and elevation (a proxy for river valleys in a region with sharp topography). The land-use regression suggests strong gradients tied to all three buffers, with coincidence (traffic and industrial sites in river valleys) showing by far the highest EC levels. Figure 8b shows OC:EC values, which are largely anti-correlated with the EC pattern because the EC pattern is primary and diminishes with dispersion.

Figure 8 shows large (order of magnitude) variations in both EC concentration and OC:EC at sub-kilometer spatial scales. This can have substantial impacts on (1) population exposures, especially if PM health effects are a function not only of particle mass, but of composition as well, and

![Figure 8](image-url)
(2) evolution of particle number, mass, and composition in near-source versus background locations. The EC emitted along roadways and near industrial areas, as shown above, are emitted at small sizes ($d_p < 100$ nm) and thus at high number concentrations. These particles can serve as cores (condensation sinks) that are subsequently coated with organics and other secondary species as they are transported away from the source areas. The sharp gradient in EC concentration in Figure 8a is driven primarily by dilution of fresh emissions, whereas the reverse gradient of OC:EC in Figure 8b is a consequence of dilution of fresh emissions, mixing with the regional background, and secondary aerosol condensation.

The entire county domain in Figure 8 is approximately one grid box in the regional model simulations discussed above, which is indicated with large dashed rectangles in the figure. Model resolution is known to influence model performance, especially resolving primary emissions with large gradients in their emissions patterns. However, even a $4 \times 4$ km fine grid centered over the greater Paris region (like the smaller squares in Figure 8) only resolved BC in the urban center as a single unstructured plume extending over $\sim 10$ km, while the results shown in Figure 8 suggest much higher resolution variability. The population distribution within the county is also highly non-uniform, with dense urban clusters and a great deal of open space. We do not yet have proxies for the secondary fraction of ultrafine particles for this area, but presuming that the results from Leipzig apply at least qualitatively, it is likely that the cumulative distributions shown in Figure 7 would show a substantially higher fraction of primary ultrafine particles, especially for individuals living near major roadways or industrial sites.

5 Implications for Human Exposure

The bottom line of the discussion above for human exposure is that atmospheric chemistry drives a significant fraction of human exposure to both particle number and fine particle mass. First, concerning particle number, particle number concentrations globally are driven by new-particle formation; however, primary emissions can be an important source of particle number in urban areas. Therefore, it is clear that both mechanisms (primary emissions and new-particle formation) contribute significantly to human exposures, but few studies have derived quantitative estimates. The distributions shown in Figure 7 overestimate the contribution of new-particle formation to particle number exposure because of the relatively coarse model grid. Simply decreasing the grid spacing in regional chemical transport models is unlikely to resolve gradients such as those shown in Figure 8. A key challenge is these strong spatial gradients in particle number near sources. For example, near-road studies show greatly elevated concentrations (factor of five or more) to particle number near roads, consistent with the strong gradients in elemental carbon shown in Figure 8. More
research is needed to quantify the fraction of human exposure to particle number driven by new-particle formation.

The picture of condensation and secondary aerosols is much clearer with respect to human exposure to fine-particle mass than to particle number. The large majority of the PM$_1$ particle mass inhaled under most circumstances (even in urban areas) is formed from the condensation of secondary products formed by atmospheric chemistry. Even near sources, the contribution of direct emissions is generally modest. For example, within a few hundred meters of a roadway the relative increase in fine PM mass is relatively modest, less than 30% over background and much less than the enhancements in particle number.$^{151}$ However, roughly half of those inhaled PM$_1$ particles will contain some sort of primary core, which in many cases will not be water soluble. Because most particles are lost to coagulation (to 100–1000 nm particles forming most of the condensation sink shown in Figure 2), it also follows that most of those particles will actually contain more than one core. Once again, the secondary coating on those particles will include highly oxidized organic compounds, whose health effects are not well studied.

Organic aerosol is a major component of fine PM mass. OA is unique in that it has substantial primary and secondary sources. Over the past decade our understanding of OA sources has dramatically changed from one based on a largely non-volatile, primary dominated aerosol to one that includes a more dynamic aerosol dominated by secondary formation.$^{49,152}$ This evolving understanding of has significant implications for human exposure. These are summarized in Figure 9: compared with traditional models of nonvolatile primary organic emissions, we expect a significant reduction in OA mass near sources as the primary emissions evaporate, followed by progressive increases in OA mass downwind as the vapors are oxidized to form SOA. This is demonstrated in model calculations with a relatively coarse grid$^{49,145}$ but to date has not been implemented in high-resolution urban dispersion models. Broadly, evaporation of the primary emissions smooths out urban:regional gradients, consistent with ambient data from urban:rural pairs in the EPA IMPROVE and EPA STN network.$^{19}$

The dynamic nature of organic emissions is predicted to lead to fairly strong seasonal and latitudinal effects. This is because the vapor pressures are temperature (and thus season) dependent and because oxidant levels (specifically OH) are also believed to have substantial seasonal trends.$^{153–155}$ For this reason, the seasonality of the primary:secondary ratio in OA may be especially strong in the northeastern US and in other locations with strong seasonal temperature fluctuations. Model simulations for the eastern US suggest a strong seasonal variation to the extent of evaporation and aging of OA owing to a combination of reduced saturation concentrations and lower OH radical concentrations in the winter compared to the summer.$^{145}$

However, the most dramatic consequence of incorporating the cycle of evaporation, oxidation, and re-condensation into chemical transport models is a dramatic shift in the composition of organics that people are predicted
to inhale. In Figure 10 we show cumulative exposure distributions to primary and secondary OA based on PMCAMx simulations for the eastern US in July 2001.\textsuperscript{49,145} Note that these are separate CPDs for POA and SOA, so one cannot simply add the exposures. This figure reveals the really dramatic change in predicted human exposure between the traditional view of relatively non-volatile and inert primary emissions and our revised view that the organic emissions are semivolatile and chemically very active. The traditional view, shown in the left-hand panel of Figure 10, predicts that people everywhere will be exposed to more primary than secondary OA and that this difference will be greatest in urban centers (where the highest exposures are found).
Figure 10  Cumulative probability distribution of human exposure in the eastern US for (a) a “traditional” organic aerosol framework dominated by non-volatile OA emissions, and (b) the revised framework including evaporation, oxidation, and recondensation of most primary emissions.
About 15% of the population is exposed to unusually high POA levels in this traditional simulation. In contrast, the revised simulation shown in Figure 10 shows that SOA dominates over POA, at almost all exposure levels, and that the peak exposure is considerably muted—only a few percent of the population is exposed to the highest SOA levels and the spike is less than a microgram above the upper quartile. If we hypothesize a different health effect from POA and SOA, consistent with toxicity studies performed in chamber aerosols,\textsuperscript{156,157} we thus conclude that the revised simulation predicts a slightly lower dose of potentially much more potent material.

It is also not at all clear whether SOA derived from different sources will have different health effects, or to what extent the highly processed OOA in ambient AMS observations has similar health effects to less oxidized SOA typically formed in smog chambers.\textsuperscript{158} Quite distinct primary sources (for example, diesel emissions and wood burning) can be aged to yield material with very similar spectra (each strongly resembling OOA) in an AMS.\textsuperscript{32,60,61} Traditional SOA sources including \textit{\textalpha{}}-pinene SOA and toluene SOA can also ultimately resemble OOA,\textsuperscript{159–161} though often the fresh SOA is a poor analogue for OOA.\textsuperscript{159,162} Unfortunately, at present we know of no epidemiological data that are suitable to test any hypothesized difference between HOA and OOA exposure.

Although there are spatial gradients in particle mass near sources, they are substantially smaller than the gradients in particle number.\textsuperscript{151} Therefore, while there is likely some bias towards secondary organic aerosol in the distributions shown in Figure 10, it is substantially less than for particle number (Figure 7). However, it is estimated that 30% to 45% of people living in large cities live within 300 to 500 m from a highway or a major road, which is the area most highly affected by traffic emissions.\textsuperscript{143} More research is needed to better quantify the contribution of HOA and OOA to human exposure in these sorts of near-source regions.

It is important to stress that the difference between the old and new paradigms for organic aerosol is not confined to some regional background haze; data from Mexico City and elsewhere show that the oxidized form of organic aerosol builds up very rapidly, and that even in areas with very high organic aerosol concentrations (such as these megacities), the OOA constituents dominate the organic aerosol within a few hours on a photo-chemically active day.\textsuperscript{62} This establishes conclusively that the local chemistry is a source (you can not build up high concentrations through dispersion—maxima unambiguously identify source regions). It is only very near sources—for example, very near roadways—that a significant gradient in very fresh versus “aged” (meaning hours, not days) particles is likely to be evident. It is therefore crucial that future human exposure studies take this chemistry into account. Epidemiological studies need to include the significant oxidation of organics likely included in most human exposure, and toxicological studies need to address the processed nature of the organics as well.
Acknowledgements

This work was supported by grant 1447056 from the US National Science Foundation.

References


111. A. Hamed, J. Joutsensaari, S. Mikkonen, L. Sogacheva, M. Dal Maso, M. Kulmala, F. Cavalli, S. Fuzzi, M. C. Facchini, S. Decesari, M. Mircea,


