Particulate matter, air quality and climate: lessons learned and future needs


1Istituto di Scienze dell’Ambiente e del Clima, Consiglio Nazionale delle Ricerche, Bologna, Italy
2Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Switzerland
3Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK
4TNO, Department of Climate, Air and Sustainability, Utrecht, the Netherlands
5Centre for Ecology and Hydrology, Edinburgh, UK
6Department of Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot, Israel
7ETH, Institute for Atmospheric and Climate Science, Zurich, Switzerland
8Department of Chemical Engineering, University of Patras, Patras, Greece
9Department of Applied Environmental Science, Stockholm University, Stockholm, Sweden
10European Commission, Joint Research Center, Institute for Environment and Sustainability, Ispra, Italy
11Environmental Research Group, King’s College London, London, UK

Correspondence to: S. Fuzzi (s.fuzzi@isac.cnr.it)

Received: 28 October 2014 – Published in Atmos. Chem. Phys. Discuss.: 8 January 2015
Revised: 21 May 2015 – Accepted: 4 June 2015 – Published: 24 July 2015

Abstract. The literature on atmospheric particulate matter (PM), or atmospheric aerosol, has increased enormously over the last 2 decades and amounts now to some 1500–2000 papers per year in the refereed literature. This is in part due to the enormous advances in measurement technologies, which have allowed for an increasingly accurate understanding of the chemical composition and of the physical properties of atmospheric particles and of their processes in the atmosphere. The growing scientific interest in atmospheric aerosol particles is due to their high importance for environmental policy. In fact, particulate matter constitutes one of the most challenging problems both for air quality and for climate change policies. In this context, this paper reviews the most recent results within the atmospheric aerosol sciences and the policy needs, which have driven much of the increase in monitoring and mechanistic research over the last 2 decades.

The synthesis reveals many new processes and developments in the science underpinning climate–aerosol interactions and effects of PM on human health and the environment. However, while airborne particulate matter is responsible for globally important influences on premature human mortality, we still do not know the relative importance of the different chemical components of PM for these effects. Likewise, the magnitude of the overall effects of PM on climate remains highly uncertain. Despite the uncertainty there are many things that could be done to mitigate local and global problems of atmospheric PM. Recent analyses have shown that reducing black carbon (BC) emissions, using known control measures, would reduce global warming and delay the time when anthropogenic effects on global temperature would exceed 2°C. Likewise, cost-effective control measures on ammonia, an important agricultural precursor gas for secondary inorganic aerosols (SIA), would reduce regional eutrophication and PM concentrations in large areas of Europe, China and the USA. Thus, there is much that could be done to reduce the effects of atmospheric PM on the climate and the health of the environment and the human population.

A prioritized list of actions to mitigate the full range of effects of PM is currently undeliverable due to shortcomings in the knowledge of aerosol science; among the shortcomings, the roles of PM in global climate and the relative roles of...
1 Introduction

Particulate matter (PM) or, more appropriately, atmospheric aerosol is currently a subject of extensive research, although it was only at the beginning of the 80s that the subject began receiving increased attention from the atmospheric science community (Fig. 1). At present there are 1500–2000 papers per year addressing research topics related to atmospheric aerosols.

Atmospheric aerosols affect air quality and, in turn, human and ecosystem well-being (WHO, 2013a; Fowler et al., 2009); and also have an important role in the Earth’s climate system (IPCC, 2013). Aerosol research efforts are therefore quite numerous on both issues, although the aerosol effect on climate is certainly the subject most widely studied (Fig. 2).

Over the last decade, the specific subjects within atmospheric aerosol science that have received the largest attention (most-cited papers, from the ISI Web of Science) deal with organic aerosols (OAs), new particle formation, aerosol sources and atmospheric budget, radiative forcing of aerosols, and aerosols and precipitation.

From an environmental standpoint, aerosols also constitute an important policy issue in air quality and climate sciences. In fact, PM pollution is probably the most pressing issue in air quality regulation worldwide, and at the same time it represents one of the biggest sources of uncertainty in current climate simulations.

This paper, rather than being a conventional review, aims to summarize the most recent results within the field of aerosol sciences, focusing on those issues on which knowledge is still rather limited and where research efforts should be increased. The paper identifies the science–policy connections within this field of science. Research, in fact, has the strong social responsibility of providing up-to-date results and ensuring their relevance to environmental policy.

The rest of the paper is organized in three main sections: aerosol sources, composition, concentrations and processes (Sect. 2); the most controversial policy issues concerning the impacts of atmospheric aerosol on air quality and climate (Sect. 3); and finally, the future drivers of changing aerosol concentration and the impact on aerosol on climate change are examined (Sect. 4). The acronyms most frequently used in the manuscript are listed in Table 1.

While the present review has a global focus, some sections are more focused towards the European situation, in particular the sections dealing with aerosol concentration levels and trends and air quality legislation and control measures.
Table 1. Definitions of acronyms used in the manuscript.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOD</td>
<td>Aerosol absorption optical depth</td>
</tr>
<tr>
<td>AMS</td>
<td>Aerodyne – aerosol mass spectrometer</td>
</tr>
<tr>
<td>AOD</td>
<td>Aerosol optical depth</td>
</tr>
<tr>
<td>AR</td>
<td>Assessment report</td>
</tr>
<tr>
<td>BC</td>
<td>Black carbon</td>
</tr>
<tr>
<td>BrC</td>
<td>Brown carbon</td>
</tr>
<tr>
<td>BSOA</td>
<td>Biogenic secondary organic aerosol</td>
</tr>
<tr>
<td>BVOC</td>
<td>Biogenic volatile organic compound</td>
</tr>
<tr>
<td>CCN</td>
<td>Cloud condensation nuclei</td>
</tr>
<tr>
<td>CMB</td>
<td>Chemical mass balance</td>
</tr>
<tr>
<td>CTM</td>
<td>Chemical transport model</td>
</tr>
<tr>
<td>EC</td>
<td>Elemental carbon</td>
</tr>
<tr>
<td>ELVOC</td>
<td>Extremely low-volatility organic compound</td>
</tr>
<tr>
<td>ERF</td>
<td>Effective radiative forcing</td>
</tr>
<tr>
<td>ERF_{aci}</td>
<td>Effective radiative forcing due to aerosol–cloud interactions</td>
</tr>
<tr>
<td>ERF_{ari}</td>
<td>Effective radiative forcing due to aerosol–radiation interactions</td>
</tr>
<tr>
<td>GCM</td>
<td>Global climate model</td>
</tr>
<tr>
<td>IN</td>
<td>Ice nuclei</td>
</tr>
<tr>
<td>IVOC</td>
<td>Intermediate-volatility organic compound</td>
</tr>
<tr>
<td>LVOC</td>
<td>Low-volatility organic compound</td>
</tr>
<tr>
<td>LVOOA</td>
<td>Low-volatility fraction of oxygenated organic aerosol</td>
</tr>
<tr>
<td>OA</td>
<td>Organic aerosol</td>
</tr>
<tr>
<td>OC</td>
<td>Organic carbon</td>
</tr>
<tr>
<td>OM</td>
<td>Organic matter</td>
</tr>
<tr>
<td>OOA</td>
<td>Oxygenated organic aerosol</td>
</tr>
<tr>
<td>PAHs</td>
<td>Polycyclic aromatic hydrocarbons</td>
</tr>
<tr>
<td>PBAPs</td>
<td>Primary biological aerosol particles</td>
</tr>
<tr>
<td>PM</td>
<td>Particulate matter</td>
</tr>
<tr>
<td>PMF</td>
<td>Positive matrix factorization</td>
</tr>
<tr>
<td>PN</td>
<td>Particle number</td>
</tr>
<tr>
<td>POA</td>
<td>Primary organic aerosol</td>
</tr>
<tr>
<td>SAPP</td>
<td>Secondary-aerosol production potential</td>
</tr>
<tr>
<td>SIA</td>
<td>Secondary inorganic aerosol</td>
</tr>
<tr>
<td>SOAs</td>
<td>Secondary organic aerosols</td>
</tr>
<tr>
<td>SSR</td>
<td>Surface solar radiation</td>
</tr>
<tr>
<td>SVOC</td>
<td>Semi-volatile organic compound</td>
</tr>
<tr>
<td>SVOOA</td>
<td>Semi-volatile fraction of oxygenated organic aerosol</td>
</tr>
<tr>
<td>TOA</td>
<td>Top-of-the-atmosphere</td>
</tr>
<tr>
<td>TSP</td>
<td>Total suspended particulate matter</td>
</tr>
<tr>
<td>UFP</td>
<td>Ultrafine particle</td>
</tr>
<tr>
<td>VOC</td>
<td>Volatile organic compound</td>
</tr>
</tbody>
</table>

distances, while larger droplets or particles will be deposited closer to their production region.

Apart from sea salt, another important feature of marine aerosols is their organic component. Two principal mechanisms for the generation of ocean-derived organic aerosol have been proposed: (1) the incorporation of organic matter (OM) into primary marine aerosols produced by breaking waves (Blanchard, 1964, 1968) and (2) gas-phase oxidation (by oxidants such as OH and ozone) of volatile organic compounds (VOCs) such as dimethyl sulfide (DMS), aliphatic amines, isoprene and monoterpenes, which can form secondary organic aerosols (SOAs) (Charlson et al., 1987). First estimates of the primary marine OAs were 14 Tg C yr\(^{-1}\), assuming a uniform organic carbon/sea salt mass ratio and sea salt emissions of 1000 Tg yr\(^{-1}\) (Duce, 1978). Other proposed methods to estimate the flux rely on a relationship between marine OA and chlorophyll in the water (O’Dowd et al., 2004). These studies provide global emission rates of varying magnitude (2–100 Tg yr\(^{-1}\)). Recently, Spracklen et al. (2011) estimated total emissions of 8 Tg C yr\(^{-1}\) (5.5 Tg C yr\(^{-1}\) in the submicron mode). Gantt et al. (2012) and Vignati et al. (2010a) showed that there is high variability among the different models and different parameterizations used. In addition, Quinn and Bates (2011) have shown that OA production is correlated with DMS emissions but not with ocean chlorophyll concentration, suggesting that the organic material in sea salt aerosols is indirectly connected to biological activity in the water column.

The organic material in the sea water contains many compounds, most of them uncharacterized (Benner, 2002). These compounds originate mostly from the degradation of marine organisms and plants and are composed of amino and fatty acids, carbohydrates, saccharides, humic substances and cell fragments (Hansell et al., 2009).

As previously stated, some marine aerosols may also form in situ, by the condensation of semi-volatile species onto existing aerosol, and hence they are considered to be of secondary origin. The CLAW hypothesis (Charlson et al., 1987) suggests that atmospheric oxidation of dimethyl sulfide will form sulfate over areas with high biological activity. Sulfate is the main precursor to secondary aerosol and cloud condensation nuclei (CCN) in the marine boundary layer; hence, its formation can modulate temperature (Bates et al., 1987; Charlson et al., 1987; Shaw, 1983). Recently, this view has been challenged by the observation that primary OA may be more predominant and may have a stronger connection to the biology of the ocean than secondary aerosol (O’Dowd et al., 2004; Quinn and Bates, 2011).
Mineral dust

Mineral dust aerosol (MDA) is often a dominant component of atmospheric aerosol in large regions of the planet. Ginoux et al. (2012) attribute 75% of the global dust emissions to natural origin, while 25% are related to anthropogenic (primarily agricultural) emissions. The largest source of natural MDA is the Sahara (Karanasiou et al., 2012). MDA affects climate through direct and indirect effects, modifies marine biogeochemistry, and impacts human health.

Iron compounds in soil dust absorb visible radiation contributing to atmospheric warming, while scattering of mineral dust particles leads to surface cooling. These effects might impact wind circulation and affect climate on a regional scale. Atmospheric warming over the Sahara region triggers an intensive heat pump effect that increases precipitation over the northern Sahel (Solmon et al., 2008). The atmospheric warming associated with dust from the Great Basin Desert in North America increases moisture fluxes in the Sierra Madre, resulting in an increase in precipitation of up to 40% (Zhao et al., 2012). Atmospheric warming and surface cooling due to MDA over northern Africa and the Arabian Peninsula reduces surface pressure leading to largescale convergence over the Arabian Peninsula. This convergence modifies wind circulation over the Arabian sea, increasing monsoon rainfall in India (Vinoj et al., 2014). Dust particles can act as IN, promoting cloud formation and precipitation and modifying cloud microphysics and the Earth’s radiative budget (Creamean et al., 2013; Lohman and Diehl, 2006; Prenni et al., 2009).

MDA contains iron, phosphorous and other micronutrients that can influence ocean productivity after dust deposition, especially in areas (about 25% of the oceans) characterized by limited availability of these substances (Schulz et al., 2012). The enrichment in micronutrients might modify the atmosphere–ocean carbon cycles, affecting the exchange of CO₂ and triggering dust–climate feedback effects (Jickells et al., 2005).

The interest in MDA in recent years has increased due to concerns about dust effects on human health. Saharan dust outbreaks over Europe are associated with the transport of biogenic particles, allergens and pathogens to the Mediterranean regions (Kellogg and Griffin, 2006). Episodes of Saharan dust transport over Europe are responsible for exceedances of PM₁₀ levels in the Mediterranean regions. For example, more than 70% of exceedances of the PM₁₀ daily limit at rural background sites in Spain are due to dust outbreaks (Querol et al., 2009). Some epidemiological and toxicological studies of coarse-particle (PM₂.₅) effects suggest a relationship between short-term exposure to coarse PM and cardiovascular effects, respiratory effects and mortality (Perez et al., 2008), while others have reported that the particles’ effects were significantly higher during non-desert dust days (Samoli et al., 2011).

Primary biological aerosol particles (PBAPs)

PBAPs contain a large range of different biological components, including microorganisms (bacteria, archaea, algae and fungi) and dispersal material such as fungal spores, pollen, viruses and biological fragments that are directly emitted to the atmosphere from their sources. Scanning electron microscopy images of some PBAPs are shown in Fig. 4. The aerodynamic diameters of PBAPs vary from tens of nanometres (viruses) of to a few hundred micrometres for pollen or plant debris (Jaenicke, 2005; Hinds, 1999; Pöschl, 2005). Typical size ranges for PBAPs are 0.05–0.15 µm for viruses and 0.1–4 µm for bacteria; fungal spores are 0.5–15 µm in size and pollen are the largest (10–30 µm; Despres et al., 2012). The atmospheric concentrations of PBAPs are not well characterized due to difficulties in measurement and identification techniques. The ambient concentrations vary with location, altitude and season but have been estimated to comprise as much as 25% of total aerosol mass globally (Jaenicke, 2005; Lang-Yona et al., 2012).

It has been shown that PBAPs can affect atmospheric processes such as ice nucleation (Christner et al., 2008; Guri-
should be noted that the resuspension of road dust is mostly in contrast to the situation regarding exhaust emissions, no policies (particulate matter from the abrasion of tyre wear, break wear, road wear and road dust suspension) are still growing in the EU. In contrast to the situation regarding exhaust emissions, no policies are in place to reduce wear emissions or resuspension. It should be noted that the resuspension of road dust is mostly not included in Fig. 5 because countries do not report these emissions. Hence the importance of total non-exhaust emissions in Europe is larger than Fig. 5 suggests, but even so, it can be seen that in the EU15 in 2010, exhaust and non-exhaust PM$_{10}$ emissions are equally important. For further discussion on the importance of non-exhaust emissions, we refer to Denier van der Gon et al. (2013). The trends for road transport in the EU new member states are slightly different, mostly because the activity growth is of greater relative strength, but here cleaner engines have a major impact. If the exhaust emissions followed the trend in non-exhaust, they would have been doubled; instead, they stabilized. In the future, however, a further decrease in exhaust emissions is expected.

Since non-exhaust particles are generated by abrasion mechanisms, this kind of particle is generally coarse and contributes mainly to the mass of the PM$_{2.5-10}$ fraction. Nevertheless, tyre wear and break wear abrasion emit particles smaller than 1 µm, contributing to the PM$_{2.5}$ mass fraction (Gietl et al., 2010). Compared to engine exhaust, non-exhaust particles are enriched in metals, metal oxides and mineral elements and contain less carbonaceous material (Visser et al., 2015). Toxicological evidence shows that non-exhaust particles are associated with negative health effects (Meister et al., 2012), but currently available data do not allow the quantification of non-exhaust particles’ impact on human health. Chemistry, morphology and hygroscopicity vary significantly among different types of non-exhaust particles, making the quantification of the toxicological properties of this particle source difficult (Amato et al., 2014).

The land-based non-road-transport emissions include categories such as rail transport, mobile machinery and inland shipping. Although emission legislation for these categories lags behind compared to road transport, the emission reduction between 2000 and 2010 is due to improved engine tech-
nologies as laid out in the directive 97/68/EC (EC, 1997) and its successors.

Partly due to the fact that it is one of the least regulated anthropogenic emission sources, emissions from international shipping are a significant contributor to air pollution and climate change (EEA, 2013a). In Fig. 5 an interesting trend in emissions from international shipping in European seas is visible. Emissions increased from 2000 to 2005 due to an increase in activity, but from 2005 to 2010 PM emissions from shipping declined. There are two main reasons. First, as a consequence of the implementation of Sulphur Emission Control Areas (SECAs) in the Baltic Sea (since 2006) and the North Sea (since 2007), the shipping fuels used on these seas have a lower sulfur content, which also results in lower PM emissions. Second, the economic crisis resulted in lower emissions mostly because ships resorted to “slow steaming” to save on fuels costs. Less fuel combustion saved money but also resulted in less PM emission.

Wood combustion

Wood combustion for residential heating has attracted increasing scientific interest in Europe and elsewhere due to its significant contribution to the OA budget. In Zurich (Switzerland) wood combustion is responsible for 45% of organic carbon (OC) in the cold season (Szidat et al., 2006). Gilarndoni et al. (2011a) reported that, in winter, wood burning accounts for 65% of total carbon (OC and elemental carbon, EC) at a rural site in northern Italy. At rural European background sites biomass burning contribution to OC ranges between 30 and 75% (Szidat et al., 2007, 2009; Gelencsér et al., 2007). Chemical and physical properties of residential wood-burning particles may vary significantly, depending on the combustion conditions, combustion appliances and wood types, and accurate emission inventories for this source are still under development. One of the main issues that needs to be solved is to what extent emission inventories should include condensable PM emissions.

As the use of wood as fuel for residential heating is spreading in developed countries, the number of studies investigating related health effects is increasing. There is evidence of an association between wood smoke exposure and health effects, including reduced resistance to infections, decreased lung function and asthma (Bølling et al., 2009). The review of Naehler et al. (2007) concluded that wood-burning particles should not be considered different from other combustion particles, regarding their health outcome. The majority of the epidemiological studies have been performed in areas affected by biomass burning where no direct marker of wood burning is available to link emissions and effects directly (WHO, 2013a).

Although the water-soluble fraction of OA emitted by wood combustion can absorb light in the visible and ultraviolet region of the spectra, its overall absorption effect is relatively small but could be relevant over bright surfaces (Chen et al., 2010).

2.1.2 Secondary aerosol

Natural and anthropogenic sources contribute to the emissions of the precursors of secondary aerosol, i.e. SO\(_2\), NO\(_x\), NH\(_3\) and VOCs and intermediate-volatility organic compounds (IVOCs). The sources of SO\(_2\) and NO\(_x\) are relatively well known, i.e. combustion of sulfur-containing fuel and fossil fuel combustion. The ability of the HNO\(_3\) produced by NO\(_x\) to form secondary aerosol depends on the availability of NH\(_3\) in the gas phase to form ammonium nitrate. Ammonia is mainly emitted by agricultural activities. NH\(_3\) emissions in the United States peak in spring in the Midwest during corn fertilization and elsewhere in summer due to manure (Paulot et al., 2014). In Europe NH\(_3\) emissions are less variable and show a maximum in spring due to fertilizer application (Paulot et al., 2014). NH\(_3\) emission control has been proposed as a cost-effective measure to control secondary inorganic aerosol (SIA), and thus PM levels, both in the United States (Pinder et al., 2007) and in Europe (Kulmala et al., 2011). In any case, the spatial distribution of the predicted reductions might be non-homogeneous, with more effective results in rural areas.

While the atmospheric processes of inorganic species are relatively well understood, the processes governing the organic fraction, particularly SOA formation, is poorly understood. Emissions of anthropogenic and biogenic precursors of SOA are quite uncertain (Hallquist et al., 2009). As a result, while SOA dominates the total OA in many locations (Zhang et al., 2007), its source attribution remains poorly understood, compromising the design of effective pollution mitigation strategies.

A good example concerning the uncertainty of SOA formation mechanisms is the correlation between biogenic SOA and tracers of anthropogenic pollution (Weber et al., 2007). The enhancement of biogenic SOA with anthropogenic pollutants is partly explained by a shift in the biogenic organics partitioning from the gas to the aerosol phase due to the presence of anthropogenic aerosol mass (Hoyle et al., 2011). NO\(_x\) might also affect SOA yields, both controlling oxidant concentration and contributing, as the NO\(_3\) radical, to SOA formation at night-time (Rollins et al., 2012). Indeed, laboratory experiments show that increased SOA yields are found in some system with increasing NO\(_3\) concentration, while the yields are lower for others (Lane et al., 2008). Finally, organics contribute with sulfuric acid to new particle formation and growth (Ehn et al., 2014; Hoyle et al., 2011; Riccobono et al., 2014). Ehn et al. (2014) identified a new class of organic species called extremely low-volatility organic compounds (ELVOCs), formed by the oxidation of biogenic VOCs (BVOCs). These species could help to explain the discrepancy between the atmospheric burden of observed SOA and that reported by some models.
SOA is not formed exclusively by biogenic organic precursors. In areas influenced by anthropogenic and terrestrial biogenic emissions, such as the Californian Central Valley, SOA from anthropogenic sources accounts for 65% of submicron organic mass (Liu et al., 2012). Petrol and diesel vehicles emit aromatic and aliphatic compounds with elevated SOA-forming potential (Gentner et al., 2012). Diesel emissions, being enriched in IVOCs, were considered to be 7 times more efficient than petrol emissions in forming secondary aerosol (Gentner et al., 2012), but there is contradictory evidence on this issue as discussed later in the paper. Residential wood burning and wildfires are also significant sources of SOA precursors (e.g. Grieshop et al., 2009; Hennigan et al., 2011).

2.1.3 Pre-industrial aerosol

The need to describe the effect on climate of atmospheric aerosol emitted by human activity triggered interest in pre-industrial aerosol properties (Carslaw et al., 2013a), i.e. the composition and the concentration of particulate matter (PM) in the absence of anthropogenic emissions. Aerosol concentrations close to pristine conditions are probably observed over the oceans, especially in the Southern Hemisphere. Over the ocean, particles are composed mainly of sea salt, organics, sulfates and aerosol emitted by wildfires transported from the continents. The determination of pre-industrial aerosol concentrations and properties in continental regions is trickier, due to the current overwhelming influence of anthropogenic activities, the lack of data from pre-industrial times and the difficulty of discriminating between the influence of human activities on emissions and natural sources over the continents (Andreae, 2007). Pristine continental conditions can be observed in the Amazon Forest (Martin et al., 2010), where the reduction of anthropogenic forest fires during the last decades is associated with a decrease in anthropogenic pressure, especially during the dry season (Gilardoni et al., 2011b). In the pristine rain forest, particles are composed mainly of primary and secondary biogenic organic aerosol and dust.

2.1.4 Source attribution

Source apportionment studies associate measured aerosol with emission sources and/or production mechanisms, utilizing numerous measurement strategies and data analysis techniques with the common aim of reducing ambient complexity to a finite number of source categories. Here we discuss both receptor-based measurements and laboratory measurements of direct emissions coupled with a simulation of their atmospheric transformations. For ease of discussion, receptor-based methods are classified as (1) “tracer-based” techniques based on molecular markers that have high source specificity but constitute a small fraction of the total apportioned mass (Sect. “Tracer-based apportionment”) or (2) “ensemble-based” studies that utilize properties corresponding to the total mass but do so at the cost of reduced chemical specificity (Sects. “Ensemble-based apportionment” and “Other ensemble-based receptor techniques”).

Overview of analysis techniques

Source-based apportionment of mass typically utilizes bilinear receptor models, where the input data matrix (i.e. composition time series) is represented as the linear combination of a set of static factor profiles and their time-dependent intensities. Bilinear models are commonly used for the analysis of ensemble-based measurements, such as chemical tracers or aerosol mass spectra. Clustering algorithms attempt to divide a measurement series into discrete groups of similar events and are frequently applied to mixing-state-sensitive measurements such as single-particle mass spectrometry. Finally, spatial mapping of probable source locations is achieved using methods such as the potential source contribution function (PSCF).

The bilinear model can be expressed as $X = GF + E$, where the $X$ matrix represents the time series (rows) of measurements (columns) and $G$ and $F$ represent the factor scores (e.g. concentrations) and profiles, respectively. The matrix product $GF$ is therefore the model reconstruction of the measured data, with $E$ as the residual. Various implementations of the bilinear model attempt to optimize the $GF$ representation of $X$ but use different methods and criteria for achieving this. The basic solution is principal component analysis (PCA; Wold et al., 1987), which resolves the components that best explain the variance in the data. However, because no constraints are applied to ensure physical interpretability (e.g. negative factor scores are allowed), source quantification is non-trivial.

The chemical mass balance (CMB; e.g. Schauer et al., 1996) implementation of the bilinear model requires as input predefined source profiles, usually chosen to represent known primary source emissions. The algorithm yields a solution that maximizes the explanatory power of the predetermined profiles, leaving a residual, which in many studies is assumed to be SOA. The selection of appropriate factor profiles is a critical and non-trivial aspect of this analysis, as order-of-magnitude variations inappportioned mass-to-tracer ratios are observed for the same source class at different sites (e.g. Elsasser et al., 2012). All major sources except the residual must be taken into account, otherwise apportionment to the residual source will be biased high. Finally, tracers contained in the fixed profiles must be chemically inert.

Positive matrix factorization (PMF; Paatero and Tapper, 1994; Paatero, 1997) is an implementation of the bilinear model that requires non-negative elements in the $G$ matrix. Unlike CMB, no a priori knowledge of factor profiles is needed. The number of factors is determined by the user, and both profiles and contributions are calculated by the algorithm such that the uncertainty-weighted residuals are minimized. PMF solutions have some degree of rotational
Figure 6. Map of the Arctic region colour-coded according to the potential source contribution factor (PSCF) for marine biogenic particles over the central Arctic Ocean during the Arctic Summer Cloud Ocean Study (ASCOS) cruise (Chang et al., 2011).

ambiguity. Most PMF analyses explore rotational ambiguity in only a single dimension; however, the multilinear engine (ME-2) implementation of PMF (Paatero, 1999) enables full control of the rotational space. The recently developed Source Finder (SoFi) tool provides an interface for configuring the ME-2 engine and analysing the results (Canonaco et al., 2013). A superior solution was obtained by this method relative to the zero rotation, as evidenced by an improved correlation of factor time series with the relevant tracers.

PSCF analysis (Ashbaugh et al., 1985) combines measurements of tracer concentrations with calculated air parcel back trajectories. The geographic area surrounding the receptor site is represented as a grid, and the probability that an air parcel passing through a grid cell will result in a tracer concentration exceeding a pre-set threshold is calculated, yielding a spatial map of likely source locations. PSCF studies are summarized by Hopke et al. (2003). PSCF has been integrated with bilinear model outputs for the identification of specific sources and/or source regions on local and regional scales (Begum et al., 2010; Heo et al., 2009) and for the evaluation of ensemble and tracer-based apportionment results (e.g. Chang et al., 2011; Gilardoni et al., 2011a; Schwartz et al., 2010). An example of PSCF analysis results is given in Fig. 6.

Cluster analysis includes algorithms such as k means (Bishop, 2006) and fuzzy c means clustering (Bezdek et al., 1984), as well as neural network algorithms such as ART-2a (Bhave et al., 2001). Generally, the aim is to divide a series of measurements into discrete sets of self-similar individual measurements. Freutel et al. (2013) developed an iterative algorithm in which known profiles were used to classify as many particles as possible, followed by cluster analysis of the residuals. The resulting residual cluster centres from this analysis were then reintroduced into the initial algorithm as known profiles. Cluster analysis has been used extensively in the analysis of single-particle aerosol mass spectrometry (Ault et al., 2009; Bein et al., 2007; Healy et al., 2009; Liu et al., 2003; Reinard et al., 2007; Snyder et al., 2009).

Tracer-based apportionment

Traditional source attribution methods utilize measurements of molecular and/or elemental markers to attribute a bulk quantity such as total OA or PM$_{2.5}$ mass. Such methods can be applied to a wide range of measurement situations, from long-term monitoring sites where routine measurements of ensemble mass and a few tracers are performed to studies involving state-of-the-art molecular speciation of aerosols.

A classic approach to the marker-based receptor model source apportionment problem involves the apportionment of OC using CMB techniques. Chemical profiles are constructed for each known primary source, and the unapportioned mass is typically assigned to SOA (e.g. El Haddad et al., 2011; Stone et al., 2008). Some studies have attempted to also include markers related to SOA tracers (Kleindienst et al., 2007). Uncertainties in the OC-to-tracer ratios for primary sources can be quite large. Because these ratios are affected by regional characteristics (e.g. fuel type, source conditions), uncertainties can be reduced if representative measurements of local sources are available. However, variations of a factor of 3 in biomass burning OC and/or levoglucosan have been observed even for measurements at a single site (Zhang et al., 2008). Because the uncertainties in each primary profile contribute to the total uncertainty in SOA, the resulting SOA uncertainties are quite large.

The CARBOSOL project (Legrand and Puxbaum, 2007; Pio et al., 2007) incorporated $^{14}$C analysis into a marker-based source apportionment conducted in the spirit of CMB (Gelencsér et al., 2007). Radiocarbon measurements combined with OC, EC, levoglucosan and cellulose showed that primary biomass burning and fossil sources dominated OC in winter, while non-fossil SOA dominated in summer (Gelencsér et al., 2007). Recent studies have expanded and adapted this approach, with statistical techniques used to select and assess uncertainties related to the selection of tracer-to-OC ratios (e.g. Gilardoni et al., 2011a; Szidat et al., 2009). These approaches all assume that tracers used in the model are chemically inert. Levoglucosan is typically used as a tracer for biomass burning; however, recent studies show that it reacts under atmospherically relevant conditions (Hennigan et al., 2010; Hoffmann et al., 2010). While such tracer instabilities could introduce large errors into CMB analyses (Roy et al., 2011), apportionment results are plausible; possibly, selection of an appropriate profile must implicitly capture the average extent of atmospheric aging experienced at the receptor site.

Until recently, only a few studies have attempted source attribution on data sets containing large numbers of organic marker compounds. The recent development of online instru-
Ensemble-based apportionment

Off-line measurements of PM$_{2.5}$ and PM$_{10}$ chemical composition have been analysed by multivariate statistical tools to identify anthropogenic and natural aerosol sources across Europe in a large number of studies (Viana et al., 2008; Larsen et al., 2012; Belis et al., 2014). More recently, ensemble-based techniques have been applied for an analysis of on-line chemical composition measurements performed with the Aerodyne aerosol mass spectrometer (AMS) and related instruments (Canagaratna et al., 2007; De Carlo et al., 2006; Drewnick et al., 2005; Fröhlich et al., 2013; Jayne et al., 2000; Ng et al., 2011) Although the chemical specificity of these spectra is reduced compared to tracer-based measurements, factor analysis of AMS spectra has nonetheless proved a powerful method for quantitative source attribution.

The first application of a bilinear factor analysis model to AMS data utilized $m/z$ 44 (CO$_2^+$) and $m/z$ 57 (C$_3$H$_5$O$^+$ and C$_4$H$_7$O$^+$) in a 2-factor custom principal component analysis model, yielding hydrocarbon-like and oxygenated organic aerosol factors, respectively denoted HOA and OOA (Zhang et al., 2005). HOA was related to primary anthropogenic emissions, while OOA was related to the secondary fraction (Zhang et al., 2007). More recent AMS source apportionment studies have in large part utilized PMF analysis. The first AMS PMF study was conducted by Lanz et al. (2007), yielding four primary and two secondary factors. The primary factors consisted of an HOA factor related to fossil fuel combustion, biomass burning organic aerosol (BBOA), charbroiling and a minor factor attributed to cooking. PMF is now used worldwide for the analysis of AMS organic mass spectra, as summarized in the review of Zhang et al. (2011). The accuracy of such analyses relies heavily on accurate calculations of measurement uncertainty (Allan et al., 2003) and protocols for evaluating PMF outputs (Ulbrich et al., 2009a). The advanced rotational control over the PMF solution offered by the ME-2 algorithm is sometimes necessary to resolve temporally or spectrally similar factors (Canonaco et al., 2013; Lanz et al., 2008).

Figure 7 shows the source attribution of OA at several European sites during three intensive operational field campaigns (Crippa et al., 2014). Successful POA (primary organic aerosol) source identification and quantification has been achieved by AMS PMF for a number of sources. HOA is frequently associated with POA from fossil fuel combustion, based on time trend correlation with chemical tracers for vehicle traffic emissions in urban areas (Aiken et al., 2009; Sun et al., 2011). Cooking organic aerosol (COA) has recently been identified as a significant POA source in urban environments (Allan et al., 2010; Crippa et al., 2013a; He et al., 2010; Lanz et al., 2007; Mohr et al., 2012; Slowik et al., 2010; Sun et al., 2011). A comparison of factor mass spectra with direct cooking emissions suggests that POA derives largely from the cooking oil rather than the food itself. Biomass burning OA (BBOA) factors have been associated with POA emissions from wildfires, controlled burns and domestic wood burning (Aiken et al., 2009, 2010; Lanz et al., 2007, 2010). A few recent studies have identified coal combustion OA at several sites in China (Hu et al., 2013; Huang et al., 2014). It should be noted, however, that the reliability of POA source apportionment by AMS is greater for near-source studies than in background areas, where, for instance, factors such as HOA are difficult to extract by factor analysis or have a less established relationship with traffic emissions.

SOA source attribution is more challenging because atmospheric aging causes profiles from different sources to tend to converge towards similar, highly oxygenated mass spectra (e.g. Capes et al., 2008; Jimenez et al., 2009; Morgan et
al., 2010; Ng et al., 2010; Zhang et al., 2011). Additionally, SOA composition is not static but continually evolves in response to gas–particle partitioning and atmospheric oxidation. Only in a few cases are clear source attributions possible (Budisolistiorini et al., 2013; Chang et al., 2011; Crippa et al., 2013b; Robinson et al., 2011; Slowik et al., 2011). AMS PMF analyses typically represent SOA as a linear combination of two OOA factors: a low-volatility, more oxygenated fraction (LV-OOA) and a less oxygenated, semi-volatile fraction (SV-OOA) (Jimenez et al., 2009; Lanz et al., 2007; Ng et al., 2010). LV-OOA is generally more characteristic of an aged, regional aerosol, while SV-OOA is more responsive to temperature-driven partitioning and/or local SOA production. Hildebrandt et al. (2010a) argued that at least in some environments the LV-OOA and SV-OOA do not correspond to different sources or processes but rather to the two extremes of the oxidation state of the SOA in the corresponding data set.

The lack of obvious source-specific features in OOA presents a major challenge for SOA source attribution studies, though some evidence does exist that unique spectral features are retained. The projection of the ambient SV-OOA spectrum on a PCA model initialized with SOA from α-pinene, diesel vehicles, mopeds and wood burning successfully resolved major contributions from SOA sources, though a quantification of these sources was not attempted (Heringa et al., 2012). AMS PMF outputs have been related to measurements from more chemically specific measurements, such as thermal desorption proton transfer reaction mass spectrometry (TD-PTR-MS), though quantitative interpretation is complicated by non-desorbing and/or thermally decomposing material in LV-OOA (Holzinger et al., 2013). Other studies have improved resolution of SOA factors (and POA) factors by incorporating both AMS and gasphase PTR-MS spectra into the same matrix for PMF analysis (Crippa et al., 2013b; Slowik et al., 2010); however, such separation has typically been process-based (e.g. photochemical age, day and night enhancement) rather than source-related. Post-analysis of AMS PMF results using radiocarbon data has successfully apportioned OOA into fossil and non-fossil fractions, as discussed in the next section (Minguillón et al., 2011). However, a true source attribution of the SOA fraction remains elusive.

Other ensemble-based receptor techniques

Radiocarbon 14C analysis facilitates source attribution by quantifying fossil and non-fossil contributions to carbonaceous aerosol (Currie, 2000; Szidat et al., 2004; Wacker et al., 2013; Zhang et al., 2012). OC and EC fractions can be analysed separately, though the distinction is not clear-cut due to method-dependent definitions and charring artefacts (Birch and Cary, 1996; Chow et al., 2001; Countess, 1990; Schmid et al., 2001). Minguillón et al. (2011) compiled EC and OC fossil fractions reported at 17 urban and rural sites in Europe and Asia, spanning different size fractions, seasons and measurement techniques. For nearly all sites, the fossil contribution to EC was higher than OC and elevated in the warmer months. In some regions, such as Alpine valleys, this seasonal difference could be clearly attributed to domestic wood burning (Szidat et al., 2007; Zotter et al., 2014). Even at urban sites, the fossil contribution to OC did not exceed 60%, and values as low as 28% were reported. Fossil contributions to OC showed less seasonal dependence; for example, decreased residential wood burning in the warmer months may be offset by an increase in biogenic SOA. The utility of such analyses is maximized when data can be segregated according to known site conditions; for example, Aiken et al. (2010) observed a decrease in the Mexico City non-fossil OC fraction from 62 to 49% when comparing low-fire and high-fire periods or when integrated with other source apportionment analyses, e.g. the determination of fossil and non-fossil SOA fractions by integrating 14C measurements with AMS-PMF results (Minguillón et al., 2011), source identification by coupling 14C measurements with molecular tracers (Gelencsér et al., 2007; Gilardoni et al., 2011a), or a combination of these approaches (Huang et al., 2014). Such integrated analyses will become even more critical with increasing use of biofuels, which will provide a 14C source from sources traditionally associated with fossil fuel combustion.

Source-dependent differences in the wavelength-dependence of light absorption by EC were exploited by Sandradewi et al. (2008a, b) in a two-component model apportioning EC to traffic and wood burning sources using aethalometer data. This model has been compared with 14C, AMS-PMF and tracer-based CMB analyses at various locations and shows good agreement near source regions (Favez et al., 2009, 2010). Assets of this model include the high time resolution of measurements relative to 14C analysis, as well as the ease of calculation and inexpensive instrumentation. Interpretative challenges include the selection of appropriate α-parameters describing wavelength-dependent light absorption (Chirico et al., 2010; Favez et al., 2010; Sciare et al., 2011). Depending on the site, significant light absorption from brown carbon (BrC) and mineral dust are also possible and must be taken into account (Yang et al., 2009).

Fourier transform infrared spectroscopy (FTIR) provides quantitative functional group concentrations of collected aerosols (Gilardoni et al., 2007; Russell et al., 2009; Takahama et al., 2013). This approach has demonstrated some promise in distinguishing between biogenic, biomass burning, marine and fuel-combustion-based sources using PMF, PSCF and cluster analysis. A threedimensional space consisting of an O/C ratio, a molar ratio of (acids+carbonyls)/alkanes and molar ratio of hydroxy/alkanes distributes factors in a manner consistent with the authors’ attributions (Russell et al., 2011). Measurements in the Finnish boreal forest suggest that the separation of biogenic and biomass burning aerosol may also be pos-
sible, though POA and SOA contributions to biomass burning fractions are indistinguishable (Corrigan et al., 2013). However, the low degree of chemical specificity means that comparisons among the profiles of related factors and/or emissions source and aging measurements are in qualitative agreement, and substantial post-analysis interpretation of ambient results is required.

Similarly to FTIR, functional group analysis by proton nuclear magnetic resonance (NMR) spectroscopy has been proposed for OA source attribution (Decesari et al., 2007). This technique was originally developed for water-soluble OM and is therefore unsuitable for targeting sources producing organic compounds with a low O/C ratio (as for fossil fuel POA). Recently, factor analysis techniques have been implemented for spectral deconvolution of NMR data sets (Finessi et al., 2012; Paglione et al., 2014), providing an additional tool for organic source apportionment, particularly useful for the determination of the biomass burning and biogenic SOA contributions. Analogously to FTIR, also the NMR-based source apportionment approaches require post-analysis interpretation.

**Source emission measurements of POA and SOA**

Source emission measurements complement ambient source apportionment studies through the determination of emission factors and source signatures. Combined with emissions inventories or usage statistics, these studies can provide a bottom-up estimate of source contributions. Although source-based measurements have traditionally investigated only the primary fraction, recent studies have used in situ oxidation techniques to simulate atmospheric aging. Methods include the introduction of source emissions into a smog chamber (Chirico et al., 2010; Grieshop et al., 2009; Heringa et al., 2011), the development of mobile oxidation techniques (mobile smog chambers and flow tube systems) (Kang et al., 2007; Platt et al., 2013; Presto et al., 2011) and the analysis of plume evolution in ambient air (e.g. Cubison et al., 2011; DeCarlo et al., 2010; Yokelson et al., 2009) A consistent feature in these emission studies is the dominance of secondary aerosol (particularly SOA) for systems and conditions having the highest emission factors.

Wood and biomass burning have long been recognized as important POA sources, but significant OA enhancement due to SOA formation has been observed in some systems. The aging of emissions from open biomass combustion yielded OA enhancement ratios ranging from 0.7 to 2.9, depending on fuel type and burn conditions (Hennigan et al., 2011; Ortega et al., 2013). Ambient estimates of aging-induced OA enhancement in wildfire plumes include no detectable enhancement (Akagi et al., 2012; Capes et al., 2008; Cubison et al., 2011; Hecobian et al., 2011; Jolleys et al., 2012), enhancements of 20–50 % (DeCarlo et al., 2010; Reid et al., 1998) and increases of a factor of 2 or more (Lee et al., 2008; Yokelson et al., 2009). Explanations for these differences include fuel type and burn conditions but also the evaporation of primary emissions following dilution (Robinson et al., 2007) and the gas-phase oxidation of repartitioning semivolatile species (Donahue et al., 2012a). These latter explanations decrease POA mass with age, causing the observed OA enhancement to serve as a lower limit for SOA production.

Road vehicles have long been recognized as a major PM source in urban areas; however, recent studies indicate that a significant fraction (and perhaps the majority) of their impact on ambient PM consists of secondary PM (Bahreini et al., 2012; Gentner et al., 2012; Hallquist et al., 2009; Nordin et al., 2013; Platt et al., 2013, 2014; Robinson et al., 2007). The relative contributions of SOA from different vehicle types to ambient aerosol remain highly controversial, with one recent study from the Los Angeles Basin concluding that SOA from petrol vehicles (Bahreini et al., 2012) dominates urban OA and a second study concluding that diesel SOA is more important (Gentner et al., 2012).

Recent studies indicate that aircraft emissions contribute significantly to particle mass and particle number (PN) concentration in the vicinity of airports (Arvanilakal et al., 2011; Dodson et al., 2009; Hsu et al., 2012; Hu et al., 2009; Levy et al., 2012; Stettler et al., 2011; Unal et al., 2005; Yu et al., 2004). POA dominates the total aerosol at low engine loads, corresponding to idle and taxiing operation (Onasch et al., 2009; Presto et al., 2011; Timko et al., 2010). However, recent experiments show that secondary aerosol dominates the total aerosol loading over all thrust conditions. Aging of aircraft emissions for a few hours in a mobile smog chamber yielded PM enhancement factors of 35 ± 4.1 (4 % engine load, corresponding to ground idle operation), 17 ± 2.5 (7 % load), 60 ± 2.2 (30 % load) and 2.7 ± 1.1 (85 % load), with SOA formation dominating the aerosol at low loads and secondary sulfate dominating at high loads (Miracolo et al., 2011).

2.2 Aerosol concentration and composition

2.2.1 Particle number concentration

Daily averages of PN concentration range between a few hundred and over 50 000 cm$^{-3}$, with lower values in the marine boundary layer and free troposphere and higher concentrations in the planetary boundary layer (Putaud et al., 2010; Spracklen et al., 2010). Putaud et al. (2010) observed an increasing gradient moving from natural background areas to urban environments and kerbside all over Europe (northern, southern and central Europe). At urban sites PN concentration exhibits a daily variability characterized by two maxima, corresponding to morning and evening rush hours, mainly due to primary particle emissions (Aalto et al., 2005; Reche et al., 2011). In southern Europe, during spring and summer, an additional maximum is observed at midday as a consequence of photochemical nucleation processes (Reche et al.,...
2011). Aalto et al. (2005) observed an increasing gradient in PN concentrations in urban areas moving from northern to southern Europe.

PN concentration is often higher at sites characterized by a higher concentration of PM$_{2.5}$. Nevertheless, PN and mass do not correlate; this is because PN is usually dominated by ultrafine particles (UFPs, particles with an aerodynamic diameter smaller than 100 nm), which usually affects particle mass very little. Conversely, particles with a diameter larger than 100 nm contribute substantially to particle mass, and their number concentration correlates with PM$_{2.5}$ values at most of the European sites (Putaud et al., 2010).

PN size distribution at urban sites and kerbside locations exhibits a mode around 20–30 nm and a second mode around 100 nm, partly due to regional background. A third mode appears at rural, near-city and urban background sites around 10–30 nm during spring and summer afternoons, which could be attributed to the nucleation of new particles (Van Dingenen et al., 2004).

Asmi et al. (2011) investigated PN concentration and size distribution in 24 European sites over a 2-year period (Fig. 8). Particles were classified according to their mobility diameter as $N_{30–50}$ (particles with a mobility diameter between 30 and 50 nm), $N_{50}$ (particles with a mobility diameter between 50 and 500 nm) and $N_{100}$ (particles with a mobility diameter between 100 and 500 nm). In northern Europe number concentrations of all size classes were often below 100 cm$^{-3}$, and the frequency distribution curves had a log-normal shape with higher concentrations in summer. In central Europe limited variability was observed at each site during the year. A larger variability within each season and among different seasons was observed in western and southern Europe. In clean continental areas the number of both Aitken nuclei and accumulation mode particles varied between 100 and 5000 cm$^{-3}$, while at polluted continental sites the number was often above 5000 cm$^{-3}$. Higher values were typically observed in spring and summer. At high-altitude sites, winter concentrations varied between 100 and 1000 cm$^{-3}$ and were representative of free-troposphere conditions. Conversely, higher PN concentrations were observed in summer, due to planetary boundary layer influence and the transport of pollutants on valley winds.

Over the decade 2001–2010, a decreasing trend in PN concentration was observed over Europe, as also in the whole Northern Hemisphere, likely due to a decrease in emissions of primary particles, SO$_2$ and co-emitted species (Asmi et al., 2013).

### 2.2.2 Particle mass concentration and trends

Measurements of PM$_{10}$ (particles with an aerodynamic diameter below 10 µm) have been performed in Europe for over 2 decades within national and international monitoring networks, including EMEP (European Monitoring and Evaluation Program), AirBase and, more recently, ACTRIS (Aerosols, Clouds, and Trace gases Research InfraStructure network). Conversely, PM$_{2.5}$ (particles with an aerodynamic diameter below 2.5 µm) has a shorter time record and lower spatial coverage, due to a more recent introduction of PM$_{2.5}$ target values by the European Air Quality directive.

Observations over the period 1994–2008 showed that PM$_{10}$ annual averages varied over 1 order of magnitude, from 5 to 54 µg m$^{-3}$ (Putaud et al., 2010; Van Dingenen et al., 2004). During the intensive observational periods of the EMEP campaign in 2009 over 32 sites, PM$_{10}$ ranged between 4 and 30 µg m$^{-3}$, in agreement with longer time records (Tørseth et al., 2012). The lowest values were observed in remote and rural sites in northern Europe, while the highest levels were reported for urban sites in southern Europe. Generally, an increasing spatial trend was observed moving from natural and rural background to kerbsides (Putaud et al., 2010). Nevertheless, the 5th percentile values of PM$_{10}$ observed near cities and at some urban sites were similar to PM$_{10}$ concentrations at natural background sites, indicating that during a limited number of days, urbanized areas might experience background aerosol concentrations, typically below 10 µg m$^{-3}$ (Van Dingenen et al., 2004).

Annual averages of urban background PM$_{10}$ over the same time frame (1994–2008) were about 30 % higher in southern Europe compared to central and northern Europe (Putaud et al., 2010). The higher PM$_{10}$ concentrations in these urban areas can be attributed to higher background PM$_{10}$ concentrations. In fact, although less pronounced, the same geographical trend was also observed at regional background sites. Querol et al. (2009) reported a north–south gradient of increasing PM$_{10}$ in the Mediterranean Basin. The gradient was attributed to Saharan dust transport, whose intensity decreases moving northwards as a consequence of dispersion, and wet and dry particle deposition.

Across Europe concentrations of PM$_{2.5}$ ranged between 3 and 35 µg m$^{-3}$ (Putaud et al., 2010; Tørseth et al., 2012). In northern and southern Europe, PM$_{2.5}$ concentration increased when moving from rural and natural background to urban background sites. This trend was less marked in central Europe (Putaud et al., 2010). A comparison of urban background and regional background concentrations suggests that 60 % of urban PM$_{2.5}$ can be attributed to the influence of the regional background (EMEP, 2011).

The ratio of PM$_{2.5}$ to PM$_{10}$ varies from site to site, ranging between 0.5 and 0.9 (Putaud et al., 2010; Tørseth et al., 2012; Van Dingenen et al., 2004). Generally, the correlation between PM$_{2.5}$ and PM$_{10}$ is very good at each individual site, indicating that meteorology has a significant role in controlling PM levels and that fine- and coarse-particle sources might co-vary (Van Dingenen et al., 2004). Typically, kerbside locations exhibit lower ratios due to the large contribution of resuspended dust to coarse-particle concentration. In rural areas, near cities and at urban sites the ratio is higher due to the increasing contribution of secondary aerosol sources to PM$_{2.5}$ concentrations (Van Dingenen et
The PM$_{2.5}$-to-PM$_{10}$ ratio ranges between 0.6 and 0.8 in northern Europe and between 0.4 and 0.6 in southern Europe (Fig. 9) (EMEP, 2012; Tørseth et al., 2012). This lower ratio is likely due to the contribution of Saharan dust transport and local dust sources to coarse-particle concentration in the Mediterranean regions.

Long-term observations of PM$_{10}$ and PM$_{2.5}$ time trends indicate a negative trend in both size fractions. From 2000 to 2009, Tørseth et al. (2012) observed an average decrease of 18 and 27% in PM$_{10}$ and PM$_{2.5}$ concentrations, respectively, corresponding to an annual trend in $-0.29$ and $-0.37$ µg m$^{-3}$ yr$^{-1}$. The average trend was calculated over 24 European sites for PM$_{10}$ and 13 for PM$_{2.5}$. Barmpadimos et al. (2012) observed a similar annual trend in PM$_{2.5}$ concentration ($-0.4$ µg m$^{-3}$ yr$^{-1}$) and a slightly higher decrease in PM$_{10}$ ($-0.4$ µg m$^{-3}$ yr$^{-1}$) at seven rural and urban background sites from 1998 to 2010. The time trend observed by K. C. Wang et al. (2012) over more than 90 sites across Europe from 1992 to 2009 was $-15\%$ of PM$_{10}$ and $-9\%$ of PM$_{2.5}$ per decade. The lower PM$_{2.5}$ decrease could likely be attributed to the difference in number and type of the investigated sites. K. C. Wang et al. (2012) relied mainly on data from urban and suburban sites, while Tørseth et al. (2012) and Barmpadimos et al. (2012) analysed mainly rural and regional background sites. Time trends in mass concentrations of PM are discussed below in Sect. 3.1.4.

Long-term measurements of PM$_{1}$ (particles with an aerodynamic diameter below 1 µm) are still limited because the Air Quality Directive does not set any target or limit values for this cut size. Tørseth et al. (2012) reported PM$_{1}$ values for six EMEP sites in 2009: daily average concentrations ranged between 3 and 12 µg m$^{-3}$. Aas et al. (2012) reported PM$_{1}$ concentrations over 2 months at 10 European sites. Values ranged between less than 1 µg m$^{-3}$ up to 12 µg m$^{-3}$. Generally PM$_{1}$ represents a significant and usually the dominant fraction of PM$_{2.5}$ (Perez et al., 2010).

### 2.2.3 Particle chemical composition

The major chemical constituents of PM are inorganic ions – mainly nitrate, sulfate and ammonium – mineral dust, sea salt and carbonaceous aerosol, i.e. OC and elemental (or black) carbon (or BC). Figure 10 shows the average composition of PM$_{10}$ and PM$_{2.5}$ across Europe.

The average daily nitrate concentration from the EMEP network in 2010 was 1.9 µg m$^{-3}$, with higher values in winter (January and February: 2.8 µg m$^{-3}$) and lower in summer...
Figure 10. Average composition of PM$_{10}$ and PM$_{2.5}$ at urban and rural sites across Europe. OM is calculated as OC · 1.4; this is why OM contribution to PM is probably underestimated, and this explains part of the unaccounted mass (adapted from Putaud et al., 2010).

(1.2 µg m$^{-3}$) (EMEP, 2012). The nitrate concentration is typically in the range of 1.7–1.9 µg m$^{-3}$ in the Mediterranean Basin, 2–4 µg m$^{-3}$ in central Europe and 0.5–1 µg m$^{-3}$ in northern Europe (Querol et al., 2009). The highest nitrate concentrations are observed in the Po Valley during winter, due to the large anthropogenic emission sources of nitrogen oxides and the large concentration of ammonia, which stabilizes nitrate in the aerosol phase (Putaud et al., 2004).

Nitrate represents 8–16 % of PM$_{10}$ and 6–16 % of PM$_{2.5}$ (EMEP, 2012; Putaud et al., 2010). The nitrate contribution to PM$_{10}$ is larger in central and western Europe compared to the other parts of the continent (EMEP, 2009; EMEP, 2011; EMEP, 2012). In PM$_{10}$ and PM$_{coarse}$ the nitrate mass fraction decreases moving from rural to kerbside. The main precursor of nitrate in urban areas is represented by nitrogen oxides mainly emitted by traffic and power generation on a regional scale, and the extended timescales for this formation are consistent with the higher nitrate mass fraction in an urban and a near-city background compared to kerbside locations (Putaud et al., 2004).

The reduction in nitrate and nitrogen oxide emissions during the last decades has been significant (on average 31 % from 1990 to 2009) although not spatially homogeneous. These reductions led to an average reduction in nitrate of 8 %, with large variability from site to site (Tørseth et al., 2012).

The average daily sulfate concentration over Europe is about 1.7 µg m$^{-3}$ (EMEP, 2011, 2012). Annual mean concentrations below 0.5 µg m$^{-3}$ are observed in northern Europe, while mean concentrations larger than 1 µg m$^{-3}$ are reported in eastern Europe (Austria, Poland, Hungary and Lithuania) (Tørseth et al., 2012). The spatial distribution of sulfate reflects that of SO$_2$ emission sources: mainly energy production and transformation processes located in eastern Europe.

The sulfate mass fraction is 9–19 % of PM$_{10}$ and 8–21 % of PM$_{2.5}$, with lower fractions at the kerbside (Putaud et al., 2010; EMEP, 2012). No clear trend is observed moving from rural to urban or near-city backgrounds, in line with the fact that the largest fraction of sulfate is produced from in-cloud processing on a regional rather than on a local scale (Putaud et al., 2004).

Starting from the late 70s, Europe adopted emission control programmes which efficiently reduced sulfate and SO$_2$ emissions by 73 % during the period 1980–2009. The change in emissions led to a reduction in sulfate concentration by 70 % ± 20 % over the same period of time (Tørseth et al., 2012).

The ammonium concentration over Europe has a daily mean of 1 µg m$^{-3}$ (EMEP, 2009, 2011, 2012) and represents 7–16 % of PM$_{10}$ and 8–12 % of PM$_{2.5}$ (Querol et al., 2009). In the period 1990–2009, ammonia and ammonium emissions have been reduced by 26 %, leading to a reduction in ammonium concentration at most of the EMEP sites equivalent to 24–25 %. Conversely, at those European rural sites where emission reductions were negligible, ammonium concentration did not show any decrease (Tørseth et al., 2012).

Analyses of nitrate, sulfate and ammonium in PM$_{1}$ have been performed only during short measurements campaigns, mainly with online techniques (aerosol mass spectrometers). These results show that sulfate accounts for 10–50 % of submicron particle mass, nitrate for 1–28 % and ammonium for 7–29 % (Zhang et al., 2007). The availability of PM$_{1}$ chemical composition data has increased recently with the deployment of ACSMs (aerosol chemical speciation monitors), valuable and robust tools for measuring chemical composition of submicron aerosol with good time resolution on a long-term basis (Petit et al., 2012; Canonaco et al., 2013; Ripoll et al., 2015).

Mineral dust contributes mainly to coarse particles, with higher concentrations in southern Europe (Putaud et al., 2010). The dust concentration in PM$_{10}$ is about 2 µg m$^{-3}$ in central Europe and varies between 4 and 14 µg m$^{-3}$ in the Mediterranean Basin. The dust concentration in PM$_{2.5}$ is less than 0.5 µg m$^{-3}$ in central Europe and varies between 1–2 µg m$^{-3}$ in southern Europe (Querol et al., 2009). The PM$_{10}$ dust mass fraction increases moving from rural to kerbside backgrounds, due to the resuspension of soil dust caused by traffic (Putaud et al., 2010).

The dust mass fraction in PM$_{2.5}$, and especially in PM$_{10}$, shows a strong seasonality with higher values during summer, due to more frequent and more intense dust transport episodes from Africa and to drier meteorological conditions, which favour dust resuspension from ground. In southern Europe, mineral dust represents about 30–40 % of PM$_{10}$ in summer and 10–15 % in winter (Aas et al., 2012).

Sea salt concentration is usually below 1 µg m$^{-3}$ and represents 2–24 % of PM$_{10}$ mass and 1–8 % of PM$_{2.5}$ (Putaud et al., 2010; Querol et al., 2009). The PM$_{10}$ sea salt concentration shows a large decreasing gradient with increasing distance from the sea (Aas et al., 2012; Querol et al., 2009).
Carbonaceous aerosol is composed of EC (or BC) and OC. The terms EC and BC refer, respectively, to the refractory and the light-absorbing carbonaceous atmospheric aerosol emitted by combustion (Bond et al., 2004). OC is a mixture of thousands of different species, whose mass concentration is commonly expressed as organic mass (OM).

Annual average EC concentration over Europe varies by a factor of almost 20, with the lowest values in Norway and the highest in northern Italy (Po Valley). In 2010 the EC annual average over the EMEP rural background stations ranged between 0.06 and 1.3 μg m\(^{-3}\) (EMEP, 2012). During the EMEP intensive operational period in 2002–2003, the range was 0.2–1.8 μg m\(^{-3}\) (Yttri et al., 2007). The EC concentration and mass fraction increases moving from natural background sites to kerbside areas and from northern to southern Europe (Putaud et al., 2004, 2010). EC represents 3–6% of PM\(_{10}\) at rural sites and 9–17% at the kerbside (Putaud et al., 2010).

Most of the EMEP sites (regional and rural background) show a maximum in the EC concentration during winter both in PM\(_{2.5}\) and PM\(_{10}\), due to high emissions from residential heating (biomass burning and fossil fuel use) and stagnant meteorological conditions (Querol et al., 2009; Tørseth et al., 2012). In 2010, the winter-to-summer EC ratio was 1.3–2.9, with higher values in areas characterized by higher EC concentrations, such as Ispra (northern Italy) and Melpitz (Germany) (EMEP, 2012).

Only a limited number of sites report long-term measurements of EC in both PM\(_{2.5}\) and PM\(_{10}\). At three EMEP Norwegian sites, EC in PM\(_{2.5}\) represents 88–94% of EC in PM\(_{10}\). High percentage values were observed in Germany as well, indicating that EC is present mainly in fine particles resulting from incomplete combustion of fossil fuels and biomass (EMEP, 2012). Size-segregated measurements of EC in urban, rural and regional background sites during shorter campaigns show that EC is predominantly associated with submicron particles (Mieiro et al., 2007; Putaud et al., 2004b).

In 2010, among the EMEP sites, the annual average of OC concentration in PM\(_{2.5}\) ranged between 0.9 μg m\(^{-3}\) in Birkenes (Norway) and 2.0 μg m\(^{-3}\) in Finokalia (Greece). During the EMEP intensive campaign in 2002–2003, PM\(_{10}\) OC measurements were also performed in more polluted regions, such as the Po Valley, and a larger variability of annual means was observed (1.2–7.8 μg m\(^{-3}\)) (Yttri et al., 2007). Yttri et al. (2007) reported a higher concentration of OC at rural background sites in central and southern Europe compared to northern Europe. At rural, urban and kerbside sites of central and southern Europe, OM accounts for 21–26% of PM\(_{10}\) and 15–26% of PM\(_{2.5}\); in north-western Europe it accounts for 15–26% of PM\(_{2.5}\) and PM\(_{10}\) (Putaud et al., 2010).

Most of the European rural and urban background sites show higher OC concentrations in winter than in summer, as a consequence of the larger influence of primary and secondary anthropogenic emissions from residential heating and the accumulation of pollutants due to stagnant meteorological conditions (Yttri et al., 2007). Winter concentration could be as high as 3 times the summer values (EMEP, 2012). The same seasonal trend in urban and most of the rural background sites indicates a strong influence of anthropogenic emissions on rural areas in Europe (Yttri et al., 2007). In Scandinavia and at Mediterranean regional background sites, summer OC concentrations were higher, likely due to a larger influence of biogenic primary and secondary organic aerosol (BSOA) (Aas et al., 2012; Querol et al., 2009; Tørseth et al., 2012; Yttri et al., 2007).

OC in fine particles is associated mainly with primary emissions from combustion sources (fossil fuel and biomass combustion) and secondary aerosol from natural and anthropogenic sources (Gelencsér et al., 2007; Gilardoni et al., 2011a). Secondary biogenic aerosol does not include only monoterpane, isoprene and sesquiterpene oxidation products but also amines as detected at several sites, including the Boreal forest and the North Atlantic (Facchini et al., 2008; Kulmala et al., 2013). OC in coarse particles derives mainly from primary biogenic aerosol particles (Yttri et al., 2007). The size distribution of OC has been measured continuously only at a limited number of sites. In Birkenes and Melpitz, the PM\(_{2.5}\)-to-PM\(_{10}\) ratio is 70 and 76%, respectively (EMEP, 2012). Tørseth et al. (2012) observed a similar ratio in Birkenes, Norway. In the Mediterranean areas, the ratio is about 1 (Querol et al., 2009).

OA in PM\(_{1}\) was most frequently analysed during short field campaigns with aerosol mass spectrometers: in Europe organic mass accounts for about 15–60% of submicron mass (Zhang et al., 2007). PM\(_{1}\) OC is dominated by oxidized OA, whose mass fraction increases moving from urban to rural and regional background sites (Zhang et al., 2007). In spring 2007 and 2008 several simultaneous field experiments were performed across Europe during the EUCAARI (European Integrated Project on Aerosol, Cloud, Climate, and Air Quality Interactions) project. Regional and rural background sites exhibited more processed and aged OA, characterized by a larger fraction of LV-OOA and highly functionalized molecules (Kulmala et al., 2011). Less OOA was observed in urban areas and characterized by less functionalized aliphatic molecules, amines, aromatics and alcohols (Kulmala et al., 2011).

### 2.2.4 Remote sensing

The short lifetime of aerosols in the atmosphere and the large variety of sources and aerosol species result in high temporal and spatial heterogeneity in aerosol distribution around the globe (Kaufman et al., 2002; Ramanathan et al., 2001). Therefore, aerosol remote sensing (RS) from space and over continental ground networks is the only means of obtaining measurements on a global scale. One of the most successful series of aerosol measurements from remote-sensing missions comes from the AERONET (Aerosol Robotic Network) surface network of stations. The network has hundreds...
of stations spread over the continents as well as measurements from ships and focused field campaigns. Each station provides well-calibrated aerosol optical depth (AOD) (with accuracy of $\sim 0.015$) as well as fundamental optical and size information such as aerosol single-scattering albedo, size distributions, fine-mode fraction, degree of non-sphericity, phase function and asymmetry factor (Dubovik and King, 2000; Dubovik et al., 2002; O’Neill et al., 2004).

Many studies discuss the differences between aerosol properties as measured locally by surface stations and the large-scale ones measured by satellites (van Donkelaar et al., 2010). Great efforts are invested in trying to find systematic links between the scales as represented by the different measurements methods. Many of the recent efforts try to merge satellite remote-sensing information with local PM measurements by ground stations and aerosol transport models. In such a way one can use the best type of information from each data source. Satellite or AERONET data are often used as the best information for the total aerosol loading in the atmospheric column and for providing constraints on the fine and coarse fractions, while the detailed aerosol composition is derived by the PM surface measurements and the transport models output.

Air over Europe is widely monitored by ground measuring systems and from space. A number of studies have been conducted to compare and evaluate the different measuring techniques and to find the best way to merge them. In such a study, Basart et al. (2012) used the CALIOPE (CALidad del aire Operacional Para Espana) modelling system combined with daily PM$_{10}$, PM$_{2.5}$ and aerosol component data from 55 EMEP/CREATE ground stations and from 35 AERONET stations. These authors showed annual correlations of $\sim 0.5$ between modelled and observed values for PM$_{10}$ and PM$_{2.5}$ and slightly higher correlations with total, coarse- and fine-mode AOD. Desert dust controlled most of the coarse-aerosol correlations, and clear underestimations of the model fine-aerosol output were attributed to underestimated levels of carbonaceous matter (EC and OC) and SIA.

According to analysis of Basart et al. (2012), the key factors affecting aerosol amount and properties over Europe are local anthropogenic emissions, the proximity to the Sahara and large-scale meteorology. They found the highest aerosol concentrations over the Po Valley and the Benelux regions with a second maximum over eastern and southern Europe. Lower PM concentrations were reported towards the north and north-west of Europe.

Robles González et al. (2003) reported on AOD and Ångström coefficients over Europe retrieved from the ATSR-2 (Along Track Scanning Radiometer; Veedvind and De Leeuw, 1998). They showed that by combining satellite data and model output, they could evaluate contributions of specific PM components to the total AOD. Their results provided information on the anthropogenic emissions of aerosol precursors such as SO$_2$ and NO$_x$ over industrial and urban areas. They combined the data with a regional chemistry transport model to show that about 15% of the total AOD in pristine areas, and as much as 70% in polluted ones, is contributed by sulfate. Nitrate contributes between 5 and 25% over most of Europe.

Bovchaliuk et al. (2013) studied aerosol properties from several AERONET stations over eastern Europe. Combined with POLDER (POLarization and Directionality of the Earth’s Reflectances) data they showed the evolution in the aerosol type and size distribution for major cities. A seasonal variability is shown, with peaks in April–May and August–September. The main contributors of the spring peak are Saharan dust, agricultural fires and sea salt transported to the Ukraine and Moldova from the Black and Azov seas. The August–September peak main contributors are forest and peat wildfires and Saharan dust. They also showed that over industrial areas, the aerosol types are probably mostly of anthropogenic origin.

Finally, a recent study (Chin et al., 2014) combined the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model with satellite and surface observations to study global aerosol variations and trends. The authors of the study showed how anthropogenic trends are modulated over the natural ones and that during the past 30 years the largest reduction in aerosol concentrations occurred over Europe, with a 40–60% decrease in AOD and reduction by a factor of 3–4 in surface sulfate concentrations. Chin et al (2014) attributed the widespread decrease in European sulfur concentration to the substantial decline in SO$_2$ emissions (Fig. 11).

### 2.3 Atmospheric evolution of aerosol

The atmospheric aerosol number and mass concentration and chemical and microphysical properties change continuously through a series of physical and chemical processes (Rudich et al., 2007). A schematic representation of aerosol sources and processes is shown in Fig. 12. Processes altering particle composition – i.e. new particle formation, condensation and evaporation, water uptake, heterogeneous chemistry, dispersion and removal – are represented as blue arrows. The coagulation process, which is relevant especially for UFPs, affects only particle number and size and is not shown in the scheme. Particle mass and number concentrations are linked to each other through the size distribution of the atmospheric aerosol population, and while UFPs typically dominate the aerosol number, the particulate mass is governed by larger particles. A consistent description of the particle mass and PN is thus critical for accurate predictions of the interactions between air quality and climate.

Atmospheric aerosol particle properties are quite inhomogeneous close to their sources, where they maintain the corresponding source characteristics (chemical composition, size, physical and chemical properties), but atmospheric processing tends to make them more homogeneous. Especially submicrometre PM gets heavily processed, and its mass distribution and composition are dominated by gas-to-particle
conversion products. Most of the sub-micrometre particle mass exists in the accumulation mode and where particles appear to be generally internally mixed away from the corresponding source regions. The new particles entering the atmosphere (either through emission or nucleation) are coated by sulfates, secondary organics, nitrates, etc., and therefore gradually become similar to each other. The conversion of primary OA to oxygenated OA discussed in Sect. 2.3.2 also contributes to this process. Most primary particles maintain in their core non-volatile elements characteristic of their source (EC, metal oxides, etc.), but for fine aerosol these components usually represent a small fraction of the particle mass. Single-particle mass spectrometers are sensitive to these elements characteristic of the particle source (Prather et al., 1994), but as the particles age, these measurements are more relevant for the determination of the PN than the particle mass.

This atmospheric processing of aerosol tends to simplify their description far from their sources. Timescales of less than 1 day are required in photochemically active environments for the conversion of complex aerosol mixtures to internally mixed populations (Wang et al., 2010). The chemical aging of OA also leads to a reduction in the variability of their hygroscopic properties. Engelhart et al. (2012) showed, for example, that even if aerosol produced during the burning of different types of biomass had very different CCN activities (the value of hygroscopicity parameter $\kappa$ varied from 0.06 to 0.6 depending on the fuel; see Petters and Kreidenweis, 2007), its chemical aging led to convergence to a $\kappa$ of $0.2 \pm 0.1$. This relatively simple behaviour enables agreement between measured and predicted CCN concentrations in environments dominated by aged aerosol within less than 1 % (Bougiatioti et al., 2009; Jurányi et al., 2011).

The evolution of BC is of special interest because of its important role as an absorbing material in the atmosphere. BC is originally emitted together with organic compounds in chain-like fractal aggregates. These particles collapse into denser particle clusters as water vapour and other gas-phase species condense on them (see Bond et al., 2013 for a review). This coating changes the morphology but also the hygroscopic and optical properties of the corresponding particles. These BC-containing particles become more spherical and can be removed more quickly from the atmosphere by wet deposition but can also absorb more solar radiation. The additional

Figure 11. Left column: monthly averaged surface SO$_2$ concentration from 1980 to 2008 at two European sites: Tange in Denmark (top two rows) and Chopok in Slovakia (bottom two rows). Observations are shown in black lines and the model results in grey bars for total amount; dotted red lines indicate the amount from fossil fuel and biomass burning. Right column: corresponding ratio of annual mean concentrations to 2000–2001 average, showing the interannual variability and trend over the 20-year period, with the observation and models as solid and dotted lines, respectively (Chin et al., 2014).
Figure 12. Sources and processes affecting atmospheric aerosol composition from near-source to continental scales. The y axis indicates the relative importance of primary vs. secondary aerosol for a given source, with primary-dominated sources at the top and secondary-dominated sources at the bottom. Representative sources are marked on the left, e.g. sea salt is a primary-only source, biogenic emissions are secondary-only sources and biomass burning is a mixed primary and secondary source. Green colours reflect the importance of particulate material relative to gaseous precursors and/or depositional losses (grey) as source emissions age over increasing temporal and spatial scales (darker shading, x axis). Processes altering particle composition are shown as blue arrows. The figure highlights the dominance of secondary aerosol across broad spatial scales and the local importance of both primary and secondary aerosol.

Figure 13. Schematic examples of the number (blue solid line) and mass (red dashed line) size distribution describing the same aerosol population. Different size-dependent dynamic processes govern aerosol number and mass loadings; the most important of these processes are outlined in the text boxes. The aerosol particles in different parts of the size distribution are in constant interaction with, on one hand, the gas phase through condensation and evaporation, and, on the other hand, with each other through coagulation.

2.3.1 In situ formation of new particles

Besides direct emissions of primary particles, the only known source of new atmospheric particles is the production of condensable vapours, their clustering and subsequent condensation. In fact, most atmospheric UFPs have been formed through this in situ new particle formation. After being formed at sizes close to 1–2 nm (Kulmala et al., 2007, 2013), the new particles are rapidly lost by coagulation to the larger end of the size distribution (Fig. 13). This causes the importance of new particle formation as a source to decrease with particle size, governed by the rate at which they grow towards larger sizes and the total aerosol loadings (see Kerminen and Kulmala, 2002; Kuang et al., 2010; Vehkamäki and Riipinen, 2012). For example, Merikanto et al. (2009) estimated that in situ particle formation is responsible for 73% of all the particles on a global scale and 45% of the CCN at 0.2%. Pierce and Adams (2009a) reached similar conclusions but reported somewhat smaller contributions for nucleation. Fountoukis et al. (2012) studied the summertime contribution of new particle formation on aerosol and CCN numbers in Europe with a regional model and predicted large increases (of up to a factor of 20 or even more at some locations) to the total PN concentration and of a factor of 7 for particles larger than 10 nm for specific regions, in particular south-eastern Europe. For particles larger than 50 nm, an increase of more than 100% was predicted for south-eastern Europe. The effect of nucleation in the lower-tropospheric PN above 100 nm was, on average, predicted to be small, with the exception of the Mediterranean area. On average, approximately 50% of particles above 10 nm were predicted...
Gas-phase bases such as amines and ammonia can enhance the sulfuric acid nucleation rate by several orders of magnitude (Almeida et al., 2013; Kirkby et al., 2011; see Fig. 14), and where present in high enough concentrations, these compounds thus probably participate in atmospheric new particle formation as well (Erupe et al., 2011; Fontoukis et al., 2012). Organic compounds of biogenic origin can also form new particles with sulfuric acid at rates that appear to explain the seasonal cycle of particle concentrations in the boundary layer (Riccobono et al., 2014). While it is known that the presence of ions can enhance particle formation rates, especially in very clean conditions (Kirkby et al., 2011), the literature suggests that the effect of changes of cosmic rays on atmospheric particle formation and CCN appears to be too small to play a significant role in climate change through their impact on particle formation (Kulmala et al., 2010; Pierce and Adams, 2009b; Snow-Kropl et al., 2011).

The concentrations of CCN formed because of nucleation are at least as much controlled by particle growth as they are by the nucleation rate (Kerminen et al., 2012). Large uncertainties regarding the growth rate of the smallest nuclei and the compounds responsible for this growth remain. In most cases, sulfuric acid levels are not sufficient to explain the ob-

**Figure 14.** The formation rates of 1.7 nm as a function of sulfuric acid vapour concentration in the CLOUD (Cosmics Leaving Outdoor Droplets) chamber in Cern (Almeida et al., 2013; Riccobono et al., 2014). (a) The effect of ammonia and dimethylamine (DMA) on the particle formation rates (see Almeida et al., 2013 for details). The markers and solid lines represent experimental data from the CLOUD chamber at 38 % RH and 278 K. Curve 1 represents nearly base-free (bases present only as contaminants) conditions, while the DMA concentration increases gradually in curves 2–5. The colour scale indicates the mixing ratio of DMA and ammonia. The small coloured squares represent field observations in the atmospheric boundary layer. (b) The effect of oxidized biogenic organics on the particle formation rate (see Riccobono et al., 2014 for details). The open markers correspond to observations without the presence of the biogenic organic precursor, while the filled markers represent the cases where organics were present. The different colours of the markers represent cases with different ionization conditions. The atmospheric observations are indicated by the green outline.
There is not, however, a general description of or the-
sible for aerosol formation in different environments, and
it has been shown that different compounds can be respon-
sition in different environments has been made in the past 2
decades, some important questions still remain. For instance,
participate in the very first steps of in situ particle formation as well
as contribute to the loss processes removing newly formed particle
from the atmosphere through increasing average sizes of the Aitken
and accumulation particles (referred to as “Background aerosol” in
the figure).

Figure 15. Schematic presentation of links between emissions of
atmospheric organic compounds, in situ particle formation and
growth (adapted from Riipinen et al., 2011). Besides growing the
particles to climatically relevant sizes, oxidized organics can par-
ticipate in the very first steps of in situ particle formation as well
as contribute to the loss processes removing newly formed particle
from the atmosphere through increasing average sizes of the Aitken
and accumulation particles (referred to as “Background aerosol” in
the figure).

served growth of fresh nanoparticles (Kuang et al., 2010; Ri-
ipinen et al., 2011). Sulfur-dioxide-rich environments are an
exception (Jung et al., 2006). Low-volatility organic vapours
appear to dominate the growth of the fresh nuclei in many
environments (Kerminen et al., 2000; Kulmala et al., 1998;
Riipinen et al., 2011; see Fig. 15). The exact identity and the
sources of these vapours are yet to be elucidated (Riipinen et
al., 2012), although recent studies have pointed to the gas-
or particle-phase formation of organic vapours with extremely
low volatilities (< 10^{-3} \text{ g m}^{-3}) accounting for an important
fraction of this growth (Kulmala et al., 1998; Donahue et al.,
2011; Hääkkinen et al., 2012; Pierce et al., 2011; Riipinen et
al., 2011; Ziemann and Atkinson, 2012; Shiraiwa et al., 2013;
Yli-Juuti et al., 2013; Ehn et al., 2014). Besides growing the
newly formed particles towards larger sizes and thus increas-
ing their atmospheric lifetime, these vapours might also be
important compounds in the very first steps of new particle
formation as well (Metzger et al., 2010; Riccobono et al.,
2014).

While significant progress towards a better understanding
of the molecular processes of in situ aerosol particle forma-
tion in different environments has been made in the past 2
decades, some important questions still remain. For instance,
its has been shown that different compounds can be respon-
sible for aerosol formation in different environments, and
their role can vary significantly also depending on the alti-
tude. There is not, however, a general description of or the-

2.3.2 Closing the gap between volatile organic
compounds and organic aerosol

OA has been traditionally viewed as a relatively inert, non-
volatile mixture of compounds from a complex array of pri-
mary sources (primary organic aerosol), coated by secondary
compounds derived from the gas-phase oxidation of volatile
precursors (SOA). The immense chemical complexity of the
OA, with thousands of complex organic compounds in typ-
ical ambient aerosol (Goldstein and Galbally, 2007), its un-
known chemical composition (less than 20% of the OA mass
has been quantified; Fuzzi et al., 2006; Hallquist et al., 2009),
the unknown physical and chemical properties of the ma-
ajority of the known OA components and the difficulty of
describing mathematically such a complex system in atmo-
spheric chemical transport models (CTMs) have seriously
limited scientific progress in both the air quality and climate
change areas. Both regional- and global-scale CTMs are of-
ten unable to reproduce the observed OA levels, their chem-
ical characteristics (degree of oxidation), their diurnal varia-
tion, etc. (Kanakidou et al., 2005; McMurry et al., 2004). As
a result, the evaluation of the effects of different strategies of
the reduction in OA concentrations in polluted areas and the
quantification of the effects of OA on the energy balance of
the planet remain challenging.

Robinson et al. (2007) suggested a possible explanation
for the observed OA mass and composition (Fig. 16). Most
of the emitted organic PM from combustion sources such as
transportation, biomass burning, etc., evaporates after emis-
sion. The resulting semi-volatile organic vapours can then re-
act in the gas phase with atmospheric oxidants forming low-
volatility oxidation products that can recondense onto the
particulate phase on timescales of several hours or even days.
This evaporation–reaction–condensation process results in
significant changes in the chemical nature of primary OA
(it becomes highly oxygenated), its size distribution, its dis-
tribution in space, and its physical and chemical properties.
This mechanism may explain why the aerosol in large urban
centres is dominated by oxygenated compounds (OOA) and
not hydrocarbon-like OA (Zhang et al., 2007).

Similar challenges exist in the efforts to understand SOA.
Initially the SOA precursors were assumed to form low-
volatility SOA with a constant yield (Pandis et al., 1992).
This description is actually still used in most climate models (Kanakidou et al., 2005). The next step was to assume the production of two surrogate products that form a pseudo-ideal solution (Odum et al., 1996). This approach is used currently by the majority of the available regional CTMs. However, the resulting models fail to reproduce observed OA concentrations in the US, Europe, Asia, etc., especially in periods (warm sunny summer days) when SOA is expected to be the dominant OA component. Identified problems include neglecting IVOCs. These compounds have saturation vapour pressures which are lower than the VOCs that dominate the anthropogenic and biogenic organic emissions to the atmosphere but are higher than the POA components (Fig. 16). Despite their relatively small emissions, they can be important sources of SOA because they can be rapidly oxidized to Semi-volatile organic compounds (SVOCs). The role of IVOCs as SOA precursors appears to be important on urban (Hodzic et al., 2010), regional (Fountoukis et al., 2011) and global scales (Pye and Seinfeld, 2010) but remains quite uncertain due to a lack of understanding of their emissions and their aerosol-forming potential.

The development and use of new methods for the quantification of the organic emissions to the atmosphere, taking into account compounds of all volatilities, is a pressing need. The measurement of only VOCs and particles (at least under the conditions of the measurements) leaves a major gap of semi-volatile and intermediate-volatility compounds. Quantification of these emissions as a function of volatility (e.g. using dilution sampling at different dilution levels) may be a major step forward. In addition, re-evaluation of yields observed in chamber studies is probably necessary. Recent studies (Matsunaga and Ziemann, 2010; Loza et al., 2010; Zhang et al., 2014) have suggested that SOA formation in chamber experiments may be substantially suppressed due to losses of SOA-forming precursors to the chamber walls. This process may lead to underestimates of SOA formation in models.

2.3.3 Aging of organic aerosol

Most organic compounds emitted in the atmosphere are highly reduced (Fig. 17). The average O:C atomic ratio of these compounds is less than 0.15 (Aiken et al., 2008). OA formed during biomass burning is an exception and has average O:C around 0.3 (Aiken et al., 2008). After their initial partitioning between the gas and particulate phases, all these compounds are oxidized via quite complex chemical pathways. Most laboratory studies have focused on the first stage of the oxidation process which takes place over a few hours. The secondary organic aerosol formed during this initial series of reactions has an O:C of around 0.4 (the exact value depends on the precursor, oxidant and concentration level of the OA formed) (Donahue et al., 2012b). The OA components at this stage are mostly semi-volatile, and they exist partially in the particulate and partially in the gas phase. These compounds continue to react in both phases in the atmosphere mainly with the OH radical but also with other oxidants. These subsequent stages of reactions are known as chemical aging of the OA (Jimenez et al., 2009).

However, most OA after a few days in the atmosphere appears to be in the form of what has been called low-volatility OOA (Fig. 17). This OOA has an O:C of around 0.9 and an effective saturation concentration smaller or much smaller than 1 µg m\(^{-3}\) (Jimenez et al., 2009; Hildebrandt et al., 2010a). The specifics of the chemical pathways of this conversion from fresh vapours and organic PM to highly oxygenated OA are still poorly understood. The timescale required for this conversion in an environment with high photochemical activity appears to be of the order of 1–2 days (Hildebrandt et al., 2010b) and gas-phase reactions of the semi-volatile vapours with OH appear to be an important pathway for the conversion.

number of studies have explored the heterogeneous chemistry in deliquesced aerosols and organic particles.

OH radicals, ozone and other atmospheric oxidants continue reacting with the organic compounds present in the particulate phase. There have been numerous laboratory studies of single component model systems where rapid oxidation (timescale of minutes to a few hours) was observed (Rudich et al., 2007). However, available laboratory measurements with multicomponent particles and field measurements suggest that these reactions proceed much more slowly in the ambient atmosphere, probably due to mass transfer delays (Shiraiwa et al., 2011b; Weitkamp et al., 2008). In this way, the particulate phase appears to provide a “safe” environment for organic compounds, prolonging their atmospheric life. Donahue et al. (2013) proposed that OAs exist throughout the troposphere in part because their heterogeneous oxidation by OH radicals is 1 order of magnitude slower than the corresponding reaction in the gas phase.

Reactions in particles can also lead to the formation of oligomers (Kalberer et al., 2004) and humic-like substances (HULIS) with a high molecular weight (Graber and Rudich, 2006). The estimated fraction of HULIS in the ambient OA is quite variable, ranging from approximately 10 % to as much as 50 %. Atmospheric HULIS, however, consist of much smaller molecules compared to terrestrial and aquatic humic substances. A number of reaction pathways for the formation of HULIS in atmospheric particles have been proposed (see Graber and Rudich, 2006, for a review) but their relative importance is not well understood. The quantification of the concentrations of oligomers in ambient OA has been difficult due to a lack of reference compounds (Iinuma et al., 2007; Hallquist et al., 2009).

Organic aerosol formation reactions take place in cloud and aerosol water (Ervens et al., 2011). The oxidation of glyoxal and methylglyoxal (oxidation products of isoprene but also aromatic VOCs) leading to the production of oxalate, among other compounds, is a major chemical pathway. While these aqueous-phase reactions clearly contribute to SOA formation in the atmosphere, the corresponding magnitude of the SOA formed remains an issue of debate. Model estimates vary from modest contributions of around 5–10 % of the global water-soluble OA burden (Myriokefalitakis et al., 2010) to contributions similar in magnitude to the gas-phase chemistry (Ervens et al., 2011; Liu et al., 2012).

Organosulfate compounds can be produced during the reactions of acidic PM and the products of the oxidation of BVOCs (Surratt et al., 2007). These reactions can increase the amount of SOA formed during the oxidation of the corresponding precursors. The corresponding compounds are both biogenic and anthropogenic in origin and are of potential interest for both air quality and climate change. The reaction pathways leading to the formation of these sulfate esters are complex. For example, Hatch et al. (2011) provided evidence that isoprene is oxidized to an epoxide which then partitions into the particulate phase, especially during high-RH peri-
ods and is then converted to organosulfate through an acid catalysed mechanism. α- and β-Pinene also form organosulfates. The contribution of organosulfates to atmospheric OA remains uncertain but has been estimated to be as much as 20% in some environments (Hallquist et al., 2009).

The aforementioned chemical pathways involve polar organic compounds dissolved in aqueous media (cloud water, deliquesced aerosols), but particulate organic compounds can also occur in organic phases. Organic phases coexisting with an aqueous phase in the aerosol were shown to exist in the sub-saturated humid atmosphere (Smith et al., 2013) and were originally predicted by the thermodynamics of simple organic–inorganic water mixtures (Clegg et al., 2001). The physico-chemical properties of particulate organic mixtures recently attracted increasing interest in aerosol science since Virtanen et al. (2010) found that SOA particles “bounce” on impactor plates like semi-rigid solid particles instead of “splashing” like liquid organic droplets. Koop et al. (2011) provided a new theoretical framework for understanding key physical properties of organic phases in the aerosol (e.g. viscosity and diffusivity) and predicted the existence of “glassy” aerosols: particles composed of metastable mixtures of organic compounds, which are highly viscous and show very long timescales for equilibration (including the diffusivity of gaseous oxidants into the particles). Recently, Booth et al. (2014) provided experimental data and a thermodynamic description of glassy aerosols composed of simple mixtures of oxygenated organic compounds (dicarboxylic acids). Saukko et al. (2012) confirmed the model of Koop et al. (2011) with observations of SOA bouncing behaviour in laboratory conditions but also found an effect of molecular composition on the apparent viscosity of the particles (aerosols made up of hydrocarbons behave much less like viscous particles than those composed of oxygenated species).

Laboratory studies have demonstrated the potential importance of many condensed-phase reactions (oligomerization, organosulfate formation, aqueous-phase SOA formation, etc.), and field studies have confirmed that these reactions take place in the ambient atmosphere at non-negligible rates. However, the importance of these pathways for atmospheric OA formation compared to the better-understood gas-phase formation pathways remains uncertain.

### 2.4 Particulate-matter wet and dry deposition

The formation and delivery of precipitation to the Earth’s surface represents a series of processes which remove particulate matter as well as water from the atmosphere very efficiently. The overall scavenging process is referred to as wet deposition and includes cloud condensation on particulate matter (nucleation scavenging) and in-cloud processes which incorporate additional particulate matter. The in-cloud processes are collectively referred to as rainout and may be distinguished from below-cloud scavenging, in which falling rain or snow intercept particulate matter and which is referred to as washout. A detailed review of wet-scavenging processes for sulfur-containing aerosols is provided by Garland (1978) and a more general review of aerosol scavenging by Fowler et al. (2009).

Wet deposition may either occur from warm clouds, i.e. through coalescence of cloud droplets, or via ice formation (see Sect. 3.2.1).

**Nucleation scavenging**, the formation of cloud droplets on particulate matter, is the most important mechanism transferring particulate matter to the wet-deposition pathway, and it accounts for the majority of the wetly deposited SO₄, NO₃ and NH₄ (Garland, 1978; Flossmann et al., 1985). The composition, size and relative humidity influence water vapour uptake by particulate matter, and above a critical size they activate to form cloud droplets (Pruppacher and Klett, 1997). Recent studies have shown the importance of organic compounds, which may coat particulate matter and influence surface activation properties (Topping et al., 2007; Kokkola et al., 2006). Particulate matter derived from biological materials, including plant debris and especially epicuticular waxes, bacteria and fungi, has been shown to influence the ice nucleation process (Deguillaume et al., 2008; Ariya and Amyot, 2004).

In addition to nucleation scavenging, cloud droplets may collect additional particulate matter through Brownian diffusion, impaction and interception and by a range of phoretic mechanisms (electrophoresis, thermophoresis and diffusiphoresis) discussed by Garland (1978). However, these processes are considered to be minor contributors to observed solutes in wet deposition.

The widespread adoption of precipitation composition networks in Europe, North America and Asia over the last 30 years has provided a valuable resource to estimate regional and global wet deposition of a range of pollutant species (Vet et al., 2014). The measurements have also been vital in order to monitor trends in the composition of precipitation in Europe (Fagerli et al., 2008) and in North America (Vet and Ro, 2008) as a check against output from regional models of emission, transport and deposition. However, the widespread use of these network measurements has not been matched by process studies to identify the contributions of the different in- and below-cloud processes to the measured solute concentrations in the precipitation samples. This process study is clearly a neglected field. Even the contribution of dry deposition of particulate matter and reactive gases to precipitation samples remains poorly understood.

The use of models to simulate the tropospheric processing and deposition of natural and anthropogenic emissions of SO₂, NO, NO₂, NH₃, VOC and the production of oxidants require wet- and dry-deposition schemes. For wet scavenging, many of the current schemes to simulate wet scavenging of particulate matter are validated against ²¹⁰Pb inventories in air and precipitation on global scales (Giannakopoulos et al., 1999).
In addition to removal by precipitation, dry deposition is another removal mechanism of PM from the atmosphere. The particles are transported towards the surface primarily by turbulence and captured by interaction with the surface (mainly vegetation) elements.

The principal physical processes that control the interaction with vegetation are thought to be fairly well understood. Transport through the turbulent part of the surface layer is by eddy diffusion and gravitational settling, while the main controls of the deposition rates are the processes by which particles overcome the quasi-laminar sub-layer that forms over all surface elements in which turbulence is suppressed. Very small particles (< 100 nm) travel through this layer by Brownian diffusion, which also governs the laminar sub-layer resistance of gas molecules, while coarse particles (> 2.5 µm) overcome this layer by gravitation. In the intermediate size range, impaction (the deposition of particles due to their inability to follow tight curvatures in stream lines, governed by particle mass) and interception (the possibility of larger particles to make contact with surface elements) are the main processes. In addition to Brownian diffusion, there are several phoretic processes by which particles may be transported across the laminar sub-layer. These include electrophoresis (driven by charge differences between the particle and the surface), thermophoresis (driven by temperature differences) and diffusiophoresis (driven by vapour gradients, especially of water). These phoretic transport mechanisms have been reviewed by Garland (1978). The processes contributing to the dry deposition of particles on vegetation are illustrated diagrammatically in Fig. 19.

Gravitational settling provides an additional, parallel non-diffusive process for the deposition of super-micron particles. Because impaction and interception are less efficient than Brownian diffusion and gravitational settling, theoretical predictions of the deposition velocity \( V_d \) as a function of particle size show a minimum in the region 0.1 to 1 µm aerodynamic diameter \( D_p \). This is qualitatively supported by the fact that this is the size range in which secondary aerosol components “accumulate” in the atmosphere and which undergoes long-range transport. Parameterizations of \( V_d(D_p) \) emerged in the 1980s and were initially developed for grass on the basis of wind tunnel experiments (Slinn, 1982) and widely applied. More recently these concepts have been developed into more generic models which simulate deposition to different layers in the canopy, an addition that has been found to be particularly important for forests (Petroff et al., 2007a; Peters and Eiden, 1992). Parameterizations of impaction and interception are highly sensitive to the choice of the size of the surface elements (“characteristic dimensions”) and in some models (Petroff et al., 2007a) also to leaf orientation. Variation within a justifiable parameter range can change the deposition rates by 1 order of magnitude. Indeed measurements of deposition rates to different grass species reflect this variability (Davidson et al., 1982). This variability means that while models can be “tuned” to reproduce measurements on the plot scale, it is unlikely that spatially disaggregated input data on vegetation characteristics and land use will become available to improve over a minimum uncertainty for use in aerosol transport models.

Several review papers have summarized existing models and measurement data (Sehmel, 1980; Petroff et al., 2007b; Pryor et al., 2008; Zhang and Vet, 2006): depending on turbulence and atmospheric stability, the deposition rates of the accumulation mode typically lie within the range 0.1 to 2 mm s\(^{-1}\) for short vegetation and 1 to 5 mm s\(^{-1}\) for tall vegetation (forests). Such moderate values are consistent with the atmospheric lifetime derived following emission spikes (e.g. volcano eruptions and nuclear accidents) and long-term integrated values, e.g. derived from \(^{210}\)Pb inventories under vegetation (Fowler et al., 2004).

In the 1990s measurements of aerosol deposition velocities emerged, which suggested much larger deposition rates for aerodynamically rough surfaces, and especially forests. The different studies included a wide range of methods. Some of the measurements were made over a Dutch forest. The measurements showed large deposition rates for nitrate aerosol derived from filter pack gradients (Wyers and Duyzer, 1997) and high deposition velocities and a steeper increase in \( V_d \) with \( D_p \) than predicted by the theoretical approaches (Gallagher et al., 1997); this has been reproduced in several studies also over shorter vegetation (Nemitz et al., 2004). In response, some Dutch chemistry and transport models switched to an empirical parameterization based on the work of Ruijgrook et al. (1997).

With the development of measurement systems for chemically resolved aerosol fluxes, based on gradient approaches using aerosol-into-liquid samplers at multiple heights and eddy-covariance approaches deploying fast aerosol mass spectrometry, new data sets have emerged that have led to a reinterpretation of the historical measurements: the results show that measured deposition rates differ greatly between chemical compounds. Thermodynamically stable compounds such as sulfate show the low deposition rates predicted by the models. Volatile compounds, in particular...
ammonium nitrate, consistently show very high deposition rates, often exceeding 50 or even 100 mm s\(^{-1}\) during daytime (Fig. 20). This can be explained (and simulated) by the evaporation of these compounds below the measurement height, within and just above the plant canopies, and by the rapid deposition of the gaseous compounds released from the particles. The increased temperature near the canopy, together with reduced concentrations of ammonia and nitric acid (the latter is always taken up by the canopy; the former is usually deposited to semi-natural vegetation but may be emitted from agricultural surfaces), provides a driver for NH\(_4\)NO\(_3\) evaporation from particles as the approach the canopy.

The ramifications of these findings are twofold:

1. Different effective deposition rates of sulfate and nitrate. Much of the NH\(_4\)NO\(_3\) flux above the canopy is deposited as gaseous NH\(_3\) and HNO\(_3\) at the surface, and its effective removal rate is much larger than that of non-volatile compounds because the deposition of the gases is not restricted by the same physical controls. Thus, the evaporation results in an enhanced sink for NH\(_4\)NO\(_3\) and increased N deposition. The representation of the thermodynamic partitioning in CTMs cannot currently resolve the strong gradients near the canopy and can therefore not simulate the process mechanistically, although work is underway to couple CTMs with in-canopy chemistry models. In the meantime, empirical parameterizations are needed to predict effective NH\(_4\)NO\(_3\) deposition rates, and a first such parameterization has been implemented into the EMEP CTM (Simpson et al., 2012). By contrast, the deposition of non-volatile compounds is correctly modelled, with the slow deposition rates currently implemented in the models.

2. Artefacts for size-segregated flux measurements. It has been shown that size-segregated PN fluxes are sensitive to equilibration with water during the deposition process: as particles grow or shrink, they move between particle size bins, creating artificial gradients (Fairall, 1984; Kowalski, 2001). Evaporation or condensation of NH\(_4\)NO\(_3\) and other compounds from or to the particle during the deposition process induce equivalent effects, and it can be shown with models that this process explains the high accumulation mode deposition rates and steep \(V_d / D_p\) relationships (Nemitz and Sutton, 2004; Ryder, 2010). As a result, it is likely that any size-segregated flux measurements of an aerosol that includes even modest contributions from NH\(_4\)NO\(_3\) is unsuitable for assessing model predictions of \(V_d(D_p)\), without explicit correction. Accurate correction becomes difficult as soon as the aerosol is externally as well as internally mixed, i.e. if some particles contain a larger contribution of volatile components than other particles in the same size class. The opposite effect, i.e. apparent emission fluxes of aerosol due to particles growing to the size range detectable by a particle counter, has been observed over a fertilized grassland, where a high NH\(_3\) concentration near the ground resulted in aerosol growth during deposition (Nemitz et al., 2009).

When measurements are carefully selected, model–measurement comparisons can be successful (Petroff et al., 2007a), suggesting that the main physical processes are incorporated into these models. Recently, Katul et al. (2010, 2011) demonstrated that the incorporation of turbophoresis into a multilayer deposition model further improved model performance. Turbophoresis describes the tendency of particles to move from an area of high into an area of low turbulent energy.

3 Particulate-matter impacts and policy

3.1 Particulate matter and air quality

3.1.1 Particulate matter and health

Ambient PM is a recognized threat to public health on a global scale, not only in highly polluted environments (WHO, 2013a). Adverse health effects due to PM exposure have already been observed at PM concentrations slightly above background levels, i.e. 3–5 \(\mu\)g m\(^{-3}\) (WHO, 2005). Adverse health effects related to PM exposure concern mainly respiratory and cardiovascular systems, but there is also limited evidence that PM affects atherosclerosis and leads to adverse birth outcome. Children and elderly are the population subgroups most sensitive to PM exposure impacts. In 2005, PM\(_{2.5}\) and ozone were responsible for 773 000 deaths from respiratory diseases, 186 000 deaths from lung cancer and 2 million deaths from cardiovascular diseases (Leelieveld et al., 2013). According to the World Health Organization, in 2012 air pollution was responsible for 7 million premature deaths, 3.7 million deaths from ambient air pollution and 4.3 million
deaths from household air pollution (WHO, 2015). Anderson et al. (2009) estimated 300 000 deaths per year in Europe due to primary PM$_{2.5}$ exposure and 245 000 due to secondary inorganic aerosol. In pollution hot spots, such as the Po Valley and the Netherlands, PM alone was responsible for a loss in statistical life expectancy of up to 12–36 months (Fig. 21, CAFE, 2005).

Epidemiological and human exposure studies show that both long- and short-term exposure to PM correlate with cardiovascular and respiratory morbidity and mortality (Anderson et al., 2012; Brook et al., 2010).

Pope and Dockery (2006) report a comprehensive review of epidemiological studies performed since 1997 which unequivocally proved the link between long-term PM exposure and mortality. The Harvard six-cities study was conducted on more than 8000 subjects, living in six US cities, and followed for a maximum of 14–16 years. The study showed that a 10 µg m$^{-3}$ increment of PM$_{2.5}$ was associated with a percentage relative risk increase (RRI) in mortality of 16 % (95 % CI 7–26) (Laden et al., 2006). The ACS (American Cancer Society) study followed more than 500 000 adults in about 150 metropolitan areas from 1982 through to 1989. The mortality RRI corresponding to a 10 µg m$^{-3}$ increment in PM$_{2.5}$ was 6.2 % (95 % CI 1.6–11) (Pope et al., 2004). Miller et al. (2007) investigated the effect of long-term PM exposure on women following more than 65 000 subjects in 36 US cities between 1994 and 1998; the increase by 10 µg m$^{-3}$ in PM$_{2.5}$ was associated with a cardiovascular mortality RR (relative risk) of 1.76 (95 % CI 1.25–2.47).

European epidemiological studies on long-term PM exposure and mortality are more limited (Pelucchi et al., 2009). A cohort study in the Netherlands on 5000 subjects aged between 55 and 69 indicated that an increase by 10 µg m$^{-3}$ in black smoke (BS) concentration was associated with an RR of 1.03 (95 % CI 0.91–1.17) for total mortality and 1.16 for respiratory mortality (95 % CI 0.91–1.48) (Hoek et al., 2002). The PAARC (Pollution Atmosphérique et Affections Respiratoires Chronique) study in France was conducted on more than 14 000 subjects aged 25–59 years and followed for 25 years: the total mortality RR associated with a 10 µg m$^{-3}$ increase in TSP (total suspended particulate matter) was 1.05 (95 % CI 1.02–1.08), while no association was found between cardiovascular and respiratory disease and TSP exposure (Filleul et al., 2005). Gehring et al. (2006) investigated PM health effects on 4800 German women aged 50–59 years and observed that total mortality RR for a 7 µg m$^{-3}$ increase in PM$_{10}$ was 1.08 (95 % CI 0.94–1.25). The strongest effects were observed for cardiovascular mortality (Gehring et al., 2006).

PM long-term exposure was also associated with diabetes and cardiovascular and respiratory diseases (Martinelli et al., 2013), including atherosclerosis (Hoffmann et al., 2007), hypertensive episodes (Brook et al., 2007), arrhythmia (Rich et al., 2005) and asthma (Gehring et al., 2010). Kassomenos et al. (2011) developed a neural network approach to estimate the significance of PM exposure in hospital admission for cardiovascular and respiratory diseases. The results for Athens showed that a 10 µg m$^{-3}$ increase in PM$_{10}$ concentration led to an 8.6 % increase in hospitalizations (Kassomenos et al., 2011). A few studies showed a correlation between PM exposure and negative reproductive outcomes, including preterm delivery, preeclampsia, cardiovascular malformations (Strickland et al., 2009; Yi et al., 2010) and neurological problems, such as mild cognitive impairment and headache (Dales et al., 2009; Ranft et al., 2009).

Lower relative risks were reported for PM short-term exposure (Pope and Dockery, 2006). One of the largest efforts to investigate PM short-term effects was the National Morbidity, Mortality, and Air Pollution Study (NMMAPS). NMMAPS investigated daily mortality data for 20 to 100 US cities. The mortality did not show a correlation with PM exposure, likely due to the sensitivity of the modelling algorithm and spatial variability (RRI for 20 µg m$^{-3}$ of PM$_{10}$ equal to 0.4 %); nevertheless, there was evidence of a link between PM$_{2.5}$ exposure and respiratory and cardiovascular illnesses (Dominici et al., 2003). The analysis of national databases in the United States revealed higher mortality RRs for PM$_{2.5}$ exposure (Zanobetti and Schwartz, 2009). The APHEA project (Air Pollution and Health: a European Approach) investigated daily mortality data from 32 European cities and observed that mortality was associated with PM exposure: the daily mortality counts associated with 10 µg m$^{-3}$ of PM$_{10}$ increased by 0.52 %, and it increased by 0.76 % and 0.71 % for cardiovascular and respiratory mortality, respectively (Analitis et al., 2006; Katsouyanni and Grp, 2006). The effects were more pronounced during the first and second day for total mortality and cardiovascular mortality, while respiratory mortality showed more prolonged lagged effects. The APHENA project (Air Pollution and Health: a European and North American Approach) confirmed previous results, with a percentage increase in daily mortality counts of 0.33 points in Europe and 0.29 in the United States when correlated with the previous day’s air pollution (Samoli et al., 2008). In Europe higher mortality risks were observed for short-term exposure to coarse particles. Perez et al. (2009) showed that the odds ratio of cardiovascular and cerebrovascular mortality per 10 µg m$^{-3}$ of PM$_{coarse}$ was 1.059 and 1.098, respectively. In Barcelona an increase in PM$_{coarse}$ of 10 µg m$^{-3}$ was associated with a mortality RR that was increased by 8.4 % during Saharan dust episodes versus an increment of 1.4 % on non-Saharan dust days (Perez et al., 2008). Other investigators did not observe such correlations (Karanaasiou et al., 2012), and Samoli et al. (2011) estimated that the particles’ effects were significantly higher during non-desert dust days.

Pope et al. (2011) investigated the PM$_{2.5}$ exposure response as a function of exposure duration (i.e. days) and exposure intensity (i.e. PM$_{2.5}$ mean concentration). Mortality RR increases with time, going from 1 to 2 days’ exposure through to 40 days’ exposure (Fig. 22). For a longer exposure time, the risk tends to increase, although with a larger scatter.
On the other hand, both short-term and long-term exposures show a linear increase in mortality risk at low and moderate estimated daily doses of PM$_{2.5}$, typical of urban environments. A higher slope was observed for long-term exposure compared to short-term experiments. At a high exposure intensity (daily dose larger than 100 µg m$^{-3}$), the mortality risk flattened off.

A few studies highlight a correlation of mortality and morbidity in urban areas with UFP concentration. UFPs (particles with an aerodynamic diameter smaller than 100 nm) represent less than 10% of the PM$_{2.5}$ mass in most urban locations, but they compose more than 90% of the PN concentration. A review of adverse health outcomes related to UFP exposure is reported by Terzano et al. (2010). The high surface-to-volume ratio of UFPs favours the accumulation of toxic air pollutants, enhancing their impact on health (Sioutas et al., 2005). A European expert panel concluded that a reduction in UFP concentration by 10000 particles cm$^{-3}$ would lead to a decrease in all-cause mortality by 0.1–2.1%, with the largest cause of uncertainty being the lack of long-term studies (Hoek et al., 2009). The experts defined as medium high the likelihood of a causal relationship of all-cause mortality and hospitalization data with short-term UFP exposure, while the link with long-term exposure was defined as medium (Knol et al., 2009). Mortality and cardiovascular morbidity data showed that UFP concentration is linked to delayed effects (Frostier et al., 2007; Stölzel et al., 2007; Wichman et al., 2000). The EPA report (2009) concluded that epidemiologic evidence is inadequate to infer a causal association between short-term exposure to UFPs and mortality.

The way PM exposure affects human health has been investigated by human exposure and toxicological studies. The
main mechanisms involve the cardiovascular and the respiratory systems, with cardiovascular diseases accounting for two thirds of PM health outcomes (Brook et al., 2010). PM exposure increases inflammatory cytokines (e.g. IL-6, IL-1β, TNF-α) and C reactive proteins (CRP), promoting inflammatory modifications, which might lead to inflammation and atherosclerosis (Hoffman et al., 2009; Rückerl et al., 2006, 2009; Schicker et al., 2009). An increase in fibrinogen instead leads to coagulation modification that eventually causes coronary artery disease (Brook et al., 2010; Rückerl et al., 2006). The damage to the respiratory system is linked to the activation of inflammatory cells triggered by PM exposure. These cells produce cytokines and reactive oxygen species (ROSs), which promote oxidative stress and inflammation, leading to asthma and chronic obstructive pulmonary disease (COPD) (Chung and Adcock, 2008; Hogg et al., 2004; Silbajoris et al., 2011; Terzano et al., 2010). ROSs are also produced directly by inhaled particles, for example by redox-active transition metals, quinones and endotoxins present on the surface of the particles or by surface-absorbed polycyclic aromatic compounds that undergo biotransformation in vivo (Kelly and Fussell, 2012). UFP, unlike larger particles, can reach different parts of the body, including the brain and the heart, with potential long-term damages (Terzano et al., 2010). Berubé et al. (2007) illustrated the different action mechanisms and effects of pulmonary toxicity associated with combustion-generated UFPs, including inflammation, carcinogenic effects and cardiovascular death.

A few epidemiological studies focused on the correlation between specific PM chemical components and adverse health effects. Chen et al. (2009) investigated the effects of heavy metals and reported a correlation of Ni and V with cardiovascular mortality and morbidity. An extended study over six counties in California showed that Cu, K, Zn and Ti in PM2.5 were related to multiple mortality categories, and especially to cardiovascular deaths (Ostro et al., 2007). Zanobetti et al. (2009) reported that hospitalization for cardiovascular and respiratory diseases in several US cities was associated with high concentration of Ni, As, Cr and Br in fine particles. The health impact of metals depends on their solubility and bioavailability, which in turn depends on the age and processing of the particle (Harrison and Yin, 2000). Together with heavy metals, polycyclic aromatic hydrocarbons (PAHs) are one of the PM components most extensively investigated for their impact on health (Dellino et al., 2010). PAHs and their derivatives (oxy-PAH and nitro-PAH) have been proven to be toxic, mutagenic and carcinogenic to humans (IARC, 2013). Motorykin et al. (2013) observed a link between lung cancer deaths and PAH emissions in upper-middle and high socioeconomic countries. PAHs in PM ambient samples correlate with tumour initial potential (Gauggel-Lewandowski et al., 2013), DNA damage (Teixeria et al., 2012) and the formation of reactive oxygen species in vitro, promoting inflammation response and cellular damage (Shiraiwa et al., 2012). A few studies showed a correlation between PM mutagenicity and PAH and nitro-PAH concentrations (de Kok et al., 2006).

PBAPs are natural particles that can be associated with negative health effects. PBAPs and protein molecules can cause allergic disease from their fragmentation (Pöschl, 2005). The transport of PBAPs, allergens and pathogens over long distances is observed, for example, during Saharan dust outbreaks over Europe (Kellogg and Griffin, 2006). Karanasios et al. (2012) reviewed specifically the effect of Saharan dust outbreaks in Europe and concluded that more studies are needed in order to understand dust effects: some studies show a link between PM exposure during Saharan dust events and mortality data, while other investigators did not observe any correlation. Giannadaki et al. (2014) estimated that desert dust was responsible for 402,000 cardiopulmonary deaths on a global scale in 2005, with a higher incidence in the “desert-belt” countries of northern Africa, the Middle East and south-eastern Asia.

Secondary inorganic ions, mainly sulfate and nitrate, account for a large fraction of PM mass. Epidemiological studies showed a causal association between sulfate and nitrate exposure and all-cause, cardiovascular and respiratory mortality (Atkinson et al., 2014; WHO, 2013a). Nevertheless, controlled exposure experiments to sulfate have shown health outcomes only at concentrations well above ambient levels (Rohr and Wyzga, 2012). Only limited toxicological evidence supports a causal relationship between sulfate and nitrate and health effects. Although ammonium sulfate and nitrate are not toxic, sulfate and nitrate might increase the bioavailability of other counter ions, such as transition metals (WHO, 2013a), or the apparent correlation may arise from cross-correlations.

Recently, attention has been given to carbonaceous material, i.e. soot carbon and OC (EPA, 2009; HEI, 2010; WHO, 2012). The parameters used to quantify soot exposure are BC, EC, black smoke (BS) and the absorption coefficient that characterizes PM (Abs), depending on the methodology employed to quantify soot particles. Most of the available data on soot particles short-term effects come from the APHEA project (Analitis et al., 2006; Katsouyanni et al., 2001; Le Tertre et al., 2002). Pooled analysis shows that all-cause mortality associated with a 10 µg m⁻³ increase in PM10 and BS was 0.48 % and 0.68 % higher, respectively; the effects were even more significant for BS exposure and cardiovascular mortality (0.90 % increase) and respiratory mortality (0.95 % increase) (WHO, 2012). Fewer studies investigated simultaneously the effects of PM2.5 and soot particles and indicated that a 1 µg m⁻³ increase in PM2.5 and EC were associated with a 0.19 and 1.45 % increase in all-cause mortality, respectively (WHO, 2012, and references therein). Four cohort studies investigated the health outcome of long-term exposure to soot particles (Beelen et al., 2008; Filleul et al., 2005; Lipert et al., 2006; Smith et al., 2009). The pooled analysis shows that the all-cause mortality RR associated with a 1 µg m⁻³ increase in PM2.5 and EC was 1.007
(95 % CI 1.004–1.009) and 1.06 (95 % CI 1.04–1.09), respectively. The WHO report on BC (WHO, 2012) concluded that there is sufficient evidence to identify an association between long- and short-term soot exposure and all-cause mortality. Nonetheless, one paper suggested that the reduction in ambient soot concentration would have a greater effect on health outcome than a reduction by the same amount of PM$_{2.5}$ or PM$_{10}$ (Janssen et al., 2011).

Toxicological and clinical studies agree on the cytotoxicity and oxidative potential of soot particles but focus almost exclusively on laboratory-generated combustion particles and freshly emitted combustion particles. BC in combustion-generated particles was found to be associated with PM oxidative potential and in vitro cytotoxicity (Garza et al., 2008; Murr and Garza, 2009). Soto et al. (2008) reported the cytotoxicity of BC in commercial nanoparticle materials. Boogard et al. (2012) observed an association between PM oxidative potential and soot carbon in ambient aerosol collected in traffic areas. Harder et al. (2005) reported the effect of BC on heart-rate increase but no inflammatory effects, while Cho et al. (2005) observed a correlation between ambient soot carbon and PM oxidative potential. Su et al. (2008) suggested that soot carbon toxicity depends on soot particle morphology, while the WHO report (2012) concluded that BC particles may not be the harmful component of PM but may act as carrier of a large variety of toxic species, such as organic molecules, which are co-emitted with soot carbon. A simultaneous investigation of soot particle morphology and toxicity is very limited (Soto et al., 2008; Su et al., 2008) and was performed only with transmission electron microscopy (TEM), which does not give information on soot-carbon–organic-carbon mixing or information on OC composition.

Since OA makes up a dominant fraction of fine-PM mass, it is difficult to isolate its health outcomes through statistical tools typical of epidemiological studies. It follows that only limited epidemiological results are available in the literature. An association between primary OC and inflammatory response was observed by Delfino et al. (2010), while Ostro et al. (2010) showed a link between OA concentration and mortality. Toxicological studies are more frequent but focus mainly on laboratory-generated particles. Biswas et al. (2009) reported a high oxidative potential for water-soluble OC emitted by heavy-duty engines, while Cheung et al. (2009) showed an association of oxidative potential with water-soluble and water-insoluble OC in diesel engine emissions. A cytotoxic response was observed for organic nanoparticles generated by laboratory flame and by diesel engines (Manzo et al., 2010; Sgro et al., 2009). Longhin et al. (2013) reported the correlation of OC and DNA damage in an urban environment. OC correlated with oxidative potential in the Long Beach harbour area (Hu et al., 2008) and in traffic-exhaust-dominated areas (Gualtieri et al., 2011; Verma et al., 2011). The limited body of evidence concerning ambient aged particles shows that SOA correlates with oxidative potential (Verma et al., 2009; Rattanavaraha et al., 2011; McWhinney et al., 2013) and with cytotoxicity (Jalava et al., 2009). Gaschen et al. (2010) and Künzi et al. (2013) observed a moderate response after exposing epithelial cells to smog-chamber SOA from different sources.

During the last decade, and following the recommendation of the US National Research Council (NRC, 2004), an increasing number of studies have investigated the health response associated with the simultaneous exposure to several pollutants, using a multi-pollutant approach (Dominici et al., 2010). A review of statistical methods available to solve the multi-pollutant problem is presented by Billionet et al. (2012). On the same issue, the US Environmental Protection Agency promoted a workshop to explore the use of source apportionment methods to resolve the contributions of different PM sources to the observed health effects in urban areas (Hopke et al., 2006). In Phoenix (AZ) secondary sulfate and traffic were the sources mainly associated with cardiovascular mortality (Mar et al., 2006), while in Washington DC sulfate and primary coal emissions were responsible for the largest increment of all-cause deaths (Ito et al., 2006). Stanek et al. (2011) reviewed a series of multi-approach studies and identified a link between cardiovascular effects and dust or combustion sources, including traffic. No single sources were unequivocally recognized, at this time, as the main drivers of negative health outcomes (Stanek et al., 2011). In Europe multi-pollutant-approach studies are more limited. Tainio et al. (2010) evaluated the number of premature deaths that could be attributed to primary PM$_{2.5}$ sources in Finland. Based on national and European emission inventories and exposure–response functions, traffic was identified as the PM source responsible for the largest number of premature deaths (Tainio et al., 2010).

The recently published WHO review (2013b) summarized the state of knowledge about the health effects of exposure to PM. The report stated that the scientific conclusions of the 2005 WHO Guidelines about the causal link between PM$_{2.5}$ and adverse health outcomes in humans are confirmed and strengthened and, hence, remain valid. However, it is now suggested that the adverse health effects do not have a threshold exposure and that in the absence of a threshold, public health benefits will result from any reduction in PM$_{2.5}$ concentrations. New studies on short- and long-term effects conclude that long-term exposures to PM$_{2.5}$ is a cause of cardiovascular mortality and morbidity, in addition to several new health outcomes (e.g. atherosclerosis, adverse birth outcomes and childhood respiratory disease). The report cites more studies that shed light on physiological effects and plausible biological mechanisms that link short- and long-term PM$_{2.5}$ exposure with mortality and morbidity. However, it is now suggested that UFPs act through different mechanisms from larger particles that dominate mass-based metrics, such as PM$_{2.5}$ or PM$_{10}$. The report indicates that BC, SOA and SIA may provide metrics for the effects of mixtures of pollutants from a variety of sources. As a result, it is concluded that both short-term (such as 24 h average) and long-term (annual

means) exposure to PM$_{2.5}$ affect health and that maintaining independent short-term and long-term limit values for ambient PM$_{10}$ in addition to PM$_{2.5}$ to protect against the health effects of both fine and coarse particles is warranted. Finally, the report concludes that there is limited epidemiological evidence for the association between short-term exposures to ultrafine PM ($<0.1 \mu m$) and cardiorespiratory health as well as the health of the central nervous system.

The HRAPIE (Health Risks of Air Pollution In Europe) report, part of the REVIHAAP (REVieu of evidence on Health Aspects of Air Pollution) project, recommended concentration–response functions to be used in quantifying the health impacts of PM (WHO, 2013b). These functions covered long- and short-term exposures to both PM$_{2.5}$ and PM$_{10}$ for a range of health outcomes including all-cause, cause-specific and post-neonatal mortality, hospital admissions, restricted activity days, and days of work lost. The reader is referred to the HRAPIE report for details.

3.1.2 Metrics for air quality and uncertainties

Current metrics

It is currently possible to measure many characteristics of the ensemble of particles present in the atmosphere. Size-specific mass, surface area, the total number of particles and the number of particles in different size ranges are all currently measurable, as is the chemical composition of particles in the atmosphere. In fact, the characterization of the chemical composition of individual particles is now feasible with aerosol mass spectrometers, and properties potentially related to toxicological activity, such as oxidative potential, are also measurable. However, the use of particular metrics or properties for legislative purposes imposes considerably more constraints than simply the ability to measure that particular property or metric. For a robust pollutant management system and legislative programme not only would one need to be able to measure the particular property or metric, but one would also need to be able to assess the damage on the basis of existing levels to predict the effectiveness of potential policy measures and assess the resulting reduction in harmful effects. Ideally one would also need to be able to undertake a cost–benefit analysis of policy measures in order to facilitate the acceptability of policy measures across governments and in civil society. Ideally, therefore, to legislate and implement policies on a pollutant or particle metric, one would need the following:

- scientific consensus on harmful effects;
- practical routine measurement methods, including a reference method;
- a monitoring database extensive in space and time;
- emission inventories and source apportionment;
- knowledge of kinetics and formation mechanisms and pathways;
- predictive models;
- abatement techniques and costs;
- dose–response functions and benefit estimates (quantitative and monetary).

When these criteria are considered, it is clear that few particle metrics or properties satisfy them. All of these criteria are satisfied for PM$_{10}$, although there is inevitably uncertainty surrounding many of them. Nonetheless, in agreeing on the first European “daughter directive” on air quality in 1999, policy makers deemed that there was sufficient information available to underpin legislation and policy on PM$_{10}$ but not on PM$_{2.5}$. Although there was a reasonable amount of health effect evidence on PM$_{2.5}$, it was judged to be insufficient for its regulation. Moreover, and probably more importantly in terms of agreeing limit values, there was virtually no routine monitoring of PM$_{2.5}$ in Europe on which to base legislation. Consequently, a requirement that member states monitor PM$_{2.5}$ was therefore included in the directive, and in the revised directive in 2008 limit values and other targets for PM$_{2.5}$ were included. By this time a considerable degree of health effect evidence regarding PM$_{2.5}$ had also emerged, in particular relating long-term exposures to mortality, as discussed in the previous section.

Other metrics, or components of ambient PM have been considered in the context of legislation and policy. In recent years BC has probably received the most attention as evidence has accumulated showing associations between BC concentrations and adverse health effects (Janssen et al., 2011; WHO, 2012). The importance of BC as a “short-lived climate pollutant” has also played a part in focusing attention on this pollutant. However, as yet, BC has not satisfied the criteria listed above and an example is afforded by the recent revision of the UNECE/CLRTAP (United Nations Economic Commissions for Europe/Convention on Long-range Transboundary Air Pollution) “Gothenburg Protocol” in 2012. This protocol sets national emission ceilings for the more important air pollutants, namely, SO$_2$, NO$_x$, VOCs and NH$_3$. The revision introduced emission ceilings for primary PM$_{2.5}$. However, while it was recognized that health effects, and climate impacts, could be reduced through reductions in BC concentrations, the absence of robust emission inventories was the principal reason why parties decided it was premature to set binding emission ceilings for BC. The revision did nonetheless require parties to develop emission inventories and data on abatement technologies, and, importantly, it required parties to give priority to reducing emissions from sources with high BC / OC ratios in order to stay below their PM$_{2.5}$ ceilings.

In the context of BC, it is interesting to note that there is currently no legislative pressure on ambient concentra-
tions of primary PM emitted by combustion sources, including road traffic. This is particularly important as evidence is beginning to emerge that some components of the ambient PM mixture are potentially more damaging than others. The WHO report (2013b) review discussed in the previous section concluded that there was evidence for increased adverse effects in those living near busy roads; it further concluded that “UFPs, carbon monoxide, NO₂, black carbon, PAH, and some metals are more elevated near roads. Individually or in combination, these are likely to be responsible for the observed adverse effects on health. Current available evidence does not allow discernment of the pollutants or pollutant combinations that are related to different health outcomes, although association with tailpipe primary PM is identified increasingly.” The review further went on to note that there was toxicological evidence that non-exhaust traffic emissions could be responsible for some of the observed adverse effects on health.

In terms of advice to the policy process regarding BC, however, the WHO report (2013) concluded that “It would be advantageous to develop an additional air quality guideline to capture the effects of road vehicle PM emissions not well captured by PM_{2.5}, building on the work on BC and/or EC (WHO, 2012) and evidence on other pollutants in vehicle emissions.” It is important to note here that the recommendation was for a guideline (such as a WHO guideline) as opposed to a legally binding limit value.

It is worth noting here that in the WHO report (2013a), papers that measured EC were also considered along with those measuring BC, but no distinction was made between the two. Clearly, if, in the future, either metric, BC or EC, is to be incorporated into legislation then a reference method will need to be defined and agreed.

In addition, the WHO report (2013a) noted that SOA has also been associated with adverse health effects, but the report further noted that evidence was insufficiently strong to distinguish between the toxicity of primary and secondary OA. There are many hazardous air pollutants and toxic air contaminants which are not explicitly regulated by EU legislation but which should also be reduced by many of the measures employed to achieve the limits on those pollutants that are regulated. Where this is not the case, other measures would need to be considered.

Some metals are currently regulated by EU Directives: lead in the 2008 directive (Directive 2008/50/EC) and arsenic, cadmium, mercury and nickel in Directive 2004/107/EC (the so-called Fourth Daughter Directive). Both directives prescribe reference methods for sampling the metals as the PM_{10} fraction, apart from mercury, which is measured in the gas phase. Possible new metals which might be included in legislation are discussed below.

A further form of particulate matter considered in legislation is the group of PAHs dealt with in the Fourth Daughter Directive. Although there are many potentially harmful PAHs, some of which occur at least partially in the gas phase, the directive requires measurements of benzo(a)pyrene as the PM_{10} fraction of PM.

### Possible future metrics

As noted above, the WHO report (2013a) review provides an authoritative starting point for the consideration of possible future metrics for the regulation of particles in the ambient atmosphere.

Another mass metric, PM_{1}, has been suggested as a possibly useful metric to manage PM levels in the atmosphere, largely on the grounds that it provides better separation of the coarse mode and accumulation mode (and ultrafine) fractions than the 2.5 micron cut-off. However, until such time as a body of evidence demonstrating associations with adverse health effects and measuring methods widely used in national monitoring networks are available, this metric is unlikely to command support in the regulatory process.

An important new area of research has emerged in the past 10 years or so, where toxicological studies have demonstrated that the mechanism leading to oxidative stress may be the route by which PM causes adverse health effects. The basic mechanism involves electron transfers in the so-called Fenton reactions (Donaldson and Bohm, 2006), leading to the body’s balance of antioxidants being disturbed and thereby increasing vulnerability to inflammatory responses. As noted earlier, the oxidative potential of ambient PM can be measured, but per se it is not particularly helpful as a metric to be included in legislation, not least because without knowing the active components, it is not obvious what policies should be employed to reduce PM concentration. Where research on oxidative potential is important, however, is where it can point to potentially toxic components of the PM mixture, which may then be easier to focus abatement measures on. Species already identified as potentially increasing oxidative stress include some transition metals, notably Cu, V, Ni, Sb and Fe (Stohs and Bagchi, 1995; Akhtar et al., 2010), as well as quinones (Kumagai et al., 2012). An interesting finding from these studies is that the non-exhaust component of vehicle emissions, deriving from brake, tyre and clutch wear, could be potentially toxic (Lodovici and Bigagli, 2011). This is important as there is as yet no regulation or policy (other than traffic reduction measures) that deals with these sources. However, SOA also contains substantial amounts of ROS (Verma et al., 2009; Platt et al., 2014) as well as of peroxides (Mertes et al., 2012).

The possible future metric with perhaps the longest research history is probably the ultrafine fraction (i.e. particles less than ~100 nm), generally measured as PN, either in total or as a size spectrum. Due to the low contribution of UFPs to the PM_{2.5} mass (Rodriguez et al., 2007), high-UFP events frequently occur under low-PM_{2.5} conditions. There are good reasons for thinking that such particles could be particularly toxic as their small size allows them to reach the brain, and “there is considerable evidence that UFPs can contribute to..."
the health effects of PM” (WHO, 2013a, Answer to Question D1). However, the WHO review felt that the data on concentration–effect functions were too scarce to allow an air quality guideline to be recommended.

In the framework of the EU FP6 project EUCAARI, a first size-resolved pan-European anthropogenic PN inventory was made (Denier van der Gon et al., 2010; Kulmala et al., 2011). Using general assumptions on the chemical composition, density and shape of the particles by source type and estimate of UFP emissions were made based on the size-resolved PN inventory. It is clear that road transport and other transport sectors dominate primary anthropogenic PN and UFP emissions in Europe. In agreement with these results, but chiefly because mass emissions from vehicles are now low and approaching the measurement uncertainty, the regulations for the Euro 5/V and Euro 6/VI vehicle emission standards incorporate limits on PN emissions. The PN emission limit value presented in Table 3 refers to solid PN, not total PN. Total PN is considerably higher than solid PN because non-refractory particles are present and/or newly formed in the exhaust gas and in the few seconds after release.

UFPs and (size-resolved) PN are not only an emerging issue in the discussions on European air quality, but they also play a pivotal role in climate discussion. In fact, aerosol PN, especially the PN > 100 nm, affects the aerosol indirect effect, one of the major uncertainties in our understanding of climate forcing. There is a general consensus that air quality and climate measures should become more integrated, avoiding measures that would be beneficial regarding one area but counterproductive for the other. Aerosols, PM and PN, are likely to be the most relevant example where policies and future scenarios will need integration to boost the effectiveness of both climate and air quality measures.

3.1.3 Local versus regional

Lenschow et al. (2001) coined the term urban increment. The kerbside increment is defined as the increase in air pollution parameters at a kerbside or street site above the urban background, whereas the urban increment is defined as the increase in concentrations in the urban background relative to the rural surroundings (see also Mues et al., 2013). The concept is based on the observation that secondary aerosols are rather homogeneously distributed, in agreement with their regional formation pattern. An aerosol of urban origin is then superimposed on this regional background, while at the kerbside an additional highly local traffic component is found (Fig. 23).

While this concept is generally accepted, the size of the urban increment is different for different regions, depending on the main emission sources in the city as well as the level of pollution in the surrounding area. Mues et al. (2013) investigated PM$_{10}$ concentrations for the years 2003–2008 in the German Ruhr area, the Dutch Randstad and the German city of Berlin and compared them to PM$_{10}$ concentrations from rural background stations. They found substantial absolute and relative urban increments for the Ruhr area (7.4 µg m$^{-3}$, 35 %) and Berlin (8.5 µg m$^{-3}$, 46 %) but quite a low value for the Randstad (3.1 µg m$^{-3}$, 12 %). These values were then compared to the results of the regional chemistry transport model LOTOS-EUROS (LOn Term Ozone Simulation – EURopean Ozone Simulation), which yielded substantially lower urban increments (3.3 µg m$^{-3}$ or 33 % for the Ruhr area, 1.5 µg m$^{-3}$ or 12 % for the Randstad and 1.7 µg m$^{-3}$ or 27 % for Berlin). Mues et al. (2013) then argued that higher-resolution models would be needed to better reflect the urban increment.

Ortiz and Friedrich (2013) developed a new regression model for estimating the urban increment in PM$_{10}$ and NO$_2$ for all cities with more than 50 000 inhabitants in Germany for the reference year 2005. They found good agreement between modelled and measured urban increments, with values between 3 and 9 µg m$^{-3}$.

The above results are in contrast to the results for the urban increment and enhanced kerbside site concentrations obtained by Mohr et al. (2011) using mobile measurements in Zurich (Fig. 24). They confirmed that secondary components were indeed rather homogeneously distributed, while primary components mostly showed enhanced concentrations close to sources. At the road site with the heaviest traffic (Rosengarten Street, Fig. 24), BC and HOA were enhanced by 11 and 2.5 µg m$^{-3}$, respectively, compared to the urban background site, while the average of all the street sites showed an enhancement of about 2.5 µg m$^{-3}$ for BC and only a very small enhancement for HOA. In contrast, the PM$_{10}$ value at the urban background site (31 µg m$^{-3}$) was virtually identical to two rural sites 100 km south-west of Zurich and 25 km north-east of Zurich, with 31 and 33 µg m$^{-3}$, respectively. This means that while road sites indeed exhibited substantially higher concentrations from primary emissions by road traffic, the urban increment in PM$_{10}$ in the case of

![Figure 23. The concept of regional background, urban increment and local traffic increment (adapted from Lenschow et al., 2001).](image-url)
Zurich was close to 0, in contrast to the original concept by Lenschow et al. (2001).

A similar result was obtained for Paris within the MEGAPOLI (Megacities: Emissions, urban, regional and Global Atmospheric POLlution and climate effects, and Inegrated tools for assessment and mitigation) project. Crippa et al. (2013a) investigated the temporal evolution of a variety of chemical components at three urban background sites in Paris and found very similar features. They concluded that particulate pollution in Paris was dominated by regional factors and that the emissions from Paris itself had a relatively low impact on the concentrations at the urban background sites as well as on the surroundings during the period of the measurements; this is in agreement with previous studies (Sciare et al., 2010).

Similarly, Keuken et al. (2013) investigated the contribution of regional, urban and traffic sources to PM$_{2.5}$ and PM$_{10}$ at a street location and up- and down-wind of the city of Rotterdam, the Netherlands. They concluded from their 1-year study that the urban background of PM$_{2.5}$ and PM$_{10}$ was dominated by the regional background and that primary and secondary PM emission by urban sources contributed less than 15%. However, they found clear differences between the street site and the urban background site, with an increase of 50% for PM$_{10}$ at the street site, which was mainly attributed to resuspension of road dust. They concluded that, in particular, people living in street canyons with intense traffic – who make up about 5% of the population of Rotterdam – are exposed to elevated levels of EC and OC from exhaust emissions, heavy metals from brake and tyre wear, and mineral particles from resuspension of road dust. Visser et al. (2015) also showed that traffic-related elements yielded the highest kerb increments.

It is important to note that traffic is also an important source of high PN concentrations, especially in winter (Bukowiecki et al., 2002). Therefore, increasing gradients in PN concentrations generally occur when moving from natural background or rural sites to urban background or kerbside sites (Putaud et al., 2010).

Some of the above discrepancies between individual studies may be related to the specific geographic environment of the cities and the periods of the measurements. As an example, Paris consists of a circular built-up region surrounded by flat, rural terrain, while Barcelona is located in a coastal depression enclosed between two river basins with urban agglomerations and industries, which will favour an accumulation of emissions within the urban area and thus an increased urban increment (Amato et al., 2011). However, we hypothesize that differences found between regional background and urban background sites may be attributed, at least partly, to a non-negligible contribution from direct traffic exposure.

However, the body of evidence lets us conclude that the PM$_{10}$ concentration as well as that of other aerosol components may not differ significantly from its regional background in most modern cities in Europe, while the concentrations of primary emissions from road traffic (mostly BC, 6.0 × 10$^{11}$ PM #km$^{-1}$

Table 3. EU EURO 5 and 6 emission standards for passenger cars$^a$ (EC, 2008, 2012).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Date</th>
<th>CO g km$^{-1}$</th>
<th>HC</th>
<th>HC + NO$_x$</th>
<th>NO$_x$</th>
<th>PM</th>
<th>PN #km$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compression ignition (diesel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euro 5a</td>
<td>Sep 2009$^b$</td>
<td>0.5</td>
<td>–</td>
<td>0.23</td>
<td>0.18</td>
<td>0.005$^f$</td>
<td>–</td>
</tr>
<tr>
<td>Euro 5b</td>
<td>Sep 2011$^c$</td>
<td>0.5</td>
<td>–</td>
<td>0.23</td>
<td>0.18</td>
<td>0.005$^f$</td>
<td>6.0 × 10$^{11}$</td>
</tr>
<tr>
<td>Euro 6</td>
<td>Sep. 2014</td>
<td>0.5</td>
<td>–</td>
<td>0.17</td>
<td>0.08</td>
<td>0.005$^f$</td>
<td>6.0 × 10$^{11}$</td>
</tr>
<tr>
<td>Positive ignition (petrol)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euro 5</td>
<td>Sep. 2009$^b$</td>
<td>1</td>
<td>0.10$^d$</td>
<td>–</td>
<td>0.06</td>
<td>0.005$^e,f$</td>
<td>–</td>
</tr>
<tr>
<td>Euro 6</td>
<td>Sep. 2014</td>
<td>1</td>
<td>0.10$^d$</td>
<td>–</td>
<td>0.06</td>
<td>0.005$^e,f$</td>
<td>6.0 × 10$^{11}$$^{e,g}$</td>
</tr>
</tbody>
</table>

$^a$ The PM and PN emission standards also apply to EURO 5 and 6 light commercial vehicles. $^b$ January 2011 for all models. $^c$ January 2013 for all models. $^d$ and NMHC = 0.068 g km$^{-1}$. $^e$ applicable only to vehicles using DI (direct injection) engines. $^f$ 0.0045 g km$^{-1}$ using the particle measurement programme (PMP) procedure. $^g$ 6.0 × 10$^{12}$ #km$^{-1}$ within first 3 years from Euro 6 effective dates.

Figure 24. Local contributions of PM$_1$ components for different sites within the city of Zurich, Switzerland (absolute values panel a, relative values panel b); averages for the whole campaign are shown. The “average all” bar represents the mean value of the local contribution of all data (Mohr et al., 2011).

heavy metals from brake and tyre wear, and resuspended road dust) do result in substantially enhanced concentrations in the street canyons themselves, such that people living in these street canyons do suffer from substantially worse air quality than at urban background sites.

3.1.4 Effect of control measures

Systematic long-term measurements of particle metrics are limited. Measurements of PM$_{10}$ have only begun to be carried out in the last 20 years or so, whereas PM$_{2.5}$ has only been measured for around 10 years or so. In many countries long-term records are even shorter than this. The assessment and understanding of long-term trends is therefore not straightforward. However, some tentative conclusions can be drawn. Figure 25 shows the trends in PM$_{10}$ and PM$_{2.5}$ since 2001 and 2005, respectively, across the EU (EEA, 2012). Despite apparent reductions in emissions, there is no significant downward trend in concentrations according to either metric over this period at all site types, including road sides. In fact, there is even an indication of an increase in the last 2 years of the time series (although this could simply be a result of meteorological changes).

Trend data are also available for some individual member states. Figure 26 shows trends in PM$_{10}$ in the Netherlands from 1993 to 2008, illustrating the decrease in PM$_{10}$ in the 1990s followed by a flattening from around 2000 onwards.

Figure 25. Trends in PM$_{10}$ (top graph, 2002–2011) and PM$_{2.5}$ (bottom graph, 2006–2011) annual concentrations (in µg m$^{-3}$) per station type (EEA, 2013b); the trends are calculated based on the data officially reported by the EU member states.

Figure 26. Measured trend in annual average PM$_{10}$ in the Netherlands (Hoogerbrugge et al., 2010).

Data for PM$_{10}$ from the UK show a similar pattern (Harrison et al., 2008), and, more recently, a report on PM$_{2.5}$ in the UK showed a similar lack of trend in PM$_{2.5}$ from around 2000 onwards (AQEG, 2012), as shown in Fig. 27.

Whilst there has been little trend in PM$_{2.5}$ in Europe over the past decade, Asmi et al. (2013) reported a negative trend in PN concentration over the Northern Hemisphere during the period 2001–2010 (up to $-7.3 \%$ yr$^{-1}$), likely due to a decrease in anthropogenic emissions. A negative trend in N$_{100}$ was also reported for Europe over the same period, although observations were limited to five locations (Asmi et al., 2013). Declining SO$_2$ emissions have also been linked to observed negative trends in sulfur dioxide concentrations, new particle formation, total PN, N$_{50}$, N$_{80}$, N$_{100}$ and N$_{150}$ at a site in northern Finland (Kyrö et al., 2014). Collaud Coen et al. (2013) found negative trends for scattering and absorption coefficients (mean $-2 \%$ yr$^{-1}$) for locations in North America but no trend over Europe during the period 1996–2010.
European emissions of primary PM and the precursors of SLAs have been reduced substantially over this period through measures taken to meet the emission ceilings in the Gothenburg Protocol and the National Emission Ceilings Directive, including the implementation of the Euro standards for vehicle emissions. Emissions of SO$_2$ have fallen substantially and so to a lesser degree have those of NO$_x$. There are few data on the response of particulate sulfate and nitrate to these emission changes, but the UK Air Quality Expert Group reported concentrations at 12 rural sites in the UK from September 1999 to December 2009. Both nitrate and sulfate showed an increase from 2000 to 2003 and then a decrease to 2009, so that, overall, there was relatively little change over the period (Fig. 28). The peaks in 2003 in both sulfate and nitrate arise from the prolonged period of very high temperatures and intense photochemical activity in Europe in that year. Reductions in sulfate, nitrate, and BC and OC mass concentrations have been reported over the United States in the past few decades, linked to reductions in anthropogenic emissions (Blanchard et al., 2012; Hand et al., 2012; Leibensperger et al., 2012; Malm et al., 2002). Over the eastern United States, observations show negative trends in surface concentrations of sulfate over the period 1980–2010, largely driven by reductions in anthropogenic emissions (Leibensperger et al., 2012). Wet deposition of sulfate decreased by 58% between 1980 and 2010; surface sulfate mass concentrations declined by 40% over 1990–2010, consistent with a 56% reduction in SO$_2$ emissions over the period 1980–2010 (Leibensperger et al., 2012). Hand et al. (2012) reported that annual mean sulfate concentrations across the United States declined at 2.7% yr$^{-1}$ over the period 1992–2010, with a linear relationship between SO$_2$ emissions and sulfate concentrations.

A comparison of trends in PM$_{10}$ concentrations and relevant emissions has been reported for the Netherlands (Hoogerbrugge et al., 2010) and is shown in Fig. 29. The comparison also shows the decline in PM$_{10}$ concentrations up to the year 2000, followed by a flattening of concentrations despite a continuing decrease in emissions, albeit at a slower rate than in the earlier part of the period.

The report on PM$_{10}$ in the Netherlands (Hoogerbrugge et al., 2010) concluded that the absence of a significant trend in this area in the 2000s could be explained by the meteorological conditions and the uncertainty in the measurements used to derive the corresponding trends. They further noted that any reductions in emissions from increasingly cleaner diesel vehicles may have been cancelled out by the increase in total distance driven and the increasing weight of the vehicles. The slower decline in PM concentrations compared with the emissions of the precursors of secondary inorganic aerosol is the significant non-proportionality between emissions of sulfur, NO$_x$ and ammonia and concentrations of sulfate, nitrate and ammonium in ambient PM (see for example Harrison et al., 2013).

Primary emissions influence observed concentrations typically over scales of the order of 1–5 km and often smaller. However, such emissions occur over large areas in major cities, so these emissions are an important source of public exposure to toxic pollutants. Dilution and the partial evaporation of primary OA contributes to this modest range of influence. On the other hand, secondary inorganic and organic aerosols can be formed over distance scales of many hundreds or even thousands of kilometres and therefore have significant effects over large areas. Recent studies in European cities (Lanz et al., 2010; Crippa et al., 2013a, b, 2014), even next to major roadways, indicate that most of the fine PM is secondary. Therefore, combined strategies of reduction in both secondary and primary PM are needed to address the air pollution problem both in major urban centres and in the rest of Europe.

**Figure 27.** Concentrations of PM$_{2.5}$ in µg m$^{-3}$ at six sites in the UK (AQEG, 2012).

**Figure 28.** Concentrations of nitrate and sulfate at UK rural sites (AQEG, 2012).
ammonia was one of the most effective ways to reduce aerosol mass concentrations in those parts of Europe characterized by high ammonium nitrate levels. The response of SIA is likely to be non-proportional. For example, Harrison et al. (2013) showed that 30% reductions in SO$_2$ and NH$_3$ separately across Europe resulted in reductions in SIA of 6 and 9%, respectively, in the UK. Moreover, there are strong reasons for reducing ammonia emissions to limit the transport and deposition of both gas-phase- and particle-phase-reduced nitrogen in Europe to minimize the impact locally, near intensive agricultural installations, and regionally in order to reduce eutrophication and minimize impacts on biodiversity. However, considering the response of SIA to emission controls and the fact that most ammonia emissions occur in rural rather than urban areas, reducing primary PM emissions may result in larger reductions in PM$_{2.5}$ in UK urban areas than in reductions from other sources. Clearly, strategies need to consider local circumstances before priority actions are taken.

The discussion above has focussed on secondary inorganic aerosol and primary PM. However, secondary organic aerosol is often the single most important component of fine particulate matter in European cities. There are still major uncertainties concerning the mechanisms of the formation of SOA, emission inventories particularly of higher carbon number organic compounds and ambient measurements against which to evaluate models (Hallquist et al., 2009; Redington and Derwent, 2013). An important point regarding SOA in terms of abatement measures is the likelihood that a significant part arises from biogenic sources, which will be much less amenable to control than anthropogenic emissions. SOA’s most important anthropogenic precursors are large VOCs (aromatics, large alkanes and olefins) but also a group of organic vapours known as IVOCs. These compounds are a relatively small fraction of the total VOCs, but they are excellent SOA precursors. Some early studies (Fountoukis et al., 2011; Murphy et al., 2012) suggest that they could be one of the most important SOA precursors in Europe. Even if these estimates are quite uncertain, the recent developments in our understanding of SOA formation suggest that policy should focus on only organic emissions (low-volatility, semi-volatile, intermediate-volatility and VOCs) from the various sources.

### 3.2 Aerosols and climate

Atmospheric aerosol effects on the climate system are a major research focus in the area of geophysics and environmental science. Twenty years of field studies have demonstrated that aerosols can impact the atmospheric radiative budget in background areas of the globe, not only in proximity to desert regions but also several hundreds to thousands of kilometres downwind of pollution sources. In the last decade, dedicated networks for the long-term monitoring of the aerosol climate-relevant properties were set up in Europe (Euro-
Aerosols affect climate directly by scattering and absorbing radiation as shown in Fig. 30. In the fifth assessment report (AR5) of the IPCC (Boucher et al., 2013), this forcing is called the radiative forcing (RF) due to aerosol–radiation interactions (RF\textsubscript{aci}). Formerly, it was called the direct aerosol effect. RF is defined as an external perturbation and is calculated as the difference between two radiative transfer simulations that only differ by the amount of that forcing agent. In addition, aerosol particles act as CCN and IN with affecting cloud properties and the radiation balance. This aerosol RF is referred to as the RF due to aerosol–cloud interactions (RF\textsubscript{ac}) and was formerly known as the indirect aerosol effect, cloud albedo effect or Twomey effect (Denman et al., 2007; Forster et al., 2007).

In addition to the pure RFs, atmospheric adjustments take place in response to the forcing. In the case of aerosol–radiation or aerosol–cloud interactions, they occur on timescales of minutes to days, so are much faster than the timescale of global warming. The sum of fast adjustments and RF is called the effective radiative forcing (ERF). In the case of aerosol–radiation interactions, ERF\textsubscript{ac} also comprises the adjustments accompanying the absorption of solar radiation by BC that affect RF. The absorbed solar radiation leads to a heating that can modify the static stability of the atmosphere and thus can impact cloud formation. Absorption of solar radiation by BC inside cloud droplets can cause droplets to evaporate and thus cause a decrease in cloud cover.

In the case of aerosol–cloud interactions, a myriad of different adjustments has been postulated as contributing to ERF\textsubscript{ac}. Changes in the cloud droplet size distribution affect drizzle and rain formation and may affect ice formation in clouds and changes in the cloud lifetime. Adjustments can either have an RF of the same sign as RF\textsubscript{ac} or have opposite sign and buffer the initial RF (Stevens and Feingold, 2009).

RF\textsubscript{ac} is caused by the increase in the cloud droplet number concentration resulting from an increase in anthropogenic aerosol particles. If the liquid-water content and cloud cover remain constant, then the polluted cloud consists of more but smaller droplets. This increases the surface area of the cloud and thus the amount of solar radiation that is reflected back to space.

Evidence for RF\textsubscript{ac} can be seen in satellite images of so-called “ship tracks”, where white lines indicate recent ship routes. These visible lines stem from an increase of up to 1 order of magnitude in the accumulation-mode aerosol concentration, causing a significant increase in the cloud droplet number concentration (up to a factor of 5) and a decrease in the cloud droplet radius (see, e.g., Durkee et al., 2000). Because of this, aerosols are hypothesized to increase the lifetime of polluted clouds (Albrecht, 1989). Smaller cloud droplets have a smaller collection efficiency, which retards drizzle and rain formation. A reduction in drizzle and rain causes less wet scavenging of aerosol particles, causing aerosol particles to accumulate in the boundary layer. By contrast, in clean, pristine conditions, where only few aerosol particles are present, the cloud has fewer droplets and consists of a rather broad cloud droplet size distribution with at least some large droplets. Such a clean cloud will more readily form drizzle or rain via the collision–coalescence process than a polluted cloud. If the drizzle or rain does not evaporate below cloud base but reaches the surface, it removes the aerosol particles within the rain and drizzle drops and below the cloud base from the atmosphere. Thus, a clean marine boundary layer tends to remain clean and a polluted boundary layer tends to remain polluted (Baker and Charlson, 1990). Rosenfeld et al. (2006) even went one step further and hypothesized that open cells in the marine boundary layer could be changed into closed cells in response to the suppression of precipitation by aerosol particles.

However, not all ship track studies show an increase in liquid water in the polluted cloud. If the air above the boundary layer is dry, then enhanced entrainment in polluted clouds can cause the evaporation of cloud droplets and subsequently lead to a reduced cloud-water content (Ackerman et al., 2004). Similarly, when tracing individual air parcel trajectories through clouds, it was found that an increase in aerosol concentration from very clean to very polluted does not increase cloud lifetime, even though precipitation is suppressed (Jiang et al., 2006). These studies show that it is not yet clear
how important the so-called “cloud lifetime effect” is, but they do show that it adds a substantial uncertainty to ERF$_{aci}$.

Another example of rapid adjustments and uncertainty processes is the aerosol effect on mixed-phase clouds. If anthropogenic activity leads to an increase in IN, this would alter the glaciation of supercooled liquid clouds due to the difference in vapour pressure over ice and water. If the vapour pressure lies between saturation with respect to ice and water, ice crystals grow at the expense of the evaporating cloud droplets and in this way quickly turn a non-precipitating mixed-phase cloud into a precipitating ice cloud. On a global scale, this causes a reduction in cloud cover, less reflected shortwave radiation and, hence, a less negative ERF$_{aci}$ (Lohmann, 2002). If, by contrast, anthropogenic activity leads to a decrease in IN efficiency of the background aerosol, this would lead to less efficient precipitation production and more reflected shortwave radiation (Hoose et al., 2008; Storelvmo et al., 2008). Which of these mechanisms dominates, and, hence, to what degree mixed-phase processes buffer part of RF$_{aci}$, is still a matter for debate.

Moreover, IN efficiency varies dramatically between particle types and was shown to be particularly enhanced for insoluble particles such as desert dust, biological particles and volcanic ash (Fig. 31). Soluble aerosols containing an insoluble core can also operate as IN, in immersion mode. Interestingly, the most efficient IN originate from natural sources, therefore, their atmospheric concentrations are affected by anthropogenic activities only through feedback processes (e.g. land use and ecosystem changes), contrary to CCN which are directly emitted or formed by anthropogenic emissions. Several laboratory studies investigated the potential of certain specific anthropogenic aerosol components to act as IN and concluded that, for instance, ice formation on soot particles is not important above $-30$ °C and below water saturation (Dymarska et al., 2006). On the other hand, in situ observations indicate an enrichment of soot in atmospheric ice particle residuals in lower-tropospheric mixed-phase clouds (Cziczo et al., 2008; Targino et al., 2009; Twohy et al., 2010); thus, there must be some mechanism for soot to enter ice clouds. Other studies showed that oxalic acid aerosol nucleates ice in the cirrus regime (Zobrist et al., 2006), suggesting that organic particles, which are ubiquitous in the atmosphere, can contribute to IN concentrations. Oxalic acid is the more inorganic among the organic compounds (it does not contain C-H chemical bonds) and is not really representative of the very wide range of chemical structures which characterize ambient organic particles. In addition, glassy SOA can also heterogeneously nucleate ice in the cirrus regime at relative humidities that are at least 10–15% below those required for homogeneous nucleation (Wang et al., 2012; Schill and Tolbert, 2013; Berkemeier et al., 2014).

IN can be either bare or mixed with other substances. As bare particles age in the atmosphere, they acquire liquid surface coatings by condensing soluble species and water vapour or by scavenging soluble particles, and they are thus transformed from deposition or contact nuclei into possible immersion nuclei. This transformation may dampen the ice-forming ability of some IN types at temperatures relevant for mixed-phase clouds (Cziczo et al., 2009; Eastwood et al., 2009; Chernoff and Bertram, 2010; Sullivan et al., 2010a). Specifically, organic coatings or oxidation by ozone tend to reduce the ice nucleation efficiency (DeMott et al., 2009; Möhler et al., 2005; Wang et al., 2011), but this depends on the ozone levels (Kanji et al., 2013). By contrast, the exposure of aerosol particles to nitric acid or ammonia can enhance IN efficiency (Salam et al., 2007; Sullivan et al., 2010b). In summary, the anthropogenic emissions of precursors of inorganic and organic secondary aerosol components can substantially alter the IN ability of natural insoluble aerosols, although the actual impact of such processes in the real atmosphere is still uncertain and requires further research.
Figure 31. The onset temperatures and relative humidities for deposition and nucleation freezing and immersion freezing for bioaerosols, mineral dusts, organics, crystalline sulfate and soot from a compilation of experimental data of sub- and supermicrometer aerosol particles in the literature (Hoose and Möhler, 2012). The large range of observed ice nucleation onset conditions is due to different experimental set-ups, particle sizes, activated fractions and chemical composition. Only those IN species for which at least three papers exists are shown. The dashed line refers to the homogeneous freezing of solution droplets (Koop et al., 2000).

A summary of ERF_{ari+aci} is shown in Fig. 32. It categorizes the different estimates of ERF_{ari+aci} according to the way they were obtained: purely from global climate model (GCM) simulation with prescribed sea-surface temperatures, from coupled aerosol–ocean GCMs (CMIP5 – Coupled Model Intercomparison Project Phase 5 – models) or if satellite estimates were involved in the estimates. Moreover, the studies were separated according to the ACI represented, i.e. according to whether they only took ACI in liquid stratiform clouds into account or also in mixed-phase or convective clouds. In the latter studies the magnitude of the ERF tends to be somewhat smaller (see Fig. 32). The physical explanation for the mixed-phase reduction in the magnitude of the ERF has been discussed above. Models that have begun to incorporate ACI in convective clouds also have a tendency to reduce the magnitude of the ERF, but this effect is less systematic (Jacobson, 2003; Lohmann, 2008; Suzuki et al., 2008) and the reasons for differences among the models in this category are less well understood. Included in this category is the study by Wang et al. (2011), which circumvented some of the difficulties of parameterizing clouds by developing a modelling framework that can explicitly represent cloud-scale circulations. It includes an array of cloud-resolving models in each model grid box and captures the spatio-temporal covariance of cloud-controlling processes. Its estimate of ERF_{ari+aci} amounts to −1.1 W m\(^{-2}\) and is thus lower than traditional GCM estimates. A less negative ERF_{ari+aci} is also found in studies that use variability in the present-day satellite record to infer ACI or that constrain GCM parameterizations to optimize agreement with satellite observations. In summary, the magnitude of ERF_{ari+aci} has been estimated as being −0.9 W m\(^{-2}\), with an uncertainty range between −0.1 and −1.9 W m\(^{-2}\) in AR5 (Boucher et al., 2013).

3.2.2 Light-absorbing carbon

Some aerosols absorb solar radiation and contribute to global warming while reducing the radiation flux at the Earth’s surface (see the following paragraph). The main light-absorbing aerosol types on a global scale are desert dust particles and soot or BC (Bergstrom et al., 2007). BC aerosols are carbonaceous particles containing EC and are characterized by significant light absorption throughout the visible spectrum (\(>5 \text{ m}^2 \text{ g}^{-1}\) at 550 nm). BC is predominantly of anthropogenic origin and is largely responsible for the top-of-the-atmosphere (TOA) positive radiative forcing of the aerosol. Upon deposition, light-absorbing aerosol particles can darken snow and ice surfaces, increasing their melting rate and eventually reducing snow cover and the extent of glaciers (Warren and Wiscombe, 1980, 1985; Flanner et al., 2009; Painter et al., 2007). Though such effects have been
known for some time, the quantification of the multiple radiative forcings of soot (or BC) is still central in current climate change research. BC is in fact the second most important anthropogenic source of global warming after CO₂ (IPCC, 2013). It is a by-product of some fundamental sectors of modern economies, but unlike CO₂, it is also produced in significant amounts by “traditional” activities, including domestic heating and, in many developing countries, agricultural and cooking practices. BC also contributes to the radiative forcing of climate, with possible impacts on precipitation and water supply in some highly populated regions of the Earth (UNEP and WMO, 2011), and BC has been related to aerosol health effects (see Sect. 3.1). It is clear that the emissions of BC in the atmosphere are an important environmental policy issue. Scientists and policy makers have realized that reducing such emissions can offer mutually beneficial opportunities to air pollution control measures and the policies aiming at slowing down global warming (Monks et al., 2009). The present section provides a short summary of the state of the art of the research on BC climate effects and presents some examples of recent assessments of climate and human health benefits that can be obtained by reducing global BC emissions (Shindell et al., 2012). A brief treatment of the emerging research topic of light-absorbing carbonaceous components other than BC (BrC) is also provided.

The state of scientific knowledge on the nature and the effects of BC particles has been recently reviewed by Bond et al. (2013). BC forms from the incomplete combustion of fossil fuel, biomass and biofuels. Measurement methods have been developed for quantifying these components on a routine basis using some characteristics of soot particles. Depending on the light-absorbing or refractory properties of soot these methods have created operational definitions for BC and EC. Although clarifications of terminology used for BC have been proposed in the literature (Bond et al., 2013; Petzold et al., 2013), the two terms are often used as synonyms of soot. This ambiguity in models or measurements is an important source of uncertainty in evaluating BC effects in the atmosphere, including climate effects (Vignati et al., 2010b).

An estimation of the net radiative forcing of BC emissions must also take into account the effects of the co-emitted species, especially OC and the precursors of inorganic secondary species, which largely counteract the warming effect of BC through the effect on single-scatter albedo. The associated water-soluble material also allows the carbonaceous particles to act as CCN, which can cause a negative albedo forcing (Spracklen et al., 2011). Most models indicate that the extent of warming is almost fully cancelled out by the cooling effect of the co-emitted organic species in the case of emissions from open burning, while a net warming characterizes the emissions from fossil fuel combustion (IPCC, 2013). Given the uneven geographical distributions of open burning and industrial fossil fuel combustion emissions, the TOA positive radiative forcing of BC is exerted mainly in the Northern Hemisphere, including the Arctic and sub-Arctic regions (Fig. 33). However, the continents where open burning is normally practiced (especially Africa and South America) experience a strong surface cooling caused by both BC and the scattering aerosol components, which can result in feedbacks on atmospheric circulation and climate changes on a regional scale (see below).

The recent assessments of the BC radiative forcing range from 0.2 to 0.76 W m⁻² (Bond et al., 2013; UNEP, 2011), with the fifth IPCC AR suggesting a much stronger posi-
tive forcing (+0.40 to +0.80 W m$^{-2}$) than the previous one (0.20 ± 0.15 W m$^{-2}$; IPCC, 2007). A large uncertainty still resides in emissions (Bond et al., 2004) independently of how BC is treated in the models. The inclusion of atmospheric reactions responsible for the conversion of BC from an initial hydrophobic to a hydrophilic state can change BC lifetime by 1 order of magnitude (Croft et al., 2005). The hydrophilic state of BC affects cloud scavenging and wet deposition, which remains the greatest source of uncertainty in models (Textor et al., 2006; Vignati et al., 2010b). As a result, the lifetime of BC has been estimated to range from a few days up to 2 weeks (Cape et al., 2012; Ramanathan and Carmichael, 2008).

Koch et al. (2009) have investigated the performance of a model ensemble (AeroCom) in predicting atmospheric BC concentrations and concluded that the simulations tend to underpredict BC concentrations only in some outflow regions, especially in Asia, but overestimate the concentrations in remote areas, especially at high altitudes. Recent aircraft observations over the central Pacific indicate very small BC concentrations, with values often less than 0.1 ng m$^{-3}$ through the depth of the troposphere (HIPPO campaign - HIAPER Pole-to-Pole Observations; Wang, 2014). Reproducing these observations requires more efficient wet scavenging of BC than is usually implemented in models. Such results suggest that the BC lifetime may be much shorter (ca. 4 days) than what is predicted in most models (6.8 ± 1.8 days in AeroCom). The constraints offered by the HIPPO observations seem to indicate that the radiative forcing of BC is less than in the IPCC AR5 best estimate. However, the availability of in situ (aircraft) measurements of BC in the free troposphere is very limited, and these data may not be representative of the large-scale concentrations.

The existing networks of sun photometers providing long-term measurements of aerosol light extinction certainly provide observations with a much better spatial and temporal coverage. AERONET radiometer data analyses, in particular, provide an estimate of the aerosol absorption optical depth (AAOD) due to BC after subtraction of the dust contribution. The comparison with the BC AAOD resulting from state-of-the-art models indicated that models underpredict the AERONET observations over vast areas of the globe (Koch et al., 2009). Conversely, semi-empirical studies relying on matching the AAOD values retrieved from the AERONET indicate a positive BC radiative forcing of the order of 0.9 W m$^{-2}$ (Ramanathan and Carmichael, 2008), i.e. much greater than predicted by the unconstrained models. Bond et al. (2013) estimated that, in order to match the observed AAOD values, the concentrations of BC determined by standard models must be scaled up by a factor of 2 to 4, depending on the regions of the globe. The authors explain such negative model biases with respect to the retrieved AAOD with uncertainties in the emissions in many areas of the planet and with changes in the optical properties of BC during its atmospheric lifetime. Specifically, the atmospheric aging of soot particles is accompanied by their coating with secondary organic and inorganic material, which can cause a 50% increase in the mass absorption coefficient (MAC) (Bond et al., 2013). The scattering material forming a shell around BC particles in fact acts as a lense directing more photons to the BC core (Jacobson, 2001). Such an effect was demonstrated in laboratory conditions for BC particles encapsulated in non-absorbing material, both inorganic and organo.
ganic (Saathoff et al., 2003; Schnaiter et al., 2005; Metcalf et al., 2013). The internal structure of ambient BC-containing particles is still highly uncertain. Measurements near source regions showed only a negligible increase in MAC with BC aging (Cappa et al., 2012), suggesting that the internal structure of ambient BC-containing particles may not fit the shell and core model. The same extent of internal mixing of BC with the secondary aerosol species is a subject of debate, with poor agreement between model simulations and the (few) available observations (Reddington et al., 2013). Another source of uncertainty is the influence of the nanoscale structure of the BC inclusions (graphitization degree, presence of voids, etc.) on refractive index and therefore on MAC (Bond and Bergstrom, 2006). Clearly, more experimental data on the actual abundance and structure of BC particles internally mixed with non-absorbing materials are required to assess the changes in BC optical properties with aging.

In conclusion, the assessment of the BC radiative forcing is still strongly dependent on the approach, with the methods relying on the AERONET AAOD observations providing a forcing for BC from fossil fuel and biofuel (excluding open burning) of +0.51 (from +0.06 to +0.91) W m$^{-2}$ (Bond et al., 2013), which is about twice that indicated by state-of-the-art models without constraints on the emissions (Myhre et al., 2009; UNEP, 2011).

The most recent IPCC expert judgment was set about halfway from the two above estimates: +0.4 W m$^{-2}$. Such a wide range in the estimated radiative forcing indicates that the quality and/or the spatial and temporal coverage of the available BC concentration and absorbtivity measurements must be improved. It also indicates that the atmospheric transformations of BC are more complex than currently implemented in models. The assumption that BC is mainly accounted for by soot material, almost chemically inert, is in fact a simplification since organic compounds can contribute significantly to aerosol light absorption, especially at short wavelengths (Kirchstetter et al., 2004). Some authors have proposed the term “light-absorbing carbon” as a more general and appropriate substitute for BC (Andreae and Gelencsér, 2006). Organic compounds absorbing in the blue region of the visible spectrum are often referred to as BrC. BrC particles were found in large amounts in biomass burning emissions and were attributed to the occurrence of aromatic compounds such as PAH, lignins (Gelencsér et al., 2003) and heteroaromatic compounds (similarly to natural pigments such as melanins) (Laskin et al., 2009). Interestingly, the BrC material can also form by (secondary) chemical reactions in aged biomass burning particles (Posfai et al., 2004; Saleh et al., 2013). Some authors have even proposed that chemical reactions starting from a variety of reactive volatile organic compounds can lead to the formation of BrC in the background atmosphere (Nozière et al., 2008). However, the optical properties (e.g. complex refractive index) of the specific secondary BrC components investigated are quite diverse (Moise et al., 2015) and their representativeness regarding ambient light-absorbing OAs is unclear. In summary, the contribution of secondary sources to BrC in ambient aerosol remains controversial (Zhang et al., 2013).

Modelling atmospheric light-absorbing carbonaceous aerosols rarely addresses the contribution of BrC to the atmospheric radiative forcing. BrC is often treated implicitly in models constrained by BC measurements, as in Ramanathan and Carmichael (2008). BC measurements are most often performed at a single wavelength; therefore, the measured light absorption of BC actually accounts also for some or even all of BrC. An attempt to extract the contribution of BrC to light absorption was recently proposed by Chung et al. (2012) based on measured AERONET AAOD values and their wavelength dependence. The results were then used to constrain a model. The authors concluded that to achieve consistency between the simulated and the retrieved spectral properties of AAOD, about 20 % of the global AAOD must be attributed to BrC. Recent model simulations treating explicitly the formation of BrC in the atmosphere estimated its radiative forcing to be in the range of +0.22 to +0.57 W m$^{-2}$, which corresponds to 27 to 70 % of the predicted BC absorption (Lin et al., 2014). In situ measurements suggest a lower contribution of BrC absorption with respect to BC; between 3 and 11 % according to Kirillova et al. (2014). These findings indicate that in situ measurements of BrC must be extended to more geographical regions and that the contribution of organic compounds to light absorption must be included explicitly in models (Andreae and Ramanathan 2013).

The radiative forcing of light-absorbing carbonaceous aerosol due to deposition on snow surfaces has been estimated in the IPCC AR5 to be of the order of +0.04 W m$^{-2}$ (+0.02 to +0.09 W m$^{-2}$), which is much lower than the atmospheric aerosol–radiation interaction. However, climate models predict a greater sensitivity of global surface temperature to changes in snow cover than changes in the aerosol light extinction. The forcing is stronger over boreal continents, with a maximum in late spring when ice and snow are exposed to strong insolation (Flanner et al., 2007). The reduction in snow cover in northern American and Eurasian regions amplifies the reduction in sea ice in the Arctic sea, with a strong positive feedback on global temperatures. The impact of BC deposition on the “third pole” (the glaciated region of central Asia) remains uncertain because global models are unable to reproduce the snowpack characteristics in this mountainous environment due to the heterogeneity of landscapes and the very complex orography (Qian et al., 2011).

The radiative effects considered so far produce an instantaneous warming of the atmosphere or, in case of BC on snow, of the Earth’s surface. However, additional forcing factors can be induced by light-absorbing carbonaceous aerosols indirectly, even on short spatial scales (“adjustments” in the IPCC AR5 terminology). For instance, BC perturbs the atmospheric temperature structure, decreasing relative humidity...
and reducing low-cloud cover; this is referred to as the semi-direct effect. Ranges of semi-direct effects are reported in reviews (Bond et al., 2013; Isaksen et al., 2009), going from negative to positive values, with high uncertainties in their evaluation. As already discussed in the previous paragraph, aerosol–cloud interactions are nonlinear and very variable with space and time; hence, they are very difficult to quantify using observations. Chen et al. (2010) investigated the effect of the change in aerosol number concentration and CCN that would result from the reduction in carbonaceous aerosol emissions using a global model. The authors reduced the primary emissions of black and OC mass and number from either only fossil fuel combustion or from all primary carbonaceous sources (fossil fuel, domestic fuel and biomass burning). The direct effect causes a cooling of about 0.1 W m\(^{-2}\) in both scenarios. This cooling is, however, compensated for by the reduction in the number of CCN. Depending on the scenario, this causes the clouds to reflect 0.13 to 0.31 W m\(^{-2}\) less radiation back to space. Thus, the net effect of these realistic combined black and organic carbon scenarios is a positive forcing of 0.1–0.2 W m\(^{-2}\). Koch et al. (2011) analysed the effect on liquid clouds of reducing black and organic carbon from biofuels on liquid clouds in a multi-model comparison. They found that this leads to a positive cloud radiative response of 0.11 W m\(^{-2}\), which is comparable in size but opposite in sign to the corresponding direct effect. Reducing diesel soot (black and organic carbon) leads to even smaller radiative effects.

The above findings indicate that, due to the complexity of interactions and feedbacks, the simple reduction in BC emissions can lead to some counterintuitive effects on the climate system. For instance, the estimation of the globally averaged BC forcing does not take into account potential climate impacts which originate from the spatial heterogeneity of light-absorbing carbonaceous particles in the atmosphere. Differential heating caused by BC between the lower and the higher troposphere or between different regions of the globe can in fact modify the atmospheric circulation with subsequent changes in precipitation regimes, as suggested for the south Asian regions (Ramanathan and Carmichael, 2008). The net warming caused by BC from fossil fuel combustion in the Northern Hemisphere with respect to the Southern Hemisphere is expected to induce a northward migration of the Hadley cell (Wang et al., 2007). Such impacts can be even more important, especially for regional-scale climate changes, than the magnitude of global TOA forcing of BC.

The recent scientific literature provides numerous examples of very complex and diverse possible climate impacts of BC aerosol (see the review by Bond et al., 2013), but there is a general consensus in considering BC as a net warming agent for climate. For this reason the reduction in BC emissions has been proposed as a basis for possible policy action aiming to contain global warming (Bond and Sun, 2005). The technological feasibility of BC emission reductions is in fact much higher than in the case of CO\(_2\). The enormous difficulties in reducing the emissions of CO\(_2\) stimulated an interest in mitigating the global warming of non-CO\(_2\) forcing agents, such as methane and BC (the so-called short-lived climate pollutants – SLCPs). On the other hand, the economical impact of climate policies relying on non-CO\(_2\) forcing agents must consider the side effects on or benefits for air quality and agriculture, since the emissions of such forcing agents are linked to the emission of pollutants (PM) or precursors of pollutants (methylene for tropospheric ozone).

The UNEP approach was integrated by further analysis presented in the work of Shindell et al. (2012). The results indicate that the eventual peak warming is largely dependent on the cumulated CO\(_2\) emissions, not on the policy based on BC and on other short-lived forcers. However, the reductions in non-CO\(_2\) forcing agents act in retarding the temperature increase, providing an option for remaining under the threshold of +2 °C throughout the 21st century (Fig. 34). Half of the warming potential reduction, accounted for by BC emission control, is evenly distributed geographically and hence has very variable national impacts on health and agriculture; these are greatest in highly populated regions (India, China) and in subtropical and tropical climates where ozone levels are very high (Middle East). The global and national benefits calculated using the GAINS (Greenhouse gas and Air pollution Interaction and Synergies) model indicate that the health benefit of BC reductions is much greater than the climate benefits and that the benefits for agriculture...
from the reduction in ozone levels are also relevant for many countries in subtropical climates. The study concluded that the BC reduction measures have clear human-health benefits, especially in highly populated countries, while helping to avoid dangerous rates of climate warming. The approach of Shindell et al. (2012) is the same as that adopted in the UNEP integrated assessment of BC and tropospheric ozone (2011) to estimate the impact of air quality policies on the atmospheric levels of short-lived climate forcing agents. The reference scenario assumes that all current air pollution control policies will be successfully implemented by 2030 and that such regulations will hold during the following decades. The scenario takes into account the very different degree of stringency of the regulation across geographical regions and between economical sectors. It also assumes that further changes in BC source emission types and strength will occur spontaneously through the technological substitution of obsolete heating and cooking practices in developing countries. Under such a scenario, the global emissions of BC will not change significantly by 2030 because the reductions gained by pollution control policies will be offset by a general increase in emissions caused by economical growth, especially in Africa and in the developing countries of other continents. Interestingly, as traditional residential combustion will decline with time and improved technologies of combustion will be adopted, the expected primary OC/BC ratios will decrease with a consequent increased global warming potential of the emissions.

### 3.2.3 Dimming and brightening

Observational and modelling studies that have been published in the past 2 decades suggest that surface solar radiation (hereafter referred to as SSR) is not constant on decadal timescales, as assumed for simplicity and due to a lack of better knowledge, but shows substantial decadal variations, often in line with air pollution patterns. Largely unnoticed over 1 decade or more, this evidence recently attracted an increasing amount of attention with the popular expressions “global dimming” and “global brightening”, which refer to a decadal decrease and increase in SSR, respectively.

### 3.2.4 Observational evidence and possible causes

The monitoring of SSR began in the early 20th century at a few locations and since the mid-century on a more widespread basis. Many of these historic radiation measurements have been collected in the Global Energy Balance Archive (GEBA; Gilgen et al., 1998) at ETH Zurich and in the World Radiation Data Centre (WRDC) of the Main Geophysical Observatory St. Petersburg. In addition, more recently, high-quality surface radiation measurements, such as those from the Baseline Surface Radiation Network (BSRN; Ohmura et al., 1998) and from the Atmospheric Radiation Measurement Program (ARM) have become available. These networks measure surface radiative fluxes at the highest possible accuracy with well-defined and calibrated state-of-the-art instrumentation at selected sites distributed worldwide.

Changes in SSR from the beginning of widespread measurements in the 1950s up to the 1980s have been analysed in numerous studies (e.g. Gilgen et al., 1998; Liepert, 2002; Ohmura and Lang, 1989; Stanhill and Cohen, 2001, and references therein; Wild, 2009, and references therein). These studies report a general decrease in SSR at widespread locations over land surfaces between the 1950s and 1980s. This phenomenon has become popularly known as “global dimming”. Increasing air pollution and the associated increase in aerosol concentrations are considered to be a major cause of the observed decline in SSR (e.g. Stanhill and Cohen, 2001; Wild, 2009). Changes in cloud amount and optical properties, which may or may not have been microphysically linked to the aerosol changes, have also been proposed to contribute to the dimming (e.g. Liepert, 2002). An attempt has been made in Norris and Wild (2007) to differentiate between aerosol and cloud impacts on radiative changes over Europe. They show that changes in cloud amount cannot explain the changes in SSR, pointing to aerosol direct and indirect effects as a major cause of these variations. Strong decadal SSR dimming since the 1950s related to increasing air pollution and associated enhanced aerosol levels have also been noted in China and India (Qian et al., 2006; Wild, 2009, and references therein). Alpert et al. (2005) argued that the decline in SSR in the 1950s to 1980s is particularly large in areas with a dense population, which also suggests a significant anthropogenic influence through air pollution and aerosols. Several studies (e.g. Dutton et al., 2006; Wild, 2009, and references therein) also noted a dimming over the 1950s to 1980s at remote sites, suggesting that the phenomenon is not purely local and air pollution may have far-reaching effects (a concept of how SSR in remote areas may be modulated by subtle changes in background aerosol levels is introduced below in Sect. 3.2.5).

More recent studies using SSR records updated to the year 2000, however, found a trend reversal and partial recovery at many of the sites since the 1980s. The term “brightening” was thus coined to emphasize that the decline in SSR and associated global dimming no longer continued after the 1980s (Wild et al., 2005). Particularly in industrialized areas, the majority of the sites showed some recovery from prior dimming, or at least a levelling off, between the 1980s and 2000. The brightening has been somewhat less coherent than the preceding dimming, with trend reversals at many different locations, but also some regions with continued decrease, such as in India (see Wild, 2009, 2012, for an overview). Brightening is not just found under all-sky conditions but often also under clear skies, pointing once more to aerosols as major causes of this trend reversal (e.g. Norris and Wild, 2007; Ruckstuhl et al., 2008; Wild et al., 2005). The transition from decreasing to increasing SSR is in line with a similar shift in atmospheric clear-sky transmission determined from

---

S. Fuzzi et al.: Particulate matter, air quality and climate change.
Wild (2009). Latest updates on global sulfur emissions indicate a renewed increase in total global sulfur emissions after the year 2000, since the rapidly growing emissions in Asia increasingly outweigh the decreasing emissions in the western world (Streets et al., 2009). This fits in with the lack of a clear overall brightening signal after 2000 and indications for renewed dimming, particularly in China (Wild, 2009).

On the other hand, the longest observational records, which go back to the 1920s and 1930s at a few sites in Europe, further indicate some brightening tendencies during the first half of the 20th century, known as “early brightening” (Ohmura, 2009; Wild, 2009).

### 3.2.5 Dimming and brightening as a function of pollution levels – a conceptual framework

Wild (2009, 2012) proposed a conceptual framework to elucidate the role of aerosols and their cloud-mediated effects for dimming and brightening, suggesting that aerosol-induced dimming and brightening can be amplified or dampened by aerosol–cloud interactions depending on the prevailing air pollution levels. In pristine regions, small changes in CCN can have a much bigger impact on cloud characteristics than in polluted environments because clouds show a nonlinear (logarithmic) sensitivity to CCN (e.g. Kaufman et al., 2005). Additional CCN due to air pollution in pristine regions may therefore be particularly effective in increasing the formation, lifetime and albedo of clouds (Kaufman et al., 2005; Rosenfeld et al., 2006), which all act towards a reduction in SSR through enhanced cloud shading.

Thus, aerosol–cloud interactions in pristine environments may cause a strong amplification of dimming (brightening) trends induced by small increases (decreases) in aerosols. This implies that dimming and brightening could be substantial even in areas far away from pollution sources, where small changes in background aerosol levels induced by long-range transport can effectively alter SSR through cloud modifications (Wild, 2009, 2012). This mechanism could potentially also be responsible for the brightening over oceans with decreasing aerosol background levels (Mishchenko et al., 2007) between the mid-1980s and 2000 consistently seen in the satellite-derived SSR records (Wild, 2009, and references therein).

In polluted regions, on the other hand, cloud microphysics effects tend to saturate with the logarithmic sensitivity to CCN, whereas the direct extinction of SSR by aerosols becomes more relevant with a linear dependency on aerosol loadings. Absorbing pollution layers further heat and stabilize the atmosphere and attenuate SSR and related surface evaporation. This generally leads to a suppression of convective cloud formation and dissolves clouds in layers heated by absorbing aerosol (known as the semi-direct aerosol effect). The associated reduction in cloud shading may partly counteract the aerosol-induced reduction in SSR in heavily polluted areas. Thus, in contrast to pristine areas, aerosol–cloud interactions may tend to dampen dimming and brightening.

**Figure 35.** Time series of annual mean atmospheric transmission under cloud-free conditions determined from pyrheliometer measurements at various sites in Russia (Moscow), Estonia (Tartu, Toravere and Tiirikoja), Switzerland (Payerne) and Japan (average of 14 sites). The records indicate an increasingly polluted atmosphere from the 1950s to the 1980s and a recovery thereafter, plus 2 distinct spikes caused by the high aerosol loadings following the volcanic eruptions of El Chichón (1982) and Mt Pinatubo (1991) (Wild et al., 2005).
trends induced by direct aerosol effects (Wild, 2009, 2012). This may explain a seemingly counterintuitive phenomenon observed in China, where under strongly increasing pollution both SSR and cloud amounts declined between the 1960s and 1990s (e.g. Qian et al., 2006).

3.2.6 Environmental implications of dimming and brightening

A growing number of studies provide evidence that the variations in SSR have a considerable impact on climate and environmental change (Wild, 2009, 2012). Wild et al. (2007) investigated the impact of dimming and brightening on global warming. They present evidence that SSR dimming was effective in masking and suppressing greenhouse warming but only up to the mid-1980s, when dimming gradually transformed into brightening. Since then, the discovered greenhouse effect has revealed its full potential, as manifested in a rapid temperature rise (+0.38 °C decade\(^{-1}\) over land since mid-1980s). More recently, Wild (2012) pointed out that the absence of global warming from the 1950s to 1980s and the subsequent reversal into rapid warming was most prominently seen in the Northern Hemisphere, while on the Southern Hemisphere rather a steady gradual warming since the 1950s was observed (Fig. 36). This fits the asymmetric hemispheric evolution of anthropogenic air pollution which strongly increased from the 1950s to the 1980s and declined thereafter in the Northern Hemisphere, while pollution levels in the Southern Hemisphere were 1 order of magnitude lower and steadily increased with no trend reversal (Stern, 2006; Wild, 2012). This again points to a possible large-scale (hemispheric) influence of aerosol-induced SSR dimming and brightening on global warming. Interestingly, the suppression of warming during the dimming period in the Northern Hemisphere was even slightly stronger over oceans than over land (a slight cooling of \(-0.03\) °C decade\(^{-1}\) over oceans between 1958 and 1985, compared to a slight warming over land with +0.04 °C per decade over the same period, based on data from the Climate Research Unit, Norwich, and the Hadley Centre, Exeter) (Wild, 2013). Even though anthropogenic air pollution sources are located over land, subtle changes in background aerosol levels over the relatively pristine oceans could have amplified SSR trends through effective cloud–aerosol interactions as outlined in the conceptual framework in Sect. 3.2.5. This may explain the lack of warming particularly also over oceans during this period.

Finally, it cannot be excluded that the current transition into a possible renewed SSR dimming, possibly triggered by a renewed increase in global pollution levels since 2000 (particularly originating from Asia) may have contributed to the lack of global warming since the beginning of the millennium (“global warming hiatus”) (Wild, 2012).

SSR is also a major energy source for surface evaporation and thereby the main driver of the global water cycle (Wild and Liepert, 2010). Wild et al. (2004) suggested that surface solar energy reductions outweighed the increasing thermal energy from the greenhouse effect from the 1960s to 1980s, resulting in a reduction in surface net radiation and associated evaporation over land surfaces, in turn causing an attenuation of the intensity of the associated water cycle. In contrast, for the more recent period of the 1980s to 2000s, Wild et al. (2008) pointed out that SSR brightening adds to the increasing energy from the enhanced greenhouse effect, leading to higher evaporation and an intensification of the global terrestrial water cycle since the 1980s. Impacts of the transition from dimming to brightening can further be seen in the more rapid retreats of glaciers and snow cover, which became evident since the 1980s as soon as the dimming ceased (Wild, 2009, and references therein).

Further research will be required to better quantify large-scale SSR dimming and brightening trends, to disentangle the contributions by aerosols and cloud microphysics and to assess the associated implications for various aspects of climate and environmental change.

4 Natural pressures on aerosol concentration trends

This section will summarize the natural drivers of changes in PM in future and explain the causes of and implications for air quality and climate policy. It is possible that PM will change in future in ways that may not be accounted for when air quality policies are set. Although, if successfully implemented, air quality policy can cause a downward trend in PM, there are many reasons why any downward trend could be reduced in magnitude because of natural factors that are not under our direct control. We will discuss issues related to PM (which is regulated by air quality policy and associated with health impacts) and quantities that are more relevant for climate change (PN, CCN, etc.).
4.1 Changes in natural aerosol emissions

This section considers external factors that may cause future changes in PM levels, irrespective of changes in anthropogenic particulate and precursor emissions. There are two main drivers that need to be considered. Firstly, a large fraction of ambient aerosol particle mass and number is derived from natural emissions from the land and ocean surfaces and the biosphere. These emissions are susceptible to modification due to changing climate. These factors have been reviewed by Carslaw et al. (2010) in terms of climate feedbacks, although PM was not explicitly considered. Secondly, climate change can modify the overall behaviour of all aerosols in the atmosphere, whether from natural or anthropogenic sources, through alterations to atmospheric chemistry, aerosol formation and deposition processes, and other meteorological drivers. These processes have previously been reviewed by Jacob and Winner (2009), Fiore et al. (2012) and Carslaw et al. (2010).

Natural aerosol is a major source of particle mass and number from both primary and secondary sources. On a global scale, by far the greatest emissions by mass are from natural sources (dust approximately 2000 Tg a\(^{-1}\) and sea spray about 8000 Tg a\(^{-1}\), versus less than 60 Tg a\(^{-1}\) for anthropogenic SO\(_2\); Dentener et al., 2006). However, the exposure of the population to these particles is less than suggested by the global atmospheric burden because people in most cases tend to live far from the major natural sources but very close to emissions of anthropogenic aerosols in urban areas. Thus, global average responses of natural aerosol emissions to climate, as reviewed in Carslaw et al. (2010), are unlikely to be directly applicable to urban populations.

Natural aerosols contribute to particle concentrations over Europe (Simpson et al., 1999), although the contribution to PM\(_{10}\) has been formally recorded only since the EU Directive 2008/50/EC on Ambient Air Quality and Cleaner Air for Europe (EC, 2008), which allows member states to compare ambient air pollutant concentrations with relevant legally binding limits after the contribution of natural sources has been subtracted. This directive recognizes that natural aerosols contribute to PM\(_{10}\) levels in Europe and therefore should not be classed as air pollutants in assessments of PM exceedances. The 2012 European Environment Agency report (EEA, 2012) provides the first assessment of the natural contributions to PM\(_{10}\) at several monitoring sites. Natural aerosols are defined as wind-blown dust, sea spray, volcanic dust and wildfires, and the directive provides guidelines on how these aerosol types should be identified in measurements, although there is ambiguity in the detection methods because there is no perfect tracer of natural versus anthropogenic aerosol sources. BSOA contributions to exceedances of the PM\(_{10}\) limit values in Europe are thought to be rare and so were not included in the legislation. This is mainly because the emissions are concentrated in cleaner regions and because BSOA makes a substantial contribution mainly to submicron aerosol mass (Zhang et al., 2007). Although not considered in PM\(_{10}\) legislation, BSOA is likely to be important for the health effects of UFPs (see Sect. 3.1.1) and so will be considered in this section.

4.1.1 Contributions of natural aerosols to PM exceedances in Europe

The contribution of natural aerosols to European PM\(_{10}\) exceedances for the years for which data are available (2008 and 2009) is as follows (EEA, 2012): Austria – 2 out of 16 days; Germany – 1 out of 40 days; Spain – 96 out of 165 days; France – 1 out of 72 days; UK – 5 out of 8 days; Greece – 5 out of 22 days; and Portugal – 6 out of 20 days. So, by country, 5–30 % of exceedance days can be attributed to natural aerosols. The mean annual contributions of natural sources to PM\(_{10}\) in 2008 were 1–3 µg m\(^{-3}\) in Italy, France, Portugal and Greece and 4–5 µg m\(^{-3}\) in Spain and the UK. The natural source that was most responsible for exceedances was transport of natural particles from dry regions outside the country (Saharan dust), followed by sea spray and wildfires. Only the UK and France reported a contribution of sea salt to exceedances and only Greece reported an exceedance (in both 2008 and 2009) due to wildfires. Dust caused exceedances in Spain, France, Greece, Italy and Portugal (Fig. 37). Thus, in terms of PM\(_{10}\) exceedances, dust is by far the most important natural aerosol over Europe. However, because of the much smaller size distribution of smoke particles from wildfires, with most of the mass below 1 µm, fires are likely to make a much larger contribution to PM\(_{2.5}\).
and PM$_1$ concentrations, but no information is available from routine measurements.

4.1.2 Soil and desert dust trends

Dust is recognized to make a major contribution to PM levels in Europe, particularly in southern countries such as Spain, Italy and Portugal (Kallos et al., 2007; Koçak et al., 2007; Pey et al., 2009; Querol et al., 2009; Rodríguez et al., 2001, 2002). Most of these dust events can be traced back to Saharan emissions, although emissions from European agricultural land during times of drought have also been documented (Bessagnet et al., 2008). The meteorological situations that favour dust uplift and transport towards Europe in different seasons are well understood. For the western Mediterranean, the main sources are the Sahara and Sahel regions (e.g. Rodríguez et al., 2001; Escudero et al., 2005), while for the central and eastern Mediterranean, northern African sources are also important (e.g. Kallos et al., 2006, 2007; Meloni et al., 2008).

There have been few studies of how climate change could affect these dust sources or the occurrence of meteorological situations that would transport the dust to Europe. It is known that northern African dust emissions are strongly related to the general meteorological situation of the North Atlantic, and in particular the Atlantic Multi-decadal Oscillation (AMO), although aerosol levels over the Atlantic (including dust) may also influence the AMO itself (Booth et al., 2012); thus, deducing the causes of any long-term change is likely to be a demanding task involving coupled global models. There is a statistically significant negative correlation between the AMO and dust emissions from Africa (C. Wang et al., 2012), with an observed decrease in dust concentration of 1.5 % yr$^{-1}$ from 1984 to 2012 (90 % confidence), and a global decrease in all major sources of 1.2 % yr$^{-1}$. However, PM exceedance events over Europe, being highly variable and episodic, are more likely to be affected by a change in the frequency of dust storms and dust transport to Europe than by changes in the mean dust concentration. However, information about the frequency of events is not available on climate change timescales.

4.1.3 Sea salt trends

Sea salt is a major component of PM$_{10}$ in the coastal regions of Europe, where it can comprise up to 80 % of the annual mean particulate mass (Putaud et al., 2004), but also contributes to PM in inland areas, with concentrations between 0.3 and 13 µg m$^{-3}$ (Manders et al., 2010). Although the dominant mass fraction of sea-spray aerosol is sea salt, measurements at Mace Head on the west coast of Ireland show that a significant fraction of submicron mass is derived from biogenic primary sources of marine organic components (O’Dowd et al., 2004). The water-insoluble organic fraction in fine marine aerosol was observed to be the most important contribution, particularly during periods of phytoplankton bloom in the North Atlantic. Future changes in sea salt emissions are projected to be small as a global mean and unlikely to be important for changes in PM over Europe. For example, Jacobson and Streets (2009) calculated decreases in global sea spray of only 0.4 % by 2030 in a warming scenario and there have not been any significant trends in sea-spray production in the North Atlantic over recent decades (Korhonen et al., 2011).

4.1.4 Wildfire trends

An assessment of the contribution of fires to PM levels in populated areas of the world is complicated by the difficulty of separating natural and human factors. In the EU legislation for natural PM$_{10}$ discussed above, fire must be demonstrated to have a natural cause to be considered a natural source (and therefore subtracted before assessing PM$_{10}$ exceedances). It is difficult to distinguish natural causes of fires, such as lighting, from human causes (such as accidents). There was an average of 95,000 fires per year in Europe in the period of 2000–2005, resulting in nearly 600 kha of burnt land per year. Most fires occur in France, Greece, Italy, Portugal and Spain (500 kha of burnt area per year), where summers are drier and hotter than in other parts of Europe (Barbosa, 2009). Studies of fires in Greece and Portugal show significant contributions to air pollution, with PM levels frequently exceeding PM$_{10}$ limits (e.g. Miranda, 2004; Hodzic et al., 2007; Miranda et al., 2008). During an intensive fire season in 2003 in Portugal, wildfire emissions caused an increase in mean PM$_{2.5}$ of 20–200 %, reaching levels of 40 µg m$^{-3}$ (Hodzic et al., 2007). Fires in Russia and eastern Europe also influence northern Europe. For example, Niemi et al. (2005) showed that increases in very low long-term background PM$_{2.5}$ levels of about 8 µg m$^{-3}$ in parts of
Finland were increased to more than 30 \( \mu g \) m\(^{-3}\) during such events.

It has been estimated that more than 90% of fires in the Mediterranean region are caused by human activities, such as negligence, accidents and deliberate ignition (Ganteaume et al., 2013). This might suggest that only about 10% of fires (those due to natural processes) are influenced by climate change. However, the situation is more complex and not completely understood. For example, Knorr et al. (2014) showed that the net effect of humans on total fire frequency (natural and human-caused) through land management, active fire suppression, or landscape fragmentation is a net suppression of fire frequency in almost all parts of the world. Thus any future upward pressures on fire frequency due to climate change over Europe would not necessarily translate into a net increase in PM emissions. Moreover, aerosol emissions are not determined solely by the number of fires but also by their duration, extent and severity, amongst other factors, and it has been shown that natural fires tend to be larger and longer-lived because they occur in less accessible locations (e.g. Vazquez and Moreno, 1998; Stocks et al., 2002).

Global emissions from wildfires vary interannually by more than a factor of 2, although only part of this variation can be attributed to weather (Generoso et al., 2003; Schultz et al., 2008; van der Werf et al., 2004). For example, fire emissions are larger during El Niño years because drought conditions associated with El Niño trigger an increase in fire activity. Forest fires in some regions clearly vary on decadal timescales, with a clear association with climate variables (Gillett et al., 2004; Westerling et al., 2006), although the global mean ~ 50% increase in wildfire emissions in the period of 1960–1990 is most likely due to increased deforestation (Schultz et al., 2008).

Future model projections of wildfires mostly have a regional focus. Increased wildfire due to climate change has been projected for North America (Flannigan et al., 2000) and Canada (Flannigan and Van Wagner, 1991; Wotton and Flannigan, 1993; Stocks et al., 1998; Flannigan et al., 2005) and the western United States in particular (Brown et al., 2004; Westerling and Bryant, 2008). Of more relevance for European air quality, increases in fire activity have been projected for Russia (Stocks et al., 1998) and the European Mediterranean area (Moriondo et al., 2006; Good et al., 2008). A global assessment of fire risk by the end of the century (2070–2100) was conducted by Liu et al. (2010), using four general circulation models (Fig. 38). They used the Keetch–Byram drought index (KBDI) (Keetch and Byram, 1968), which assesses fire risk in terms of a parameterization of soil moisture deficit (based on precipitation and maximum temperature). At present, southern Europe is assessed to have the lowest fire risk out of all the regions studied (United States, South America, Asia, South Africa and Australia) but is projected to have the greatest increase by 2070–2100 (a 300% change in KBDI), leading to a change in fire risk from low to moderate, with such conditions extending from June to November. Whether an increased fire risk actually leads to more fires and greater smoke emission depends on the complicating human factors discussed above. Again, although European fires mostly have a human cause, the increased fire risk in the future may result in greater spreading of deliberate and accidental fires, but this has not been demonstrated.

Given the many complex factors that control fire emissions, projections of changes in emissions over Europe in a future climate, where the cause of fires is dominated by human factors, have not been made. In the western US, interannual variability in wildfires is the dominant driver of observed variability in summertime organic carbon concentrations (Jaffe et al., 2008; Spracklen et al., 2007). Spracklen et al. (2009) calculated that increases in temperature could cause the annual mean area burnt in the western United States to increase by 54%, resulting in a 90% increase in wildfire carbonaceous emissions by the 2050s relative to present-day conditions and an increase in OC concentrations of 40%. This change represents an increase in summertime mean PM\(_{2.5}\) of about 1 \( \mu g \) m\(^{-3}\) in limited regions. Similar increases were projected for the same period by Yue et al. (2013) based on 15 GCMs in CMIP3. They concluded that 75% of the projected change in OC in the western United States would be caused by changes in wildfire. No such estimates exist for Europe.

### 4.1.5 Biogenic secondary organic aerosol trends

SOA makes a substantial contribution to aerosol mass at submicron sizes (Zhang et al., 2007). As described above, SOA is not expected to contribute significantly to PM\(_{10}\), but it does make a substantial contribution to submicron particles, particularly below a few hundred nanometres. A large fraction of this organic mass has been attributed to natural emissions of BVOCs. In forested areas, biogenic SOA has been directly linked to the OA mass concentration. For example, over Scandinavia, the aerosol mass is proportional to the length of time the air has spent over forested land (Tunved et al., 2006, 2008). There are large uncertainties in the contribution of BSOA to global OA mass, with estimates using different approaches ranging from 12 to 1820 Tg production of SOA per year (Kanakidou et al., 2005; Goldstein and Galbally, 2007; Hallquist et al., 2009) (assuming a constant OA/C mass ratio equal to 2.0 to convert reported values to OA mass). Spracklen et al. (2011) used a pseudo-inverse modelling approach and global AMS measurements to derive an optimized SOA source of 140 ± 90 Tg yr\(^{-1}\), although a large and uncertain fraction of this is reported to be “anthropogenically controlled” and may be derived either from anthropogenic VOC emissions or from the effect of anthropogenic pollution on the production of SOA from natural BVOCs. Over Europe, the optimized model predicts SOA concentrations of 2–4 \( \mu g \) m\(^{-3}\).
Future changes in BVOC emissions and the effect on global SOA were reviewed in Carslaw et al. (2010). There have been several subsequent studies applying updated knowledge regarding the response of BVOC emissions to climate change and SOA formation chemistry, including the effects of temperature on particle size distributions and CCN (Paasassen et al., 2013) and the direct and indirect radiative effects (Scott et al., 2014). Jiang et al. (2010) showed that regional increases in BSOA by 2050 reached a maximum of about 26%. Megeritis et al. (2013) used a regional CTM driven by biogenic emissions from MEGAN (Model Emissions of Gases and Aerosols from Nature). In a scenario with temperatures everywhere across Europe increased by 2.5 K, they calculated increases in summertime biogenic SOA of 20% over northern parts of Europe, amounting to a change in PM$_{2.5}$ of less than about 0.5 µg m$^{-3}$. The contribution elsewhere in Europe was smaller. The effect of the changing emissions on CCN may be different because organic compounds can also act as a source of new particles from nucleation (Carslaw et al., 2013b; Riccobono et al., 2014; Scott et al., 2014), which could amplify the response.

4.2 Changes in climate and the impact on aerosol and air quality

The impact of climate change on air quality is difficult to assess because it is not driven by a dominant factor. In the case of air quality impacts on climate change, this dominant factor is the emissions of pollutants that have fairly well-understood direct effects on aerosols and trace gases — so-called short-lived climate forcers. In contrast, changes in air pollutants in response to climate change depend upon how the multiple complex interactions among the chemical species, the land surface and other factors respond to changes in climate (temperature, rainfall, humidity, etc.). For example, temperature affects the chemical rates that determine pollutant concentrations; changes in precipitation can alter aerosol concentrations; atmospheric circulation changes can affect pollutant distributions (Pausata et al., 2015); and emissions from the biosphere (such as VOCs) are temperature and moisture dependent.

Jacob and Winner (2009) reviewed knowledge of the effect of climate change on air quality with a focus on 21st-century projections. The review included an analysis of multiple approaches, including observed correlations of PM with meteorological variables, model perturbation studies using CTMs and projections using GCMs. They concluded that “GCM–CTM studies of the sensitivity of surface PM to 21st-century climate change find annual mean effects of the order of 0.1–1.0 µg m$^{-3}$ for North America and Europe, with no consensus between studies as to the sign of the effect.” This assessment included changes in natural emissions already discussed in the previous section.

As part of the EU Pan-European Gas–AeroSOls–climate interaction Study (PEGASOS) project, an expert elicitation was carried out to identify and rank the most important ways in which climate change could affect PM, and the results are given in Table 4. The 37 participants were all members of the PEGASOS project with varying levels of expertise in atmospheric composition and climate. The time horizon of interest was defined to be 2050 with a focus on European air quality and climate. Participants were asked to rate the importance of each factor in terms of the need to include the process in a coupled air-quality–climate model. Twenty-eight model processes were identified that could link climate and air quality, including the emissions discussed above. Further interactions were suggested by the participants during the process. Of these 28, 9 were estimated to be of medium to high importance for air-quality–climate coupling over a time period to 2050 — that is, their inclusion in a model would lead to an important difference in the prediction of air quality and climate over that time period compared to neglecting the process. The top two processes were identified as the impact of climate change on BVOC emissions, leading to changes in OA, with other chemical processes such as reaction rates and particle volatility being ranked as having low to medium importance.

Many of these processes have been studied in regional and global models (Carslaw et al., 2010; Jacob and Winner, 2009), although not in a consistent way that enables a direct comparison and quantification of PM changes under a particular scenario and not always with a focus on Europe.

A study using the PMCAM-x (PM Comprehensive Air quality Model) regional CTM over Europe (Megeritis et al., 2013) showed that an increase in temperature of 2.5 K would cause a small increase in PM of about 1% averaged over Europe. In the summer, increases in biogenic SOA occur in the model due to increased BVOC emissions (described above) but are compensated for by a decrease in ammonium nitrate, which becomes more volatile at higher temperatures. In winter, similarly small increases occurred but due instead to the vaporization of primary OA and an increase in sulfate. Thus, over Europe, changes in PM in a warmer climate may be moderated by compensating effects, although there could be decreases of up to 1 µg m$^{-3}$ in regions where ammonium nitrate aerosol dominates. This effect is similar to estimates for those parts of the US where ammonium nitrate dominates (Dawson et al., 2007; Jacob and Winner, 2009). Further work is needed to determine whether these changes in PM could have a feedback effect on aerosol radiative forcing that has so far not been considered in models, but the net effects on PM appear to be small.

The next two most important processes were identified as the impact of changing atmospheric circulation and precipitation on chemical species and aerosol. The risk of extreme heat events over Europe is projected to increase as a consequence of greenhouse gas emissions (Barnett et al., 2006; Meehl and Tebaldi, 2004; Schär et al., 2004). Such events are associated with circulation patterns that also cause the stagnation of air and the build-up of high levels of pollution.
Table 4. Results of the expert elicitation in the PEGASOS project. Thirty-seven participants rated the importance of a list of factors that could influence the coupling of air quality and climate by 2050. This table summarizes results for PM. High importance implies “must include in the model”, medium importance “fairly important for the model”, low importance “desirable, but not essential, to have in the model” and negligible “can be neglected in the model”. The rank is based on a weighted mean of the scores (negligible: 1; low: 2; medium: 3; high: 4).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Discussed here</th>
<th>Effect</th>
<th>Neg</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
<th>Uncertain</th>
<th>Ave. score/5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>Changes in temperature cause changes in BVOC emissions and SOA</td>
<td>0</td>
<td>11</td>
<td>18</td>
<td>71</td>
<td>2</td>
<td>3.6</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Changes in precipitation frequency and/or intensity affect aerosol removal</td>
<td>0</td>
<td>7</td>
<td>45</td>
<td>48</td>
<td>1</td>
<td>3.4</td>
</tr>
<tr>
<td>3</td>
<td>Y</td>
<td>Changes in circulation patterns affect periods of meteorological stagnation</td>
<td>0</td>
<td>4</td>
<td>60</td>
<td>36</td>
<td>5</td>
<td>3.3</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>Changes in climate cause changes in forest fires</td>
<td>0</td>
<td>10</td>
<td>57</td>
<td>33</td>
<td>0</td>
<td>3.2</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>Changes in the land surface (human-induced or natural) affect BVOC emissions and SOA</td>
<td>0</td>
<td>11</td>
<td>61</td>
<td>29</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>Changes in climate cause changes in dust uplift and/or transport</td>
<td>7</td>
<td>19</td>
<td>48</td>
<td>26</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>7</td>
<td>Y</td>
<td>Changes in temperature cause changes in SOA partitioning into aerosol</td>
<td>0</td>
<td>27</td>
<td>54</td>
<td>19</td>
<td>4</td>
<td>2.9</td>
</tr>
<tr>
<td>8</td>
<td>Y</td>
<td>Changes in temperature cause changes in nitrate partitioning into aerosol</td>
<td>0</td>
<td>31</td>
<td>46</td>
<td>23</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Changes in climate cause changes in water vapour, oxidants and hence sulfate</td>
<td>0</td>
<td>32</td>
<td>48</td>
<td>20</td>
<td>5</td>
<td>2.9</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Changes in tropospheric composition affect photolysis rates and hence PM</td>
<td>0</td>
<td>33</td>
<td>46</td>
<td>21</td>
<td>5</td>
<td>2.9</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>Changes in air pollutants such as ozone impact the terrestrial biosphere and BVOC emissions and therefore feedback on ozone and PM</td>
<td>0</td>
<td>41</td>
<td>33</td>
<td>26</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>12</td>
<td>Y</td>
<td>Changes in climate cause changes in marine aerosol production</td>
<td>15</td>
<td>19</td>
<td>42</td>
<td>23</td>
<td>4</td>
<td>2.7</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Changes in temperature cause changes in methane emissions, which impact OH and aerosol production</td>
<td>7</td>
<td>36</td>
<td>39</td>
<td>18</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>14</td>
<td>Y</td>
<td>Changes in climate and land surface affect dry deposition</td>
<td>7</td>
<td>37</td>
<td>37</td>
<td>19</td>
<td>3</td>
<td>2.7</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>Changes in stratospheric ozone affect tropospheric UV and hence tropospheric ozone and PM</td>
<td>9</td>
<td>30</td>
<td>52</td>
<td>9</td>
<td>7</td>
<td>2.6</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Changes in climate affect vertical mixing and/or boundary layer height and hence ozone and PM</td>
<td>4</td>
<td>40</td>
<td>52</td>
<td>4</td>
<td>5</td>
<td>2.6</td>
</tr>
<tr>
<td>17</td>
<td>Y</td>
<td>Changes in temperature cause changes in reaction rates affecting sulfate formation</td>
<td>11</td>
<td>54</td>
<td>32</td>
<td>4</td>
<td>2</td>
<td>2.3</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>Changes in surface radiation caused by changes in aerosols impacts BVOC emissions and SOA</td>
<td>11</td>
<td>56</td>
<td>30</td>
<td>4</td>
<td>3</td>
<td>2.3</td>
</tr>
</tbody>
</table>
and are responsible for a large fraction of European PM exceedances. At present, while it is known that extreme heat events will increase, there is doubt about the relative contribution of changes in the frequency and duration of the events versus changes in the intensity (of heat) during an event (Clark and Brown, 2013), although heat intensity seems to be more important. Elevated PM levels are susceptible mostly to the duration of such events, although heat may also play a role because biospheric emissions increase with temperature and heat is also important for the risk of fire (see previous section). A further circulation effect of potential importance for European PM is the North Atlantic Oscillation (NAO). Pausata et al. (2013) showed that phases of NAO since 1980 are correlated with wintertime PM anomalies of up to several micrograms per cubic metre. Any long-term change in the NAO would therefore impact PM levels.

The status of the understanding of precipitation impacts remains poor because of disagreement among climate models regarding the effect of climate change on precipitation (IPCC, 2007, 2013). Using a GCM, Avise et al. (2009) projected a change in summertime PM$_{2.5}$ of $-1.0$ µg m$^{-3}$ across the US by 2050 due to changes in climate, compared to a $+3.0$ µg m$^{-3}$ due to projected changes in anthropogenic emissions over the same period. Changes in precipitation were found to have a very large effect of up to $-3.0$ µg m$^{-3}$ in some regions. However, aerosol is affected more by the frequency of precipitation rather than by the amount, which is a model quantity that is not well predicted at present, and model projections of changes in continental precipitation are highly uncertain. Nevertheless, the response of PM to changes in precipitation remains a significant uncertainty.

The impact of circulation changes on PM for the US was assessed by Tai et al. (2012a). They found that the frequency of cold fronts was strongly associated with the observed interannual variability in PM$_{2.5}$. Because GCMs project a decrease in the frequency of midlatitude cyclones with global warming (Bengtsson et al., 2006; Christensen et al., 2007; Lambert and Frye, 2006; Pinto et al., 2007; Ulbrich et al., 2009b), Tai et al. (2012b) assessed the potential degradation of air quality between 2000 and 2050 by analysing IPCC AR4 data from 15 models. In 1999–2010, observations over the US found robust correlations ($r > 0.5$) of annual mean PM$_{2.5}$ with temperature, with the dominant mode representing frontal passages. They then projected changes in PM$_{2.5}$ between 2000 and 2050 based on changes in simulated temperature and estimated a likely 2000–2050 increase of about 0.1 µg m$^{-3}$ in annual mean PM$_{2.5}$ in the eastern US. The main cause of this small increase was less frequent frontal ventilation. Smaller decreases in PM were predicted over the north-western US due to more frequent maritime air flow. Thus, the response of PM to synoptic-scale dynamical changes is small, at least in terms of the annual mean.

Changes to terrestrial ecosystems can also cause changes in aerosol. Wu et al. (2012) projected a 0.1–0.2 µg m$^{-3}$ increase in the organic fraction of PM in parts of the US due to climate-driven changes in ecosystem type. Globally, they found that a climate-driven increase in SOA burden by 2100 was negated by a similar fall in SOA when changes in land use were taken into account. A similar effect was found by Heald et al. (2008). Changes over Europe, where land use is not projected to change substantially, are very small. Both of these studies simulated changes in SOA, but changes in land cover will also affect all aerosols and could lead to a small net decrease in some parts of the world.

Furthermore, climate change is driving changes in natural species composition but also alters agricultural and silvicultural practice. Associated impacts on land cover will, mainly through changes in surface roughness, alter the net deposition of aerosol. In addition, aerosol deposition rates are in part regulated by turbulence: thus, a more turbulent climate may reduce atmospheric lifetimes. The situation is different for the volatile aerosol compounds, whose dry deposition removal rate is effectively governed by the likelihood of them evaporating during deposition (see above). For example, wetter vegetation surfaces would provide a more efficient sink for NH$_3$, increasing the potential for NH$_4$NO$_3$ evaporation, and stronger temperature gradients would also increase the evaporation and the effective deposition rate of nitrate. Higher air temperatures are more likely to shift the gas–aerosol partitioning towards the gas phase throughout the boundary layer, resulting in lower NH$_4$NO$_3$ concentrations throughout but not necessarily in an increased surface removal rate.

Overall, climate change exerts complex and sometimes competing pressures in PM levels in Europe due to changes in natural aerosol and precursor gas emissions, changes in meteorology affecting aerosol transport and removal, and the effect of climate change on the physical and chemical behaviour of all aerosol particles in the atmosphere. The conclusions for Europe are similar to those of, e.g., Jacob and Winner (2009). Changes in PM$_{2.5}$ of about 1 µg m$^{-3}$ are possible locally due to the effect of changes in the chemical production of SOA and the partitioning of the organic compounds and nitrate between the gas and aerosol phases. According to current our understanding, increased emissions of VOCs from the terrestrial biosphere in a warmer climate could account for about a 20% increase in this source of OA, amounting to about 0.5 µg m$^{-3}$ in parts of Europe. The effects of changes in meteorology are poorly understood, particularly with regard to the effect of changes in precipitation and aerosol wet deposition. Regional-scale decreases in PM$_{2.5}$ of 3 µg m$^{-3}$ have been projected for 2050 for the US, but the confidence in such projections is very low due to the poor handling of aerosol wet scavenging in models and the uncertainty in how precipitation frequency, distribution and intensity will change. Synoptic-scale changes in circulation leading to more extreme heat waves are projected in many models and are likely to affect PM extremes, but more research is needed specifically on the PM response, which, although related, is not controlled by the same factors as
heat extremes. Dust and wildfires are an important source of PM over southern parts of Europe and are now recorded operationally as a contribution to PM$_{10}$ exceedances. Future changes in dust events are more likely to be controlled by changes in the transport of the dust to Europe than by net changes in dust emission. Although the transport pathways are well understood, it is not known how they will respond, if at all, to climate change. Future changes in wildfires over Europe depend in a complex way on climate-driven changes in fire risk, but they are strongly influenced by human interference in the causes and suppression of fires. At present, there are no predictions of changes in fire PM emissions for the European area that take all these factors into account.

5 Conclusions and outlook

There are many open questions in atmospheric aerosol science that require basic and fundamental studies that can help explain observed phenomena, provide the fieldwork with new directions and enable better model predictions for future conditions and scenarios.

The development of innovative measurement techniques and data analysis methods for the apportionment of PM sources has been effective in the development and implementation of air quality policy, especially targeting primary emission sources. For example, the introduction of EURO standard for vehicle emissions and the reduction in SO$_2$ emissions contributed to the decrease in PM$_{10}$ and PM$_{2.5}$ during the last 20 years. Still, identifying and quantifying SOA sources in ambient air remains challenging. A dichotomy exists between source attribution techniques that rely on source-specific tracers that lead to highly uncertain scaling to ensemble mass and techniques addressing the full aerosol composition but with reduced chemical specificity (and thus reduced source resolution power). There is an urgent need for analytical and measurement techniques capable of bridging this gap. Further, it is clear that both the magnitude and composition of emissions depend strongly on location or time-specific source conditions (burn or driving conditions, fuel type, etc.) and that this affects both primary aerosol and secondary precursors. Source attribution models must be able to adapt to this variability and to the dynamic nature of the emitted species in the atmospheric system.

It has been shown that describing SOA formation using traditional precursors is inadequate, and the importance of SVOCs and IVOCs in SOA has become apparent. Given the partitioning-driven description of SOA formation employed by the volatility basis set (VBS; Robinson et al., 2007), the distinction between primary and secondary emissions is blurred. A new paradigm for source emission characterization and reporting is needed that encompasses not only primary aerosols and gases but also SVOCs, IVOCs and secondary aerosol production potential (SAPP). Furthermore, the spatial and temporal representativeness of aerosol emission inventories is often limited and the consistency across scales needs improvement.

The WHO (2013a) report recognized that there is consistent evidence that the decrease in air pollutant concentration during the last 10 years around the world, due to policy measures or unplanned reductions, has led to an improvement in health. This conclusion is supported by a large body of evidence from short-term and long-term studies of human exposure to air pollution. Epidemiological, toxicological and controlled exposure studies during the last 10 years have strengthened the causal association between PM$_{2.5}$ exposure and adverse health outcome. In addition, more insight into physiological mechanisms that explain PM$_{2.5}$ short-term and long-term health effects is available. So far, there is no conclusive evidence to pinpoint a single pollutant or a limited number of species as the main harmful components of PM, and further research in this area is urgently needed to inform policy priorities. It has not been possible either to conclusively show that specific PM components, at relevant outdoor concentrations, are harmless.

A major challenge in this field is, therefore, relating chemical constituents to health effects and speciation of reactive species in particles (e.g. peroxides, radicals, metals). In addition, relating aerosol processes to their health impacts (such as the aging of biomass burning and pollution aerosols, characterization of suspended material and understanding the detailed chemistry of trace toics that partition to particles) will help the assignment of specific health effects to aerosol components. The fact that the toxicological properties of aerosol particles are affected by the presence of reactive, labile functional groups (such as peroxides and radicals) means that the toxic potential of aerosol particles is not conserved upon emission in the atmosphere and that it can be either enhanced or depressed by secondary processes, hence challenging any simple approach to PM toxicity source apportionment.

To better understand the role of PM in public health and to optimize the available tools for air quality improvement, there is a need for (i) a better focus on a multi-pollutant approach, which should integrate PM sources, chemical properties, physical properties and gas pollutants, (ii) a better understanding of the mechanisms responsible for PM health effects, (iii) improving exposure estimates integrating micro sensors, hybrid models, local-scale observations and remote sensing and (iv) promoting an interdisciplinary approach that integrates atmospheric chemistry, global modelling, exposure science, toxicology, epidemiology and statistics.

The future legislation for PM reduction should not only control primary aerosol emissions and known precursors but also regulate SAPP. This will require the development of standardized methods for SAPP measurement and could be conceptually similar to ozone production potential from individual VOCs. Policy measures to improve air quality need support from air quality models, but several models have limitations in predicting secondary aerosols. New model approaches describe better atmospheric aerosol loadings, but
vertical distribution and temporal variability are still poorly accounted for.

The UNEP report (2011) confirms that the current scientific understanding of the aerosol–radiation effects is sufficient to promote the evaluation of measures to limit emissions of the light-absorbing fraction of the aerosol (the BC) that has now been recognized as the second most important warming agent on a global scale beside CO₂ (Bond et al., 2013). Such measures would not only mitigate global warming in the medium term but also would lead to substantial benefits in terms of air quality improvement (Shindell et al., 2012). The approach of limiting the aerosol impacts on climate based on controlling both the emitted warming compounds (BC) and the co-emitted cooling agents (OC, SO₂, etc.) means that priority should be given to the reduction in or substitution of the emissions characterized by a high BC ratio with respect to the (cooling) co-emitted combustion products, and the approach is able to provide guidelines for the future technological improvements of combustion practices both in the industrialized and the developing countries.

As far as the role of aerosol in climate change is concerned, the RF due to aerosol–radiation interaction (or direct forcings) that is directly linked to aerosol chemical composition and size distribution can be diverse and also work in opposite directions. Nevertheless, the recent evaluation of the aerosol direct forcing uncertainty reported by the IPCC AR5 is more robust than what was previously reported, based on much evidence from models, remotely sensed data and ground-based observations (IPCC, 2013). The quantification of the RF due to aerosol–cloud interaction is still unsatisfactory due to the complexity of the processes involved (Boucher et al., 2013) and requires further advances in modelling and observing systems.

As clouds contribute more than half of the total planetary albedo (Trenberth et al., 2009) and precipitation is the only natural process that brings naturally desalinated sea water to the continents, understanding anthropogenic effects on clouds and rain formation is also critically important. The natural complexity of the problem poses great challenges in modelling and measuring the relevant components, making aerosol–cloud interactions the source of the largest climate uncertainties. The outcome of aerosol–cloud interactions depends on the aerosol size distribution and chemistry and on the cloud regime and environmental conditions. The potential magnitudes of effects and the large uncertainties make aerosol research one of the most important field of climate sciences.

Although rain is the end result of many complex processes within the cloud, it is clear that an increase in CCN number will delay the onset of warm rain formation (Albrecht, 1981. Andreae et al., 2004; Tao et al., 2007). Many high-resolution cloud-resolving models and observational studies that are sensitive to processes on smaller scales suggest various aerosol effects that are not likely to be buffered (Stevens and Feingold, 2009). These effects are complicated, and sensitive to the model microphysical and dynamical schemes as well as to the remote-sensing methods; therefore, they are less understood. For example, a phenomenon defined as aerosol invigoration of convective clouds (Altaratz et al., 2014; Andreae et al., 2004; Koren et al., 2005; Rosenfeld et al., 2013; Tao et al., 2012) suggests a link between the cloud’s vertical and horizontal development and the CCN concentration. The mechanism is based on the fact that when the cloud’s convection is significant, microphysical processes tend to be more closely coupled to dynamical ones. Cloud invigoration, if it exists, might have significant effects on both the radiation and water budget.

This outlook is certainly not comprehensive and other aerosol-related issues could equally be mentioned. In the authors’ opinion, however, the above subjects represent the most pressing challenges to be addressed in the near future within the aerosol science field at large in order to be able to provide policy makers with relevant support to implement wise policies to mitigate air pollution (and human health) and climate change, which are definitely “two sides of the same coin” (Swedish EPA, 2009).

Acknowledgements. This review was prepared as part of the European Commission project ACCENT Plus (Grant FP7 265119). Leone Tarozzi is gratefully acknowledged for his contribution in the final editing of the manuscript.

Edited by: U. Pöschl

References


Berubé, K., Balharry, D. C., Hicks, M., Sexton, K. J., and Jones, T. P.: Inter-omic analysis of particulate air pollution, 10th Annual UK Review Meeting on Outdoor and Indoor Air Pollution Research, 1–2 May 2007, Institute of Environment and Health, Cranfield University, UK, 2007.


S. Fuzzi et al.: Particulate matter, air quality and climate


Kroll, J., Smith, J., Che, D., Kessler, S., Worsnop, D., and Wilson, K.: Measurement of fragmentation and functionalization pathways in the heterogeneous oxidation of oxidized...


S. Fuzzi et al.: Particulate matter, air quality and climate


Miethe, P., Pfaffenberger, L., Dommen, J., Kalberer, M., and Baltensperger, U.: Development of a sensitive long path absorption

www.atmos-chem-phys.net/15/8217/2015/


Pausata, F. S. R., Gaetani, M., Messori, G., Kloster, S., and Dentener, F. J.: The role of aerosol in altering North Atlantic atmospheric circulation in winter and its impact on air quality,


Ryder, J.: Emission, deposition and chemical conversion of atmospheric trace substances in and above vegetation canopies, Ph.D., School for Earth, Atmospheric and Environmental Sciences University of Manchester, Manchester, 241 pp., 2010.


Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Posi-


