Comparison of Measurement-Based Methodologies to Apportion Secondary Organic Carbon (SOC) in PM$_{2.5}$: A Review of Recent Studies

Deepchandra Srivastava$^{1,2,3}$, Olivier Favez$^1$, Emilie Perraudin$^{2,3}$, Eric Villenave$^{2,3}$ and Alexandre Albinet$^{1,*}$

$^1$ INERIS, Parc Technologique Alata, BP 2, 60550 Verneuil-en-Halatte, France; deepchandra.srivastava@gmail.com (D.S.); olivier.favez@ineris.fr (O.F.)

$^2$ CNRS, EPOC, UMR 5805 CNRS, 33405 Talence, France; emilie.perraudin@u-bordeaux.fr (E.P.); eric.villenave@u-bordeaux.fr (E.V.)

$^3$ Université de Bordeaux, EPOC, UMR 5805 CNRS, 33405 Talence, France

* Correspondence: alexandre.albinet@gmail.com or alexandre.albinet@ineris.fr; Tel.: +33-3-44-55-64-85

Received: 21 September 2018; Accepted: 11 November 2018; Published: 16 November 2018

Abstract: Secondary organic aerosol (SOA) is known to account for a major fraction of airborne particulate matter, with significant impacts on air quality and climate at the global scale. Despite the substantial amount of research studies achieved during these last decades, the source apportionment of the SOA fraction remains difficult due to the complexity of the physicochemical processes involved. The selection and use of appropriate approaches are a major challenge for the atmospheric science community. Several methodologies are nowadays available to perform quantitative and/or predictive assessments of the SOA amount and composition. This review summarizes the current knowledge on the most commonly used approaches to evaluate secondary organic carbon (SOC) contents: elemental carbon (EC) tracer method, chemical mass balance (CMB), SOA tracer method, radiocarbon ($^{14}$C) measurement and positive matrix factorization (PMF). The principles, limitations, challenges and good practices of each of these methodologies are discussed in the present article. Based on a comprehensive—although not exhaustive—review of research papers published during the last decade (2006–2016), SOC estimates obtained using these methodologies are also summarized for different regions across the world. Conclusions of some studies which are directly comparing the performances of different methodologies are then specifically discussed. An overall picture of SOC contributions and concentrations obtained worldwide for urban sites under similar conditions (i.e., geographical and seasonal ones) is also proposed here. Finally, further needs to improve SOC apportionment methodologies are also identified and discussed.

Keywords: aerosols; particulate matter; SOA; SOC; source apportionment

1. Introduction and Objectives

Organic matter (OM) constitutes a major fraction, approximately 20–60%, of fine airborne particles [1]. Besides their abundance, the ambient composition of atmospheric particulate organic matter (POM) remains poorly understood due to its chemical complexity and large measurement uncertainties [2,3].

Atmospheric POM has both primary (directly emitted) and secondary (formed in the atmosphere) sources, which can be either natural or anthropogenic. Primary biogenic aerosols include pollen, bacteria, fungal and fern spores, viruses, and fragments of plants [4,5]. Such particles mainly belong to the coarse aerosol fraction and their global emissions on Earth reach up to 1000 Tg year$^{-1}$ [6].
Anthropogenic primary sources include fuel combustion from transportation (road, rail, air and sea), energy production, biomass burning, industrial processes, waste disposal, cooking and agriculture activities [7,8]. Emitted particles are mainly associated with the fine aerosol fraction and a global emission rate of about 50 Tg year$^{-1}$ has been estimated for anthropogenic POM [9,10].

Secondary particles are formed in the atmosphere by gas-particle conversion processes such as nucleation, condensation and heterogeneous chemical reactions [11–13]. The contribution of secondary organic aerosol (SOA) to organic aerosol (OA) reaches up to 80% under certain atmospheric conditions [13]. Most of organic aerosols (OA) in urban and rural atmospheres are speculated to be secondary in nature but their exact chemical composition remains uncertain [12,14,15]. Both biogenic and anthropogenic gaseous emission sources contribute to the SOA production [13,16]. Air quality models still have difficulties reproducing the observed particulate matter (PM) concentration levels due to a poor representation of the OA fractions—primary and notably secondary—reinforcing the need for improving the knowledge on SOA formation processes and on their contribution to the total OA [17–19].

In addition to the difficulties in chemically characterizing POM due to its high complexity and diversity, there is a real challenge to identify relevant criteria to distinguish primary organic carbon (POC) from secondary organic carbon (SOC). A clear information on SOC formation and a right approach to apportion SOC are highly needed to apply efficient air quality strategies.

Several data treatment methodologies have been developed in the last decades to evaluate the contribution of SOC to total OM or PM. Offline methods, usually based on filter measurements, together with emission inventories data, include elemental carbon (EC) tracer approach [20–22], statistical receptor models such as chemical mass balance (CMB) [23] or positive matrix factorization (PMF) [14,24–26], SOA-tracer method [27], radiocarbon ($^{14}$C) measurements [28–30], water soluble organic carbon (WSOC)-based method [31] and regression approaches [32]. Moreover, thanks to recent advances in online aerosol mass spectrometry (AMS) [33,34], real time measurements of aerosol chemical composition have improved the knowledge on OA sources over the last 15 years, and successfully enforced the classification of their primary and secondary origins [35–37].

Among all the methods mentioned above, CMB and EC-tracer methodologies are the most commonly used worldwide. Increases in the use of AMS combined with positive matrix factorization (PMF) data analysis as well as SOA-tracer method, filter PMF approach and $^{14}$C measurements is also noticeable, due to their unique feature to establish direct connections between different SOA fractions and the nature of their precursors and/or their formation processes [30,38–41].

The present paper aims at presenting a comprehensive—although not exhaustive—summary of results obtained worldwide on SOC estimates during the last decade. After a brief synthesis of the current knowledge on SOA precursor emission inventories, the principles, limitations and challenges of each of the five methods mentioned above are discussed. The results obtained from each of these different methodologies are documented for different regions across the world (America, Asia, Europe and Middle East). The studies comparing directly the performances of the different methodologies are reviewed and a comparison of the results obtained worldwide under similar conditions (i.e., geographical and seasonal ones) is also proposed. Finally, good practices to apportion SOC, depending on the methodology used, are suggested along with the discussion on future needs for SOC apportionment approaches.

2. Major Sources of SOA Precursors: Current Knowledge from Emission Inventories

SOA is formed in the atmosphere by oxidation reactions of hydrocarbons leading to the generation of volatile or non-volatile organic compounds involved in gas phase oxidation processes to form new particles either by nucleation or through condensation on pre-existing particles [15,42]. Global SOA production from biogenic volatile organic compounds (BVOCs) ranges from 2.5 to 44.5 Tg year$^{-1}$, whereas the global SOA production from anthropogenic VOCs (AVOCs) ranges from
3 to 25 Tg year\(^{-1}\) [9,43]. The major classes of SOA precursors are volatile and semi volatile-alkanes, alkenes, aromatic hydrocarbons, and oxygenated compounds (Figure 1).

![Distribution of the major classes of secondary organic aerosol (SOA) precursors](image)

**Figure 1.** Distribution of the major classes of secondary organic aerosol (SOA) precursors (adapted from Ziemann and Atkinson, [11]).

Biogenic SOA precursors are mostly alkenes; with ~50% isoprene and ~40% monoterpenes, the rest being other reactive alkenes, such as sesquiterpenes, oxygenated and unidentified VOCs [11]. Isoprene and monoterpenes have always been associated with a major fraction of the total BVOC emissions. Isoprene has the largest global atmospheric emissions of all non-methane VOCs, estimated to be 500 Tg year\(^{-1}\), with a range of 440–600 Tg year\(^{-1}\) [44]. Despite low SOA yield, it contributes to about 4.6 Tg year\(^{-1}\) to SOA mass [43]. The production of SOA from photo-oxidation of terpenes is speculated to make up to 13–24 TgC year\(^{-1}\) from global emission of monoterpenes of about 140 Tg year\(^{-1}\) [45]. Other terpenoid compounds, such as sesquiterpenes, have lower emission than isoprene or monoterpenes, with a global emission of 26 TgC year\(^{-1}\) [46]. However, they may substantially contribute to SOA formation as they are very reactive and show high secondary aerosol formation yields [16]. BVOCs come also largely from oceans, particularly dimethylsulfide (DMS), which is oxidized into methane sulfonic acid aerosol (MSA) [47]. Other identified marine SOA components are dicarboxylic acids [48], dimethyl- and dimethylammonium salts [17,49]. Estimations also suggest that the global production of SOA from marine isoprene is insignificant in comparison to terrestrial sources [50].

Anthropogenic SOA precursor emissions consist of ~40% alkanes, ~10% alkenes, and ~20% aromatics (mostly trimethylbenzenes, xylenes and toluene), the remaining part being oxygenated and unidentified compounds [11]. Most of the anthropogenic SOA is formed from the oxidation of substituted monoaromatic compounds and long-chain alkenes [31,51] emitted from sources such as fossil fuel burning, vehicle emissions, biomass burning, solvent use and evaporation [52–60]. Global emission of aromatic compounds is about 18.8 Tg year\(^{-1}\) [61] and results in an estimated range of SOA production of 2–12 Tg year\(^{-1}\) [61]. The oxidation of evaporated primary OA (POA) vapors
has also been observed as a potential source of SOA in the atmosphere [62–64]. Approximately 16 TgC year\(^{-1}\) (9–23 TgC year\(^{-1}\)) of the traditional POA could remain permanently in the condensed phase, while 19 TgC year\(^{-1}\) (5–33 TgC year\(^{-1}\)) may undergo gas-phase oxidation before re-condensing onto pre-existing particles [17,65]. POA includes compounds with lower volatilities than traditional SOA precursors, such as long chain \(n\)-alkanes, polycyclic aromatic hydrocarbons (PAHs), and large alkenes, which are therefore partitioned in the atmosphere between the gaseous and particulate phases. PAHs have been identified as a major component in emissions from diesel engines and wood burning sources [66,67]. Photooxidation of these compounds in the gas phase has been shown to yield high molecular weight oxygenated compounds [68,69], which may partition with the particle phase and lead to significant SOA formation [70]. PAHs are estimated to yield 3–5 times more SOA than light aromatic compounds and account for up to 54% of the total SOA from oxidation of diesel emissions, representing a potentially large source of urban SOA [71–73]. Other anthropogenic precursors also lead to the formation of SOA, notably phenolic compounds and furans, largely emitted by biomass burning [74,75] and could account substantially to the SOA formed during winter periods.

Finally, on a global scale, BVOC emissions are expected to be one order of magnitude larger than anthropogenic ones [76]. Isoprene has the largest global emission (3–5 times higher than monoterpenes) resulting in dominant SOA production on a global scale. At regional and urban scales, anthropogenic sources are also believed to account for a significant fraction of SOA [9,53,54,77,78] with an order of magnitude similar to biogenic SOA. However, it should be noted that estimations presented in this section remain highly uncertain due to possible biases from model simulations based on data from laboratory oxidation experiments and due to the uncertainties linked to emission inventories for some species. A better understanding of the complex physico-chemical mechanisms involved in SOA formation is still required to better evaluate SOA fluxes. This notably implies further field/laboratory studies and subsequent relevant methodologies to estimate the SOA fraction. Some of these methodologies are described and compared in the following sections.

3. Description of the Main Approaches to Apportion SOC Fraction

This section proposes a comprehensive, but not exhaustive, review on recent applications of the most commonly used methods for SOC estimation from field measurements with a presentation and discussion of their principles, limitations and challenges. The review proposed here concerns the studies reported from 2006 to 2016, focusing on the most recent information available on SOC estimations. Only annual data and those related to the spring–summer period are considered. Data availability and statistical representativeness of all the world regions explain this choice. Besides, due to enhanced biogenic emissions and photo-chemical activities, the spring–summer period is the most favorable season to observe high SOA concentrations.

In the first place and as a first limitation, it should be noted that all offline filter-based methods may suffer from sampling artifacts, leading to overestimation and/or underestimation of atmospheric concentrations of the target compounds. On one hand, sorption of gas species on the filter or formation of secondary compounds by chemical reactions on the collection support between particulate compounds and atmospheric oxidants (\(O_3\), \(NO_x\), \(OH\)) induce an overestimation of particulate phase concentrations (i.e., a positive artifact). On the other hand, volatilization of particulate compounds collected on the filter or chemical degradation due to reactions between collected compounds and atmospheric oxidants lead to an underestimation of particulate phase concentrations (i.e., a negative artifact). Such artifacts are highly dependent on temperature, compound vapor pressures and sampling flow rates [2,79–86].

3.1. EC-Tracer Method

3.1.1. Principle

The EC-tracer method has been an extensively used approach since the 1980s [20–22,87–93].
The main advantage of this method is that it uses only ambient measurements of organic carbon (OC) and EC, which are readily available. Since EC and primary OC are mostly emitted by the same combustion sources (either modern or fossil fuel), EC can be used as a tracer for primary combustion generated OC \[ 21,22,94\]. The ratio of ambient concentrations of particulate OC to EC includes information about the extent of SOC formation. Ambient \([OC/EC]\) ratios larger than those specific to primary emissions illustrate SOA formation. In this method, OC primary can be expressed according to Equation (1).

\[
[OC]_p = \left[ \frac{OC}{EC} \right]_p [EC] + [OC]_{\text{non-comb.}}
\]  

(1)

The SOC fraction can be estimated using the following Equation (2).

\[
[OC]_s = [OC] - [OC]_p
\]  

(2)

where \([OC]\) is the measured total OC concentration, \([OC]_p\) is the POC concentration, \([OC/EC]_p\) represents the ratio of OC to EC concentrations for the primary sources affecting the site of interest, and \([OC]_{\text{non-comb.}}\) is the non-combustion contribution to the POC \[21,94,95\], \([EC]\) is the measured EC concentration, and \([OC]_s\) is the SOA contribution to the total OC. Sources of \([OC]_{\text{non-comb.}}\) include soil and road re-suspended PM, biogenic sources (i.e., plant detritus, resuspension of other biogenic material), etc. \[89,95,96\]. Cooking activities are generally considered as a non-combustion source \[87\]. All of these parameters are time-dependent meaning they may be substantially influenced by meteorological conditions and emission scenarios \[22,96\]. Details regarding the calculation of \([OC/EC]_p\) and \([OC]_{\text{non-comb.}}\) are provided in the supplementary material (SM).

3.1.2. Limitations and Challenges

As detailed in the supplementary material (SM), ambient \([OC/EC]\) ratios substantially fluctuate with time and locations (Table SA1, SM) and the EC-tracer method may suffer from major issues linked to the choice of constant values used in Equation (1). The assumption of constant primary \([OC/EC]_p\) and \([OC]_{\text{non-comb.}}\) values, assumed to be representative of the period of the study, may not always be fully relevant. In particular, day to day \([OC/EC]_p\) values are function of the nature of the emission sources, the meteorological conditions and the influence of atmospheric pollutant transport, leading to significant uncertainties \[94,95\]. Therefore, a constant ratio might not be appropriate for the application of the EC-tracer method on a long-term basis, such as yearly timescale \[97,98\]. In addition, the application of the EC tracer method is not straightforward to data collected during cold periods. In this last case, datasets must be thoroughly examined to determine the days when secondary formation of particulate OC is expected to be negligible. Usually, parameters used for the determination of “primary emission predominance” conditions are low solar radiation, low temperature and low \(O_3\) concentration levels, and/or the occurrence of high NO and low NO\(_2\) concentrations. Based on these criteria, Lonati et al. \[97\] obtained a \([OC/EC]_p\) ratio of 9.5, for the subset of data collected in Milan (Italy) during several cold seasons. However, this value seemed too high to be assumed as representative of primary ratio compared to previously reported values in several other studies \[21,94,99\]. This result suggested that the selection of a constant \([OC/EC]_p\) value may not be adequate under all meteorological conditions. Besides \([OC/EC]_p\), the estimation of \([OC]_{\text{non-comb.}}\) is another important parameter in the application of the EC tracer method. \([OC]_{\text{non-comb.}}\) is usually assumed to be small \[88\] or negligible \[91,100\] and often estimated by the intercept of the regression line of Equation (1). This leads to artificially higher values of SOC especially for the smaller values of EC \[89\]. Emission inventories can also be used to estimate both \([OC/EC]_p\) and \([OC]_{\text{non-comb.}}\) parameters, but the accuracy of this approach, especially for \([OC]_{\text{non-comb.}}\) is questionable.

Under the significant influence of local sources (e.g., wood combustion), presenting higher OC and lower EC emission rates, higher values of \([OC/EC]\), not necessarily due to the existence of SOC derived from photochemical reactions, may be observed \[101\]. Consequently, qualitative estimation
of SOC using [OC/EC] ratios should be applied only after a careful inspection of local sources of OC and EC. Moreover, the presence of a significant fraction of semi-VOCs (SVOCs) in the aerosol content could induce significant variations of the [OC/EC] ratio, depending on the change in ambient air temperature [87]. For instance, an increase of temperature from winter to summer would result in a decrease of the minimum [OC/EC] ratio due to the evaporation of primary SVOCs at higher temperatures in summer.

Another issue arises from the EC/OC analysis. The most commonly used thermal protocols are from NIOSH (National Institute for Occupational Safety and Health), IMPROVE (Interagency Monitoring of PROtected Visual Environment) and EUSAAR 2 (European Supersites for Atmospheric Aerosol Research) [102–105]. They differ mostly in their temperature programs and optical correction types for charring based on transmittance or reflectance. The three protocols are comparable for total carbon (TC) concentrations but the results may significantly vary concerning EC–OC split [106–108]. Moreover, depending on the protocol used, very low EC loading can be difficult to measure, for instance in the case of samples from remote locations.

In addition to the above limitations, it should also be noted that a variety of linear regression techniques and simple slope estimators can also present a considerable variation in the [OC/EC] ratio. For instance, significant differences have been witnessed in SOC estimates made for several Mexican cities in different studies [109–111], as well as other locations around the world, and the selection of the approach to calculate [OC/EC] ratio is probably one of the main reasons explaining the differences observed. Details on the use of different regression techniques and associated issues are provided in the SM.

In general, several authors indicated a [OC/EC]_p ratio of approximately 2, used then as a threshold for interpreting ratios exceeding this value as an indicator of the presence of SOA [21,22,94,99,112]. Higher [OC/EC]_p ratios may be due to the different approaches adopted to determine the dominant primary emission period by taking the above-mentioned conditions into account though there is no way to avoid the contribution of secondary formation processes. To account for the limitations of the EC-tracer method, several authors proposed to estimate the method associated uncertainties considering the EC/OC measurement and the assumptions inherent to the EC tracer method itself. Lim and Turpin [91] suggested a value of ±10% of uncertainty in the SOC estimation while Pachon et al. [93] estimated uncertainties of about of 80% on SOC values in winter and 47% in summer.

All of these results further suggest the strong need of a standard procedure to select primary [OC/EC] ratio and [OC]_{non-comb}.

3.1.3. Review of Recent Studies Based on the EC-Tracer Method

Figure SA1 (SM) synthesizes the locations of the studies discussed in this section. Detailed references about these results are given in Tables SA2 and SA3 in the SM. It is important to note that no (or very few) data are displayed here for Africa, Oceania, Central Asia, Russia, Central and South America. This does not imply that no EC tracer studies have been conducted so far at these places but means that they globally fall outside the specific criteria of this review (PM$_{2.5}$ OC fraction and studies from 2006 to 2016).

3.1.3.1. America

Seasonal and regional variations of SOA have been examined thoroughly over the North American continent using the EC-tracer method [1,38,89,92,93,109,113–121]. Overall, annual SOC levels ranged from 1.1 to 2.7 µgC m$^{-3}$, contributing to 27–70% of PM$_{2.5}$ OC (Figure 2; Table SA2 (SM)). The highest contributions (~63% on average) were obtained for rural locations, while contributions were in the range between 30% to 50% at urban sites. SOC estimates in warm periods in the US obtained from the literature showed that 29–43% (1.0–1.8 µgC m$^{-3}$) of PM$_{2.5}$ OC was secondary, with the highest contribution (43%, 1.2 µgC m$^{-3}$) at a rural location (Centreville) (Figure 3, Table SA3 (SM)).
Figure 2. Annual secondary organic carbon (SOC) (green) and primary organic carbon (POC) (grey) contributions to PM$_{2.5}$ organic carbon (OC) estimated using the elemental carbon (EC) tracer method for all the monitored sites from 2006 to 2016. Results are presented in increasing order of OC concentration levels (µgC m$^{-3}$). In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites.

Few examples of SOC estimations are also available in Central and South America [109,118–121] (Figure 2). The annual average SOC contribution was approximately 57% and reached up to 87% of PM$_{2.5}$ OC in Mexico [109] (Figure 2, Table SA2 (SM)). Some other previous studies showed lower SOA contributions [110,111], though the estimations were not carried out using the same
approach. In Santiago (Chile), no substantial differences have been observed between the annual and spring–summer SOC contributions (annual: \(29 \pm 6\%\) and spring-summer: \(31 \pm 6\%\) [121]. These results reflect that SOC may be a significant contributor to fine OC through the year as well as in the warm period due to favorable meteorological conditions.

### Figure 3.

Spring–summer SOC (green) and POC (grey) contributions to PM\(_{2.5}\) OC estimated using the EC tracer method for all the monitored sites from 2006 to 2016. Results are presented in increasing order of OC concentration levels (\(\mu\text{gC m}^{-3}\)). In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites.

#### 3.1.3.2. Europe and the Middle East

Similarly to the North American continent, SOC estimates using the EC tracer method have been extensively performed throughout Europe [96,97,122–134]. The annual and spring–summer

<table>
<thead>
<tr>
<th>Site Description</th>
<th>SOC Contribution (%)</th>
<th>POC Contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mount Tai (China)</td>
<td>1.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Mount Abu (India)</td>
<td>1.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Athens (Greece)</td>
<td>2.0</td>
<td>9.4</td>
</tr>
<tr>
<td>Ponghu (Taiwan)</td>
<td>2.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Veneto region (Italy)</td>
<td>2.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Birmingham (UK)</td>
<td>2.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Brisbane (Australia)</td>
<td>2.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Birmingham (UK)</td>
<td>2.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Condell (US)</td>
<td>2.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Zizy Potok (Poland)</td>
<td>2.6</td>
<td>9.4</td>
</tr>
<tr>
<td>San Pietro Capofiume (Italy)</td>
<td>3.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Mount Heng (China)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Bologna (Italy)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Chiba Prefecture (Japan)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Krakow (Poland)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Pittsburgh (USA)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Madrid (Spain)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Xuhan (Taiwan)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Cape Fugue (Taiwan)</td>
<td>3.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Gogou (Korea)</td>
<td>4.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Atlantic (USA)</td>
<td>4.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Yokohama (Japan)</td>
<td>4.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Hong Kong (China)</td>
<td>4.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Birmingham (USA)</td>
<td>4.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Qingdao (China)</td>
<td>5.0</td>
<td>9.4</td>
</tr>
<tr>
<td>Zhaoqing (China)</td>
<td>5.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Milan (Italy)</td>
<td>5.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Thessaloniki (Greece)</td>
<td>5.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Back Garden (China)</td>
<td>5.7</td>
<td>9.4</td>
</tr>
<tr>
<td>Seoul (Korea)</td>
<td>6.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Incheon (Korea)</td>
<td>6.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Tainan (Taiwan)</td>
<td>6.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Taipei (Taiwan)</td>
<td>7.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Yellow river delta (China)</td>
<td>7.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Wangdangsha (China)</td>
<td>7.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Monterey (Mexico)</td>
<td>7.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Nanjing (China)</td>
<td>8.0</td>
<td>9.4</td>
</tr>
<tr>
<td>Jinchang (China)</td>
<td>8.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Taichung (Taiwan)</td>
<td>8.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Ahmedabad (India)</td>
<td>8.7</td>
<td>9.4</td>
</tr>
<tr>
<td>Shanghai (China)</td>
<td>8.8</td>
<td>9.4</td>
</tr>
<tr>
<td>Fuzhou (China)</td>
<td>9.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Xiamen (China)</td>
<td>9.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Shandong (China)</td>
<td>9.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Quanzhou City (China)</td>
<td>9.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Pinglung (Taiwan)</td>
<td>9.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Ganzhou (Korea)</td>
<td>10.4</td>
<td>9.4</td>
</tr>
<tr>
<td>Guangzhou (China)</td>
<td>11.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Beijing (China)</td>
<td>11.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Shenzhen (China)</td>
<td>11.9</td>
<td>9.4</td>
</tr>
<tr>
<td>Chengchun (China)</td>
<td>12.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Taishan (China)</td>
<td>12.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Puyan City (China)</td>
<td>13.0</td>
<td>9.4</td>
</tr>
<tr>
<td>Yulin (China)</td>
<td>13.7</td>
<td>9.4</td>
</tr>
<tr>
<td>Wuhan (China)</td>
<td>14.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Tianjin (China)</td>
<td>14.3</td>
<td>9.4</td>
</tr>
<tr>
<td>Santiago (Chile)</td>
<td>14.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Neiliang (China)</td>
<td>14.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Chongqing (China)</td>
<td>15.0</td>
<td>9.4</td>
</tr>
<tr>
<td>Hangzhou (China)</td>
<td>17.1</td>
<td>9.4</td>
</tr>
<tr>
<td>Chengdu (China)</td>
<td>17.2</td>
<td>9.4</td>
</tr>
<tr>
<td>New Delhi (India)</td>
<td>17.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Pune (India)</td>
<td>17.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Xian (China)</td>
<td>27.5</td>
<td>9.4</td>
</tr>
</tbody>
</table>

(%)
average of SOC levels varied in the ranges 1.6–4.9 µgC m⁻³ and 0.90–4.4 µgC m⁻³, contributing to 36–73% and 41–84% of PM₂.⁵ OC, respectively (Figures 2 and 3, Tables SA2 and SA3). As expected, SOC contributed in warmer periods at higher rates to PM₂.⁵ OC than annually at all locations (61% on average), with the highest contribution observed in Milan (84%, 4.4 µgC m⁻³). Surprisingly, the annual and spring-summer SOC contributions to PM₂.⁵ OC at both urban and rural background sites in Birmingham did not show the expected pattern of higher contribution in summer than in winter [123,133]. This has been also spotted in other studies from other continents where the data obtained did not show strong SOC seasonality (i.e., Pittsburgh, PA, USA [113]). These results suggest that a lower mixing layer height in winter favors SOC precursor stagnation, subsequently leading to the SOC formation and notably anthropogenic SOA [73]. Processes that can also contribute to SOC formation are the adsorption of semi-volatile OCs onto existing solid particles and the dissolution of soluble gases that may undergo reactions on particles [135,136].

The number of monitoring studies are limited in the Middle East region. Estimations of SOC contribution at 11 sites in Palestine (Nablus, East Jerusalem and Hebron), Jordan (Amman, Aqaba, Rahma and Zarka) and Israel (West Jerusalem, Eilat, Tel Aviv and Haifa) have been recently reported [137]. The average SOC levels at all sites, excluding East Jerusalem and Nablus, were about 2.7 µgC m⁻³, corresponding to a contribution of 55% to the PM₂.⁵ OC. Results suggested that the significant formation of SOA at urban locations was due to the gas/particle conversion of gaseous hydrocarbon precursors and reinforced the importance of identifying SOA precursors for effective reduction of aerosol loadings [137].

3.1.3.3. Asia

The environmental behavior of particulate bound SOC has been investigated in Asia in places such as Korea, Japan, Taiwan, China and India [138–144] (Figures 2 and 3; Tables SA2 and SA3).

The annual and spring–summer SOC levels at urban locations of Korea varied in the ranges 1.1–4.6 µgC m⁻³ to 1.1–5.7 µgC m⁻³, contributing to 15–57% and to 45–83% of PM₂.⁵ OC, respectively. In the case of Japan and Taiwan, the annual and spring–summer SOC, at urban and suburban locations, accounted for >35% of PM₂.⁵ OC. Carbonaceous aerosol in China has drawn special attention in recent years due to the very high PM and SOA concentration levels observed notably during haze events [145–151]. All in all, the annual and spring–summer SOC levels varied in the ranges 2.5–23.9 µgC m⁻³ to 1.8–12.2 µgC m⁻³, contributing to 25–63% and 21–67% of PM₂.⁵ OC, respectively [152–179] (Figures 2 and 3). SOC estimations were especially performed in the Sichuan basin, one of the most populated regions in China (about 100 million people) [180]. The annual average SOC levels were 5.1, 4.5 and 11.4 µgC m⁻³ in Chengdu, Neijiang and Chongqing, respectively, contributing to 27%, 25% and 34% of PM₂.⁵ OC levels. The SOC fractions presented in these studies were very homogeneous and significantly lower than in previous studies using the EC tracer method, with e.g., about 40% reported by Cao et al. [164] and 57% by Zhang et al. [170] in several urban regions of China. The rapid urbanization and industrialization of the region had substantially modified the emissions in the ambient air inducing a difference in the observed OC/EC ratios and preventing the availability of a stable [OC/EC]p in the EC tracer method. Besides, to understand the secondary processes in rural and mountainous areas of South and North China, measurements were also conducted at Mount Heng and Mount Tai [181,182]. High contributions of SOC to PM₂.⁵ OC (61–67%) were observed in spring–summer period indicating the influence of long-range transport of carbonaceous aerosol from the PRD (Pearl River Delta) and Eastern China; highly urbanized and industrialized regions. At both locations, the occurrence of in-cloud SOA formation and the role of heterogeneous chemistry was highlighted.

As in China, air quality monitoring in India is being undertaken more rigorously than ever due to the very high PM concentration levels observed [183–190]. The annual SOC levels varied in the range 6.0–26.4 µgC m⁻³, contributing to 42–58% of PM₂.⁵ OC in major urban cities such as New Delhi, Gurgaon, Ahmedabad, Pune and Mumbai (Figure 2; Table SA2 (SM)). Significant seasonal
patterns have been observed at all of these locations with the highest SOC contribution noticed at Pune in summer (70%). Results were consistent with another study conducted at Pune by Pipal and Gursumeeran Satsangi [186], which showed higher effective carbon ratio values (ECR, defined as the ratio of SOC to the sum of POC and EC) in summer, indicating the larger formation of SOC in the warm period [187]. At a high-altitude site (Mount Abu) in western India, studies have shown very low SOC concentrations (0.8 µgC m$^{-3}$) with a contribution of about 51% to PM$_{2.5}$ OC [183].

3.2. Chemical Mass Balance (CMB)

3.2.1. Introduction

In CMB, as in other source–receptor models, ambient concentrations of chemical species are expressed as the sum of the products of source compositions and contributions [23]. CMB is based on an effective variance least square approach to establish a balance between the source and the receptor site and finally to estimate the source contributions [191–193]. Fingerprints of the source emissions (source profiles) are used to calculate the atmospheric concentrations of the chemical species $i$ at the receptor site $k$, $C_{ik}$, as follows (Equation (3)):

$$C_{ik} = \sum_j a_{ij}S_{jk}$$

where, $a_{ij}$ is the relative concentration of the chemical species $i$ in the OC emissions from the source $j$, and $S_{jk}$ is the contribution to OC at the receptor site $k$ originating from the source $j$. The source profile abundances (i.e., the mass fraction of a chemical species in the emissions from each source type) and the receptor concentrations, with appropriate uncertainty estimates, are the input data of the model. The conservative nature of the chemical species with no significant removal through dry and wet depositions and/or degradation/formation by chemical reactions over time during the transport from the source to the receptor is one of the major assumptions considered in this model [194]. Another assumption is the non-colinear nature of the source profiles [192]. In addition, the number of species must be larger than the number of sources to produce significant results.

The CMB applied to OC using molecular markers (Table SB1 (SM)) usually only considers primary sources because the determination of profiles for secondary sources is difficult to be obtained. Thus, the OC not apportioned (un-apportioned OC) refers to SOC. SOC is then defined as the difference between the measured OC concentration and the aggregated OC concentration from all primary sources resolved by CMB. The OC concentration obtained from all primary sources is referred as source contribution estimates (SCEs) (Equation (4)):

$$SOC = \text{Measured OC} - \sum \text{SCEs (Primary sources)}$$

CMB is successfully applied if source profiles consistent with the measurements performed at the receptor locations are used. Source profiles are the mass abundances of the chemical species in source emissions and are regarded as a category of sources rather than individual emitters [192,195]. Relevant source profiles should be used as an input to evaluate correct SCEs, otherwise it can lead to ambiguity in the obtained results and could be considered as a major limitation of the method.

3.2.2. Limitations and Challenges

Several factors make the use of CMB analysis strenuous. OC is not necessarily completely fitted by the model as markers and source profiles for SOC do not exist in all cases.

The selection of appropriate source profiles is one of the critical steps to obtain a good fit between the CMB model results and PM total concentrations. Such profiles are generated using emission samples from a range of emitters of a particular source category which are analyzed to determine their chemical composition and identify specific molecular source markers [192] (Table SB1, SM). They are used for the identification and the quantification of the contributions of the different sources
to PM. To a large extent, the CMB model results rely on the accuracy of the source profiles used as an input. However, in the absence of locally relevant source profiles, the SCEs can be prone to produce erroneous results. While the typical components of any source profiles are found to be more-or-less similar, the relative mass abundances vary with specific feature locations and emitter characteristics. Thus, the use of different combinations of source profiles in CMB can provide statistically valid but completely different solutions. Generation of SOC source profiles is difficult due to the complexity of the chemistry describing SOA formation and the diversity of its composition [196]. Therefore, the major challenge for the real SOC estimation from CMB is the lack of source profiles for the secondary sources [197,198].

As already explained, the difference between the measured OC and the sum of all apportioned primary sources is usually attributed to SOC. However, this approach presents a major issue when unknown primary sources contribute significantly to OC or are missed out in the available source profiles. In that case, SOC estimation using CMB could be overestimated or underestimated.

As specified in the preamble of this section, sampling artifacts could also influence the results obtained using CMB, although El Haddad et al. [41] showed that such artifacts appeared to marginally influence the amount of un-apportioned OC.

Another important issue is the atmospheric stability of the molecular markers used. Currently, all the source–receptor models used in the source apportionment studies assume that marker compounds are chemically stable in the atmosphere (and therefore called tracer compounds) [193]. However, their photodegradation may occur in the atmosphere, causing an underestimation of the contributions of sources, especially in summer [62]. As an example, levoglucosan has been used as the specific molecular marker of biomass burning aerosol for a long time, based on its high emission factors and on its assumed chemical stability [199]. Studies showed that significant atmospheric chemical degradation of levoglucosan could occur on a timescale similar to that of atmospheric transport and deposition [200–203]. This could induce an underestimation in the contribution of the biomass burning source in the source apportionment results whatever the source–receptor model used.

3.2.3. Review of Recent Studies Based on CMB Approach

Figure SB1 (SM) shows a summary of the application of CMB in source apportionment studies performed worldwide. Detailed references about all the results considered here are presented in Tables SB2 and SB3. Note that, even if many studies are reported, no data or very few are proposed for Africa, Oceania, East and Southeast Asia, Russia, Central and South America, as they did not satisfy the specific criteria of this review (PM$_{2.5}$ fraction and studies from 2006 to 2016).

3.2.3.1. North America

CMB has been extensively used in the US to apportion the SOC fraction in PM$_{2.5}$ [93,111,204–215].

For all the different sampling type locations studied, annual SOC contributions presented an extremely wide range of values with SOC contributions to PM$_{2.5}$ OC from 2 to 67% and concentrations from 0.0 to 4.6 µgC m$^{-3}$ (Figure 4, Table SB2 (SM)). All the urban locations showed SOC contributions $>$33% while rural and remote sites could show either very low SOC contributions like in Northern and Southern Minnesota (2%) or quite high contributions like in Texas and Centreville (63%) with both, low and high OC concentration levels (from 0.8 to 6.3 µgC m$^{-3}$). Authors showed that Minnesota is often downwind of major Midwest regional sources (from Illinois, Wisconsin, Iowa and Missouri) in summer and could be also influenced by local coal-fired electricity generation emissions explaining probably the low SOC amounts observed [204].

During the warm period, the observed SOC contributions and concentrations at all urban locations ranged from 21% to 79% and 0.7 to 2.8 µgC m$^{-3}$, respectively (Figure 5, Table SB3 (SM)). Biogenic SOA, from isoprene and pinene precursors, accounted as an important SOA source in the eastern US where up to 80% of land is covered by forests for instance in the south [216–220]. In the summer
period, when abundant biogenic VOC emissions exist, favorable atmospheric conditions lead to high SOA formation.

Figure 4. Annual SOC (green) and POC (grey) contributions to PM$_{2.5}$ OC using chemical mass balance (CMB) for all the monitored sites from 2006 to 2016. Results are presented in increasing order of OC concentration levels ($\mu$gC m$^{-3}$). In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites.
Figure 5. Spring-summer SOC (green) and POC (grey) contributions to PM$_{2.5}$ OC using CMB for all the monitored sites from 2006 to 2016. Results are presented in increasing order of OC concentration levels (µg C m$^{-3}$). In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites.

3.2.3.2. Europe and the Middle East

There are few CMB studies that apportion the SOC fraction in PM$_{2.5}$ in Europe [41,124,221–226]. The average annual SOC contributions to PM$_{2.5}$ OC ranged from 10% to 44% in London, Birmingham and Milan (urban and rural sites) and were in the concentration ranges of 0.7 to 1.2 µg C m$^{-3}$ [124,162,226] (Figure 4, Table SB2 (SM)). Higher CMB SOC estimates were observed in the warm period, i.e., 33–78% (1.4–3.7 µg C m$^{-3}$) with the highest contribution (78%) at Marseille (France) and the lowest (33%) at Milan (Italy) [41,223] (Figure 5, Table SB3 (SM)). Interestingly, at Summit (Greenland, Denmark), a remote site far from any biogenic or anthropogenic sources and activities, the SOC contribution reported was about 95%. These results highlight the long range transport and SOA formation until high latitude regions [227].

The Middle East also witnesses a lack of studies to estimate the SOC fraction in PM$_{2.5}$ using CMB [228–230]. The annual SOC concentrations reported in the literature were in the range of 1.4–5.5 µg C m$^{-3}$, contributing to about 42–67% of PM$_{2.5}$ OC (1.8–5.5 µg C m$^{-3}$) for sites located in Jordan, Palestine, Iraq and Israel (Figure 4, Table SB2 (SM)). For the spring–summer period, data are only available for West Jerusalem (Israel) and East Jerusalem (Palestine) with similar SOC contributions and concentrations to the annual average i.e., 39% (1.9 µg C m$^{-3}$) and 48% (2.5 µg C m$^{-3}$) of PM$_{2.5}$ OC, respectively (Figure 5, Table SB3 (SM)). These results showed that SOA formation in the Middle East is significant in the region year-round. The intensity of the sunlight and high temperatures might result in high reaction rates of SOA precursors and, in combination with drier conditions (lower wet deposition), would induce a larger SOA formation [229].

3.2.3.3. Asia

In the Asian continent, several studies have been performed during the last decade and notably in China [231–244]. The annual and spring–summer SOC concentrations in Asia were in the ranges 0.7–7.7 µg C m$^{-3}$ and 0.8–4.8 µg C m$^{-3}$, respectively, accounting for 12–67% and 11–80% of SOC in PM$_{2.5}$ OC, respectively (Figures 4 and 5; Tables SB2 and SB3).
Only two sites have been investigated in Central Asia at Bishkek and Teploklyuchenka, in Kyrgyzstan where large SOC contributions (>72%) were observed in the summer period and at both sites (urban and rural). In India (Kanpur, Agra), Pakistan (Lahore) and Korea (Gwangju), the annual SOC concentrations were in the range 1.9–7.7 µgC m\(^{-3}\) corresponding to SOC contributions of about 32\% [238–240]. In China, the annual SOC concentrations were in the range 1.2–3.3 µgC m\(^{-3}\) and contributions about 14\% to 50\% of PM\(_{2.5}\) OC (Figure 4). Interestingly, SOC concentrations at the rural site Hok Tsui in Hong Kong showed a SOC contribution larger than 50\% throughout the year [244]. As expected during the warm period, the SOC concentrations and contributions were larger (4.8 µgC m\(^{-3}\) and >50\% of PM\(_{2.5}\) OC) (Figure 5).

3.3. SOA-Tracer Method

3.3.1. Introduction

The SOA-tracer method was developed by Kleindienst et al. [27] to estimate the SOA contributions from several biogenic and anthropogenic hydrocarbon SOA precursors to ambient OC concentrations using a series of organic molecular compounds called tracer (or marker) compounds (see also Section 3.2.2). For each hydrocarbon precursor, SOA tracer compounds were first identified and measured during irradiation experiments, performed in a smog chamber in the presence of NO\(_x\). The SOA mass fraction, \(f_{SOA, hc}\), defined as the ratio of the sum of the organic tracer concentrations to the mass concentration of aerosol formed in the smog chamber, is equal to the total SOA concentration in this case (Equation (5)).

\[
f_{SOA, hc} = \frac{\sum_i [tr_i]}{SOA}
\]  

with \([tr_i]_m\), the mass concentration of the tracer \(i\) in µg m\(^{-3}\). The SOA mass fraction is obtained using gravimetric measurements to convert the SOA mass fractions into SOC mass fractions using SOA/SOC mass ratios (Equation (6)).

\[
f_{SOC, hc} = f_{SOA, hc} \frac{SOA}{SOC}
\]

Laboratory experiments were conducted to determine SOC mass fractions for isoprene, \(\alpha\)-pinene, \(\beta\)-caryophyllene, and toluene. Details of the laboratory generated SOC mass fractions are provided in SM (Table SC1). All the other details about the smog chamber experiments are well described in the paper by Kleindienst et al. [27].

3.3.2. Limitations and Challenges

The main advantage of the SOA-tracer approach lies in the direct attribution of SOC concentrations and contributions to specific gaseous organic precursors. However, the method also suffers from several limitations.

Mass fractions, calculated considering the sum of tracer compounds, have been determined from single hydrocarbon smog chamber irradiations under varying SOA precursor and NO\(_x\) concentrations. Due to the complexity of atmospheric photo-oxidation chemical mechanisms, the wide range of organic and inorganic compounds present in the atmosphere, and the myriad of possible actual meteorological and photochemical conditions, considerable uncertainty may be associated with the use of a single-value mass fraction for each precursor. The main systematic uncertainty is probably the representativeness of smog chamber processes compared to those occurring in the atmosphere with differences in relative humidity, precursor and particulate matter concentrations, nature and oxidant concentrations, and irradiation conditions. Thus, the mass fractions derived from chamber experiments may be different from ambient air conditions. This has already been noticed when the SOA tracer method was applied for the first time to ambient air samples collected at Research Triangle Park (RTP), North Carolina [27]. Uncertainties in the results were estimated to about 25\% for isoprene, 48\% for \(\alpha\)-pinene, 22\% for \(\beta\)-caryophyllene, and 33\% for the toluene SOA mass fractions, respectively.
As defined, a tracer should be stable in the atmosphere [199]. However, as already specified above in Section 3.2.2, one of the most important limitations is related to the stability of these molecules in the atmosphere and, in this case, the term marker would be more appropriate. The atmospheric lifetimes of SOA tracers (markers) have been theoretically estimated based on their volatility [42]. The exact values are only available for few of them (e.g., cis-pinonic acid ~2.1–3.3 days and MBTCA (3-methyl-1,2,3-butanetricarboxylic acid) ~1.2 days) [245,246].

Molecular tracer species are not necessarily produced from single precursors but could originate from the oxidation of many other molecules. For instance, α-pinene SOA tracer compounds have also been observed in the laboratory as by-products of β-pinene and d-limonene oxidation [247]. Thus, the estimation of α-pinene derived SOC contribution may also contain some contributions from other monoterpenes. Similarly, Kleindienst et al. [27] also showed that toluene SOA tracer could also be formed from the photo-oxidation of xylene.

So far, the main limitation of the method to properly evaluate total SOC is probably the limited number of SOA tracers identified for specific known gaseous organic precursors and the few SOA/SOC data available in the literature. For instance, toluene is the only anthropogenic VOC considered in the SOA tracer method. Laboratory studies have shown considerable SOA yields from several aromatic VOCs, such as xylenes, ethyl-benzene, ethyl-toluene, trimethyl-benzenes, benzene [51,248,249]. Currently, there is a lack of SOA tracer compounds for these precursors [75]. In addition, recent research showed that semi-volatile (SVOCs) and intermediate-volatility organic compounds (IVOCs) including cyclic, linear, branched alkanes, PAHs etc., are important classes of SOA precursors [56,62,72,250–256]. SVOCs and IVOCs are abundant in gasoline- and diesel-powered vehicle exhausts as well as other anthropogenic sources (e.g., biomass burning), but little is known about their molecular composition and no SOA markers with proper SOC mass fractions have been reported for such classes of compounds.

Phenol, cresols, furans and methoxy-phenols account for a significant fraction of pollutants emitted by biomass burning. The oxidation of these compounds is likely to form SOA, contributing significantly to OA loadings in the atmosphere [74,75,257]. Therefore, neglecting SOA derived from biomass burning would also underestimate the SOC fraction, especially in winter when this source of energy is largely used.

Besides, the presence of organosulfates, a class of organic compounds reported to be formed from the oxidation of aromatic and polyaromatic compounds, isoprene, and monoterpenes in the presence of acidic sulfate seed particles [254,258,259], could account for a significant fraction of total SOC but they are not yet included in the SOA tracer method. Similarly, no marker compounds have been identified to account for the organonitrate fraction that constitutes a significant part of OA in urban environments [260].

In addition, the tracer-based approach does not consider SOA formed through aqueous phase processes. Similar to the reactions in the gaseous phase, soluble organic compounds in the aqueous phase (cloud, fog mist and rain) can react with OH radicals and/or by direct photolysis. Enhanced production of acids, such as oxalic acid, and oligomers has been observed [261–265]. The low volatile species produced may remain in the particle phase after water evaporation, leading to the formation of new SOA called aqueous SOA (aqSOA) [266,267]. Multiphase photooxidation chamber experiments have shown that the presence of a liquid water cloud lead to a faster and higher SOA formation than under dry conditions. For instance, in the case of isoprene and methacrolein, only a single cloud evaporation–condensation cycle under irradiation led to 2–4 times higher SOA mass yield than those observed for photooxidation experiments carried out under dry conditions [268–270]. Finally, in-cloud processing of isoprene contributes to at least 1.6 Tg year$^{-1}$ for a global biogenic SOA production of 8–40 Tg year$^{-1}$ [271]. However, SOA mass fraction from such formation pathways has never been evaluated from any already-studied precursor. Thus, the actual SOA production in the real atmosphere could be higher than that estimated using mass fractions proposed by Kleindienst et al. [27] and used in the SOA-tracer method.
Another issue is related to the quantification of these markers in PM samples. Most SOA tracers are of well-known atmospheric biogenic and anthropogenic VOC precursors such as isoprene, monoterpenes, sesquiterpenes, and toluene are not commercially available. In that scenario, proxy compounds have been used frequently [27,38,272]. The lack of authentic standards critically implies a bias in the quantification of these species. Hu et al. [272] noted that the quantification uncertainty caused by using different surrogates other than ketopinic acid (a surrogate for all tracer compounds to derive mass fractions of SOA tracers in the SOA tracer method by Kleindienst et al. [27]) was estimated to be within a factor of three. The use of different surrogate standards could be another reason for observing large uncertainties in the SOC estimation. Several research groups have resorted to synthesize these compounds. However, the synthesis remains a quite expensive and/or time-consuming task. Only very recently, a few papers have reported the use of authentic standards that are commercially available or synthesized on purpose by worldwide suppliers [73,273].

Finally, compared to other available methods, the SOA tracer method is presently the only approach which procures individual SOA contributions from different precursors. However, further laboratory and field evaluations of the method are needed to improve the SOC contributions by considering a large variety of SOA precursors.

3.3.3. Review of Recent Studies Based on the SOA-Tracer Method

The SOA tracer approach has been applied in many studies across the world including several sites in the US [27,38,78,197,274–279], the Canadian Arctic [280–282], Europe [39,41,283,284], and Asia [28,40,64,178,179,198,272,285–292] (Figure 6).

Figure 6. Spring–summer SOC contributions to PM$_{2.5}$ OC over the world for all the selected sites from 2006 to 2016 using the SOA tracer method. OC concentrations (µgC m$^{-3}$) for each site are indicated into brackets. Box-plots represent the SOC concentrations (µgC m$^{-3}$) for USA and China. Box plots include minimum (10th), first quartile (25th), median (50th), third quartile (75th) and maximum values (90th). The number of data points considered is presented into brackets on the box-plot graphs. α Toluene SOC was not reported at these locations. β β-caryophyllene SOC was not reported at these locations. γ Toluene SOC and β-caryophyllene SOC were not reported at these locations. δ Only isoprene SOC was reported at these locations. Details about the unaccounted SOC is given in the main text (Section 3.3.3).
Note that the annual SOC estimates are only available for a few sites in the US and China. The analysis and the quantification of all SOA tracers needed to apply the SOA tracer method for the SOC estimation is a difficult and labor-intensive task and could explain the limited number of available results in the literature. Figure 6 shows the results obtained over the world in the PM$_{2.5}$ fraction only for the spring–summer period. Detailed references about all the results considered here are presented in Tables SC2 and SC3. Total SOC discussed below represents the sum of SOC$_{\text{isoprene}}$, SOC$_{\alpha\text{-pinene}}$, SOC$_{\beta\text{-caryophyllene}}$ and SOC$_{\text{toluene}}$. Unaccounted SOC refers to the SOC from the other class of precursors not accounted in the SOA tracer method or to the SOC from a given precursor which is normally included in the SOA tracer method but not in the total SOC calculation due to a lack of data in the present case.

### 3.3.3.1. North America

The annual SOC contributions observed at Bondville, Northbrook, RTP, Cincinnati, Detroit, East St. Louis in the US varied in the range 8–53%. SOC concentrations from the different precursors ranged from 0.2 to 0.9 µgC m$^{-3}$ for isoprene, from 0.1–0.3 µgC m$^{-3}$ for α-pinene, from 0.1–0.2 µgC m$^{-3}$ for β-caryophyllene and from 0.2–0.3 µgC m$^{-3}$ for toluene [27,277,279]. In the spring–summer period, the average SOC levels in the US including urban, suburban, rural and remote locations varied in the range 0.2–2.6 µgC m$^{-3}$, contributing to 5–98% to PM$_{2.5}$ OC (Figure 6). Among the urban and suburban locations, Northbrook exhibited the highest SOC contribution (61%) and Bakersfield the lowest one (5%). Rural and remote sites (Medina, Bondville, RTP) showed that more than 60% of OC is from a secondary origin.

Results obtained at all sites located in Eastern US (Atlanta, Birmingham and Pensacola, East St Louis), have shown that α-pinene (59% of total SOC) and isoprene (26% of total SOC) SOAs were the main SOC contributors even at urban and suburban locations. Model simulations, as mentioned in Section 3.2.3.1, have shown significant emissions of isoprene and monoterpene in the Eastern part of the North American continent [44,293–295]. The land surface in these areas is characterized by rolling or hilly terrain with heavy vegetation, mainly consisting of mixed coniferous (mainly loblolly pines) and deciduous (mainly oak and hickory) forests. Then, during spring and summer, biogenic emissions drive the production of SOA. In addition, specific experiments have also been performed in the region to understand the impacts of anthropogenic pollution on biogenic SOA formation [219,220,296]. Finally, in highly urbanized and industrialized cities such as Mexico, Detroit, Cleveland, Riverside and Pasadena, SOC$_{\text{toluene}}$ appeared as a major contributor to the total SOC (40 to 80%).

### 3.3.3.2. Europe

Only few studies have been performed in Europe and only biogenic SOC was accounted. Anthropogenic SOC was unfortunately neither measured nor detected. The average SOC in Europe ranged from 2% to 12% (0.02–0.49 µgC m$^{-3}$) of PM$_{2.5}$ OC (Figure 6). It is worth noting that at all European sites, SOC$_{\text{isoprene}}$ contributions were significantly lower than at the sites monitored in North America due to large isoprene emissions in the US. Monoterpene emissions are dominant in Europe and account on average from 40% to 60% of total emitted VOCs [297]. The remaining fraction is dominated by isoprene and other VOCs. This fact can be easily observed at Marseille (France), where SOC$_{\alpha\text{-pinene}}$ contributed to 3.5% of the total OC while SOC$_{\text{isoprene}}$ contribution was only about 0.6%. On contrary, at remote or rural sites, Jülich (Germany) and K-puszta (Hungary), the contribution of both isoprene and α-pinene SOC were similar.

### 3.3.3.3. Asia

In Asia, several studies have been conducted in India, Japan but most especially in China (Figure 6). Studies conducted at urban and suburban background sites in Shanghai showed average annual SOC contributions of about 3% corresponding to SOC concentrations of about 0.3 µgC m$^{-3}$ [179].
In the spring–summer period, the average SOC concentrations at urban sites in Beijing, Shanghai, Hong Kong and Wangqingsha (Pearl River Delta; PRD) were 3.3, 0.6, 3.0 and 3.1 µgC m$^{-3}$, respectively. Results from all urban and suburban locations highlighted the major contributions of SOC$_{\text{isoprene}}$ (20–50%) and SOC$_{\text{toluene}}$ (50–80%) to total SOC except at Hong Kong where SOC$_{\alpha\text{-pinene}}$ (48%) and SOC$_{\beta\text{-caryophyllene}}$ (33%) were dominant. The very high urbanization and industrialization of these cities explain such SOC$_{\text{toluene}}$ contributions as well as the very high isoprene emissions in North Eastern China [298]. In Hong Kong, the observation of larger amounts of SOC attributable to monoterpenes rather than isoprene is consistent with high emissions of monoterpenes in the region [272], in addition to the higher SOA formation yields from monoterpenes oxidation [16]. The only study conducted in India at Mumbai [286] showed a very minor contribution of SOC to fine OC of about 2% equivalent to SOC concentration of 0.47 µgC m$^{-3}$. Authors suggested that high ambient temperatures and relative humidity in tropical regions may affect the SOA yields, the gas/particle partitioning, and aging processes explaining such low SOC contribution values [285,286]. Similarly, only one study has been performed in Japan so far at a rural site in Hokkaido [288]. The SOC contribution to total OC was about 16% (0.69 µgC m$^{-3}$) and mainly due to SOC$_{\text{isoprene}}$.

Finally, several studies have been performed at remote sites in China and on the Tibetan Plateau. Overall, very low SOC contributions to the total OC were observed (1–12%) except at the Tibetan plateau where the SOC contribution observed in spring–summer period was about 38% with very low total OC concentrations (1.90 µgC m$^{-3}$) while, on annual basis, the contribution was about 13% [64]. At other locations in China, total OC was significantly higher (4.8–18 µgC m$^{-3}$) and probably impacted by direct emissions explaining the low SOC contributions observed.

3.4. Positive Matrix Factorization (PMF) (Including AMS Data Analysis)

3.4.1. Introduction

As other receptor models, the goal of PMF is to solve the chemical mass balance between the measured chemical species concentrations and source profiles as a linear combination of factors $p$, species profile $F$ of each source, and the amount of mass $G$ contributing to each individual sample (Equation (7)):

$$x_{ij} = \sum_{k=1}^{p} G_{ik} F_{kj} + E_{ij}$$

where $x_{ij}$ represents the measured data for species $j$ in sample $i$, and $E_{ij}$ represents the residual of each sample/species not fitted by the model.

Thus, PMF is a multivariate factor analysis tool that decomposes the matrix $x$ ($n \times m$), where $n$ is the number of samples and $m$ is the number of chemical species, into both matrices, factor contributions $G$ ($n \times p$) and factor profiles $F$ ($p \times m$), that need to be ascribed to a specific source.

The best model solution is obtained by minimizing the function $Q$ (Equation (8)):

$$Q = \sum_{i} \sum_{j} \left( \frac{e_{ij}}{s_{ij}} \right)^{2}$$

where $s_{ij}$ represents the measurement uncertainty of each data point. The $Q$ value can be used to determine the optimal number of factors. The theoretical $Q$ value should be approximately equal to a value of $n \times m$, the number of values in the data matrix or the degree of freedom of the datum in the data set.

PMF does not rely on information from the correlation matrix but utilizes a point-by-point least squares minimization scheme and differs from the other factor analysis models such as principal comment analysis (PCA) by the property to consider standard deviations of observed data values and to introduce the constraint of non-negativity of all the factor matrices $G$ and $F$ to get physically meaningful solutions. The input data matrix contains the measured species concentrations and their corresponding uncertainties. One of the main features of the PMF results is their quantitative nature.
It is then possible to obtain the composition of the sources determined by the model [24,25]. This is the distinctive advantage of PMF over other multivariate factor analysis approaches.

An estimation of the data uncertainties could be performed using known concentrations and the limit of detection values. The estimation of uncertainties is calculated using Equation (9) [299]:

\[
\sigma_{ij} = \begin{cases} 
\frac{5}{6} LD_j & \text{if } X_{ij} < LD_j \\
\sqrt{(LD_j)^2 + (CV_j X_{ij})^2 + (aX_{ij})^2} & \text{if } X_{ij} \geq LD_j 
\end{cases}
\]  

(9)

where \( LD_j \) is the detection limit for compound \( j \) (defined as the lowest concentrations of the compound that can be measured with a signal to noise ratio of 3), \( CV_j \) is the coefficient of variation for compound \( j \) (calculated as the standard deviation of repeated analyses divided by the mean value of the repeated analyses), and \( a \) is a factor that could be applied to account for additional sources of uncertainty [300].

Alternatively, the PMF program can compute heuristic error estimates, \( s_{ij} \), for each \( x_{ij} \) based on the data point or on its analytical error. This is done by means of define codes within the model [301].

3.4.2. Limitations and Challenges

PMF has been widely applied to apportion the sources of PM based on the offline chemical analysis of filter samples. Such datasets generally include major chemical species, such as OC, EC, water-soluble inorganics, and metals. However, many of these species are not source specific and/or of secondary origins, making it difficult to directly link PMF factors with the different OA sources and/or (trans-)formation processes. In such cases, SOC is commonly calculated as the sum of OC loadings in sulfate- and nitrate-rich PMF factors or individually [93,205,302,303]. Few studies also reported the use of species such as WSOC and humic-like substances (HuLiS) to characterize SOC [174]. By comparison, molecular organic markers (tracers) are mostly source-class specific and may provide a more definitive link between factors and source classes. Molecular markers for SOA and POA can be directly included in the PMF model providing an insight into the primary–secondary split of OA sources [14,40,73,214,273,304–306]. The effectiveness of the method depends on the molecular markers used. In addition, and as already mentioned in Sections 3.2.2 and 3.3.2, both the stability of molecular markers in the atmosphere and their limited number for known precursors of SOA, can hamper the PMF filter-based source apportionment. Finally, as PMF analysis requires a larger number of samples (preferably >100) to get a statistically robust solution, it has not been extensively performed to apportion OA sources based on filter measurements.

As already discussed in Section 3, the long sampling duration and ageing of OA of collected filter samples could lead to possible sampling artifacts. In addition, filter samplings rely on measurements performed over several hours to days, making it difficult to capture the fast-atmospheric chemical processes related to OA. By comparison, online aerosol chemical characterization techniques, such as AMS, are faster, less labor-intensive and allow the quantitatively determination of non-refractory PM with high time resolution [34,307–311]. The cost and complex maintenance requirements of the AMS make its deployment barely practicable for long-term monitoring. Consequently, most available datasets are often limited to a few weeks of measurements. This hinders the determination of the regional and seasonal aerosol characteristics and the identification of changes in the pollution trends representing a significant limitation to validate atmospheric chemistry model validation and for the evaluation of air quality policies. For these reasons, the aerosol chemical speciation monitor (ACSM) has been developed for routine monitoring purposes allowing a similar discrimination of OA sources but with low resolution mass spectrometry [312–316].

Several types of OA obtained from the mass spectra can be apportioned by applying PMF to AMS/ACSM data. The oxygenated fraction of OA (OOA) identified by PMF-AMS/ACSM is usually associated with SOA [35]. This OOA fraction is commonly sub-divided in semi-volatile or low oxidized (SV-, and LO-OOA, respectively) and low volatile or more oxidized (LV-, and MO-OOA, respectively) [311]. By comparison to filter-based PMF with SOA markers, the explicit characterization
of OOA is not possible with PMF-AMS/ACSM because the mass fragmentation obtained is not specific. In some cases, detailed analyses have been done and highlighted additional sub-SOA fractions. For instance, biogenic SOA from isoprene epoxydiols (IEPOX-SOA), organic-nitrate aerosols, and/or marine SOA can be quantified using online aerosol mass spectrometry [298,317–320].

PMF-AMS/ACSM analyses are expected to provide results with relatively low uncertainties (i.e., below 20%) for both the aerosol chemical characterization and the apportionment of the different OA (about 6%) [321,322]. However, as for all mass spectrometry techniques, the interpretation of mass spectra from natural samples can be complicated by several interferences [34]. Other factors identified using PMF-AMS/ACSM such as hydrocarbon-like OA (HOA), which in urban areas shows correspondence with fossil fuel POA, could potentially include other primary sources such as biomass burning (BBOA) or cooking-like organic aerosol (COA) [307,323]. Some studies also suggested that OOA should not be always equated to SOA in places where the direct emission of oxygenated aerosol species from sources like biomass burning, charbroiling, cooking etc., are dominant. For instance, OOA was shown to also contain a significant amount of primary emissions in Zurich during the summer season [323]. Thus, the inclusion of oxidized primary particles in OOA cannot be ruled out and could cause an elevated SOA contribution under certain circumstances.

Drawbacks of online AMS or ACSM measurements also exist in the investigation of the submicron fraction only, and at a limited number of monitoring stations. To extend the spatial and temporal coverage of AMS measurements, the application of the AMS to nebulized water extracts of filter samples has been developed (offline AMS) [324–326]. This approach facilitates the investigation of specific events and can extend the measurements to the PM$_{10}$ or coarse aerosol fraction while it is not possible using online aerosol mass spectrometry [325–327].

3.4.3. Review of Recent Studies Based on the PMF Approach

3.4.3.1. Filter-Based PMF Studies

Only few studies reported in the literature are based on the use of filter-based PMF to estimate the SOC fraction in the PM$_{2.5}$ fraction. These annual based studies covered a wide geographical region in North America but not for the rest of the world (but in Hong Kong and Shanghai, China) (Table SD1, SM). As mentioned before, PMF analysis is usually performed using “traditional” speciation data and the use of specific POA and SOA molecular markers is still rare (including for other aerosol fractions like PM$_{10}$) [14,26,40,73,179,214,273,304–306,328].

Overall, the SOC in the US ranged from 12% to 57% (0.1–0.9 µg C m$^{-3}$) of PM$_{2.5}$ OC and 24% to 66% (0.7–6.8 µg C m$^{-3}$), in Hong Kong and Shanghai. In some cases, contributions from anthropogenic and biogenic sources have not been differentiated by PMF [40], especially when no molecular markers have been used (SOA based on nitrate or sulfate factors) [93,205].

The biogenic SOA fraction was apportioned by using SOA markers from the oxidation of isoprene (2-methylglyceric acid, 2-methylthreitol, 2-methylyxanthithiol and C5-alkene triols) [27], α-pinene (MBTCA, pinonic acid, norpinonic acid, pinic acid, 3-(2-hydroxy-ethyl)-2,2-dimethylcyclobutane-carboxylic acid and 3-acetyl-hexanediolic acid, 3-hydroxyglutaric acid, 2-hydroxy-4,4-dimethylglutaric acid) and β-caryophyllene (β-caryophyllinic acid) [14,26,40,214,304,329]. Shrivastava et al. [14] have observed that biogenic SOA contributed to more than 50% of the summertime OC in Pittsburgh (USA). Similarly, the relative average contributions of isoprene SOC, α-pinene SOC and β-caryophyllene SOC in the midwestern US, were 20%, 5% and 19% of PM$_{2.5}$ OC, respectively [26]. In other PMF studies, the biogenic SOA factor followed the same pattern with high contributions during the warm period [214,304].

The anthropogenic SOA fraction was mostly characterized in filter-based PMF analysis using di- or tri-carboxylic aliphatic or aromatic carboxylic acids (e.g., 1,2-benzenedicarboxylic acid, 1,3-benzenedicarboxylic acid, 4-methyl-1,2-benzenedicarboxylic acid, benzene tricarboxylic acid, benzene tetracarboxylic acid, phthalic acid, succinic acid, 2,3-dihydroxy-4-oxopentanoic acid (DHOPA))
and phthalates [40,179,214,305,328]) but also very recently, in the PM$_{10}$ fraction, using oxy-PAHs, nitro-PAHs or methylnitrocatechols [73,273].

3.4.3.2. PMF-AMS/ACSM Based Studies

By comparison to filter-based PMF, extensive studies on PMF-AMS/ACSM have been reported by many authors in the literature with an estimate of the SOC contributions and concentrations in PM$_1$ (OOA factors). Several previous papers already reviewed the results obtained worldwide [12,36,310,330]. Here, the overall average SOC distribution has been obtained by compiling results, only for the spring–summer period, from these review papers and from some other more recent papers [12,35,330–335] (Figure 7). Detailed references about all the results considered in the present study for PMF-AMS/ACSM SOC estimates are presented in Table SD2 (SM). Major conclusions obtained for different world areas are briefly presented hereafter.

3.4.3.3. North America

In North America, the average spring–summer SOC concentrations in urban locations ranged from 1.3 to 3.2 µgC m$^{-3}$ contributing to 33–74% of PM$_1$ OC. The possible sources for the OOA observed in these studies included SOA (from either anthropogenic or biogenic precursors), the oxidation of HOA, and/or BBOA [12,330]. SOC contributions to PM$_1$ OC at rural/remote locations were significantly
larger i.e., about 67–91%, corresponding to SOC concentrations in the range 0.4–2.8 µgC m⁻³. As explained above (Sections 3.2.3.1 and 3.3.3.1), biogenic emissions of SOA precursors (isoprene and monoterpenes) play a major role in the SOC concentrations observed in the spring–summer period especially at rural/remote locations [44,219,220,293–296].

3.4.3.4. Europe

The studies to apportion the SOC fraction using PMF-AMS/ACSM performed in Europe cover a wide geographical area providing a good overview of the SOC distribution over this continent. The average SOC estimates in the warm period varied at urban locations (or urban downwind) in the range 0.6–4.1 µgC m⁻³, contributing to about 27–83% of PM₁ OC, with the highest contribution at Zurich (Switzerland) (83%) and the lowest at Edinburgh (UK) (26%).

Several rural, remote and altitude sites have also been investigated in Europe using AMS or ACSM measurements. SOC concentrations in such locations ranged from 2.1 to 4.0 µgC m⁻³, corresponding to SOC contributions to PM₁ OC of about 46% to 93%. In such cases, the highest SOC concentrations observed (>3.5 µgC m⁻³) in Hyytiälä (Finland), Jungfraujoch (Switzerland), Payerne (Switzerland), Finokalia (Greece) and Melpitz (Germany) corresponded to the highest SOC contributions (>74%) indicating the long range transport of air masses with efficient oxidation processes leading to the formation of high amounts of SOA [335].

3.4.3.5. Asia

In Asia, SOC concentrations at urban and suburban sites (or urban downwind) ranged from 2.2 (Tokyo, Japan) to 8.6 µgC m⁻³ (Shenzhen, China) corresponding to SOC contributions to PM₁ OC of about 45–73%. Higher SOC concentrations were observed in China in link with the growing industrialization of this country and large VOC emission from anthropogenic activities [12,36,330,332,334]. At rural/remote locations SOC concentrations observed were in the range 0.2 to 2.4 µgC m⁻³ for sites located in Japan and Korea with contributions from 43% and up to 100% in Okinawa Island (Japan) which is located about 400–500 km from the Chinese coast and the main Japanese island [12,330]. Long range transport could explain such SOC contributions observed [12]. Finally, the growing number of studies currently conducted in Asia, and notably in China and in India, will soon provide a much more comprehensive picture of OA fractions in these world areas.

3.5. ¹⁴C (Radiocarbon) Measurements

3.5.1. Introduction

Besides the use of specific organic marker species, isotopic abundances can also help to discriminate OA sources. The study of radiocarbon (¹⁴C) itself does not allow the direct discrimination of SOC and POC but combined with another SOC apportionment method, it gives an insight for the distinction between the fossil fuel and the non-fossil fuel SOC origins.

¹⁴C is present in small amount in living (or contemporary) materials but at an approximately constant level and it is nearly absent in fossil fuels, which are much older than the ¹⁴C half-life of about 5730 years. In addition, radiocarbon remains in its original state throughout chemical processes. Thus, a radiocarbon measurement provides a unique possibility of distinguishing quantitatively the relative contributions of both, fossil and contemporary carbon sources [336]. The radiocarbon content of a carbonaceous sample is expressed as the fraction of “modern carbon” (f_M) based on the ¹⁴C/¹²C ratio observed in the atmospheric CO₂ of the year 1950, as a reference, following Equation (10) [337].

\[
    f_M = \left( \frac{¹⁴C}{¹²C} \right)_{\text{samples}} \bigg/ \left( \frac{¹⁴C}{¹²C} \right)_{\text{AD1950}}
\]
Values of $f_M$ approach 0 for fossil sources, to values larger than 1, for contemporary sources. The $f_M$ value for contemporary sources exceeds unity due to the atmospheric nuclear weapon tests in the 1950s and 1960s that significantly increased the radiocarbon content of the atmosphere [338]. The term “modern carbon” only refers to measurements relative to the 1950 standard, and the terms “contemporary” or “non-fossil,” and “fossil” carbon refer to quantities after correction.

Based on this specificity, carbon isotope data have been used in several studies to estimate the fossil fuel and contemporary (non-fossil) contributions to SOC by combining results from different SOC apportionment methodologies such as CMB and EC tracer method with radiocarbon measurements. None of the other methodologies used and detailed here can do so far something similar except the SOA-tracer method. However, as specified previously in Section 3.3.2, the lack of SOA tracer compounds for many precursors does not allow a complete description and discrimination of anthropogenic and biogenic SOCs. Thus, $^{14}$C data can be useful to improve such discrimination and to understand the evolution mechanism of biogenic SOC, assumed as equivalent to the non-fossil fraction during the warm periods, and fossil SOC, which is still missing using other methodologies.

3.5.2. Limitations and Challenges

First, the radiocarbon analysis requires large sample quantities to be quantitative and reproducible. In addition, this kind of analysis is quite expensive and then difficult to be performed routinely.

Second, emissions from nuclear power plants and incinerators of waste medical or biological products, containing $^{14}$C used as a radioactive tracer, can significantly bias the estimated $f_M$. This could induce an overestimation of the true proportion of contemporary carbon. Such cases have been reported previously, suggesting that a $^{14}$C contamination is uncommon but not impossible and could impact about 10% of the PM sampling sites [339]. The occasional artificially inflated value of the modern carbon fraction needs to be always considered sincerely because its extent is not predictable.

Finally, another common problem is the true estimation of the biogenic fraction. $^{14}$C measurements only allow the discrimination of fossil from non-fossil emissions. As both sources contribute to the contemporary carbon fraction, biofuel combustion and biomass burning emissions, which are mostly anthropogenic, cannot be separated from biogenic emissions using this methodology. The contemporary fraction is then considered as representative of the biogenic fraction only in summer. However, it can be also influenced by biomass burning emissions in case of forest fires or green waste burning.

3.5.3. Review of Recent Studies Based on $^{14}$C Measurements

Across the world, the application of $^{14}$C measurements, in combination with another methodology to determine SOC, has been reported in the literature only in a limited number of studies for the spring–summer period [29,30,41,340–345] (Figure 8). To the best of our knowledge, data on an annual basis have never been reported. Details about all the results considered here are presented in Table SE1 (SM).

3.5.3.1. North America

Overall, the application of radiocarbon measurements in the US has been done at different sites, including urban, suburban, rural and remote locations, with SOC concentrations in the range of 0.3 to 5.2 $\mu$gC m$^{-3}$, contributing to more than 40% of PM$_{2.5}$ OC [342,343].

Very high contributions of SOC$_{\text{contemporary}}$ have been observed at all sites (about 72% of total SOC on average, 50–100% equivalent to 0.3–3.7 $\mu$gC m$^{-3}$). As explained before, SOC$_{\text{contemporary}}$ could account from both biogenic and biomass burning sources. As the contribution from biomass burning was probably negligible in warm periods of the year (except during forest fire events [346,347]), these results suggested that SOC$_{\text{contemporary}}$ could mostly be attributed to biogenic sources. Moreover, at rural locations, the estimated contemporary SOC was even higher and varied from 82% to 99%.

As mentioned in Sections 3.2.3.1 and 3.3.3.1, significant emissions of isoprene and monoterpene in the Eastern part of the US, and the formation of biogenic SOA, have been reported [44,219,220,293–296].
SOC_fossil exhibited high contributions at all urban and suburban sites (in the range 12 to 54% of total SOC equivalent to 0.0–2.6 µgC m\(^{-3}\)). This contrasted with results obtained at rural locations where SOC_fossil accounted for about 7% of total SOC on average (3–18% equivalent to 0.0–0.7 µgC m\(^{-3}\)). These results showed the importance of fossil carbon emissions on the formation of SOA in urban environments.

![Summer-Spring PM\(_{2.5}\)](image)

**Figure 8.** Spring–summer SOC contributions to PM\(_{2.5}\) OC over the world for all the selected sites from 2006 to 2016 using \(^{14}\)C (Radiocarbon) measurements. OC concentrations (µgC m\(^{-3}\)) for each site are indicated into brackets. % refers to the sum of SOC_fossil + SOC_contemporary. In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites. * Contemporary SOC was not reported.

### 3.5.3.2. Europe

In Europe, the average SOC contributions to PM\(_{2.5}\) OC at sites where radiocarbon measurements have also been performed were larger than 21% and up to 80% of PM\(_{2.5}\) OC with a concentration range of 0.7–4.7 µgC m\(^{-3}\) [30,41,341,344]. The average SOC levels observed for contemporary and fossil fractions were in the range 0.6–3.4 µgC m\(^{-3}\) and 0.1–1.3 µgC m\(^{-3}\). Results showed that for all site typologies, SOC from non-fossil sources (SOC_{contemporary}) was dominant during the warm period (72–95% of total SOC) with slight to moderate contributions of fossil SOC (5–28% of total SOC). Note that, in Goteborg (Sweden), no contemporary SOC was reported.

The site to site contributions of contemporary SOC were consistent with the wide geographical area covered across south-central Europe [30]. Rural and remote sites showed very high SOC_{contemporary} contributions related most probably to biogenic SOA. High contribution of SOC_fossil (28% of total SOC, 1.3 µgC m\(^{-3}\)) was noticed in the Po Valley (Italy) known to be highly impacted by anthropogenic activities from the surroundings urban and industrialized areas. In addition, it has been shown that the concentrations of biogenic SOA observed in warm periods were also enhanced by anthropogenic primary emissions [344]. Both urban locations (Goteborg (Sweden), Marseille (France)) investigated, showed SOC_fossil contributions of about 21% of total SOC.
3.5.3.3. Asia

In Asia, only two studies (both in China) have been reported, using $^{14}$C measurements for SOC apportionment purposes. SOC contributions and concentrations levels at both urban locations investigated were about 50% of PM$_{2.5}$ OC; 3.7–3.9 $\mu$gC m$^{-3}$ (Beijing) and 59% of PM$_{2.5}$ OC; 2.0–4.1 $\mu$gC m$^{-3}$ (Guangzhou), respectively [340,345].

On one hand, the differences observed in SOC composition between both sites (49 vs. 67% of total SOC, in Beijing and Guangzhou, respectively) may be linked to the differences of climate and land cover. In spring, the land of North China (Beijing) is bare and trees are still leafless, whereas, in South China (Guangzhou) lush vegetation is present, emitting large VOC quantities including species such as isoprene, $\alpha$-pinene, $\beta$-caryophyllene, etc., which are efficient biogenic SOA precursors [340]. On the other hand, the larger SOC$_{\text{fossil}}$ contributions observed in Beijing (51% of total SOC) than in Guangzhou (33% of total SOC) was probably due to the primary anthropogenic emissions. In addition, the differences in meteorological conditions between both cities may be another reason to explain such different SOC compositions.

4. Review of the Studies Directly Comparing Different Methodologies

The above discussions highlight the advantages and limitations of the different methodologies usually used to estimate SOC. They have been applied worldwide, showing many interesting features in terms of the influence of site typology, SOA precursor origins, SOC seasonality, etc. If the differences observed on the SOC contributions and concentrations can be related to the geographical origins and/or the meteorological conditions, they can also be linked to the uncertainties of the methodologies applied themselves. There are only few examples available in the literature where these methodologies have been compared thoroughly. This section summarizes the results reported in the literature where CMB, EC tracer method, SOA tracer method and/or PMF approach have been directly compared (Table 1).

4.1. EC-Tracer vs. CMB

A direct comparison between CMB and EC tracer methods has been reported in at least six studies [1,14,93,124,205,207] (Table 1).

Overall, in terms of both correlations and absolute concentrations, a good agreement between both methodologies has been observed in summer while the differences in winter were significantly larger. As explained before (Section 3.1.2), the determination of [OC/EC]$_p$ in the EC tracer method might be challenging in winter due to the significant influence of local primary emissions. Contrarily, as observed in Pittsburgh (US), under the influence of regional air masses with high SOA background concentration levels, [OC/EC]$_p$ and/or the intercept used in the EC-tracer method may be overestimated leading to an underestimation of the SOC concentrations [14,207]. The impact of the determination of [OC/EC]$_p$ has also been observed for the study of the SOC diurnal cycles in Riverside (CA, US). The reported diurnal cycles were similar using both EC tracer and CMB methodologies (and PMF too) but slightly higher SOA/OA ratios have been observed throughout the evening/night using the EC-tracer method. This probably resulted from lower [OC/EC]$_p$ estimated during the night due to reduced diesel traffic at that time [1]. Overall, the CMB approach seemed to overestimate SOC by comparison to the EC tracer method and again especially in cold period. This overestimation could be due to the significant contributions to OC of unknown and/or unresolved primary sources (e.g., cooking activities and natural gas combustion) then accounted as SOC [205] (see Section 3.2.2). Besides, SOC estimates seemed comparable at urban sites while larger values from CMB were observed at rural sites [205]. Again, the estimation of [OC/EC]$_p$ is quite challenging especially in remote locations which are far from primary sources (see Section 3.1.2). Finally, higher uncertainties have been noticed using the EC tracer method and can be explained by the different approaches used for the estimation of [OC/EC]$_p$ [1,93].
Table 1. List of the studies reporting a direct comparison of SOC contributions evaluated using different methodologies.

<table>
<thead>
<tr>
<th>Locations</th>
<th>Sampling Periods</th>
<th>Methodologies</th>
<th>Main Conclusions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pittsburgh, Pennsylvania (USA)</td>
<td>Annual (July 2001–July 2002)</td>
<td>X X</td>
<td>Good agreement ($r^2 = 0.71$, slope = 0.75). Better in summer ($r^2 = 0.81$, slope = 0.91) than in winter ($r^2 = 0.45$, slope = 0.75).</td>
<td>[207]</td>
</tr>
<tr>
<td>London, Birmingham, Birmingham (UK) *</td>
<td>Summer (June 2010–August 2010)</td>
<td>X X</td>
<td>SOC estimates were in broad agreement (urban sites, $r^2 = 0.70–0.92$, slope = 0.80–0.92; rural site, $r^2 = 0.69–0.92$, slope = 0.73–0.88).</td>
<td>[124]</td>
</tr>
<tr>
<td>Atlanta, Birmingham, Centreville, Yorkville (USA)</td>
<td>Annual (January 2000–December 2002)</td>
<td>X X</td>
<td>Lower SOC estimates using PMF. Good correlation between SOC estimates using CMB and EC tracer method (CMB vs. EC, $r = 0.67–0.84$; CMB vs. PMF, $r = 0.58–0.74$; EC vs. PMF, $r = 0.40–0.78$). Comparable SOC levels (CMB vs. EC) at urban sites while at rural sites, larger SOC levels estimated using CMB.</td>
<td>[205]</td>
</tr>
<tr>
<td>Atlanta (USA)</td>
<td>Summer/Winter (February 1999–December 2007)</td>
<td>X X</td>
<td>Higher SOC estimates using CMB especially in winter. CMB vs. EC vs. PMF: summer: $2.0 \pm 0.9$ vs. $1.5 \pm 1.4$ vs. $1.4 \pm 0.8 \mu g m^{-3}$; winter: $1.8 \pm 1.0$ vs. $0.8 \pm 2.0$ vs. $0.9 \pm 0.9 \mu g m^{-3}$. The highest uncertainty was obtained using the EC tracer method. The PMF uncertainties were significantly higher than the uncertainties in the CMB method.</td>
<td>[93]</td>
</tr>
<tr>
<td>Pittsburgh (USA)</td>
<td>Annual (July 2001–August 2002)</td>
<td>X X</td>
<td>All methods (EC tracer, CMB, PMF-filter and PMF-AMS) provided the same seasonal pattern with more SOA in summer than in winter. Summer, EC tracer vs. other approaches: $55–70%$ vs. $30–40%$ SOC in PM$<em>{2.5}$ OC; winter, CMB vs. other approaches: $50$ vs. $10%$ SOC in PM$</em>{2.5}$ OC. PMF-filter vs. CMB: Non-winter: $r^2 = 0.55$, slope = 0.72; in winter, poor correlation and low slope.</td>
<td>[14]</td>
</tr>
<tr>
<td>Riverside (USA)</td>
<td>Summer (July 2005–August 2005)</td>
<td>X X</td>
<td>SOA estimates were consistent for all the methods. Diurnal cycles of SOA/OA ratios were similar with maximum ratios observed during the early afternoon. However, the EC-tracer method apportioned SOA slightly differently throughout the evening/night.</td>
<td>[1]</td>
</tr>
<tr>
<td>Locations</td>
<td>Sampling Periods</td>
<td>Methodologies</td>
<td>Main Conclusions</td>
<td>References</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Birmingham, Centreville, Atlanta (USA)</td>
<td>Spring/Summer (August 2003–August 2005)</td>
<td>X  X</td>
<td>For Atlanta, SOC estimates were similar (EC vs. SOA tracer: spring: 1.4 vs. 1.3 µg m(^{-3}); summer: 1.2 vs. 1.4 µg m(^{-3})). For Birmingham and Centreville, the differences were significantly larger (on average, 1.8 vs. 2.8 µg m(^{-3}); 1.2 vs. 2.7 µg m(^{-3}), respectively).</td>
<td>[38]</td>
</tr>
<tr>
<td>Wangqingsha, Pearl River Delta (China)</td>
<td>Summer (August 2008–September 2008), Fall–Winter (November 2008–December 2008)</td>
<td>X  X</td>
<td>Good agreement in summer ((r = 0.57), slope = 0.91, EC vs. SOA tracer SOC: 3.2 vs. 3.1 µg m(^{-3})) and better than in fall–winter (EC vs. SOA tracer SOC: 6.7 vs. 2.0 µg m(^{-3})). The minimum OC/EC ratio could be not representative of (OC/EC)(_p) in winter season (biomass burning impacted). In fall–winter, other SOA precursors (“non-traditional” SOA) were probably significant but not considered in the SOA tracer method.</td>
<td>[178]</td>
</tr>
<tr>
<td>Marseille (France)</td>
<td>Summer (June 2008–July 2008)</td>
<td>X  X</td>
<td>Both methods followed different temporal trends (only biogenic SOC was considered for the SOA tracer method) (r^2 = 0.18). CMB vs. SOA tracer: 2.1–8.5 vs. 0.0–0.6 µg m(^{-3}).</td>
<td>[41]</td>
</tr>
<tr>
<td>Hong Kong, (China)</td>
<td>Summer (July 2006–August 2006)</td>
<td>X  X  √(^1,2)</td>
<td>SOC estimates showed very similar time evolutions throughout the sampling period. The average SOC from CMB, PMF-filter and SOA tracer method were: 7.8, 6.8 and 5.0 µg m(^{-3}) during high pollution episodes (regional transport) and 1.2, 0.7 and 0.5 µg m(^{-3}) under the influence of local emissions (local days).</td>
<td>[40]</td>
</tr>
<tr>
<td>Detroit, Cincinnati, East St. Louis, Northbrook, Bondville (USA)</td>
<td>Annual (March 2004–February 2005)</td>
<td>X  √(^2)</td>
<td>SOA estimates were highly consistent except for few months with high secondary contributions (r^2 = 0.76), slope = 1.01). Underestimation by PMF when the secondary contributions were very low.</td>
<td>[26]</td>
</tr>
<tr>
<td>Shanghai (China) 2 sites: 1 urban + 1 suburban</td>
<td>Winter (January 2010, January 2011)/Spring (April 2010–May 2010)/Summer (July 2010)/Autumn (October 2010–November 2010)</td>
<td>X  X  √(^1,2)</td>
<td>SOA contributions might be underestimated with the SOA tracer method (only terpenes and aromatic compounds considered). EC vs. SOA: fall–winter: 2.8–8.8 vs. 0.1–0.4 µg m(^{-3}); spring–summer 1.5–2.2 vs. 0.1–0.6 µg m(^{-3}). PMF (2.1–2.8 µg m(^{-3})) with no variation: a large part of the SOC was associated with nitrate and sulfate but not with the measured SOA tracers. SOA tracer and PMF-filter SOC estimates were significantly correlated (r^2 = 0.68). As commercial standards for many of the tracers are not available, large uncertainty in the quantification of the SOA tracers.</td>
<td>[179]</td>
</tr>
<tr>
<td>Locations</td>
<td>Sampling Periods</td>
<td>Methodologies</td>
<td>Main Conclusions</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Atlanta (USA)</td>
<td>Summer (July 2001–August 2001)/Winter (January 2002)</td>
<td>X</td>
<td>Good correlation between SOC estimates (CMB vs. PM, $r^2$ = 0.43–0.50, slope = 3.2–7.4). Larger SOC estimates using CMB may be due to the unresolved primary OC that would attribute to the CMB (high bias) and the SOA from the resolved primary sources that have not been included in the PMF SOA (low bias).</td>
<td>[303]</td>
</tr>
<tr>
<td>Hong Kong (China)</td>
<td>10 sites: 9 urbans + 1 traffic</td>
<td>1998–2002</td>
<td>The SOC estimates by the EC tracer method were consistently higher than PMF method. Overestimation by 70–212% for the summer samples and by 4–43% for the winter samples. The overestimation by the EC tracer method resulted from the inability of obtaining a single OC/EC ratio that represented a mixture of primary sources varying in time and space.</td>
<td>[98]</td>
</tr>
<tr>
<td>Mexico City (Mexico)</td>
<td>Spring (March 2006)</td>
<td>X</td>
<td>Better agreement using CMB estimates corrected from PMF LOA factor (local low nitrogen OA): 49% (PM$_{2.5}$ OC) vs. 46% (PM$_1$ OC); $r^2$ = 0.46, slope = 1.01.</td>
<td>[348]</td>
</tr>
<tr>
<td>Guangzhou (China)</td>
<td>Summer (July 2006)</td>
<td>X</td>
<td>Good correlation between SOC and OOA from PMF ($r^2$ = 0.60) but low regression slope (0.31) indicating that there was a substantial amount of noncarbon elements (e.g., O, N) in OOA.</td>
<td>[168]</td>
</tr>
<tr>
<td>Jiaxing, Yangtze River Delta (China)</td>
<td>Summer (July 2015)/Winter (December 2015)</td>
<td>X</td>
<td>Good agreement in summer with similar time trends (PMF vs. EC tracer: 7.2 vs. 6.8 $\mu$g m$^{-3}$). In winter, strong biomass burning events led to overestimate SOA using the EC tracer method (3.9 vs. 7.0 $\mu$g m$^{-3}$).</td>
<td>[349]</td>
</tr>
<tr>
<td>Beijing (China)</td>
<td>Fall (November 2013)</td>
<td>X</td>
<td>SOC estimates obtained using EC tracer method were consistent with those from PMF ($r^2$ = 0.69) with very similar concentration levels.</td>
<td>[153]</td>
</tr>
</tbody>
</table>

Used methods: X; Filter-based PMF: √; PMF-AMS: √√; Filter-based PMF and PMF-AMS: √√√; ¹ SOC based on sulfate and nitrate factors; ² SOC factors based on SOA molecular markers; * underlined: rural sites; italic: suburban sites; normal: urban sites.
4.2. SOA-Tracer vs. Other Methodologies (EC-Tracer, CMB and PMF-Filter)

A comparison of SOC evaluated data using the SOA tracer method with those obtained using other methodologies such as the EC tracer method, CMB and PMF has been reported in few studies [26,38,40,41,178,179] (Table 1).

Results obtained using the SOA tracer approach agreed well with data reported from the above other methodologies. However, the SOA tracer method tended to underestimate the SOC concentrations. As explained before (Section 3.3.2), the number of SOA tracers and so, of SOA precursors, considered in the method is limited. This is especially true for anthropogenic SOA precursors (i.e., only toluene SOA is considered) explaining that the agreement observed was better in summer [38,178], with a stronger biogenic SOA impact, than in the fall–winter period [178,179]. SOA derived from biomass burning in such periods could be significant but is not considered in the “traditional” SOA tracer method. Besides, the accuracy of both CMB and EC tracer methodologies, could be affected by local primary sources in winter (see Sections 3.1.2, 3.2.2 and 4.1). However, the agreement between the SOA tracer method with CMB could also be low even in summer when only biogenic SOA tracers are considered, highlighting the importance of anthropogenic SOA precursors even in warm period [40,41,179]. The comparisons made with PMF-filter were quite consistent [26]. It is also worth to note that when the secondary contributions were very low, the PMF-filter method underestimated the SOC concentrations, by comparison with the SOA tracer. This may be due to the non-linearity of the yield curve or to errors introduced by PMF-filter because of the error structure of data at these very low levels [26]. Finally, large uncertainties in the SOA tracer method could be linked to the fact that the quantification of the SOA markers was not achieved with the use of authentic standards [179] (Section 3.3.2).

4.3. PMF vs. Other Methodologies (EC-Tracer, CMB)

There are several examples in the literature where the PMF-filter or PMF-AMS [1,14,26,40,93,98,153,168,179,205,303,348,349] methodologies have been compared with the CMB and/or EC tracer methods (Table 1).

A good agreement between PMF, for both PMF-filter and PMF-AMS, and the other methodologies, has been usually observed with significant correlations and similar SOC estimates. As explained before, both CMB and EC tracer methods tended to overestimate SOC, notably in the winter period, due to the inherent limitations of both methodologies when the sampling site is locally influenced by primary emissions [14,93,98,179,205,303,349] (see Sections 3.1.2, 3.2.2 and 4.1).

The impact of local sources has been highlighted by Aiken et al. (2009) who observed, in spring in Mexico, a better agreement between CMB and PMF-AMS after subtracting the local low nitrogen OA fraction (LOA) from the non-apportioned OC, assumed to be SOC, initially obtained with CMB. In addition, the difference between PMF-AMS and the EC tracer method could also be significant even in summer when a substantial amount of non-carbon elements (i.e., O, N) were probably included in the OOA [168]. However, the good agreement between PMF-AMS and the other methodologies was supported by the fact that nice correlations (r² = 0.9) and regression slopes (0.8–0.9) were usually obtained between POC and HOA (in summer) [168,307,348].

SOC estimated from the PMF-filter approach was generally lower than that from the CMB or EC tracer methodologies [14,98,205]. PMF could over-attribute OC to certain factors because the model fits the measured OC while CMB only fits the molecular marker data. This occurred if markers for some important categories of OC (typically SOC ones) were not included in the PMF model [14]. Without the use of SOA markers, SOC was apportioned based on the OC content of the nitrate and sulfate factors (see Section 3.4.3) and the results obtained could not provide further information on SOC because of the co-linearity of OC sources [93]. As mentioned before for the SOA tracer method (Section 3.3.2), the identification of new molecular markers from other SOA precursors is strongly needed and would allow a better description of the SOC associated with nitrate and sulfate factors [179].
Finally, the different SOC apportionment approaches provide consistent results and they can achieve the same conclusions especially in the warm period. A good SOC apportionment may be more difficult to achieve when the influence of local (anthropogenic) emissions, like in the cold period, is significant (EC tracer method and CMB) or due to the lack of SOA molecular markers to account SOA from, notably, anthropogenic precursors and/or sources (SOA tracer method and PMF-filter).

5. Comparison Based on the Overall Picture Obtained from the Review of Recent Studies

The above discussions show that results obtained using different SOC apportionment methodologies are comparable under certain conditions. Thus, a comparison of the SOC estimates obtained worldwide, and reviewed before, is proposed in this section. To get comparable results, only urban and suburban locations are considered and both, annual and spring–summer SOC estimates are discussed.

5.1. Comparison of the Annual SOC Estimates Obtained Worldwide

A comparison of annual SOC estimates obtained in North America, Europe, the Middle East, India and China using CMB and EC tracer method is presented on Figure 9 (data from Tables SA2 and SB2). Results from other methodologies are not included due to the lack of data points on an annual scale.

Figure 9. Comparison of annual SOC estimates (PM$_{2.5}$) obtained using CMB and EC tracer method for North America, Europe, the Middle East, India and China. Only urban and suburban locations are included. Box-plots show the SOC and POC concentrations (µgC m$^{-3}$) for each continent and each method. Into brackets, number of data points considered. Box plots include minimum (10th), first quartile (25th), median (50th), third quartile (75th) and maximum values (90th).

Overall, SOC estimates evaluated using the EC tracer method are always larger than the ones obtained with CMB whatever the geographical region considered. In North America, the agreement observed is quite good, but the differences are particularly notable for Europe and China where a
factor of about two, in terms of both contributions and concentrations, between both methodologies can be observed. Similarly, in India, a factor of two can also be noted but only for SOC concentrations. Both estimated SOC contributions are quite similar in this case. Previous discussions have already accentuated the fact that the selection of \([\text{OC/EC}]_p\) is quite challenging in cold period (fall–winter) under the influence of local primary sources (notably biomass burning) leading to an overestimation of SOC (see Sections 3.1.2 and 4). This is particularly true in Europe, China and India where biomass burning for residential heating, and to a lesser extent, for cooking purposes, is one of the major sources of OA in winter [30,350–353]. However, this is also true for North America but the factor of two difference is not observed. This is probably because the sites/locations where the CMB and EC tracer studies have been performed are not necessarily the same and are geographically distant (Figures SA1 and SB1, SM). Interestingly, the SOC estimates in the Middle East using CMB and EC tracer method are very similar. The stable climate and meteorological conditions through the year with no proper heating period explain such observations with low bias between both methodologies.

Despite the limitations of both methodologies (see Sections 3.1.2, 3.2.2 and 4), clear patterns can be observed with annual SOC and POC concentration levels, comparable in India and China and twice higher than the ones estimated in the other regions of the world. These results agree with modelling evaluations in terms of both annual SOA concentration levels and world distribution [354]. They also highlight the significant influence of anthropogenic SOA in the developing countries [145,354]. It seems more difficult to conclude in terms of SOC contributions due to the inherent limitations of both methodologies especially when the cold period is also considered. A probably better overview can be made based on spring–summer SOC estimates only, as presented in the next section.

5.2. Comparison of the Spring–Summer SOC Estimates Obtained Worldwide

The comparison of SOC estimates obtained for the spring–summer period in North America, Europe and China using CMB, EC tracer, SOA tracer and PMF-AMS methodologies is presented on Figure 10 (data from Tables SA3, SB3, SC3 and SD2). The results from the SOA tracer method are not included for Europe to avoid any kind of false representation because only biogenic SOC has been accounted in the reviewed studies. Results from \(^{14}\text{C}\) measurements are also not included because they arise from other methodologies such as EC tracer or CMB.

When focusing on the warm period, a good agreement can be seen between all methodologies for both SOC contributions and average SOC concentrations. However, the SOA tracer method tends to underestimate the SOC estimates and it is particularly obvious for China. The rapid urbanization and industrialization in China have led to high emission of anthropogenic SOA precursors [145]. However, and as detailed before (Sections 3.3.2 and 4.2), only one anthropogenic marker (DHOPA) from toluene photooxidation is accounted for the SOA tracer method to estimate the SOA from anthropogenic sources. SOC from the oxidation of many other VOCs and SVOC/IVOCs (i.e., alkanes, PAHs, furans, phenolic compounds) is not considered using such methodology. It is true that biogenic emissions of isoprenoids (including both isoprene and monoterpenes) are significantly larger than the emissions of anthropogenic VOCs in China during daytime in summer [355]. Nevertheless, several studies have also shown a high regional background of anthropogenic VOCs in summer that could lead to significant secondary SOA formation [179,198] explaining the observed SOC underestimation using the SOA tracer method. Besides, biogenic emissions, including SOA precursors such as isoprene and monoterpenes, significantly impact the organic loading during summer in North America [44,219,220,293–296] (Sections 3.2.3.1, 3.3.3.1 and 3.5.3.1). As seen before, the biogenic SOA fraction is well documented in the SOA tracer method (see Section 3.3) explaining the comparable SOC estimates obtained with the other methodologies for North America.

Finally, the SOC contributions to PM$_{2.5}$ or PM$_{1}$ OC in the spring–summer period are comparable in all the considered regions with values ranging from 39% to 64% (excluding the SOA tracer method for China). Average spring–summer SOC concentration ranges are similar in North America and Europe (1–3 \(\mu\text{gC m}^{-3}\)) and are significantly higher in China (4–6 \(\mu\text{gC m}^{-3}\)). The relatively good
consistency between the different methodologies in North America, for both SOC concentration levels and contributions, may suggest that forest fire events, which are usually episodic in summer, are not a major contributor to SOA, although they could play an important role on the PM concentration levels [346,347]. Biogenic SOA should dominate in such seasons. It is also worth to note the wide range of SOC concentrations observed in China notably from the PMF-AMS results. The country experiences recurrent severe haze pollution events even in summer time, with fine PM reaching very high concentration levels across many cities [145,356] and SOA is a significant contributor to such PM concentrations [145,148]. Thus, these specific events, especially characterized using AMS/ACSM instrumentation, could explain, to a certain extent, the observed variability of SOC concentrations in China.

![Figure 10. Comparison of spring-summer SOC estimates obtained using CMB, EC tracer method, SOA tracer method (PM$_{2.5}$) and PMF-AMS (PM$_{1}$) for North America, Europe and China. Only urban and suburban locations are included. Box-plots show the SOC and POC concentrations (µgC m$^{-3}$) for each continent and each method. Into brackets, number of data points considered. Box plots include minimum (10th), first quartile (25th), median (50th), third quartile (75th) and maximum values (90th).](image)

6. Conclusions and Future Implications

This review has discussed the results reported on SOC estimates obtained by different methodologies reported in the literature. Although the SOC estimates are method dependent, the results reviewed here confirm that SOC constitutes a significant fraction of OC.

The different methodologies to apportion SOC are subject to a certain number of uncertainties but they provide quite consistent results especially in the warm period (spring–summer). Some discrepancies can be observed due to their inherent limitations when the cold period is considered. In such cases, the selection of the appropriate methodology should be done very carefully because each approach has its own advantages and disadvantages and the use of a single approach can cause severe misinterpretation. Thus, this review hardly recommends the use of a combination of different SOC estimation methodologies to apportion the SOC concentrations/contributions in order
to get the highest level of confidence in the results obtained. This further suggests that a combination of approaches could enhance the understanding about subtle differences existing in OA contents from different measurements.

As it requires only the determination of the concentrations of EC and OC and generally provides good SOC estimates, the EC tracer method is clearly the easiest methodology to use and should be applied as a base one. However, it is also strongly limited in winter due to the difficulty to estimate primary OC-to-EC ratio without any ambiguity.

CMB also provides SOC concentrations and contributions in good agreement with the other approaches. The lack of emission profiles for uncommon and thus un-apportioned sources inducing an overestimation of SOC, is the principal limitation of this methodology. This is especially the case when the sampling site is strongly affected by local emission sources. Thus, CMB is probably one of the most difficult approaches to apply, as source profiles for SOC do not exist and this may lead to an over/underestimation depending on the conditions. Therefore, CMB is not the first recommended approach. The current development of emission profile databases, such as SPECIATE or SPECIEUROPE, is of great value to get the appropriate input to the CMB model for each case study and can clearly improve the evaluation of SOC using such approaches [357,358].

In combination with such methodologies, the use of radiocarbon measurements (\(^{14}\)C) may provide detailed information on the SOC origins and can be helpful to improve the understanding of the atmospheric involved processes. Nevertheless, the main limitations are related to the analytical costs and the amount of material required. Thus, this methodology is difficult to implement routinely. Such a method should be applied only on short-term campaign datasets in combination with other approaches (EC tracer or CMB) to refine and improve the SOC estimates.

Similarly, PMF-AMS/ACSM provides a good estimation of SOC, together with additional information about the SOA oxidation state and/or the volatility range. Detailed analyses can highlight additional sub-SOA fractions such as IEPOX-SOA or marine SOA [298,317–320]. More sub-SOA factors are then expected but they are difficult to be resolved due to the non-specific nature of the obtained mass fragments.

Only a detailed chemical characterization at the molecular level and the use of key species, i.e., molecular markers, can allow a definitive link between SOC content and SOA origins to be obtained. Both the SOA tracer and PMF-filter methods have been shown to be in good agreement with the other SOC apportionment methodologies, especially in the warm period in the case of the SOA tracer method. In such periods, the larger biogenic emission rates and biogenic SOA fraction is precisely considered in the SOA tracer method. This is not true for anthropogenic SOA precursors, explaining the underestimation often observed for other periods of the year, especially in winter when the SOA derived from biomass burning could be substantial but not considered in the “traditional” SOA tracer method. The lack of markers for several SOA sources or SOA precursors, notably to account for anthropogenic SOA, together with the SOA/SOC mass ratio data, are probably the main limitations of the SOA tracer method. The SOA tracer method is, in the current state of research, fully applicable in the case of biogenic influenced sampling sites but probably not recommended in the winter period or when the locations are highly impacted by anthropogenic activities. When SOA molecular markers are used in the PMF approach, the primary/secondary OA split depends on the species used. As PMF analysis requires a larger number of samples and molecular markers to get a statistically robust and meaningful solution, an expensive and exhaustive sample analysis work is required. Finally, further works are needed to identify SOA molecular markers from VOC and SVOC/IVOCS precursors and notably anthropogenic ones including markers for organosulfates, organonitrates, and multi-phase (aqueous) processes. Additionally, an experimental evaluation of the atmospheric stability of the main molecular markers is still necessary [42].
The full distinction and comprehension of OA sources are still difficult to achieve due to the complexity and the variability of the processes involved. In comparison to AMS/ACSM, the recent development of new instruments provides an online aerosol chemical characterization at the molecular level which, combined with PMF analysis, allows a better description of the OA and SOA sources. The shortcomings of low time resolution with filter measurements can be overcome using this kind of instruments to elucidate more information about the chemical processes involved. Such instrumentations include, for instance, TAG-AMS (thermal desorption aerosol gas chromatograph-AMS) [359–361], EESI-TOF-MS (extractive electrospray ionization-time of flight-MS) [362–366], or different inlets associated to PTR-MS (proton-transfer reaction-MS) or CIMS (chemical ionization MS) such as FIGAERO (filter inlet for gases and aerosols) [367], CHARON (chemical analysis of aerosol online) or thermo-desorption (TD) systems [37,368–371], which have been successfully applied to apportion SOC in details. However, these instruments are still costly and only appropriate for short term studies.

In parallel, like in the case of $^{14}$C measurements, several authors have reported the combination of different datasets from several measurement systems to refine the source apportionment of SOA using PMF. Slowik et al. [372] were the first to combine AMS and PTR-MS data measurements, highlighting the capability of PMF to resolve more SOA (OOA) factors and improving the interpretations of their sources and photochemical processes [333,372]. This kind of approach has been explored in other studies, with the combination of AMS or ACSM data with other measurements, such as ambient and thermally denuded OA spectra (TD-PMF-AMS) [373] or with single particle mass spectrometry [374], by merging high resolution mass spectra of organic and inorganic aerosols from AMS measurements [375,376] or combining offline AMS data and organic markers or $^{14}$C measurements [145,377]. In addition, the combination of PMF-ACSM outputs with inorganic species and black carbon (BC) measurements [378] or ACSM mass spectra with metal concentrations [379], allowed the source apportionment of PM rather than only OA. This further suggests that a combination of approaches could enhance the understanding about subtler differences existing in OA and SOA content by accounting for different measurements.

These improvements in the understanding of SOA origins will facilitate to bridge the existing gap between atmospheric models and field measurements [19,380–382].

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4433/9/11/452/s1, Figure SA1. Location of the monitored sites of the studies considered in this review reporting the use of the EC tracer method for the evaluation of SOC PM$_{2.5}$. In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites, Figure SB1. Location of the monitored sites of the studies considered in this review reporting the use of the CMB approach for the evaluation of SOC in PM$_{2.5}$. In black, urban sites; in blue, suburban sites; in red, rural sites and in green, remote sites, Table SA1. Reported [OC/EC] ratios in PM$_{2.5}$ for some locations worldwide, Table SA2. List of the studies considered for the annual SOC contribution to PM$_{2.5}$ OC estimated using the EC tracer method for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$, Table SA3. List of the studies considered for the spring–summer SOC contribution to PM$_{2.5}$ OC estimated using the EC tracer method for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$, Table SB1. List of the molecular markers commonly used to apportion primary aerosol sources using CMB, Table SB2. List of the studies considered for the annual SOC contribution to PM$_{2.5}$ OC estimated using CMB for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$, Table SB3. List of the studies considered for the spring–summer SOC contribution to PM$_{2.5}$ OC estimated using CMB for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$, Table SC1. Details of the SOA tracers considered in the SOA tracer method and laboratory generated aerosol mass fractions, Table SC2. List of the studies considered for the annual SOC contribution to PM$_{2.5}$ OC estimated using the SOA tracer method for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$Table SC3. List of the studies considered for the spring–summer SOC contribution to PM$_{2.5}$ OC estimated using the SOA tracer method for all the monitored sites from 2006 to 2016. SOC and POC concentrations in µgC m$^{-3}$Table SD1. List of the studies considered for the annual SOC contribution to PM$_{2.5}$ OC estimated using the PMF approach (filer based) for all the monitored sites from 2006 to 2016. SOC and OC concentrations in µgC m$^{-3}$, Table SD2. List of the studies considered for the spring–summer SOC contribution estimated using the PMF approach (AMS/ACSM based). SOC and OC concentrations in µgC m$^{-3}$, Table SE1. List of the studies considered for the spring–summer SOC contribution to PM$_{2.5}$ OC estimated 14C measurements in combination with another methodology. SOC and POC concentrations in µgC m$^{-3}$. 

Atmosphere 2018, 9, 452
Author Contributions: All authors contributed to the research and to the collaboration of this manuscript. D.S. performed the literature review and prepared the first draft of this paper. D.S. and A.A. analysed the review results and wrote the manuscript. A.A. and O.F. supervised the overall research work. A.A., O.F., E.P. and E.V. provided guidelines for the write up of the manuscript and contributed to its editing and finalization.

Funding: This was supported by the French Ministry of Environment (MTES).

Acknowledgments: The authors wish to thank Patrick Bodu for the design of the figures.

Conflicts of Interest: The authors declare no conflict of interest.

References


Atmosphere 2018, 9, 452


39. Kourtchev, I.; Warnke, J.; Maenhaut, W.; Hoffmann, T.; Claeys, M. Polar organic marker compounds in PM$_{2.5}$ aerosol from a mixed forest site in western Germany. *Chemosphere* 2008, 73, 1308–1314. [CrossRef] [PubMed]


54. Chen, J.; Ying, Q.; Kleeman, M.J. Source apportionment of wintertime secondary organic aerosol during the California regional PM$_{10}$/PM$_{2.5}$ air quality study. *Atmos. Environ.* 2010, 44, 1331–1340. [CrossRef]

55. Marta, G.V.; Manuel, S. Secondary organic aerosol formation from the oxidation of a mixture of organic gases in a chamber. *Air Qual. Tal.* 2010. [CrossRef]


78. Rutter, A.P.; Snyder, D.C.; Stone, E.A.; Shelton, B.; DeMinter, J.; Schauer, J.J. Preliminary assessment of the anthropogenic and biogenic contributions to secondary organic aerosols at two industrial cities in the upper Midwest. *Atmos. Environ.* 2014, 84, 307–313. [CrossRef]  
84. Tsapakis, M.; Stephanou, E.G. Collection of gas and particle semi-volatile organic compounds: Use of an oxidant denuder to minimize polymeric aromatic hydrocarbons degradation during high-volume air sampling. *Atmos. Environ.* 2003, 37, 4935–4944. [CrossRef]  
90. Yu, S.; Dennis, R.L.; Bhave, P.V.; Eder, B.K. Primary and secondary organic aerosols over the United States: Estimates on the basis of observed organic carbon (OC) and elemental carbon (EC), and air quality modeled primary OC/EC ratios. *Atmos. Environ.* 2004, 38, 5257–5268. [CrossRef]  


97. Lonati, G.; Ozgen, S.; Giugliano, M. Primary and secondary carbonaceous species in PM$_{2.5}$ samples in Milan (Italy). *Atmos. Environ.* 2007, 41, 4599–4610. [CrossRef]


109. Mancilla, Y.; Herckes, P.; Fraser, M.P.; Mendoza, A. Secondary organic aerosol contributions to PM$_{2.5}$ in Monterrey, Mexico: Temporal and seasonal variation. *Atmos. Res.* 2015, 153, 348–359. [CrossRef]


120. Seguel, R. Estimations of primary and secondary organic carbon formation in PM$_{2.5}$ aerosols of Santiago City, Chile. *Atmos. Environ.* 2009, 43, 2125–2131. [CrossRef]


122. Favez, O.; Cachier, H.; Sciare, J.; Le Moullec, Y. Characterization and contribution to PM$_{2.5}$ of semi-volatile aerosols in Paris (France). *Atmos. Environ.* 2007, 41, 7969–7976. [CrossRef]

123. Laongsri, B.; Harrison, R.M. Atmospheric behaviour of particulate oxalate at UK urban background and rural sites. *Atmos. Environ.* 2013, 71, 319–326. [CrossRef]


141. Choi, J.-K.; Heo, J.-B.; Ban, S.-J.; Yi, S.-M.; Zoh, K.-D. Chemical characteristics of PM$_{2.5}$ aerosol in Incheon, Korea. *Atmos. Environ.* **2012**, *60*, 583–592. [CrossRef]


143. Khan, M.F.; Shirasuna, Y.; Hirano, K.; Masunaga, S. Characterization of PM$_{2.5}$, PM$_{2.5-10}$ and PM$_{10}$ in ambient air, Yokohama, Japan. *Atmos. Res.* **2010**, *96*, 159–172. [CrossRef]


Atmosphere 2018, 9, 452


175. Wang, F.; Guo, Z.; Lin, T.; Rose, N.L. Seasonal variation of carbonyl pollutants in PM\textsubscript{2.5} at an urban ‘supersite’ in Shanghai, China. *Chemosphere* 2016, 146, 238–244. [CrossRef] [PubMed]
176. Duan, J.; Tan, J.; Cheng, D.; Bi, X.; Deng, W.; Sheng, G.; Fu, J.; Wong, M. Sources and characteristics of carbonaceous aerosol in two largest cities in Pearl River Delta Region, China. *Atmos. Environ.* 2007, 41, 2895–2903. [CrossRef]
183. Kumar, A.; Ram, K.; Ojha, N. Variations in carbonaceous species at a high-altitude site in western India: Role of synoptic scale transport. *Atmos. Environ.* 2016, 125 Pt B, 371–382. [CrossRef]
186. Pipal, A.S.; Gursumeeran Satsangi, P. Study of carbonaceous species, morphology and sources of fine (PM\textsubscript{2.5}) and coarse (PM\textsubscript{10}) particles along with their climatic nature in India. *Atmos. Res.* 2015, 154, 103–115. [CrossRef]
187. Safai, P.D.; Raju, M.P.; Rao, P.S.P.; Pandithurai, G. Characterization of carbonaceous aerosols over the urban tropical location and a new approach to evaluate their climatic importance. *Atmos. Environ.* 2014, 92, 493–500. [CrossRef]
191. Watson, J.G.; Cooper, J.A.; Huntszicker, J.J. The effective variance weighting for least squares calculations applied to the mass balance receptor model. *Atmos. Environ.* 1984, 18, 1347–1355. [CrossRef]
208. Minguillon, M.C.; Arhami, M.; Schauer, J.J.; Sioutas, C. Seasonal and spatial variations of sources of fine and quasi-ultrafine particulate matter in neighborhoods near the Los Angeles–Long Beach harbor. *Atmos. Environ.* 2008, 42, 7317–7328. [CrossRef]


217. Xu, L.; Pye, H.O.T.; He, J.; Chen, Y.; Murphy, B.N.; Ng, N.L. Large contributions from biogenic monoterpenes and sesquiterpenes to organic aerosol in the Southeastern United States. *Atmos. Chem. Phys. Discuss.* 2018, 2018, 1–47. [CrossRef]


224. Pirovano, G.; Colombi, C.; Balzarini, A.; Riva, G.; Gianelle, V.; Lonati, G. PM$_{2.5}$ source apportionment in Lombardy (Italy): Comparison of receptor and chemistry-transport modelling results. *Atmos. Environ.* 2015, 106, 56–70. [CrossRef]

225. Yin, J.; Harrison, R.M.; Chen, Q.; Rutter, A.; Schauer, J.J. Source apportionment of fine particles at urban background and rural sites in the UK atmosphere. *Atmos. Environ.* 2010, 44, 841–851. [CrossRef]


261. Carlton, A.G.; Turpin, B.J.; Altieri, K.E.; Seitzinger, S.; Reff, A.; Lim, H.-J.; Ervans, B. Atmospheric oxalic acid and SOA production from glyoxal: Results of aqueous photooxidation experiments. *Atmos. Environ.* 2007, 41, 7588–7602. [CrossRef]


263. Lim, Y.B.; Tan, Y.; Turpin, B.J. Chemical insights, explicit chemistry, and yields of secondary organic aerosol from OH radical oxidation of methylglyoxal and glyoxal in the aqueous phase. *Atmos. Chem. Phys.* 2013, 13, 8651–8667. [CrossRef]


Atmosphere 2018, 9, 452


291. Hantson, S.; Knorr, W.; Schurgers, G.; Pugh, T.A.; Armet, A. Global isoprene and monoterpenes emissions under changing climate, vegetation, CO$_2$ and land use. Atmos. Environ. 2017, 155, 35–45. [CrossRef]


327. Zhang, Y.; Sheesley, R.J.; Bae, M.-S.; Schauer, J.J. Sensitivity of a molecular marker based positive matrix factorization model to the number of receptor observations. Atmos. Environ. 2009, 43, 4951–4958. [CrossRef]
331. Li, Y.J.; Lee, B.P.; Su, L.; Fung, J.C.H.; Chan, C.K. Seasonal characteristics of fine particulate matter (PM) based on high-resolution time-of-flight aerosol mass spectrometric (HR-ToF-AMS) measurements at the HKUST Supersite in Hong Kong. Atmos. Chem. Phys. 2015, 15, 37–53. [CrossRef]


375. Sun, Y.; Zhang, Q.; Schwab, J.; Yang, T.; Ng, N.; Demerjian, K. Factor analysis of combined organic and inorganic aerosol mass spectra from high resolution aerosol mass spectrometer measurements. Atmos. Chem. Phys. 2012, 12, 8537–8551. [CrossRef]


