UNIVERSITÀ DEGLI STUDI DI MILANO

Department of Physics

PhD School in Physics, Astrophysics and Applied Physics

Cycle XXXIII

DEVELOPMENT AND OPTIMISATION OF EXPERIMENTAL AND MODELLING APPROACHES TO CHARACTERISE HIGH-TIME RESOLUTION ATMOSPHERIC AEROSOL AND ITS SOURCES

Disciplinary Scientific Sector FIS/07

PhD Thesis of:

Alice Corina Forello

Director of the School: Prof. Matteo Paris

Supervisor of the Thesis: Prof. Roberta Vecchi

A.Y. 2020-2021

To my parents

Contents

List of Figures

List of Tables

Abstract

Atmospheric aerosol impacts on local, regional, and global scale causing adverse effects on human health, affecting visibility, and influencing the climate. For this reason, the scientific community is strongly interested in the physical-chemical characterisation of aerosol and its emission sources. Thanks to technological improvements in this field, high time resolution measurements and analyses have become increasingly important since processes involved in aerosol emission, transformation and removal in the atmosphere are subject to short time scales (in the order of one hour). The research presented in this PhD thesis mainly focuses on the implementation of modelling and experimental approaches in order to expand the knowledge about properties of atmospheric aerosol and its sources with high time resolution.

Main PhD activities are shortly summarised in the following:

- A source apportionment study was performed on a dataset with different time resolutions (24, 12, and 1 hour) collected in Milan (Italy) in 2016. This advanced multi-time resolution approach – implemented through the Multilinear Engine algorithm – is still scarcely available in the literature, although it allows to get rid of the limited chemical characterisation typical of high-time resolution data and the poor temporal details of low-time resolution samples. In addition, as an original contribution, in this source apportionment study chemical variables were joined to the aerosol absorption coefficient measured at different wavelengths as input to the model. This original approach was proved effective in order to (1) strengthen source identification; (2) retrieve source-dependent optical absorption parameters, i.e. source-specific absorption Ångström exponents and mass absorption cross sections at different wavelengths, as results of the model. It is noteworthy that, at the state of the art, in source apportionment models based on optical absorption data (e.g. Aethalometer model) values for the absorption Ångström exponents are fixed a priori by the modeller, thus carrying a large part of uncertainties in the model results. Results from this activity are reported in [P4] [IO1] [IP1] [NO4].

- In the frame of the international collaborative project CARE (Carbonaceous Aerosol in Rome and Environs), a high time resolution (one and two hours) dataset collected in Rome (Italy) in 2017 was used as input in an advanced receptor model. Different measurement techniques provided the optical (absorption and scattering coefficients) and chemical characterisation (elements, elemental and organic carbon, non-refractory components such as organic aerosol, nitrate, sulphate, ammonium) of atmospheric aerosol. In particular, an ACSM (Aerosol Chemical Speciation Monitor) detected the organic aerosol (OA) fraction. Results from the source apportionment analysis of this high time resolution dataset were a posteriori compared to ACSM separation of the organic fraction in terms of HOA (hydrocarbon-like organic aerosol), BBOA (biomass burning-like organic aerosol), and OOA (oxygenated organic aerosol) provided in a previous literature work. In this study, the original contribution consisted in analysing the whole dataset with a multi-time resolution and a multi-variable approach, by the application of the Multilinear Engine algorithm. This approach based on receptor modelling resulted to be effective in relating primary and secondary OA contributions to their emission sources, highlighting the possibility to obtain a source-dependent separation of the OOA fraction, which is typically associated in the literature to not-well specified secondary processes. This is of particular interest for the receptor modelling community, since the assessment of the origin of secondary compounds is one of the main limitations of this type of models. Additionally, since in this study also the optical absorption coefficient retrieved at 7 wavelengths by an Aethalometer was used as input to the model, the methodology proposed in [P4] was further tested on a different site impacted by different sources. It allowed e.g. to retrieve optical absorption contribution from mineral dust, besides the typical fossil fuel and biomass burning contributions retrieved by more widespread models based on optical absorption data such as the Aethalometer model. Results from this

activity can be found in [P2].

- Contribution to the INFN (National Institute of Nuclear Physics) experiment TRACCIA (Time Resolved Aerosol Characterization Challenging Improvements and Ambitions), devoted to the realisation of the new high time resolution sampler STRAS (Size and Time Resolved Aerosol Sampler), in collaboration with other Italian research groups (INFN-Florence, INFN-Genoa, INFN-Lecce). The contribution of this PhD activity was in the sampler design phase (e.g. sizing of the sampler characteristics to obtain the proper cut-off diameter), and in the preliminary testing phase to verify the collection efficiency of the sampler. Preliminary tests were performed both on field and in the atmospheric simulation chamber ChAMBRe (Chamber for Aerosol Modelling and Bio-aerosol Research, partner of the H2020 EUROCHAMP2020 project and member of the Joint Research Unit ACTRIS-IT), where particles with certified dimensions were injected to study STRAS experimental cut-off diameter. Details can be found in [NO1].

It is worth mentioning that during the PhD research, six months (from March to August 2019) were spent at the EGAR (Environmental Geochemistry and Atmospheric Research) group of the Institute of Environmental Assessment and Water Research of the Spanish National Council (IDAEA – CSIC) in Barcelona (Spain), with a fellowship obtained within the Erasmus+ Traineeship programme. The aim was the collaboration with the EGAR research group (in particular with Dr. Fulvio Amato) on receptor modelling approaches, on which Dr. Amato can be considered one of the leading international experts. The shared expertise was useful in the source apportionment analyses performed during my PhD activity (see e.g. [P2]), and a collaborative work on a high time resolution dataset collected in different traffic sites in Barcelona has been started and it is still in progress. Moreover, an application of the original approach first proposed in publication [P4] to data collected with daily resolution is now under investigation in the frame of a collaboration with the Air Quality research group of the Department of Engineering and Nuclear Science of the Instituto Superior Técnico of Lisbon (Portugal).

In addition to the above mentioned activities, collaboration to other works of the Environmental Physics research group of the University of Milan was carried out. Examples are the study of aerosol optical absorption properties by online and offline instrumentation, the analysis of size-segregated samples collected by impactors, the assessment of aerosol residence time by means of radioactive tracers; some of them were carried out in the frame of various national and international projects. Main results are reported in [P1] [P3] [V1] [S1] [S2] [IO2] [IO3] [IO4] [IP2] [NO2] [NO3] [NO5].

Introduction

Motivation

The abatement of pollutant emissions is one of the most important principles of the Thematic Strategy of Air Pollution of the European Commission policy. Source apportionment (SA), i.e. identification and quantification of pollution sources, is mandatory to implement Air Quality Directives and to establish mitigation strategies on concentration levels of pollutants. The harmonisation of SA methods is the aim of the FAIRMODE (Forum for Air quality Modeling) network chaired by the European Commission Joint Research center. Among atmospheric components, aerosol is of particular interest due to its impact on environment, climate, and human health. Receptor modelling approaches based on factor analysis have been extensively used to perform SA of atmospheric aerosol; this methodology has the advantage of exploiting information from real-world measurements, but one of its main issues is that the origin of secondary compounds in the atmosphere is difficult to assess.

Routine sampling of atmospheric aerosol on filters is usually carried out with a time resolution of 24 hours, collecting enough material to perform full aerosol chemical characterisation and quantify also aerosol components with very low concentrations (e.g. trace elements). However aerosol emission, transformation, and removal processes are related to shorter time scales ($\sim$hour); reducing sampling duration was proved effective to improve the ability of receptor models to resolve sources, and allow the identification of episodic sources that might not negligibly impact on a short time interval. Over the last decades, the development of different instrumentations with very high temporal resolution (down to one minute) has provided an increasing detailed chemical and physical characterisation of atmospheric aerosol. From the experimental point of view, reliable instrumentation is crucial for advancements in atmospheric aerosol studies.

Large and fully chemically characterised datasets contain valuable information to be exploited in SA studies. However, the combination of heterogenoeus data obtained from different types of instrumentation to be used as input to the receptor models is not straightforward, and it might be critical since it impacts on the output of the modelling procedure. Advanced receptor models have been proposed in the literature, although they are not commonly used yet. Aim of these advanced approaches is to make the most of available datasets, e.g. exploiting fully chemically characterised low time resolution data together with high time resolution ones as input to the model (this is the case of the multi-time resolution model implemented in this work).

At the state of the art, the majority of SA studies of atmospheric aerosol are based on dataset comprising only chemical composition. Very few examples of successful applications using joint matrices, i.e. input datasets containing also variables related to aerosol properties different from the chemical composition such as e.g. the particle number size distribution, are available in the literature. Aerosol optical properties have a crucial role in the Earth radiative budget and on visibility, and thanks to recent advancements in instrumentation, they can be continuously measured with high time resolution; however, their combination with chemical composition data in a joint matrix to retrieve additional information on emission sources is lacking at present.

In this thesis, experimental and modelling approaches were developed and optimised in order to improve the knowledge on atmospheric aerosol and its emission sources. A new high time resolution sampler was characterised, and datasets comprising chemical and light absorption properties of aerosol provided by different instrumentations were explored, in order to gain new insights on aerosol sources and to advance in the formulation of new modelling approaches.

Thesis overview

In Chapter 1 a brief introduction about atmospheric aerosol is outlined, in particular regarding its main emission sources and properties such as size, chemical composition, and light absorption.

Chapter 2 thoroughly describes the principles of receptor modelling, starting from its basic bilinear formulation. The chapter is then focused on the Positive Matrix Factorization (PMF) and in particular on the Multilinear Engine (ME-2) program, that is used in this thesis. ME-2 was developed in the late 1990s to solve general multilinear problems, and this feature can be exploited to implement advanced receptor modelling approaches through scripts. Theoretical generalisation to multilinear problems is given, and the rotational indeterminacy of factor analysis is discussed. The possibility to add a priori information in the source apportionment analysis, and the different methods for uncertainty estimation in the modelled solution are explored, since these two features will be exploited in the case-studies developed in this thesis and reported in Chapter 3. An overview of the general modelling procedure with PMF is then provided, concerning preliminary checks on data robustness, data preparation for model input, ME-2 script preparation, basic evaluation and interpretation of the modelled solution. In Sect. 2.6 advanced receptor modelling approaches are treated, and equations at the basis of the advanced multi-time resolution model - that is the one used in this thesis - are reported together with specific details about additional procedures for input data and ME-2 script preparation. Finally, the state of the art on the task of modelling joint matrices, i.e. input datasets containing also variables different from the chemical composition, is discussed.

Chapter 3 reports the implementation and application of the advanced receptor modelling approaches developed in this thesis. Results of two case-studies are presented: one from a measurement campaign performed during summertime and wintertime 2016 in Milan, and the other carried out during February 2017 in Rome in the frame of the CARE (Carbonaceous Aerosol in Rome and Environs) international project. Input datasets consist in aerosol chemical composition data measured with different time resolutions and different instrumentations, and joint with aerosol light absorption data at different wavelengths.

Chapter 4 focuses on the development and characterisation of the new high time resolution atmospheric aerosol sampler STRAS (Size and Time Resolved Aerosol Sampler), that was carried out in the frame of the INFN (National Institute of Nuclear Physics) experiment TRACCIA (Time Resolved Aerosol Characterization Challenging Improvements and Ambitions) in collaboration with other Italian research groups. STRAS development and main experimental results from its preliminary characterisation in the atmospheric simulation chamber ChAMBRe (Chamber for Aerosol Modelling and Bio-aerosol Research) are presented. Theoretical background about sampler design is reported in Appendix A.

Chapter 1

Atmospheric aerosol: basic properties

Atmospheric aerosol (in the following also referred to as particulate matter, PM) is the polydisperse ensemble of solid and liquid particles suspended in the atmosphere; they are supposed to maintain their physical-chemical properties long enough to be observed and measured. Literature studies about atmospheric aerosol have increased enormously over the past decades [1] due to evidence of its impact on local, regional, and global scale on environment, climate, and human health. In particular [2]:

- aerosol particles influence visibility and radiative forcing of the climate system through direct effects, i.e. by scattering and absorption of light, and indirect effects on clouds and precipitation resulting from aerosol activity as cloud condensation nuclei (CCN) or ice nuclei (IN);

- besides air quality, PM has an impact on soil and water through dry (i.e. convective transport, diffusion, and adhesion) and wet (i.e. precipitation) deposition processes; moreover, aerosol particles can threaten cultural heritage;

- aerosol particles can catalyse chemical reactions in the atmosphere, and play a role in heterogeneous chemical reactions;

- epidemiological studies have shown correlation between PM and severe health effects including enhanced mortality, cardiovascular, respiratory, and allergic diseases. Very small particles are suspected to be particularly hazardous since they can penetrate the membranes of the respiratory tract, enter the blood circulation, and be transported in other parts of the body. Moreover, particles play an important role in the spreading of pathogens such as viruses and bacteria.

Atmospheric aerosol is a very complex system, since it is generated by a variety of sources and it is characterised by heterogeneous chemical composition, size, and shape; moreover, its properties are highly variable from a spatial and temporal point of view, and strongly interconnected (Fig. 1.1).
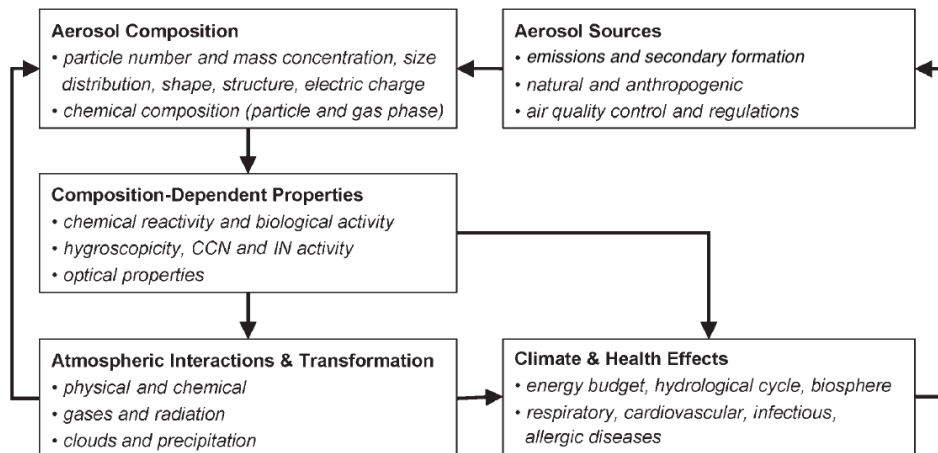
Figure 1.1: Overview of interconnection and feedback between atmospheric aerosol sources, composition, properties, interactions and transformation, and climate and health effects. Figure modified from [2].

Processes involved in aerosol emission, transformation and removal take place on short time scales, of the order of the hour; for this reason, high time resolution measurement and physical-chemical characterisation are primarily important to study in detail particle properties. In this research frame, it is thus mandatory to develop experimental and modelling methodologies aiming at identifying main sources and their impact in the atmosphere.

1.1 Size

Particle size is one of the most important physical quantity in aerosol science, since the laws governing particle behaviour in the atmosphere depends on it [3]. Atmospheric particles are characterised by different size and shape, typically dependent on the formation mechanism. In general, particles emitted by natural sources (e.g. biogenic emissions) have a variety of shapes and large size, while particles originated from anthropogenic emissions such as combustion processes at high temperature, have more regular shape and smaller size (some examples are reported in Fig. 1.2).
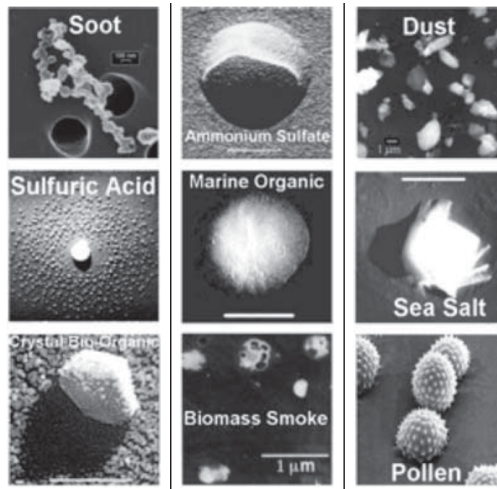
Figure 1.2: Morphology of various aerosol particles. Figure adapted from [4].

In aerosol science, theories are often developed in the approximation of spherical particles; the use of "equivalent diameters" enables to consider also particles with different shape and density. An equivalent diameter is the diameter of a spherical particle that is characterised by the same value of a particular physical property as that of an irregular particle [3] e.g. the aerodynamic diameter d_{ae} is the equivalent diameter of a spherical particle with unit density ($1 \ g \ cm^{-3}$) and the same inertial properties (i.e. the same terminal velocity, see Sect. A.1 in the Appendix) of the real particle.

In the atmosphere particles are polydisperse, with size spanning over different orders of magnitude (from few nm to tens of μm); for this reason, it is necessary to characterise their size distribution from a statistical point of view. Atmospheric size distribution can be empirically represented by the sum of log-normal distributions (also called modes in the literature). A log-normal distribution is defined as:

$$\frac{dN}{dlnd_p} = \frac{N}{\sqrt{2\pi}ln\sigma_g}exp[-\frac{1}{2}(\frac{lnd_p - ln\bar{d_{pg}}}{ln\sigma_g})^2] \qquad (1.1)$$

where N is the total number of particles for cm^3, d_p is the particle diameter, d_{pg} is the geometric mean diameter, and σ_g is the geometric standard deviation. This expression gives the number of particles having diameters lying between lnd_p and $lnd_p + dlnd_p$. Another example is the size distribution by volume (V), that can be calculated starting from the size distribution by number, considering that (in the approximation of spherical particles):

$$\frac{dV}{dlnd_p} = \frac{dN}{dlnd_p}\frac{\pi}{6}d_p^3 \qquad (1.2)$$

In Fig. 1.3, a comparison between size distribution by number and by volume is reported. As can be seen, atmospheric aerosol size distribution can be characterised by a structure with four main modes (and two sub-modes) [5]:

- Nucleation mode ($\simeq 0.001 - \simeq 0.01\mu m$): particles in this size range are freshly emitted and produced in situ from the gas phase by nucleation. Their diffusion is mainly due to Brownian motion, and their removal consists in coagulation into larger particles;

- Aitken mode ($\simeq 0.01 - \simeq 0.1\mu m$): particles are mainly produced by high temperature combustion processes, and after coagulation of smaller particles. Diffusion is mainly due to Brownian motion, and coagulation into larger particles is still an important mechanism;

- Accumulation mode ($\simeq 0.1 - \simeq 2\mu m$): particles have different origin, as they can result from primary emissions, condensation, and coagulation. They have the highest residence time in the atmosphere, since neither Brownian motion nor turbulent diffusion nor gravitational settling are particularly efficient in this size range. Their removal is mainly due to wet deposition. The accumulation mode might consist of two overlapping submodes: the condensation mode and the droplet mode;

- Coarse mode ($\simeq 2 - \simeq 50\mu m$): particles have usually (primary) natural origin and are produced by mechanical processes (e.g. erosion). Their diffusion is usually negligible, and their removal is mainly due to gravitational settling.
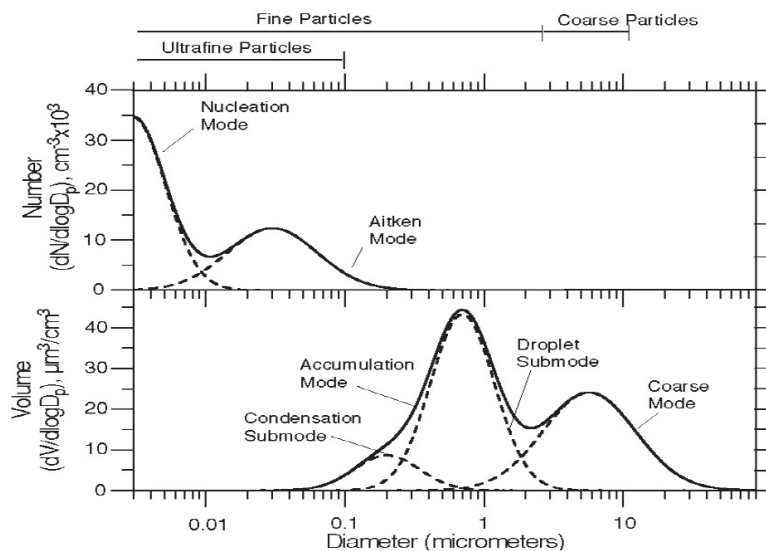
Figure 1.3: Example of particle size distribution by number and volume. Dashed lines refer to individual modes, solid lines to their sum. Figure from [6].

1.2 Chemical composition and sources

Aerosol chemical composition is strongly influenced by emission sources. Particles can be directly emitted in the atmosphere (these are referred to as primary aerosol), or formed in the atmosphere through gas-to-particle conversion processes (these are called secondary aerosol) [5]. Emission sources can be classified as natural sources and anthropogenic sources; in the following, the most relevant atmospheric aerosol sources on a global scale are described (the following description is based on [1] [7] and references therein).
Among natural sources:

- *Marine aerosol.* Aerosol deriving from marine environment accounts for a large fraction of primary natural aerosol. Primary marine aerosol is generated by bubble bursting from breaking waves, and by the action of surface wind (in fact its production depends on wind speed). Marine aerosol generally contributes to the coarse aerosol fraction, but smaller particles are also present; while large particles likely deposit close to the production site, small particles (in the range 0.1 - 1 μm, also called "film drops") can have a long atmospheric lifetime and they can be transported over large distances. Marine aerosol is of interest also because of its direct (through scattering) and indirect (acting as cloud condensation nuclei - CCN - and ice nuclei - IN) impact on the atmospheric radiative budget.

Table 1.1: Main constituents of bulk sea water composition. Table adapted from [5].

Species	Percentage by weight
Cl	55.04
Na	30.61
SO_4^{2-}	7.68
Mg	3.69
Ca	1.16
K	1.1
Br	0.19
C (noncarbonate)	$3.5 \cdot 10^{-3}$ - $8.7 \cdot 10^{-3}$

Marine aerosol is formed mainly by Na and Cl, with minor contributions from SO_4^{2-}, Mg, Ca, and K (Table 1.1). Part of the Cl in the fine particles of NaCl can transform into gaseous phase due to reactions with sulphuric acid (in gaseous or aqueous phase) and nitric acid (in gaseous phase):

$$NaCl(s) + HNO_3(g) \rightarrow HCl(g) + NaNO_3(s)$$

$$2NaCl(s) + H_2SO_4(g + aq) \rightarrow 2HCl(g) + Na_2SO_4(s)$$
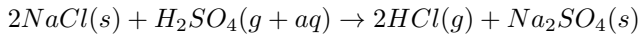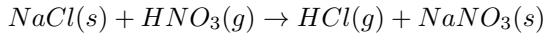
These atmospheric transformations lead to chloride deficit in the particulate phase, and they might decrease the atmospheric Cl/Na ratio respect to the composition of bulk sea water. Apart from sea salt, marine aerosol comprises also organic components mostly originated from degradation of marine organisms and plants. Two different processes for the generation of organic components in the atmosphere have been proposed: (1) incorporation of organic matter into primary marine aerosol produced by breaking waves; (2) gas-phase oxidation (by e.g. $\cdot$OH and ozone) of volatile organic compounds such as e.g. dimethyl sulphide (DMS). DMS is emitted by phytoplankton activity and is considered one of the most significant precursors of atmospheric sulphate in oceanic regions.

- *Mineral dust.* "Mineral dust" is used here to denote the fraction of airborne aerosol deriving from the soil and consisting of inorganic material [8]. The aerosol mineral fraction (or crustal fraction) is mainly generated by the action of wind on Earth's surface. On a global scale, main emission sources are deserts (the largest source of mineral dust in the atmosphere is the Sahara), but any type of soil can be a potential source. Factors influencing dust emission are soil surface (i.e. texture and roughness), soil moisture, vegetation cover, and meteorological parameters such as wind speed and precipitation. Mineral dust aerosol impacts the climate through direct (mineral dust particles scatter light, although some of their iron compounds absorb visible radiation) and indirect (action as CCN and IN) effects. In general, these particles are mainly constituted by calcite, quartz, dolomite, clays (especially kaolinite and illite), feldspar and smaller amounts of calcium sulphate and iron oxides such as e.g. hematite (elements associated to these compounds are Si, Al, K, Na, Ca, Fe, and trace elements such as Ti, Ba, Sr, Rb, Li, Zr); anyway, the chemical composition strongly depends on soil characteristics. Even if most of the mineral dust aerosol is expected in the coarse fraction, it has been estimated that between 7% and 20% of mineral dust particles have a diameter lower than 1 μm; in particular, particles in the size range 0.1 - 5 μm are characterised by long atmospheric residence time, with the possibility to be transported for hundreds of kilometers. In the Mediterranean region, Saharan dust transport is responsible for exceedances of PM10 daily limit (e.g it was estimated that more than 70% of exceedances at rural background sites in Spain are due to dust outbreaks).

- *Biogenic emissions.* Vegetation and some types of microorganisms contribute to primary and secondary biogenic emissions. Primary biogenic aerosol (PBA) is characterised by diameters up to 100 μm including pollen, fern spores, and fungal spores; particles with diameters smaller than 10 μm comprise small fragments and excretions from plants, animals, bacteria, viruses, carbohydrates, proteins, waxes, etc.. Typical size range for some PBA is $0.05 - 0.15$ μm for viruses, $0.1 - 4$ μm for bacteria, $0.5 - 15$ μm for fungal spores, and 10 - 30 μm for pollen. Examples of PBA tracers are mannitol and arabitol. PBA may also

act as CCN and IN, and they can be vectors for the trasmission of plant, animal, and human diseases. In addition, volatile organic compounds (e.g. isoprene and terpene) emitted by the biosphere may act as precursors of secondary organic aerosol.

- *Volcanic eruptions.* Volcanic emissions contribute to tropospheric and stratospheric concentration increase of greenhouse gases, sulphur, and aerosol, possibly causing acid rain and affecting the climate, air traffic, and public health. Main gaseous emissions are H_2O, CO_2, SO_2, and HCl, but emissions are difficult to predict since they depend on the eruption characteristics. Volcanic emissions comprise primary and secondary aerosol; the latter is e.g. sulphate resulting from SO_2 oxidation. Variable concentration of Al, Si, S, Cl, K, Ca, Ti, Mn, Fe, Cu, and Zn are part of the chemical composition of particles originated by volcanic eruptions. The size interval of particles released by volcanic plumes is very large, but volcanic ashes generally fall in the range $1 - 10 \ \mu m$. During strong eruptions aerosols may reach the stratosphere, where average residence time is 1-2 years, and strongly impact the climate worldwide; when injection impacts only the troposphere, the atmospheric average residence time is about 1 week.

Among anthropogenic sources:

- *Traffic.* Especially in urban areas and densely populated regions, transport-related aerosol is one of the main sources of primary and secondary anthropogenic emissions. Traffic emissions can be divided in exhaust, i.e. tailpipe emissions, and non-exhaust, i.e. from vehicle brake wear, tyre wear, road surface abrasion and resuspension. Although the correct assessment of the different contributions to non-exhaust emissions in the atmosphere is still an open issue, in the literature they are estimated to contribute in total as much as exhaust emissions. Opposite to exhaust emissions, non-exhaust ones have not been subject to regulation policies up to now. Exhaust particles result from fossil fuels combustion, and are mainly constituted by carbonaceous compounds; among them, black carbon (BC) is the most efficient absorber of visible light. Non-exhaust particles are enriched in metals, mineral elements, and contain less carbonaceous material. For example, elements found in brake and tyre wear are Cu, Fe, Ba, Cr, Mn, Zn, Sr, Cd, Mo, Sb. Road dust resuspension also contributes to non-exhaust emissions, especially during winter months in Northern Europe where sand and salt are used to avoid pavement freezing. Non-exhaust particles are generally in the coarse fraction, but particles smaller than $1 \ \mu m$ were also found in emissions from brake and tyre wear.
  It is noteworthy that traffic is a relevant source of nitrogen oxides that act as precursors e.g. of nitrate in aerosol particles. Air traffic and maritime traffic are also contributors to aerosol loading in the atmosphere; for example ship emissions - besides their contribution to secondary sulphate formation - can be traced by products of oil combustion (e.g. V and Ni).

- *Biomass burning.* Biomass burning is both a natural (wildfires) and anthropogenic (agricultural burning, domestic heating) source of atmospheric aerosol, and its emissions strongly depend on combustion conditions. Main aerosol components are carbonaceous compounds (mainly organic carbon, OC, and in smaller amounts elemental carbon, EC), trace of various inorganic compounds originating by dust and ashes, e.g. soluble salts of potassium, ammonium, sulphate, and nitrate. Other compounds are formed from the pyrolisis of cellulose, such as levoglucosan. A fraction of organic aerosol emitted by biomass burning, i.e. Brown Carbon (BrC), can efficiently absorb light in the visible and ultraviolet spectra. Household stoves for domestic heating account for an important fraction of atmospheric biomass burning emissions; factors influencing this type of emission are stove design, operating conditions, combustion conditions, wood characteristics (e.g. more sophisticated stoves release less aerosol mass respect to fireplaces, but with a higher percentage of elemental carbon and inorganic compounds). Most of the particles produced during biomass burning are in the accumulation mode.

- *Industrial activities and combustions.* Industrial activities emitting primary aerosol or precursors of secondary aerosol in the atmosphere are highly variable; emissions strongly depend on the production process, the technology, and the materials used. For example, steelworks are important emitters of metal pollutants such as Fe, Zn, and Mn. Energy production worldwide is also still dependent on fossil fuels combustion; e.g. coal burning mainly emits sulphur, carbon, chloride, and metals. Residential coal combustion is currently a serious problem especially in developing countries, due to the emission of toxic components such as polycyclic aromatic hydrocarbons (PAHs).

Secondary aerosol forms after emission of precursors, i.e. SO_2, NOx, NH_3, and volatile organic compounds ($VOCs$), by both natural and anthropogenic sources. Main sources of SO_2 and NOx are sulphur-containing fuel and fossil fuels combustion occurring at high temperature. The production of ammonium nitrate depends on the availability of NH_3, which is mainly emitted by agricultural activities (e.g. fertilisation processes). NH_3 emission reduction has been proposed as a measure to control the formation of secondary inorganic aerosol (SIA); on the other side, the formation of secondary organic aerosol (SOA) is still poorly understood in terms of anthropogenic and biogenic precursors and formation processes. As a result, although SOA generally accounts for a relevant fraction of total organic aerosol concentration in the atmosphere, its source attribution is still an open issue [1].

1.3 Light absorption

Interaction between aerosol and radiation has been thoroughly studied, and the complete theoretical background can be found in [3] [5] [9]. To the purpose of this thesis work, theory of aerosol light absorption will be briefly outlined in the following [10].

Aerosol light extinction, i.e. the removal of light from a beam, is caused by both aerosol scattering, i.e. redistribution of light into different directions, and aerosol absorption, i.e. conversion of light into thermal energy. To study interaction of a spherical particle with radiation, the size parameter x is introduced:

$$x = \frac{2\pi r}{\lambda} \tag{1.3}$$

where r is the particle radius and λ is the radiation wavelength. Particle light scattering and absorption properties are defined through the complex refractive index m:

$$m = n - ik \tag{1.4}$$

where n and k are respectively the real and imaginary part of the refractive index. All particles scatter light, but particles with an imaginary part k almost equal to zero (i.e. $k < 10^{-5}$ [11]) do not absorb light significantly.

Absorption by an individual particle can be characterised through the absorption cross section σ_{ap} (common unit: m^2) and the absorption efficiency Q_{ap}, that is the (adimensional) ratio between absorption and geometric (σ_{geo}) cross sections:

$$Q_{ap} = \frac{\sigma_{ap}}{\sigma_{geo}} \quad (for\ spherical\ particles\ \sigma_{geo} = \pi r^2) \tag{1.5}$$

For an ensemble of N_{tot} particles in a volume V, the absorption coefficient b_{ap} (common unit: Mm^{-1}) can be defined as the sum of the absorption cross section $\sigma_{ap,i}$ of each particle i, divided by the volume:

$$b_{ap} = \frac{\sum_{i=1}^{N_{tot}} \sigma_{ap,i}}{V} \tag{1.6}$$

For identical particles, Eq. (1.6) can be written as $b_{ap} = N \cdot \sigma_{ap}$, where N is the particle number density $N = \frac{N_{tot}}{V}$.

The single scattering albedo ω of an individual particle is the ratio between the scattering cross section σ_{sp} and the extinction cross section σ_{ep}, that is $\sigma_{ep} = \sigma_{ap} + \sigma_{sp}$. For an ensemble of particles, the scattering, absorption, and extinction coefficients (b_{sp}, b_{ap}, and b_{ep}, respectively) are taken into account:

$$\omega = \frac{b_{sp}}{b_{ep}} = \frac{b_{sp}}{b_{sp} + b_{ap}} \tag{1.7}$$

ω ranges from 0 for completely absorbing (or black) particles to 1 for completely scattering (or white) particles. Power law relations can be used to express the wavelength dependence of aerosol optical properties such as scattering, absorption, or ω. For example, for the absorption coefficient b_{ap} yields:

$$\frac{b_{ap}(\lambda_1)}{b_{ap}(\lambda_2)} = \left(\frac{\lambda_1}{\lambda_2}\right)^{-\alpha} \tag{1.8}$$

where α is the absorption Ångström exponent. In the literature, α values are used to derive information about the type of absorbing aerosol; atmospheric aerosol absorbing electromagnetic radiation in the ultraviolet (UV)

and visible region is constituted by black carbon, brown carbon, and mineral dust.

Black carbon

BC is defined as carbonaceous material characterised by a significant non-zero imaginary part k of the complex refractive index (see Eq. (1.4)) that is wavelength independent over visible and near-visible spectral regions ($k \simeq 0.79$). The constant value of k results in an absorption coefficient b_{ap} that is inversely proportional to the wavelength - i.e. α in Eq. (1.8) is equal to 1 - in the case of small BC particles in the Rayleigh regime (i.e. when in Eq. (1.3) $x << 1$). These particles tend to agglomerate into a fractal-like morphology, and the connection of this morphology to optical properties is not straightforward. BC is formed during incomplete combustion of gaseous hydrocarbons in high-temperature combustion. In the literature, carbonaceous material that partially corresponds to the BC fraction can be indicated with other names depending on the analytical technique considered; for example, elemental carbon (EC) is operationally quantified through thermo-optical methods exploiting its thermally refractory nature (instead of its absorption optical properties as for BC).

Brown carbon

BrC is defined as light-absorbing carbonaceous material with an imaginary part k of the complex refractive index (see Eq. (1.4)) that increases towards shorter visible and UV wavelengths, resulting in an α value in Eq. (1.8) higher than 1. BrC is a fraction of the organic carbon, and its quantitative impact on radiative forcing is still characterised by high uncertainty due to gaps to be filled about its origin, chemical composition, optical properties, and mixing state. Optical properties of BrC may be related to water-soluble organic carbon compounds and humic-like substances (HULIS). Initial observations of BrC were linked to its origin during biomass burning, but more recent studies have suggested also other sources and formation processes such as e.g. secondary organic aerosol formation processes.

Mineral dust

Detailed description of the mineral dust source has been already given in Sect. 1.2. On a global scale, mineral dust aerosol is associated to heavy loading but it is less efficient in aerosol light absorption respect to BC. Mineral dust morphology is generally non-spherical and not-well characterised, making calculations of its optical properties more challenging. Among major components of mineral dust, hematite (Fe_2O_3) was found to have the largest absorption at UV and visible wavelenghts. In the literature, α values associated to mineral dust are usually higher than 2 (see e.g. [12] and references therein), i.e. higher than for BC but in the same range as BrC.

Chapter 2

Modelling methodologies: source apportionment of atmospheric aerosol

2.1 Introduction

The abatement of pollution in the atmosphere is one of the target of the European Union policy [13]. Source apportionment (SA) refers to the identification and quantification of emission sources that impact atmospheric concentrations of pollutants; this practice is mandatory to develop strategies with the aim of reducing atmospheric pollutants levels. Focusing on atmospheric aerosol, in the literature SA studies can be performed through two main approaches: source-oriented models (SMs) and receptor-oriented models (RMs).

Briefly, SMs are air quality models based on data from emission inventories, meteorological fields and atmospheric concentrations at the boundary of the considered modelled area [14]. Among them, some of the most used are the Eulerian chemistry-transport models (CTMs), that require as input an extensive dataset (e.g. 3D meteorological data, 3D emission data) to be able to simulate both primary and secondary PM in the atmosphere. One feature of SMs is the possibility to theoretically predict air quality changes in relation to emission changes; anyway, results from these models are clearly limited by the formulation of the CTM used and by the quality of input data in terms of emission inventories and meteorological fields.

This thesis focuses on RMs: these models need as input concentrations of different PM components, with the advantage of exploiting information on pollutants from real-world measurements. This chapter deals with basic theory of RMs and specifically with the Positive Matrix Factorization (PMF) model, comprehending the investigation of advanced applications.

2.2 Receptor modelling

Receptor models (RMs) are based on the principle of mass conservation between the emission source and the receptor site (i.e. the location where measurements were taken), and the input data are constituted by atmospheric concentrations of different aerosol chemical species in different samples. The mathematical formulation of RMs is based on factor analysis and it can be introduced as a bilinear problem: the input data matrix $\mathbf{X}$ (matrix elements x_{ij}) is decomposed in the product of the two factor matrices $\mathbf{F}$ (matrix elements f_{kj}) and $\mathbf{G}$ (matrix elements g_{ik}), related to factors chemical profiles and factors temporal contribution, respectively; factors can be subsequently interpreted as the main sources impacting the investigated area (factor-to-source assignment will be explored in Sect. 2.5.4). The basic general equation to be solved is the following:

$$x_{ij} = \sum_{k=1}^{P} g_{ik} f_{kj} + e_{ij} \tag{2.1}$$

where the indices i, j, and k indicate the sample, the species, and the factor, respectively; P is the total number of factors; the matrix $\mathbf{E}$ (matrix elements e_{ij}) is composed of the residuals, i.e. the difference between measured and modelled values. Graphical matrix notation for Eq. (2.1) is reported in Fig. 2.1. As it will be outlined in the following, also measurement uncertainties are needed as input and they play an eminent role in the search for the model solution.

Figure 2.1: Matrix formulation of the bilinear problem in Eq. (2.1). Pedix n represents the total number of samples, m the total number of variables, P the total number of factors.

All RMs rely on some main assumptions:

- factors chemical profiles do not vary significantly over time;

- factors are independent of each other (not collinear);

- chemical species do not chemically react during the transport from the source to the receptor site;

- measurements uncertainties are random, uncorrelated, and normally distributed.

The validity of these assumptions depends largely on the input dataset; RMs can generally tolerate some reasonable deviations, considering that larger deviations from these assumptions may increase uncertainties in model estimations [15]. One of the most critical issues of RMs is that the formation of secondary compounds in the atmosphere cannot be taken into account, so that the estimation of their origin in terms of sources is still difficult to assess through this methodology alone. Equation (2.1) is solved differently depending on the available a priori information regarding the sources of the investigated area; from this point of view, the Chemical Mass Balance (CMB) model and the multivariate models can be considered somehow opposite approaches (Fig. 2.2).

Receptor models

Multivariate receptor models
- P is unknown
- f_{kj} are unknown
- multiple samples *(i = 1,…,n)* are required

$$x_{ij} = \sum_{k=1}^{P} g_{ik}\, f_{kj} + e_{ij}$$

CMB model
- P is known
- f_{kj} are known
- solution on a sample by sample basis

Knowledge required about pollution sources prior to receptor modelling
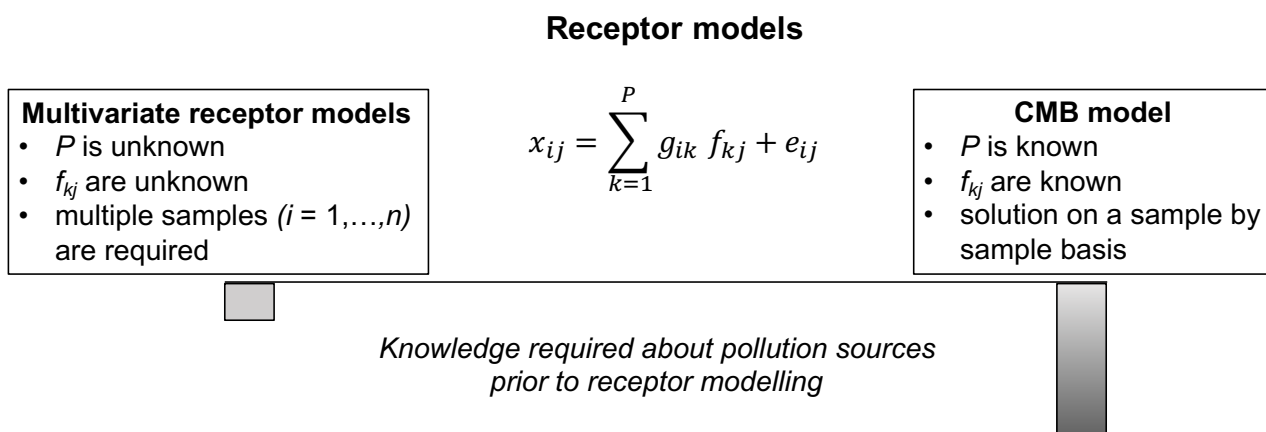
Figure 2.2: Two main receptor modelling approaches used to solve Eq. (2.1) (figure modified from [16] and [17]). The input data matrix $\mathbf{X}$ (matrix elements x_{ij}) is decomposed in the product of the two factor matrices $\mathbf{F}$ (matrix elements f_{kj}) and $\mathbf{G}$ (matrix elements g_{ik}), related to factors chemical profiles and factors temporal contribution, respectively; the matrix $\mathbf{E}$ (matrix elements e_{ij}) is composed of the residuals, i.e. the difference between measured and modelled values. The indices i, j, and k indicate the sample, the species, and the factor, respectively; P is the total number of factors; n represents the total number of samples.

The CMB model assumes a complete knowledge about the nature of emission sources and their chemical characteristics: in Eq. (2.1) the factor matrix elements f_{kj} are a priori known and fixed, so that the factor

matrix elements g_{ik} can be retrieved solving the equation on a sample by sample basis [18]. This approach involves the further approximation that source profiles remain the same between the emitter and the receptor.

On the other side, in multivariate models g_{ik} and f_{kj} in Eq. (2.1) are both unknown and they have to be derived from the application of the model exploiting internal correlations and relationships between input concentration data [16]. In the past, different multivariate methods have been proposed in the literature [18]; this thesis is focused on the application of the widely used Positive Matrix Factorization (PMF) in the custom version by Paatero [19].

2.3 Positive Matrix Factorization (PMF)

The Positive Matrix Factorization (PMF) model solves the factor analytic problem in Eq. (2.1) through a weighted least squares fit (WLSF) method [20] [21]. Any WLSF relies on the minimisation of a sum of squared expressions multiplied by a weight:

$$\theta = min_\theta \sum_l w_l (x_l - y_l(\theta))^2 \tag{2.2}$$

where θ represents the parameter(s) to be estimated and respect to which the minimum min_θ is calculated; l is an index used for enumeration; w_l values represent the weights; x_l values represent the observations; y_l values represent the fitted values predicted by the physical model y. The best possibile estimation (i.e. minimum variance respect to true values) for θ is obtained when the weight $w_l = (std.dev(x_l))^{-2}$, where $std.dev(x_l)$ is the standard deviation of the measured quantity x_l. For PMF, referring to the nomenclature of Eq. (2.1):

- $\theta = \{\mathbf{G}, \mathbf{F}\}$

- $x_l = x_{ij}$

- $y_l(\theta) = \sum_{k=1}^P g_{ik} f_{kj}$

- $w_l = (\sigma_{ij})^{-2}$

where the elements σ_{ij} of the matrix $\boldsymbol{\sigma}$ are estimations of the standard deviations of x_{ij}, and they are given as input data together with x_{ij}. Instructions for the best assessment of this matrix will be outlined in Sect. 2.5.2. Substituting in Eq. (2.2), the "object function" Q to be minimised can be defined as:

$$Q = \sum_i \sum_j \left(\frac{x_{ij} - \sum_{k=1}^P g_{ik} f_{kj}}{\sigma_{ij}} \right)^2 = \sum_i \sum_j \left(\frac{e_{ij}}{\sigma_{ij}} \right)^2 \tag{2.3}$$

PMF performs this minimisation under the constraint that elements of $\mathbf{G}$ and $\mathbf{F}$ are non-negative [22]:

$$g_{ik} \geq 0, f_{kj} \geq 0 \tag{2.4}$$

in order to take into account the physical observation that negative concentrations from emission sources are not possible.

The task can be solved iteratively starting from pseudorandom values $\{\mathbf{G_0}, \mathbf{F_0}\}$ for matrices $\mathbf{G}$ and $\mathbf{F}$, even if its solution is not straightforward since two different non-linear terms are present: inequalities (Eq. (2.4)) and products of unknown in the residual matrix $\mathbf{E}$ (Eq. (2.3)). The object function Q may show local minima; in general, algorithms that can be implemented for PMF are able to find a minimum of Q, but they do not indicate if it is a local or global one. To overcome this issue, the model can be run with different sets of $\{\mathbf{G_0}, \mathbf{F_0}\}$ in order to be reasonably sure that the lowest computed Q value corresponds to the global minimum.
Originally, the algorithm PMF2 was developed to solve specifically bilinear PMF problems, starting from a generalisation of the traditional Alternating Least Squares (ALS) algorithm (with both factors $\mathbf{G}$ and $\mathbf{F}$ changed together) and exploiting the Gauss-Newton approach to omit second order terms in the residual. The Gauss-Newton scheme was then solved by Cholesky decomposition and back-substitution; non-negativity of g_{ik} and f_{kj} was achieved through logarithmic penalty functions [22] [23]. As a further step, in the late 1990s a more flexible algorithm was developed to solve general multilinear problems: the Multilinear Engine.

2.4 Multilinear Engine (ME-2 program)

The Multilinear Engine program ME-2 [19] [24] was developed to solve general multilinear and quasi-multilinear problems; before giving insights into the program, basic definitions are given.

In multilinear problems, the mathematical expression to model the data is assumed to be linear for each group of variables, i.e. the expression for each data value can be written as sums of products of a number of unknown variables. The expression is homogeneous, which means that each product consists of the same number of variables (this is not the case for quasi-multilinear models). The mathematical expression in Eq. (2.1) can be considered the simplest and most common multilinear model; as we have already mentioned in Sect. 2.2, the model is called "bilinear" because the unknown variables are divided in two groups (factors $\mathbf{G}$ and $\mathbf{F}$). If the number of groups of unknowns is three, the model is "trilinear"; if it is unspecified, the model is generally called "multilinear". In source apportionment studies of atmospheric aerosol, deviations from a bilinear problem may appear due to particular characteristics of the input dataset. For example, data values might consist of concentrations of different chemical species, measured at different times and in different size fractions, and we might be interested in retrieving information about source chemical profiles, temporal patterns and also contributions in each size fraction.

The general form of a multilinear problem is the following:

$$x_m = \sum_{k=1}^{K_m} \prod_{n \in I_{mk}} f_n + e_m \tag{2.5}$$

where:

- m enumerates the equations to be solved; it identifies data to be fitted by the model. E.g. for the particular case of a bilinear problem in Eq. (2.1), $m = (ij)$

- x_m represents the data to be fitted; considering the application to receptor modelling, data correspond to concentration values

- k enumerates the products; e.g. for Eq. (2.1), and in general for receptor models, k identifies the factors

- K_m is an integer that indicates the total number of products for each equation; for Eq. (2.1), and in general for receptor models, all K_m equal P, the total number of factors

- f_n is the collection of all the unknown factor elements in the model; e.g. for Eq. (2.1), g_{ik} and f_{kj}

- I_{mk} are the index sets to define all the products in each equation; e.g. in Eq. (2.1) I_{mk} are the P sets $\{ik,kj\}$ obtained considering $k = 1, 2, ..., P$ for each equation

- e_m represents the residuals between x_m and modelled data.

Starting from the WLSF method in Eq. (2.2) applied to the model in Eq. (2.5), the definition of the object function Q to be minimised in this case is the following:

$$Q = \sum_m \left(\frac{x_m - \sum_{k=1}^{K_m} \prod_{n \in I_{mk}} f_n}{\sigma_m} \right)^2 = \sum_m \left(\frac{e_m}{\sigma_m} \right)^2 \tag{2.6}$$

where σ_m are the uncertainties connected to each equation of the model, and they are given by the user as input data together with x_m; more practical information about uncertainties will be given in Sect. 2.5.2.

It is known that a trivial scaling indeterminacy is present in the solution of all multilinear models, i.e. the model can be rescaled by a value $r \neq 0$. Considering as an example the bilinear case in Eq. (2.1), we can define the rescaled product $\bar{g}_{ik}\bar{f}_{kj} = (rg_{ik})(f_{kj}/r) = g_{ik}f_{kj}$ that produces the same modelled outputs of the original model. The risk of runaway scales of factors is present, i.e. values of $\mathbf{G}$ increasing without limit and values of $\mathbf{F}$ converging toward zero, and viceversa. Factor scale indeterminacy may be removed by including in the model a special a priori information about norms of columns of factor matrices; for the bilinear model, the common practice in ME-2 is the normalisation of columns of the $\mathbf{G}$ matrix (even if different examples are present in the literature [23]). In fact, when considering environmental data, elements in the columns of $\mathbf{G}$ (representing temporal patterns) vary less than a factor of 100, while elements in the rows of $\mathbf{F}$ may vary by six orders of

magnitude [25]. Thus, normalisation on $\mathbf{F}$ might cause problems in the algorithm since it might be based only on the few largest values of the matrix, ignoring smaller values. Anyway, this is not a restriction from the point of view of the user, since desired renormalisation can always be performed a posteriori.

The general formulation in Eq. (2.5) is at the basis of the ME-2 program. This equation is used to model atmospheric concentration data, but also a priori information (e.g. a known ratio between concentrations of two species in the chemical profile of a source) can be introduced in ME-2 exploiting the same structure; equations related to a priori information are called "auxiliary equations" while equations related to data values are called "main equations" [19]. Expliciting the contribution of auxiliary equations in the object function Q, Eq. (2.6) becomes:

$$Q = \sum_m \left(\frac{e_m}{\sigma_m}\right)^2 + \sum_v \left(\frac{r_v}{s_v}\right)^2 = Q_{main} + Q_{aux} \qquad (2.7)$$

where v specifies the subscript for auxiliary equations, r_v represent the residuals of auxiliary equations, and s_v denote the "softness" of auxiliary equations [25]. s_v are chosen by the user in order to give the desired weight to the auxiliary information and to indicate the acceptable tolerance to satisfy it [19]. Since normalisation of columns of factor matrices is always required in order to avoid factor scale indeterminacy (see previous discussion), in pratical cases Q_{aux} is always higher than zero in the minimisation process.

The possibility to introduce additional a priori information as a target to be accomplished in the model is one of the main features of ME-2 respect to PMF2; this and other features will be better investigated in Sects. 2.4.1 and 2.4.2.

The flexibility of ME-2 allows more advanced users to set up new kinds of models by implementing ME-2 scripts. The program is controlled by an initialization file or .ini file, written in a special-purpose script language [26]; the .ini file is the interface between the user and the underlying ME-2, no graphical interface is supported (neither as output). The .ini file constitutes the first part of ME-2, containing the instructions written by the user; the second part computes the solution to the problem specified by the user.

The approach used by PMF2 algorithm (see the end of Sect. 2.3) to find the solution was not the best choice to solve general multilinear problems that might be extremely large. ME-2 solves the PMF problem applying the conjugate grandient algorithm, that exploits properties of the Jacobian matrix of multilinear models; non-negativity of factor elements is achieved by inverse preconditioning. More insights about the conjugate grandient algorithm are given in [19].

One widespread, freely available and ready-to-use application of ME-2 is the EPA (Environmental Protection Agency) PMF 5.0 [27]. In this application, the solution of a bilinear problem is implemented, and the user has to give parameters specification through the EPA PMF 5.0 graphical user interface; then, ME-2 solves the problem minimising the object function Q (Eq. (2.7)). In addition, output from ME-2 are formatted by EPA PMF 5.0 to facilitate the interpretation by the user.

This thesis will mainly focus on the implementation and optimisation of advanced receptor modelling approaches, starting from script files, implementing them and exploiting the flexibility of ME-2.

2.4.1 Rotational ambiguity and implementation of a priori information

Receptor models are based on factor analysis. Factor analytic problems are ill-posed, which means that the problem can be solved, but it does not produce a unique solution due to "rotational indeterminacy" [25] [28]. To better understand this statement, the following discussion will be focused on the particular case of a bilinear model, but the generalisation to multilinear problems is straightforward. Let's consider the system of equations (2.1) in a matrix notation:

$$\mathbf{X} = \mathbf{GF} + \mathbf{E} \qquad (2.8)$$

As already discussed in Sect. 2.3, $\{\mathbf{G}, \mathbf{F}\}$ are determined minimising the object function Q under non-negativity constraints on factor elements. The factor matrices $\{\mathbf{G}, \mathbf{F}\}$ can be transformed into another pair $\{\bar{\mathbf{G}}, \bar{\mathbf{F}}\}$ corresponding to the same Q value of $\{\mathbf{G}, \mathbf{F}\}$. In the literature, it is traditionally said that $\{\mathbf{G}, \mathbf{F}\}$ can be "rotated" into $\{\bar{\mathbf{G}}, \bar{\mathbf{F}}\}$, even if the real procedure is not a rotation but a linear transformation. Considering a nonsingular square matrix $\mathbf{T}$ and its inverse as $\mathbf{T}^{-1}$ ($\mathbf{TT}^{-1} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix), the following transformation can be written:

$$\mathbf{X} = \mathbf{GF} + \mathbf{E} = \mathbf{GTT}^{-1}\mathbf{F} + \mathbf{E} = \bar{\mathbf{G}}\bar{\mathbf{F}} + \mathbf{E} \qquad (2.9)$$

This rotation can be accepted only if non-negativity constraints on factor elements are still satisfied. Sometimes non-negativity constraints are enough not to allow rotations; in this case, the solution is unique. It is well known that rotational ambiguity is not present if a sufficient number of elements of **G** or **F** are zeroes [28]. In all other cases, the domain of possible solutions should be explored in order to evaluate if one solution appears more plausible than others from a physical point of view.

The formal definition in Eq. (2.9) is a "pure" rotation, but it does not take into account the flexibility of factors, so that in real situations "approximate" or "distorted" rotations are performed. In distorted rotations the product of factor matrices change, and consequently the Q value changes; these changes in factors are acceptable as long as the Q value does not increase too much. Unfortunately, there is not a fixed rule about the acceptable increase in the Q value: generally speaking, increases in the Q value of thousands appear questionable, of hundreds may be acceptable, of a few tens may be acceptable without question [25].

Evaluation of the plausibility of a solution is performed exploiting available a priori information. Conceptually (without giving details about the ME-2 script language here), a priori information can be introduced in the ME-2 script as follows [25]:

1. Setting some factor elements equal to given fixed values, both zero or non-zero values. In the script, the user declares that a certain factor element should not vary; thus, it retains its initial fixed value during the WLSF procedure and it is handled as a constant, not as a variable.

2. Setting individual lower and/or upper limits (inequality constraints) for some factor elements. By default, all factor elements have lower limits equal to zero; in the script, these limits are accessible and they can be modified with small negative values or as needed. Also upper limits can be specified in a similar way.

3. Pulling some factor elements toward specific target values; this is done introducing "pulling equations" in the model. Pulling equations contribute to the Q_{aux} in Eq. (2.7), with r_v corresponding to the difference between the considered factor elements and their target values. Pulling equations have an advantage over points (1) and (2): if the Q value increases too much when the pulled factor element reaches its target value, this can be considered an indication of the fact that the constraint is incompatible with the model.

4. Pulling functions of factor elements (e.g. a ratio of factor elements) toward specific target values. Also in this case a pulling equation is introduced in the model (see point (3)), but r_v in Eq. (2.7) corresponds to the difference between known relationships of factor elements and their target values.

A practical case of a priori information about relationships of factor elements (point (4)) is a known ratio between the concentration of two species in a specific source as singled out in the chemical profile (factor **F**). For example, $\frac{f_{k2}}{f_{k3}} = t$, where k indicate the specific source, 2 and 3 represent two different species, and t is the numerical value of the ratio; it can be e.g. the stoichiometric ratio between sulphate and ammonium, or the ratio between Cl and Na when fresh marine air masses impact over the investigated area. This expression must be introduced as a pulling equation, giving a contribution to Q_{aux}:

$$0 = f_{k2} - t f_{k3} + r_v \tag{2.10}$$

$$Q_{aux,v} = \left(\frac{r_v}{s_v}\right)^2 = \left(\frac{0 - (f_{k2} - t f_{k3})}{s_v}\right)^2$$

Another example of a priori information that can be introduced in the model is a balance between different factor elements, e.g. a mass balance. In this case, each factor makes sense from a physical point of view if the measured mass concentration introduced as an input variable in the model is equal or larger than the sum of concentrations of individual chemical species:

$$0 = \sum_j c_j f_{kj} + r_v \tag{2.11}$$

where c_j are numerical coefficients: for the mass $c_{mass} = +1$; for species different from the mass $c_j = -1$ (more insights about the use of c_j can be found in [25]). In this case the inequality (the mass is larger than the sum of species) can be implemented, so that negative values of the residual r_v are allowed, which means that the equation do not contribute to Q if the residual r_v is negative:

$$Q_{aux,v} = \left(\frac{r_v}{s_v}\right)^2 \qquad\qquad if \quad r_v > 0 \tag{2.12}$$

$$Q_{aux,v} = 0 \qquad\qquad if \quad r_v \leq 0$$

Obviously, before the introduction of a mass balance equation in the model, it is necessary to check if the measured data values fulfill the condition; if individual samples do not obey the balance, one should not expect something different from the computed factors. In general, to control the impact of a priori information in the model, it is better to introduce them one at a time.

2.4.2 Methods to estimate uncertainty in the solution

In PMF modelling, uncertainty analysis of the solution aims at estimating a range of plausible values around computed factor elements [29]. There are three main causes of uncertainty in PMF solutions:

- random errors in input data values;

- rotational ambiguity;

- modelling errors.

Random errors in input data values are intrinsically associated with the measurement process, even when no systematic bias is present. Uncertainty due to rotational ambiguity is peculiar of all factor analytic models; rotational ambiguity has been extensively treated in Sect. 2.4.1. Modelling errors arise because the model is always a simplification of the real world; modelling errors take into account e.g. variation of source chemical profiles with time, incorrectly estimated input data uncertainties, correlated (i.e. systematic or non-random) errors in input data values, incorrectly specified total number of factors P. The relative importance of the three types of uncertainty mainly depends on the size of the dataset. As the size of the dataset increases, the significance of random errors decreases due to the law of large numbers, and the significance of rotational ambiguity decreases too since the number of zeroes in the factors often increases. Opposite, modelling errors are associated to non-random disturbances that are not subject to the law of large numbers, thus their relative importance can be assumed to be the highest for large datasets. Anyway, large datasets usually contain much more information respect to small datasets and this feature can enhance the model performance.

Three uncertainty estimation methods can be implemented in ME-2 scripts (they are already available in EPA PMF 5.0): bootstrap analysis (BS), dQ-controlled displacement of factor elements (DISP), and bootstrap enhanced with displacement (BS-DISP) [29] [30].

Bootstrap analysis (BS): BS is a typical statistical method to estimate uncertainty. It is based on perturbation of the original dataset in order to produce a "perturbed dataset" that is subsequently analysed through PMF modelling as the original one; this is done for a certain number of runs (e.g. 100). BS can be implemented in ME-2 in three steps: re-sampling, re-weighting, and random rotational pulling [23]. Re-sampling and re-weighting deal with the perturbation of the original dataset: re-sampling is performed generating random weights for each main equation (2.5) in order to consider only equations with a positive weight; during re-weighting, previously calculated positive weights are used to change the uncertainties associated to the main equations. The perturbed dataset becomes the input for new PMF runs, and the perturbed solution is subject to random rotational pulling: randomly chosen **G** elements are pulled up or down, as an attempt to take into account rotational ambiguity in BS analyses. Different BS types can be implemented, considering different ways of generating weights and different input data treatment for the perturbed analysis. For example, in the EPA PMF 5.0 three main modalities are defined and controlled by the ME-2 script variable *bsmode*:

- *bsmode = 11*: traditional and default BS. Positive and negative random weights are generated, and all the three steps (re-sampling, re-weighting and random rotational pulling) are perfomed.

- *bsmode = 14*: no weights are generated and all main equations are taken into account, but 50% of the original dataset is substituted by modelled values (which means residuals = 0), and the remaining 50% is substituted by (2 · original values - modelled values) (which means residuals = 2 · original residuals).

- *bsmode = 15*: two positive weights are generated, one lower than 1 and one higher than 1, and associated to the main equations. Since generated weights are all positive, no re-sampling is performed (just re-weighting and random rotational pulling are applied). It is supposed to be statistically equivalent to *bsmode*=11, because the distribution of weights is implemented in order to share the same properties in terms of expected value and variance.
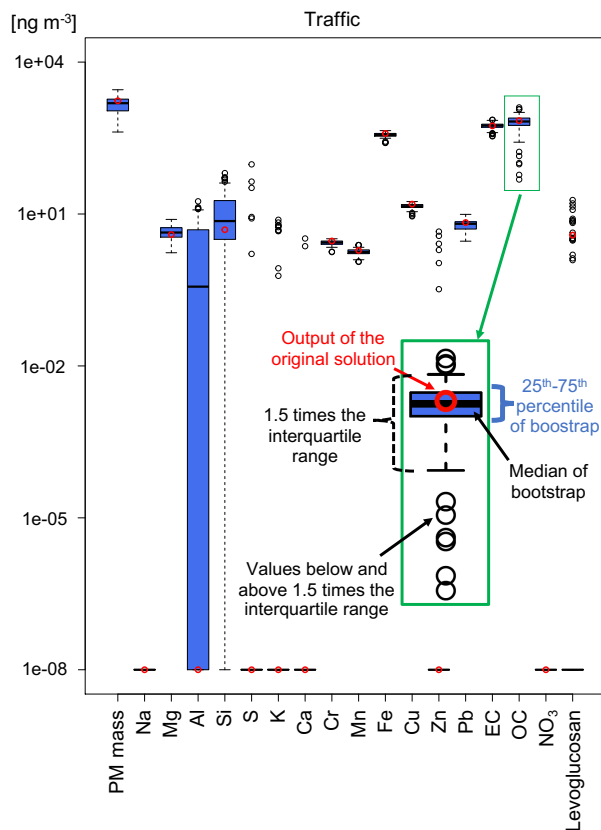
Figure 2.3: Example of a box plot resuming outputs from a bootstrap analysis; statistics for the chemical profile of a factor identified as traffic are reported. Figure modified from [31].

Variations in the perturbated source chemical profiles (factor **F**) can be used to estimate the uncertainty in the original solution (Fig. 2.3).

Through the comparison of factors contributions (factor **G**) between the original and the perturbed solution, the perturbed factor is assigned to the original one with which it has the highest correlation, above a user-specific threshold (e.g. Pearson correlation coefficient $r > 0.60$ is considered as default in EPA PMF 5.0). If no correlation coefficient is above the threshold, the factor is classified as "unmapped"; any factor with a percentage of unmapped cases of approximately 80% or above (i.e. 80% or less mapping) should be better investigated [27].

BS can be perfomed to obtain reliable estimation of uncertainty due to random errors in input data; it is not specifically designed to explore rotational ambiguity, even if some rotational uncertainty is catched in the procedure. Re-sampling may lead to the omission of some or all zero values in the factors, thus giving an overestimation of the variation in the solution if the zero values were reliable; to reduce the impact of these "too much" perturbed solutions, percentiles of the obtained distribution of results can be used for error limits (e.g. 5th and 95th percentiles). This is one of the reason why $bsmode$=15 was developed besides traditional $bsmode$=11: all main equations are considered in $bsmode$=15, and this might lead to more stable replications. Regarding modelling errors, it is not known how well BS captures them.

Displacement analysis (DISP): Starting from the original solution, DISP repeatedly fits the model so that each variable in factor **F** in turn is perturbed (or "displaced") from its original fitted value; in this way, uncertainty estimates for individual variables of **F** are obtained. Displacement of each variable is extended until the object function Q increases by a maximum allowed change dQ_{max}; this extension is interpreted as the upper or lower interval estimate of the considered variable. DISP captures the uncertainty caused by random errors in input data, but underestimation of uncertainty in the solution might appear if data errors are correlated or if they are underestimated in the original dataset; on the other hand, overestimation of uncertainty in the solution might appear if data errors are overestimated in the original dataset (e.g. for species that were downweighted

in the original PMF analysis, uncertainties estimated by this methodology are known to be too large). Occasionally, DISP causes a significant (tens or hundreds) decrease of Q: it means that the original solution was not a global minimum. In this case, it is necessary to solve the original model again, using more random starts to find the global minimum; then, the DISP analysis can be performed another time. DISP is also able to catch uncertainty from rotational ambiguity, but it is not known how well it captures modelling errors.

Bootstrap enhanced with displacement analysis (BS-DISP): BS-DISP is a combination of bootstrap and displacement analyses: the original dataset is perturbed through re-sampling and re-weighting (no random rotational pulling is performed in this case), and then each variable in factor **F** in turn is displaced. This process may be viewed as follows: BS gives a perturbed solution, then the DISP analysis determines an approximation for the rotationally accessible space around that perturbed solution. Taken together, they represent both the random uncertainty and the rotational uncertainty for the modelled solution. As for the BS analysis alone, re-sampling at the beginning of the BS-DISP procedure may lead to the omission of some or all zero values in the factors, thus giving an overestimation of the variation in the solution if zero values were reliable; percentiles of the obtained distribution of results can be used for error limits (e.g. 5th and 95th percentiles). As for the DISP analysis alone, decrease of the Q value respect to the original one may occur; in this case, it is suggested to reject the resample corresponding to the decrease in Q. The overall BS-DISP analysis remains valid, though currently there is no way to a priori quantify the number of allowed rejections. BS-DISP results are more robust than DISP ones, since the displacement phase of BS-DISP does not displace as strongly as DISP by itself.

For further details, please refer to [29] where important recommendations about how to document uncertainty estimations for PMF modelling are reported.

2.5 General modelling procedure for source apportionment with PMF

2.5.1 Preliminary checks

Preliminary knowledge about the study area is necessary in order to define a clear methodological framework, i.e. to define the objectives of the modelling analysis and to plan the experimental work. Different pieces of information can be retrieved, studying e.g. previous source apportionment studies, local measurements of atmospheric concentrations of pollutants available in the literature, prevailing local wind directions, the geography as well as the topography of the area in order to understand source-receptor relationships [13].

Multivariate techniques such as PMF work properly with large number of samples as input. Suggestions about a reliable number of samples can be found in the literature [13], but in general the minimum number of samples cannot be established a priori, as it depends on the amount of information contained in the dataset: there should be enough samples to catch the variability of sources, including samples where contributions from some sources are absent or negligible.

The chemical species to be included in the analysis should be selected on the basis of work objectives, site characteristics, and expected sources, taking into account available resources. For modelling purposes, increasing the number of species may help to reduce collinearity problems and to potentially resolve a higher number of sources. Another consideration regards the inclusion of chemically redundant species, e.g. sulphur and sulphate; usually, literature studies include only one of them, in order to avoid double counting of sulphur atoms [32]. In general, if two species are proportional to each other - as it should be if all sulphur is present as sulphate, with a stoichiometric ratio equal to 3 - then it does not matter which one is used as input in the analysis [13] (Fig. 2.4). Other common examples of double counting between elemental and ionic species are e.g. Na and Na^+, K and K^+, Ca and Ca^{2+}, Mg and Mg^{2+}, Cl and Cl^-. In the literature, soluble potassium K^+ is a better indicator of biomass burning emissions respect to K, thus a good choice might be the inclusion as input data of K^+ and of insoluble K (= K - K^+). Redundant species can be compared in the phase preceding the modelling analysis, in order to check measurements carried out through different analytical techniques and to introduce the most reliable variables in the model. In this phase, ionic balance between measured atmospheric positive and negative ions might be checked, and a preliminary mass closure study based only on PM composition might be performed to have a rough idea about the impact of possible sources [33].
Anyway, introduction of redundant species is not a priori harmful for the analysis but it should be taken into account at a later stage, e.g. retaining only one of the two species to verify the mass closure for each factor.
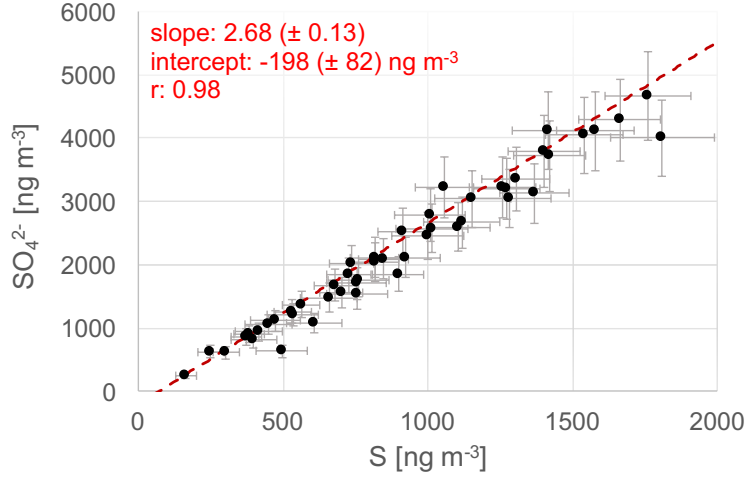
Figure 2.4: Example of comparison between sulphate (SO_4^{2-}, y axis) and sulphur (S, x axis) concentrations measured on different samples during summer and winter 2016 in Milan (Italy). The slight difference (of the order of 10%) respect to the expected stoichiometric coefficient 3 can be ascribed to the use of different analytical techniques.

Since one of the main purposes of receptor modelling applications is to apportion PM mass, its measurement should be considered. PM mass can be taken into account in the modelling procedure through two general approaches. The first is to include PM mass as an input species; in this case, the model will apportion PM mass to each factor as it apportions the other species. In the literature, it has been suggested that the uncertainties of PM mass concentrations should be substantially increased (e.g. set at four times their values [34]) to decrease the weight in the model fit. PM mass might be also considered an example of double counting, because all other aerosol components concentrations are comprised in PM mass [32]. The second method of apportioning total PM mass is through an external regression of the factor contributions (factor **G**) obtained from the modelling procedure onto the PM mass measurements, as in the following equation:

$$PM_i = \sum_{k=1}^{P} g_{ik} a_k \qquad (2.13)$$

where PM_i is the total PM mass from sample i, and a_k is the regression coefficient for factor k resulting from regressing the factor contributions g_{ik} on PM_i. This method assumes that g_{ik} are error free, that is clearly not true; the advantage is that negative values of a_k can be considered an indication that too many factors were specified in the PMF modelling, and that the analysis should be performed with fewer factors [13].

2.5.2 Data preparation

In source apportionment studies of atmospheric aerosol through PMF, typical input data are:

- x_{ij} (elements of matrix **X**): concentrations of different chemical species (denoted by subscript j) in different samples (denoted by subscript i). x_m will be also used, referring to the general nomenclature of Sect. 2.4.

- σ_{ij} (elements of matrix $\boldsymbol{\sigma}$): estimations of the standard deviations of x_{ij}. The general nomenclature σ_m will be also used.

Occasionally, other variables have been used in the literature (e.g. gaseous species, meteorological parameters), which will be better investigated in Sect. 2.6.2. As it will be further discussed, σ_{ij} are associated to experimental uncertainties u_{ij} because u_{ij} are typically used as a starting point to estimate σ_{ij}, but the two quantities do not necessarily have to coincide. To distinguish them, in the following the different nomenclature will be maintained.

Signal-to-noise ratio (S/N). A statistical parameter to be evaluated for each species before its insertion in the model is the signal-to-noise ratio (S/N), which indicates if variability in the measurements is real or due to the noise associated to the data [27]. The signal related to the data can be approximated by the difference between concentration x_{ij} and uncertainty u_{ij}, since in general concentrations are constituted by the sum of

signal and noise. When x_{ij} is equal to or lower than u_{ij}, the datum is considered to have no signal. Referring to n as the total number of samples, S/N is defined as:

$$(S/N)_j = \frac{1}{n} \sum_i d_{ij} \qquad (2.14)$$

$$where: \quad d_{ij} = \frac{x_{ij} - u_{ij}}{u_{ij}} \qquad \qquad if \quad x_{ij} > u_{ij}$$

$$d_{ij} = 0 \qquad \qquad if \quad x_{ij} \leq u_{ij}$$

For example, species with concentrations always below their uncertainty have S/N equal to 0; species with concentrations that are twice the uncertainty values have S/N equal to 1. Species can be categorised as "bad" if S/N ratio is less than 0.5; "weak" if S/N ratio is between 0.5 and 1; "strong" if S/N ratio is higher than 1 [27], even if a slightly larger limit might be considered, e.g. S/N higher than 1.2 [35]. Even a small overweighting (i.e. too small uncertainty respect to the true error value) of weak and bad variables can be harmful for the analysis and should be generally avoided; anyway, it can be considered acceptable if such variables can represent markers for sporadic sources (see e.g. [36]). For weak variables, it is recommended a routine downweight of uncertainties by a factor of 2 or 3; bad variables should be omitted from the analysis or, if not desirable, they should be strongly downweighted by a factor from 5 to 10 [37].

Missing data and data below minimum detection limit (MDL). MDL denotes the lowest concentration level that can be determined to be statistically significant from an analytical blank [38]. Missing data might be present in the dataset, e.g. due to technical problems in the instrumentation, in the analytical technique, or due to filter contaminations. For both missing and below MDL data, it is not straightforward to define a concentration value to be inserted in the model, but they are associated to different pieces of information that can be exploited.

Missing data. Empty entries in input corresponding to missing data are allowed by the ME-2 program. In this case the equation contains no terms and it is ignored, i.e. it is not used to retrieve the modelled value. The variable *missdatlim* in ME-2 script controls settings for missing data: all entries that are below the value of *missdatlim* are ignored. For example if -999 denotes missing data, *missdatlim* can be imposed equal to -990. Empty entries might create problems with advanced receptor modelling approaches which seem to be more sensitive to missing data [39] and they cannot be introduced as input in EPA PMF 5.0. To deal with missing data in the model, different approaches are present in the literature: (1) eliminate samples with missing data, generally applied when key or several species are missing; (2) eliminate species with missing data, typically used when a large percentage of species observations are missing (e.g. higher than 50% [13]); (3) set a value with large associated uncertainty. Among the various approaches reported in the literature [32], Polissar et al. [40] proposed a method dealing with what is mentioned in point (3). It considers the geometric mean between concentrations of the chemical components in the dataset, with associated uncertainties usually equal to four times the concentration values estimated; other similar methods consider the use of mean or median values (in some cases seasonally averaged).

Below minimum detection limit (MDL) data. Species containing too many concentration values below MDL (e.g. higher than 50% [13]) can be neglected in the analysis. When this is not the case, data below MDL can be introduced in the model according to Polissar et al. [40] (as for this thesis work), where they are replaced with $\frac{MDL}{2}$ and the associated uncertainty with $\frac{\overline{MDL}}{2} + \frac{MDL}{3}$, where $\overline{MDL}$ corresponds to the arithmetic mean between the MDL values for the considered species; sometimes, the value $\frac{5}{6}MDL$ is used in the literature instead of $\frac{\overline{MDL}}{2} + \frac{MDL}{3}$ for the associated uncertainty. Other recommended literature approaches replace data below MDL with statistically intensive method (e.g. Kaplan-Meier, Maximum Likelihood Estimation, Robust Regression on Order Statistics [41]) or consider "uncensored" values, i.e. negative or below MDL values given as output from analytical methods. The aim of all these procedures is to avoid the introduction of biases in the analysis, that may result in the creation of artificial factors (called "ghost factors") when substitution procedures are not carefully checked. A special treatment of uncertainties has been developed for data below MDL, exploiting in the ME-2 .ini file the errormodel code *em*=-16, that will be better explored in the following.

Setting the standard deviations σ_m. Standard deviations of the main equations are an important link between the factor analytic model and the physical reality, and they should be used to "communicate" to the model various kinds of information. For example, high uncertainties can be associated to missing and below

MDL data; influence of outliers (i.e. very large values) can be controlled increasing their standard deviations in the model respect to their experimental uncertainties (see next paragraph). In general, what should be kept in mind is that standard deviations σ_m act as weights for input variables x_m in the minimisation process (see Sect. 2.3): some of the variables might be less important than others, so that one can subjectively increase the associated uncertainties (i.e. decrease the associated weight) in the model for such variables. Influence of similar procedures over the solution should always be carefully checked, since loosing the objectiveness of the analysis is a real risk; anyway, increasing the uncertainty of an irrelevant variable is less drastic than omitting it from the analysis or not measuring it at all [21]. Setting uncertainties should be perfomed before the modelling phase, and then these uncertainties can be further treated through ME-2. The error model function is used to specify how the final uncertainties to be used are set, and it is identified with the variable em in the .ini file. Together with em, three parameters $C1$, $C2$, and $C3$ have to be defined. $C1$ is usually identified with the input uncertainty given by the user, while $C2$ and $C3$ are used to introduce extra modelling uncertainties. Different error model codes are defined (an exaustive list - also for auxiliary equations - is given in [24] [26]), but the generally suggested one for environmental data is $em = -14$:

$$\sigma_m = C1 + C2\sqrt{max(|x_m|, |y_m|)} + C3 \cdot max(|x_m|, |y_m|) \tag{2.15}$$

where y_m denotes the values fitted by the model.

$em = -15$ is used for missing data (empty entries). $em = -16$ do not consider any contribution to Q if the fitted value y_m is below the data value x_m (i.e. residual e_m is positive); this error model can be used when x_m corresponds to a below MDL datum. $em = -17$ is the opposite of -16: no contribution to Q is considered if the fitted value y_m exceeds the data value x_m (i.e. residual e_m is negative).

Robust mode and outliers. Outliers can be defined as large atypical values respect to the average behaviour of data. Outliers can arise due to e.g. contaminations or errors in the analyses, a local source visible only occasionally (e.g. fireworks), higher impact of one source during a specific episode (e.g. plumes, that are undiluted transported from the source to the receptor) [22]. When needed, influence of these data on solutions can be diminished by the user increasing their standard deviations, both a priori or a posteriori after the fitting procedure, through the identification of data values corresponding to high scaled residuals (that are defined as $\frac{e_m}{\sigma_m}$, see Sect. 2.5.3) [21]. Handling of outliers can also be performed through the ME-2 algorithm using the robust mode, that is controlled by the special variable $robust$ in the .ini file. When $robust=1$ the robust mode is active (otherwise, $robust=0$) and uncertainties of input data for which the absolute value of scaled residuals $\frac{e_m}{\sigma_m}$ is higher than a fixed parameter, are increased. The parameters that can be fixed are $posoutdist$ for positive scaled residuals and $negoutdist$ for negative ones, and they are usually imposed equal to 4 [32]. In this case, the object function Q to be minimised is defined as:

$$Q_{robust} = \sum_m \left(\frac{e_m}{h_m \cdot \sigma_m} \right)^2 \tag{2.16}$$

$$where: \quad h_m = 1 \qquad\qquad\qquad if \quad |\frac{e_m}{\sigma_m}| \leq \alpha$$

$$h_m = |\frac{e_m}{\sigma_m}| \cdot \frac{1}{\alpha} \qquad\qquad if \quad |\frac{e_m}{\sigma_m}| > \alpha$$

considering $posoutdist=negoutdist=\alpha$ (but the generalisation for different values of $posoutdist$ and $negoutdist$ is straightforward).

Other relevant parameters in ME-2 script. The preparation of a ready-made ME-2 script requires the initialisation of other relevant special parameters respect to what already considered in the previous paragraphs.

$n1$: the number of samples (number of rows of matrix $\mathbf{X}$).

$n2$: the number of input variables, i.e. the number of chemical species (number of columns of matrix $\mathbf{X}$).

np: the total number of factors P. A major consideration in searching for the solution is to find the best number of factors to fit the dataset, and different np values should be considered in different analysis. Evaluation of the best number of factors is based on statistical and physical considerations (which are better investigated in Sect. 2.5.3 and 2.5.4).

$contrun$: a constant used to identify the starting values $\{\mathbf{G_0}, \mathbf{F_0}\}$ for the analysis:

... =0 pseudorandom values are considered; this type of analysis is called "base run".

... =1 the analysis starts from previously computed results (from a base run); this type of analysis is called "continuation run". This modality can be used to introduce a priori information through the implementation of auxiliary equations.

... =2 the analysis starts from previously computed results while also some special procedures are specified, e.g. estimation of uncertainties in a previously computed solution.

numtasks: the number of runs of the model to be performed. Different runs are always suggested when contrun=0 (i.e. when the model starts from different pseudorandom values), in order to find the global minimum of the object function Q. At least 20 random starts should be computed [27] [42], but also higher numbers can be found in the literature. When contrun $\neq$ 0, different criteria are used (e.g. for uncertainty estimation with bootstrap analysis, 100 runs are suggested).

numoldsol: a number used when results from previously computed analysis are considered (i.e. with contrun=1 or 2). For example, if we want to perform a continuation run starting from the third random start of the base run, numoldsol=3.

seed1: the initial value used to generate pseudorandom starting values $\{\mathbf{G_0}, \mathbf{F_0}\}$ of factors matrices. When several starts are computed (numtasks>1), then *seed1* varies in the script automatically so that the same result is not computed again and again.

alowlim and *blowlim*: the low limits for factor matrices. For example, alowlim represents the low limit for the factor $\mathbf{G}$ of temporal contribution (identified by AA in the .ini file) and it might be imposed slightly negative (as in the EPA PMF 5.0); in this way, true rotations can be accepted even in the presence of a large number of zero values [30].

2.5.3 Basic evaluation of the solution

The total number of factors P is a fundamental parameter to model the reality; choosing a representative value for this parameter is an essential part of the professional skills of the modeller [42]. This choice is influenced by preliminary evaluation of the study area, physical and chemical properties of real-world data, statistical and physical evaluation of the modelled solutions. Even the final purpose of the modelling procedure influences the final choice of P: for example, P might be smaller if the aim is to better determine the strongest sources at a receptor site, and larger if also weaker or episodic sources are of interest.

The optimum number of factors should not be evaluated through comparison between the obtained Q value and the theoretical Q value (that can be calculated considering the degrees of freedom of the problem). In fact, the former strongly depends on the assumed standard deviations of data values. A suggested way to exploit Q as an indicator of the reliability of the chosen P, is to use differences in obtained Q values [42]. When the number of factors P is increased by one, then the number of free parameters in the model increases approximately by $(n+m)$, where n is the total number of samples and m the total number of input variables; an increase of the number of free parameters corresponds to a similar decrease of the Q value. If the increase of P by one is deemed useful, the decrease of Q should be significantly higher than $(n+m)$. More specific criteria to evaluate how much the decrease of Q is significant cannot be given, because it is mainly dependent on the statistical properties of residuals for the specific considered problem.

After the fitting procedure, evaluation of the goodness of the solution should start from checking if criteria for convergence have been satisfied over different runs of the model. In ME-2 script, the special variable giving information about convergence is *endingnow*: a value of 4 indicate that convergence has been achieved, while other values might indicate different problems during the procedure. After this check, one model run is chosen (the run corresponding to the global minimum of the object function Q is chosen in the standard procedure) and further inspected as the final solution.

The first method to assess the goodness of the fit is the inspection of the distribution of the scaled residuals $\frac{e_m}{\sigma_m}$ for each variable. When scaled residuals are within certain limits (usually the range [-3,+3] is considered [27]) and normally distributed, the variable can be considered well-modelled. Distributions of scaled residuals with too large spread might indicate that the variable is not well-modelled or that uncertainties for the considered variable are too low. Distributions concentrated near zero might indicate that uncertainties are too high (i.e. high σ_m leads to too small $\frac{e_m}{\sigma_m}$) [32], but also that they are too small and a unique factor - i.e. a factor where only the considered variable has a significantly non-zero value - appeared in the analysis. Anyway, not all unique factors are physically meaningless and due to noise [21]. Information gained from scaled residuals can also be used as a diagnostic tool for adjusting species uncertainties in subsequent runs of the model [32], even if this

type of procedures should always be performed with caution and clearly reported. Examples of distribution of scaled residuals are reported in Fig. 2.5.
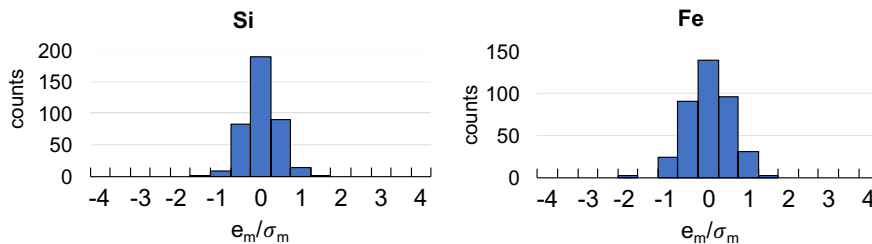
Figure 2.5: Example of distribution of scaled residuals $\frac{e_m}{\sigma_m}$ for two variables: silicon (Si) and iron (Fe). Figure modified from [31].

A more detailed assessment of the fit can be done by comparing all modelled variables with measured ones. Correlations (with statistics such as the correlation coefficient r and slope) and differences in temporal patterns between predicted and input variables should be evaluated, to identify problematic variables or not fitted episodes present in the data (e.g. high dust concentrations due to episodic Sahara dust advections). Examples are reported in Fig. 2.6.
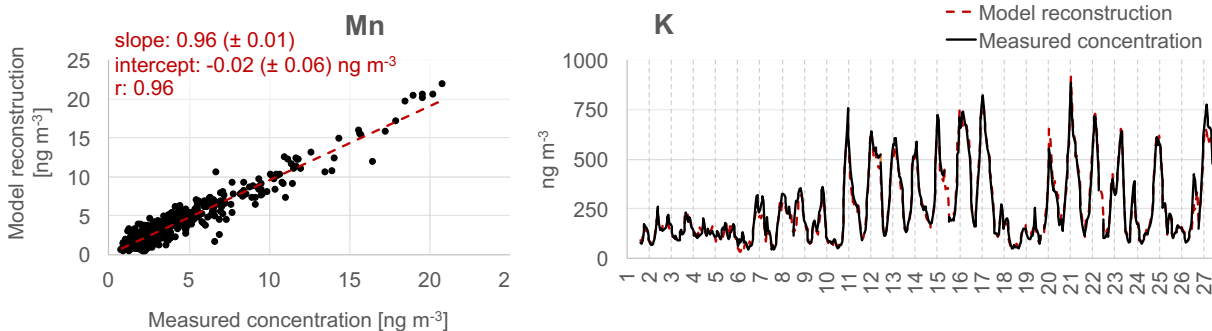
Figure 2.6: Examples of comparison between model reconstruction and measurements during February 2017 in Rome (Italy). On the left, linear regression between modelled (y axis) and measured manganese (x axis). On the right, temporal patterns of modelled and measured potassium; vertical lines show midnight for each day of February 2017.

When mass is available, sum of all chemical species for each factor should be lower than the mass apportioned to the factor itself; this can be verified both when the mass is used as an input variable and when the mass is apportioned a posteriori through a multilinear regression. In the latter case, negative contributions to the mass of some factors are a good indication that too many factors have been considered in the modelling procedure [32].

Introduction of a priori information in the obtained solution (see Sect. 2.4.1) allows the evaluation of the compatibility of model results with the physical reality. Methods to estimate uncertainty in the solution (see Sect. 2.4.2) can be used to evaluate the reproducibility and stability of the solution itself.

2.5.4 Interpretation of the solution

Factor-to-source assignment. Assignment of identities to the P factors should be based on different strategies. An initial qualitative step might consist in the identification of the main tracers for each factor. Tracers can be defined as those variables that are present with a high percentage in the considered factor, and they can be associated to specific sources on the basis of available information, e.g. previous literature studies at the same site or in similar locations [16] or targeted samplings nearby the source to get experimental profiles about it. Another parameter than can be used to identify the main tracers is the Explained Variation for factor **F** (EVF), which is another way to measure the contribution of each chemical species in each factor [43] [44]. EVF

for the chemical species j in the factor h is defined as (similar equations yields for factor $\mathbf{G}$):

$$EV_{hj} = \frac{\sum_i \frac{|g_{ih}f_{hj}|}{\sigma_{ij}}}{\sum_i \left(\frac{\sum_{k=1}^{P}|g_{ik}f_{kj}|+|e_{ij}|}{\sigma_{ij}}\right)} \qquad if \quad h = 1,...,P \tag{2.17}$$

$$UEV_{hj} = \frac{\sum_i \frac{|e_{ij}|}{\sigma_{ij}}}{\sum_i \left(\frac{\sum_{k=1}^{P}|g_{ik}f_{kj}|+|e_{ij}|}{\sigma_{ij}}\right)} \qquad if \quad h = P+1$$

where the indices i indicate the samples, k the factors, and UEVF for factor $P+1$ represents the unexplained residuals. For example, a source with relatively high EVF values for Al and Si might be associated to a crustal origin. The unexplained residuals give information on the goodness of the reconstruction of each variable, which can be considered an advantage of EVF over the percentage species. An example of comparison between percentage species and EVF for a factor identified as mineral dust is reported in Fig. 2.7.
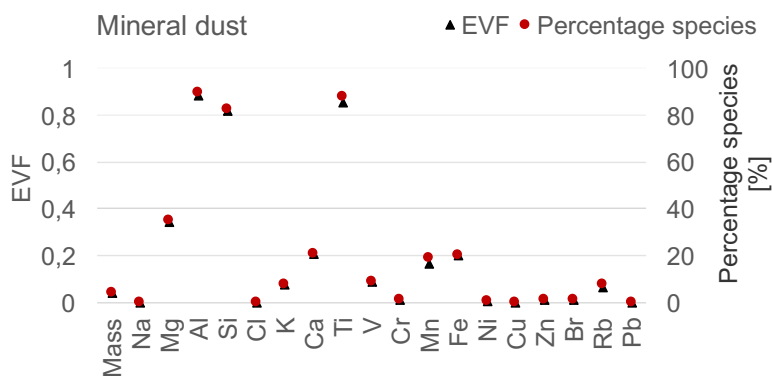
Figure 2.7: Example of use of the percentage species and explained variation for factor $\mathbf{F}$ (EVF) for the identification of main tracers.

Another common strategy is to search the literature for works containing measured source profiles with characteristics similar to the ones obtained from the analysis (factor $\mathbf{F}$). Databases of source profiles such as SPECIATE [45] [46] and SPECIEUROPE [47] are also available for such comparisons, and the DeltaSA tool [48] was developed to assist in source identification. Typical ratios between chemical species in the obtained factor profiles can help in the identification. For example, literature ratios between Na, Mg, and Cl from measurements of bulk seawater composition [5] might help in the assignment of a factor with similar ratios as a marine source, even if atmospheric transformations and enrichments can partially modify these values. Local source sampling can be perfomed (e.g. sampling of local road dust [49]) because sampled sources should resemble PMF profiles more strongly than the ones obtained in other locations.

Temporal patterns exhibited by source contributions (factor $\mathbf{G}$) can help in the factor-to-source assignment. For example, seasonality is expected from sources such as biomass burning, which likely shows largest contributions during cold winter months, or secondary sulphate, because photochemical activity is typically high during the summer. Differences in weekdays and weekends can be examined as well, since some sources like vehicular traffic are generally expected to be more active during week days due to working activities. If input data are characterised with high time resolution, more detailed information can be obtained exploiting for example average diurnal cycles (Fig. 2.8).

Auxiliary analyses. Ancillary measurements of meteorological parameters, such as wind speed and wind direction, might help in associating factors with specific directions or areas from which emissions likely originate. Local wind pattern can be analysed, also in conjunction with atmospheric concentrations of main tracers or with results from the modelling analysis. For example, the Conditional Probability Function (CPF) [50] is very useful for showing which wind directions are dominated by high concentrations. Techniques to analyse back trajectories of air masses can be useful to identify sources related to long-range transport episodes; some examples are the Potential Source Contribution Function (PSCF) or the Concentration Weighted Trajectory (CWT) [41] [51].

Correlation of factor temporal contributions (factor **G**) with time series of external tracers can be used in source identification [13], exploiting also gaseous species (e.g. different literature works consider CO as a reliable tracer for traffic vehicular emissions).
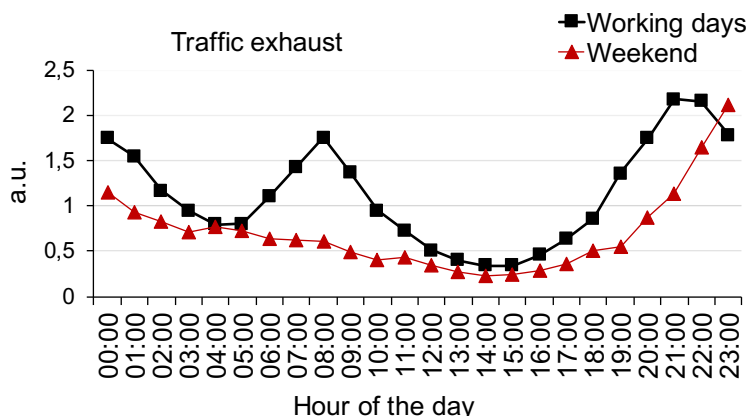
Figure 2.8: Example of differences in the average diurnal cycles for the traffic exhaust source during working days and weekends. Averages are calculated over normalised factor **G** (arbitrary unit).

2.6 Advanced receptor modelling approaches

2.6.1 Multi-time resolution model

In traditional receptor modelling approaches, input data have to be introduced at the same time resolution; if data with different time resolutions are present, high time resolution data are usually averaged on the time scale of the low time resolution ones. This practice can lead to a significant loss of information, since high time resolution data contain very detailed information about temporal patterns, with the possibility to retrieve episodic sources; unluckily, they usually lack of a complete chemical characterisation and they are often collected over intensive campaigns only. In contrast, low time resolution data can be completely speciated from a chemical point of view and they can be routinely collected over long periods of time, giving additional and complementary information respect to high time resolution data.

The multi-time resolution receptor model was developed exploiting the flexibility of ME-2, in order to use each data value in its original time schedule [39] [52]. To this aim, the basic main equation (2.1) has to be modified to take into account different time resolutions of samples:

$$x_{sj} = \frac{1}{t_{s2} - t_{s1} + 1} \sum_{k=1}^{P} f_{kj} \sum_{i=t_{s1}}^{t_{s2}} g_{ik} \eta_{jm} + e_{sj} \tag{2.18}$$

where the indices s, j, and k indicate the sample, the species, and the factor, respectively; P is the total number of factors; t_{s1} and t_{s2} are the start and end times for the sth sample in time units (i.e. the shortest sampling interval in the considered dataset); i represents one of the time units of the sth sample. η_{jm} are called "adjustment factors": they are introduced because some chemical species might be replicated with different time resolutions and measured with different analytical methods (represented by the subscript m). When η is close to unity, species concentration measured by different analytical techniques can be considered in good agreement; non-replicated species have adjustment factors set to unity by default.
As can be deduced from Eq. (2.18), factors temporal contributions g_{ik} have the temporal resolution of one time unit, i.e. the shortest sampling interval in the dataset considered. This feature can lead to physical unrealistic time variation in those factors mainly characterised by few or no species measured with high time resolution; for this reason, a regularisation equation is introduced:

$$g_{(i+1)k} - g_{ik} = 0 + \epsilon_i \tag{2.19}$$

where ϵ_i represent the residuals of this equation. The problem is solved minimising the object function Q in Eq. (2.7), where Q_{aux} comprises contributions from the regularisation equation too.

At the state of the art, source apportionment studies carried out by the multi-time resolution model are still scarce in the literature. Nevertheless, the few available studies have already tested and applied the model in different ways: in combination with ancillary analyses based on wind direction [39]; using datasets with very different time resolutions (30 minutes and 24 hours) [52]; exploiting synthetic datasets and sensitivity analyses to examine how profile variations, measurement errors, and source collinearities affect the model performance [53]; using a composite dataset including volatile organic compounds and particulate species [54]; implementing the possibility to add a priori information and using model results to apportion risk in the exposed population [55]; implementing the bootstrap technique to evaluate solution uncertainties [56]; exploiting measurements from online aerosol mass spectrometers [57] [58].

The general modelling procedure outlined in Sect. 2.5 is valid also when data aim to be analysed through advanced receptor modelling approaches. Some differences may arise in the preparation of input variables and ME-2 script. Basic ME-2 script for the multi-time model can be found at

https://datalystica.com/userdata/filemanager/data/ME2_engine.zip or

https://www.psi.ch/sites/default/files/import/lac/ModelSoFiEN/ME2_engine.zip.

In the following, details about data preparation for the multi-time resolution model are given; the paragraph can be considered complementary to Sect. 2.5.2.

Data preparation

Missing data. Attention must be paid when missing data have to be introduced in the multi-time resolution model; in fact, the model appeared to be more sensitive to missing data than the traditional PMF since its first applications. Imputation of missing data as empty entries or as treated according to Polissar et al. [40], might give large artifact peaks in factors where a key variable is missing. For this reason, a different approach is proposed in the literature [39]: x_{sj} for missing data is obtained through linear interpolation; the associated uncertainty σ_{sj} is calculated as $k \cdot x_{sj}$, where k is a coefficient whose values is between 0.75 and 5. k values are chosen depending on the total range of concentrations of the input variable j considered.

Time units. Information about the time interval of each sample must be given as input, since samples with different time resolutions are introduced. This piece of information is implemented in the multi-time model through the "time units". A time unit is defined as the shortest sampling interval in the considered dataset; e.g. if both hourly and daily samples are present, a time unit corresponds to one hour. Example of an input data file for the joint analysis of hourly and daily samples is given in Fig. 2.9; information on time units for each sample are given from column 6 to 8. The column labelled as "Length" contains the time interval of each sample in terms of time units: for hourly samples, Lenght=1; for daily samples, Length=24. "Begin" and "End" correspond respectively to t_{s1} and t_{s2} in Eq. (2.18), i.e. the start and end time units of each sample. Time units are assigned starting from the oldest sample (in Fig. 2.9, the daily sample starting at 20/06/2016 h9) in ascending order. It is noteworthy that the total number of samples is generally different from the total number of time units.

Adjustment factors. Adjustment factors (η_{jm} in Eq. (2.18)) are introduced in the modelling procedure because some chemical species might be replicated with different time resolutions and measured with different analytical methods. They can be implemented in the .ini file and defined as free factor elements (*freefact* in the script language), i.e. elements to be determined in the iteration, like factors $\mathbf{G}$ (identified by AA in the script) and $\mathbf{F}$ (identified by BB). In this case, a consistency check between η_{jm} values calculated by the model and differences in experimental data characterised by high and low time resolution should always be performed a posteriori. An alternative method consists in the homogenisation of data prior to their use as input in the analysis.

Other relevant parameters in multi-time script. Relevant script parameters that are characteristic of the multi-time model are given in the following:

n1: total number of time units (this value corresponds to the number of rows of factor $\mathbf{G}$).

ns: total number of samples (number of rows of matrix $\mathbf{X}$).

n2: total number of input variables (number of columns of matrix $\mathbf{X}$).

np: total number of factors P.

XXX: matrix with dimensions $ns \times n2$. In the script, it is calculated multiplying each element x_{sj} of the input data matrix $\mathbf{X}$ (identified with XX in the .ini file) with the corresponding time interval of the sth sample. Considering Eq. (2.18), XXX contains the product $x_{sj} \cdot (t_{s2} - t_{s1} + 1)$. In the script XXX is defined as *maindata*, i.e. it is the matrix used as target in the main equation.

Datestart:	Time	Dateend:	Time	Tzone	Begin	Length	End	Na	Na_std	Mg	Mg_std	Al	Al_std	Si	Si_std
11/07/16	10.00.00	11/07/16	11.00.00	EST	434	1	434	76	14	84	12	243	27	992	94
11/07/16	11.00.00	11/07/16	12.00.00	EST	435	1	435	66	13	78	12	247	27	1023	95
11/07/16	12.00.00	11/07/16	13.00.00	EST	436	1	436	80	14	78	12	296	31	1145	102
11/07/16	13.00.00	11/07/16	14.00.00	EST	437	1	437	70	13	83	12	302	33	1192	113
11/07/16	14.00.00	11/07/16	15.00.00	EST	438	1	438	62	13	78	12	298	31	1167	107
11/07/16	15.00.00	11/07/16	16.00.00	EST	439	1	439	85	15	82	13	306	34	1130	106
11/07/16	16.00.00	11/07/16	17.00.00	EST	440	1	440	88	15	67	11	295	32	1042	97
...						...								...	
20/06/16	09.00.00	21/06/16	09.00.00	EST	1	24	24	11	19	4	6	107	11	394	29
21/06/16	09.00.00	22/06/16	09.00.00	EST	25	24	48	11	19	48	49	181	16	740	54
22/06/16	09.00.00	23/06/16	09.00.00	EST	49	24	72	11	19	53	51	248	21	1045	75
23/06/16	09.00.00	24/06/16	09.00.00	EST	73	24	96	11	19	82	61	321	26	1273	91
24/06/16	09.00.00	25/06/16	09.00.00	EST	97	24	120	46	35	128	76	515	40	1793	127
25/06/16	09.00.00	26/06/16	09.00.00	EST	121	24	144	11	19	68	54	344	27	1112	80
26/06/16	09.00.00	27/06/16	09.00.00	EST	145	24	168	23	28	45	47	219	19	670	49
...						...								...	

Figure 2.9: Example of an input data file for the joint analysis of hourly and daily samples; this structure is the one needed for the basic multi-time ME-2 script. The first five columns are used for the identification of the sample; information about time units are given from column 6 to 8. Concentrations and associated uncertainties of input variables are then reported (in $ng\ m^{-3}$).

AAA: matrix with dimensions $ns \times np$. Considering Eq. (2.18), AAA contains $\sum_{i=t_{s1}}^{t_{s2}} g_{ik}$ and in the script it is defined as a *sefact*, i.e. it is calculated through a sub-expression. Sub-expressions are expressions of constant factors, and/or free factors (as in this case), and/or other *sefact*.

Smoothcols: name of the subroutine (i.e. a function that is already defined in the library) through which Eq. (2.19) is implemented in the model.

In this thesis work, the multi-time model implemented by the research group of Environmental Physics of the University of Milan [56] was used as a basis for the modelling procedure. Respect to the basic script, the model was expanded in order to allow the imposition of constraints on modelled factors, and to perform a bootstrap analysis to estimate solution uncertainties.

2.6.2 Modelling joint matrices

The term "joint matrices" is here used to denote input datasets adding variables different from the aerosol chemical composition to the usual dataset comprising the aerosol chemical components. In the literature, examples are joint analyses with gaseous species (such as volatile organic compounds) or with variables characterised by different dimensional units (e.g. particle chemical concentrations in $\mu g\ m^{-3}$ and number size distributions in *particle number* m^{-3}). At the state of the art, the most general discussion about the task of modelling variables with different dimensional units is given by Paatero through the open access review of an article in the journal *Atmospheric Chemistry and Physics Discussions* [59]. In that comment, it is stated that the problem has not been studied in depth in the literature so far, and the discussion about it is usually based on a widespread misunderstanding, as it will be outlined in the following.

Based on the nomenclature given in the review [59], joint matrices can be represented as [**X Z**], where **Z** indicates the matrix containing variables with different dimensional units respect to **X**, which contains aerosol chemical composition. [**X Z**] indicate the input data matrix in Eq. (2.8).

In contrast to what generally believed, analyses of matrices with different dimensional units are not a priori harmed by these differences; anyway, attention must be paid to the normalisation of factors in order to preserve the quantitative nature of the model. It is important to scale factors so that the average of each column of **G** is normalised to unity. In this way, elements g_{ik} of **G** are pure numbers and elements f_{kj} of **F** represent the average contribution of source k to variable j, carrying the same dimensional units of variable j in the input data matrix [**X Z**]. Contributions of aerosol species over the total mass can be retrieved "off-line", considering

only the part of the input matrix related to the aerosol chemical composition.

A parameter defined to model joint matrices is the "total weight". The total weight is defined for both matrices **X** and **Z** as:

$$total\ weight(X) = \sum_{i,j} \left(\frac{x_{ij}}{\sigma_{ij}}\right)^2 \tag{2.20}$$

$$total\ weight(Z) = \sum_{i,j} \left(\frac{z_{ij}}{\sigma_{ij}}\right)^2$$

This parameter can be used to evaluate the weights of the matrices in the modelling process: if **X** and **Z** are equally important but have different sizes, all σ_{ij} for the larger matrix should be increased, so that total weight of **X** and **Z** become approximately equal. This implies a deviation from the general principle of determining standard deviations σ_{ij}. In general, the matrix with the highest total weight will "drive the model".

Analyses of joint matrices often lead to disappointing results when some factors only (or mainly) fit variables of **X** while other factors only (or mainly) fit variables of **Z**; this result is worthless when the aim is to relate all the variables to the same emission sources. Anyway, it must be stressed that these failures must not be ascribed to the different dimensional units of the two matrices. Success of this modelling approach depends e.g. on the validity and rotational ambiguity of models computed separately for **X** and for **Z**, and on the total weight assigned to **X** and **Z**.

At the state of the art, one of the most widespread examples of a joint matrix characterised by different dimensional units is the combination of aerosol chemical composition and number size distribution. The assumption of constant profiles at the basis of receptor models (see Sect. 2.2) implies that the size distribution of particles emitted from a source does not vary significantly with time, and that measured number concentrations have a linear relationship with number and mass contributions from all sources [60] [61]. Combination of number size distribution and chemical composition data into a joint multivariate analysis allows to obtain source characteristics in both size distributions and chemical composition simultaneously (see Fig. 2.10).
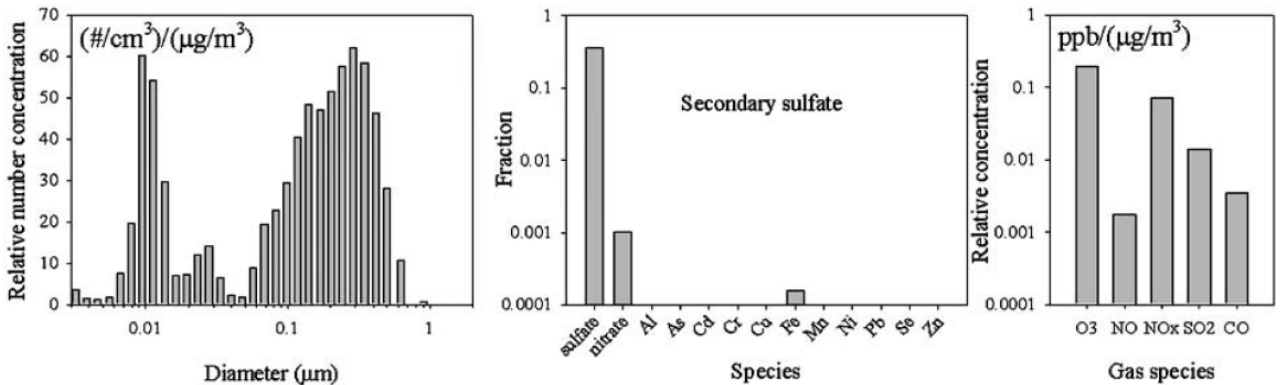
Figure 2.10: Example of a source profile that can be obtained from a joint analysis of size distribution and chemical composition (in this case, of both particles and gases). Figure adapted from [60].

It is not trivial to derive meaningful results from the analysis of joint matrices; for this reason, sometimes number size distribution and particle chemical composition are modelled separately and results from the modelling process are compared at a later stage [62] [63] [64]. A comparison between separate and joint analyses is reported in [65]. A new approach proposed in [66] attempts to exploit modelling results (factor **G**) obtained from the analysis of aerosol chemical composition as input to the joint analysis with number size distribution.

Incorporation of variables different from the particle chemical composition can be useful to reduce the rotational ambiguity of the solution. Meteorological parameters, especially wind direction and wind speed, represent valuable additional information in receptor modelling but these data cannot be used as dependent variables (i.e. as part of a joint matrix) because their relation with observed concentrations is far from linear. The "Expanded" model was developed to allow the introduction of independent variables [67] [68] [69] [70] [71]; this different modelling approach will not be investigated in this thesis. Examples of other independent variables considered in the literature are e.g. days of the week, seasons, and Radon concentrations [72] (in a less recent work [73], Radon concentrations were also tested as input variable in a joint matrix).

Chapter 3

Modelling applications

3.1 Introduction

In this chapter, implementation and application of advanced receptor modelling approaches reported in Chapter 2 are presented. During experimental campaigns data are usually collected with different time resolutions by different instrumentations, in order to completely characterise aerosol physical-chemical properties. The main aim of this PhD thesis was the implementation of a unified approach for the analysis of this type of data.

Data measured with different time resolutions were introduced as input to an advanced model, which is still scarcely used in the literature. In Sect. 3.2 a new methodology is proposed to exploit aerosol optical absorption data together with more traditional chemical composition measurements in receptor modelling. Results from this approach gave insights on absorption optical properties of aerosol sources, overcoming the main limitations of optical source apportionment methods (i.e. the fixed number of sources and a priori values for source-specific absorption Ångström exponent). In Sect. 3.3 a further investigation concerns the identification of primary and secondary organic contributions apportioned to the different emission sources - which is very challenging although being paramount in aerosol science - exploiting an original approach to data analysis based on receptor modelling. This approach gave promising results with a view to identifying the origin of secondary aerosol, which is an open issue for receptor models due to the hypotheses of the models themselves. Results of the case-studies presented here are reported in papers [31] and [74].

3.2 Exploiting multi-wavelength aerosol light absorption coefficients in receptor modelling

In this work, the multi-time resolution receptor model for source apportionment (see Sect. 2.6.1) implemented by the research group of Environmental Physics of the University of Milan [56] was used as a basis for the modelling process. Input data with different time resolutions (24 hours, 12 hours, and 1 hour) were organised in a joint matrix (see Sect. 2.6.2) consisting of aerosol chemical composition data (in $ng\ m^{-3}$) and multi-wavelength aerosol light absorption coefficient b_{ap} (in Mm^{-1}). Besides the traditional source apportionment, the impact of different sources on b_{ap} was estimated. First of all, this piece of information allowed a more robust identification of sources; moreover, this approach led to the assessment of source-dependent mass absorption cross section (MAC) and absorption Ångström exponent (α) values as a result of the model, without any a priori assumption. It is worth noting that currently in source apportionment models based on optical absorption data - e.g. the Aethalometer model [75], or the multi-wavelength absorption analyzer (MWAA) model [76] [77] - values for α related to fossil fuel emissions and biomass burning are fixed by the modeller, thus carrying a large part of the uncertainties in the model results as also reported by literature works (see e.g. [78] [79]).

3.2.1 Site description and aerosol sampling

Two measurement campaigns were performed during summertime (June – July) and wintertime (November – December) 2016 in Milan (Italy). Milan is the largest city (more than 1 million inhabitants, doubled by commuters everyday) of the Po Valley, a very well-known hotspot pollution area in Europe due to both large emissions from a variety of sources (i.e. traffic, industry, domestic heating, energy production plants, and

agriculture) and low atmospheric dispersion conditions [80] [81] [82] [83] [84]. The sampling site is representative of the urban background and it is situated at about 10 m above the ground, on the roof of the Physics Department of the University of Milan, less than 4 km from the city centre [85]. It is important to note that during the sampling campaigns, a large building site was active next to the monitoring station.

Aerosol sampling was carried out using instrumentation with different time resolutions. Low time resolution PM10 data, with sampling durations of 24 and 12 hours during summertime (20 June – 22 July 2016) and wintertime (21 November – 22 December 2016), respectively, were collected in parallel on PTFE (Whatman, 47mm diameter) and pre-fired (700°C, 1h) quartz-fibre (Pall, 2500QAO-UP, 47mm diameter) filters. Low-volume samplers with EPA PM10 inlet operating at 1 m^3h^{-1} were used. High time resolution data were collected during shorter periods (11 – 18 July and 21 – 28 November 2016) by a streaker sampler [86]. Briefly, the streaker sampler collects the fine and coarse PM fractions (particles with aerodynamic diameter $d_{ae} < 2.5$ μm, and $2.5 < d_{ae} < 10$ μm, respectively; see Sect. 1.1 for d_{ae} definition) with hourly resolution. Particles with $d_{ae} > 10$ μm impact on the first stage and are discarded, while the coarse fraction deposits on the second stage consisting of a Kapton foil; finally, the fine fraction is collected on a polycarbonate filter. The two collecting supports are kept in rotation with an angular speed of about $1.8°h^{-1}$ to produce a circular continuous deposit on both stages. The streaker sampler operates at 1 l min^{-1}.

Meteorological data were available at a monitoring station belonging to the regional environmental agency (ARPA Lombardia) which is less than 1 km away.

3.2.2 PM mass concentration and chemical characterisation

PM10 mass concentration was measured on PTFE filters by a gravimetric technique. Weighing was performed by an analytical balance (Mettler, model UMT5, 1 μg sensitivity) after a 24 hours conditioning period in an air-controlled room as for temperature (20 ± 1°C) and relative humidity (50 ± 3 %). These filters were then analysed by energy dispersive X-ray fluorescence (ED-XRF) analysis to obtain the elemental composition. In this work, the ED-XRF spectrometer ED2000 by Oxford Instruments was used; it is equipped with a Coolidge tube with an Ag anode, and a solid-state Si(Li) detector. Micromatter^{TM} standard reference samples were used to retrieve the sensitivity curve for the quantitative analysis. Spectra deconvolution was performed with the software WinQxas (software developed by International Atomic Energy Agency). More details about the set-up can be found elsewhere (see e.g. [87]). For most elements and samples, concentrations were characterised by relative uncertainties in the range 7% – 20% (higher uncertainties for elements with concentrations next to MDLs) and minimum detection limits of 0.9 – 30 ng m^{-3} (depending on the element) with the above-mentioned sampling conditions. All the analyses on PTFE filters were performed at the Department of Physics of the University of Milan.

Quartz-fibre filters were analysed at the Department of Chemistry of the University of Milan. For each quartz-fibre filter, one punch (1.5 cm^2) was extracted by sonication (1 hour) using 5mL ultrapure Milli-Q water and levoglucosan and inorganic anions concentrations were quantified. Levoglucosan concentration was determined by high-performance anion exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD) [88] only in winter samples. Indeed, as already pointed out by other studies at the same sampling site [89] and routinely assessed at monitoring stations in Milan by the regional environmental agency (private communication), levoglucosan concentrations during summertime are lower than the MDLs (i.e. about 6 ng m^{-3}), due to both lower emissions (no influence of residential heating and negligible impact from other sources) and higher OH levels in the atmosphere depleting molecular marker concentrations [90] [91]. Uncertainties on levoglucosan concentration were about 11%. The quantification of the main water-soluble inorganic anions (SO_4^{2-} and NO_3^-) was performed by ion chromatography (IC) [92]; MDLs were 25 ng m^{-3} and 50 ng m^{-3} with summertime and wintertime sampling conditions, respectively, and uncertainties were about 10%. Unfortunately, due to technical problems, no data on ammonium were available.

Another punch (1.0 cm^2) of each quartz-fibre filter was analysed by thermal optical transmittance analysis (TOT, Sunset Inc., NIOSH-870 protocol) [93] in order to assess OC and EC concentrations. MDLs were 75 ng m^{-3} and 150 ng m^{-3} with summertime and wintertime sampling conditions, respectively, and uncertainties were in the range 10% – 15%.

Hourly elemental composition was assessed by particle-induced X-ray emission (PIXE) technique at the 3 MV Tandetron accelerator of INFN-LABEC laboratory (Sesto Fiorentino, Florence, Italy), where an external beam facility is dedicated to measurements of atmospheric aerosol. A properly collimated proton beam is extracted from vacuum to air through a thin extraction window, and samples are scanned in steps corresponding

to 1 hour aerosol deposit. During the analysis the charge is measured by integrating the beam current on a graphite Faraday cup positioned just behind the sample. To obtain efficient simultaneous detection of all elements, different Silicon Drift Detectors (SDDs) optimised for low and medium–high X-ray energies are used. Micromatter[TM] standard reference samples were used to obtain the calibration curve for the quantitative analysis. PIXE spectra were fitted using the software GUPIXWIN [94]. More details about the experimental set-up can be found elsewhere (see e.g. [95] [96]). As low time resolution PM10 samples were also available, fine and coarse elemental concentrations determined by PIXE analysis were added up to obtain PM10 concentrations with hourly resolution. PM10 hourly concentrations for most elements and samples were characterised by relative uncertainties in the range $10\% - 30\%$ (higher uncertainties for elements near MDLs) and MDLs ranged from a minimum of 0.1 $ng\ m^{-3}$ to a maximum of 15 $ng\ m^{-3}$ (higher MDLs typically detected for Z<20 elements).

3.2.3 Aerosol light absorption coefficients measurements

The aerosol light absorption coefficient (b_{ap}) at four wavelengths λs = 405 nm, 532 nm, 635 nm, and 780 nm, was measured on both low and high time resolution samples with the in-house polar photometer PP_UniMI, which is described in detail in Vecchi et al. [97] and Bernardoni et al. [98]. Briefly, a laser beam impinges perpendicularly on the filter so that it hits the particle layer first. Light scattered at angles from 0° to 173° (with steps of about 0.4°) is continuously collected by a photodiode located on a rotating arm whose centre of rotation coincides with the centre of the sample. The angular distributions of light transmitted and scattered by the blank and by the sampled filter are used to calculate integrals of the light scattered in the forward and in the backward hemispheres. Finally, a radiative transfer model is applied in order to retrieve b_{ap} from these quantities. The radiative transfer model here considered was firstly developed by Hänel [99] and it is analogous to the one implemented in the Multi-Angle Absorption Phomoter (MAAP) [100]. It is based on two steps: the "adding method" takes into account interactions (such as multiple scattering effects) between particles deposited on the filter and the filter itself; the "two stream approximation" [101] is used to describe radiative processes in the particles layer.

Low time resolution optical measurements were performed on PTFE filters since their physical characteristics can be considered more similar to polycarbonate filters used by the streaker sampler. Moreover, previous works reported a bias on b_{ap} measured by instrumentation using fibre filters ([102] [103] [104] and references therein). Vecchi et al. [97] found that b_{ap} at 635 nm was up to 40% higher when measured on a quartz-fibre filter compared to parallel samples collected on PTFE. This effect was ascribed to sampling artefacts due to organics in aerosol samples collected in Milan.

As for high time resolution samples, b_{ap} was measured only in the fine fraction collected on polycarbonate filters since absorption of the Kapton foil on which the coarse fraction was collected did not allow b_{ap} assessment. Anyway, b_{ap} values in PM2.5 and PM10 were expected to be fairly comparable, as aerosol absorption in the atmosphere is mostly due to particles in the fine fraction at heavily polluted urban sites like Milan. To verify this assumption, high time resolution b_{ap} data in PM2.5 were averaged over the timescale of low time resolution b_{ap} in PM10 and compared (Fig. 3.1); the agreement was good, between 11% and 13% depending on the λ, except for b_{ap} at λ = 405 nm for which b_{ap} values in PM10 were significantly higher (27%) but with most data (83%) within experimental uncertainties. To take into account this difference, b_{ap} data at λ = 405 nm were homogenised before using them in the model, following the criterion used for chemical species (further details about the homogenisation procedure can be found in Sect. 3.2.5).

Uncertainties on b_{ap} were quantified in 15% and MDL was in the range $1 - 10$ Mm^{-1} depending on sampling duration and wavelength [97] [98]. The pre-treatment procedure for experimental uncertainties and MDLs was the same used for chemical variables in order to create suitable input matrices required by the multi-time resolution model. Optical system stability was checked during the measurement session, evaluating the reproducibility of the measurement on a blank test filter. Laser stability was also checked at least twice a day and the recorded intensities were used to normalise blank and sampled filter analysis.
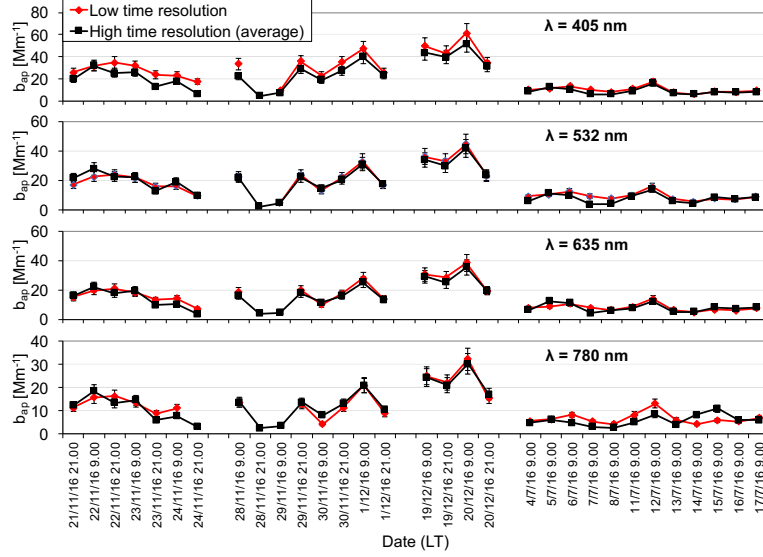
Figure 3.1: Comparison between low time resolution (12 and 24 hours) and average on the same time period of high time resolution (originally 1 hour) aerosol absorption coefficient at different wavelenghts measured by the polar photometer PP_UniMI. Low time resolution measurements were performed on PM10 collected on PTFE filters, while high time resolution ones were performed on PM2.5 collected on polycarbonate filters.

3.2.4 Concentration values

Basic statistics on mass and chemical species concentrations at different time resolutions are given in Table 3.1.

Table 3.1: Basic statistics (mean, median, 10th percentile, 90th percentile) on mass and chemical species concentrations at different time resolutions. Note that high time resolution samples were collected during shorter periods (about two weeks) respect to low time resolution ones (about two months). NA represents Not Available data [31].

	Low time resolution samples				High time resolution samples			
$\mu g\ m^{-3}$	mean	median	$10^{th}perc$	$90^{th}perc$	mean	median	$10^{th}perc$	$90^{th}perc$
PM10 mass	38.4	30.9	16.2	76.6	NA	NA	NA	NA
EC	1.39	1.18	0.54	2.63	NA	NA	NA	NA
OC	8.1	6.2	3.4	14.7	NA	NA	NA	NA
NO_3^-	5.19	2.63	0.71	13.9	NA	NA	NA	NA
SO_4^{2-}	2.95	2.55	0.89	5.56	NA	NA	NA	NA
Levoglucosan	0.68	0.63	0.33	1.15	NA	NA	NA	NA
$ng\ m^{-3}$	mean	median	$10^{th}perc$	$90^{th}perc$	mean	median	$10^{th}perc$	$90^{th}perc$
Na	89	70	27	165	89	60	21	196
Mg	59	59	24	94	44	40	20	73
Al	178	161	78	287	111	91	38	218
Si	700	635	331	1104	413	340	146	737
S	1320	1004	411	2908	696	630	167	1246
K	467	349	198	871	318	257	80	611
Ca	912	954	418	1475	512	420	211	948
Cr	10	9	4	15	7	7	4	11
Mn	14	12	5	24	8	7	4	13
Fe	1120	940	577	2040	794	692	335	1363
Cu	45	36	21	92	34	29	12	63
Zn	90	83	35	148	66	56	16	133
Pb	24	17	10	50	20	13	6	35

Most variables showed higher mean and median concentrations during the winter campaign, when atmospheric stability conditions influenced the monitoring site; exceptions were Al, Si and Ca, which had lower

median concentrations (as detected in low time resolution samples). This was expected as they are typical tracers of soil dust resuspension [16] that can be more relevant during summertime due to drier soil conditions and stronger atmospheric turbulence. Moreover, the good correlation between these elements - Al vs. Si: $R^2 = 0.94$ (see Fig. 3.2) and Ca vs. Si: $R^2 = 0.78$ - suggested their common origin.

High time resolution
slope: 0.261 ± 0.004
intercept: 2.1 ± 2.0
R²: 0.93

Low time resolution
slope: 0.268 ± 0.009
intercept: -9.3 ± 7.2
R²: 0.94
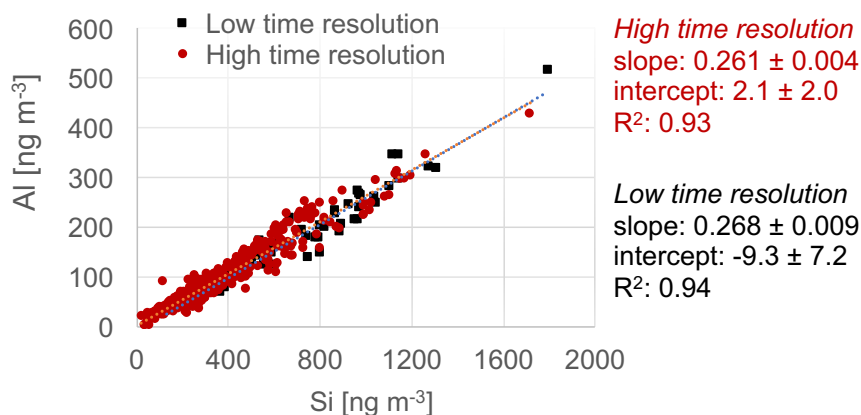
Figure 3.2: Linear regression between Al and Si concentrations. Note that high time resolution samples were collected during shorter periods (about two weeks) respect to low time resolution ones (about two months).

Potassium showed the clearest seasonal behaviour in concentration values going from 284 $ng\ m^{-3}$ (10th–90th percentile: 151–344 $ng\ m^{-3}$) to 660 $ng\ m^{-3}$ (10th–90th percentile: 349–982 $ng\ m^{-3}$) in summer and winter, respectively, in low time resolution samples. K is an ambiguous tracer, since it is emitted by a variety of sources, among which there are crustal resuspension and biomass burning. In our dataset, wintertime K values showed a good correlation with levoglucosan concentrations ($R^2 = 0.71$), suggesting the impact of biomass burning as levoglucosan is a well-known tracer for biomass burning emissions in winter samples [105]. Also looking at the K-to-Si ratio (where Si was taken as a soil dust marker), significant seasonal differences came out (Fig. 3.3); it was 0.35 ± 0.15 in high time resolution summer samples and 2.0 ± 2.2 in winter ones, to be compared with the much more stable Al-to-Si ratio (i.e. 0.26 ± 0.04 and 0.28 ± 0.09 in summer and winter, respectively) indicating a soil-related origin (Al-to-Si ratio equal to 0.3 is reported e.g. by [106] for earth crust composition).
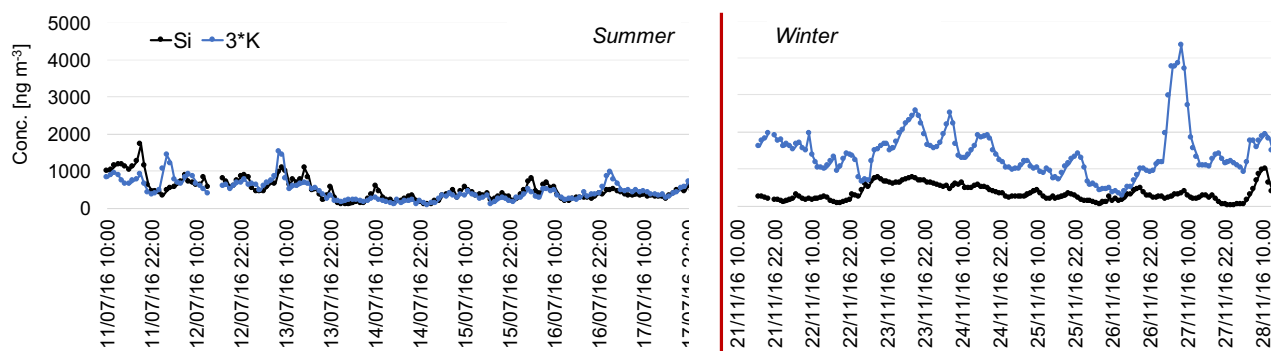
Figure 3.3: Temporal patterns of Si and K concentrations during summer and winter campaign.

Among the elements typically associated with anthropogenic sources, Fe and Cu showed a good correlation (e.g. $R^2 = 0.72$ on hourly resolution samples) as well as Cu and EC (Cu vs. EC: $R^2 = 0.84$, on low time resolution data); in addition, the diurnal pattern of Fe and Cu showed traffic rush hours peaks (07:00–09:00 LT and around 19:00 LT) as shown in Fig. 3.4. These results were suggestive of a common source; in the literature these aerosol chemical components are reported as tracers for vehicular emissions. In particular, EC is associated to exhaust emissions from tailpipe, while Fe and Cu to non-exhaust emissions as reported e.g. in [16] [107].
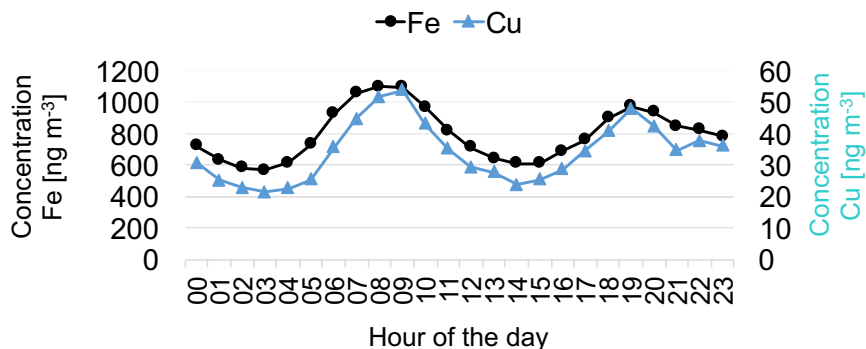
Figure 3.4: Diurnal profile of Fe and Cu concentrations [31].

Diurnal mean temporal patterns for b_{ap} at different wavelengths (retrieved from hourly resolved data) are displayed in Fig. 3.5.
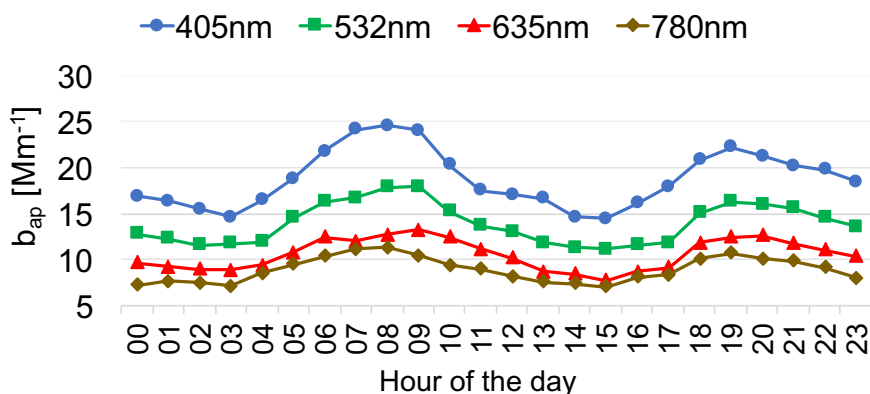
Figure 3.5: Diurnal profile of the aerosol absorption coefficient b_{ap} measured at different wavelengths [31].

3.2.5 Modelling a joint matrix with the multi-time resolution model

Input data

Concentrations of replicated species with different time resolutions were pre-homogenised and then inserted into the input matrix $\mathbf{X}$, instead of using adjustment factors in the model (see Sect. 2.6.1 for more information about adjustment factors). Concentration data with longer sampling intervals (24 and 12 hours in this work) were considered to be the benchmark, since analytical techniques usually show a better accuracy on concentration values far from MDLs (i.e. samples collected on longer time intervals) [39] [52].

Variables were then classified as weak and strong according to the signal-to-noise ratio (S/N) criterion (see Sect. 2.5.2). For hourly data, only strong variables (S/N > 1.2) were considered; for low time resolution data, weak variables such as Na, Mg, and Cr (with S/N equal to about 0.8) being strong variables in hourly samples were also included, although with associated uncertainties comparable to concentration values in order to avoid the exclusion of too many data. Indeed, excluding these low time resolution variables from the analysis gave rise to artificial high values in the time contribution matrix $\mathbf{G}$ for sources traced by these species (in this case it was an issue for aged sea salt traced by Na and Mg); this oddity has been already reported in the literature [39].

Ranges of experimental uncertainties and MDLs are reported in Sect. 3.2.2 and 3.2.3 for chemical and optical absorption analyses, respectively. Variables with more than 20% of the concentration data below MDL values were omitted from the analysis [52]. Uncertainties and data below MDL were pre-treated according to Polissar et al. [40] (for further details see Sect. 2.5.2). In general, missing concentration values were estimated by linear interpolation of the measured data and their uncertainties were assumed to be 3 times this estimated value [39] [52]. As for summertime levoglucosan data (always below MDLs), the approach was to include them

as below MDL data and not as missing data following Zhou et al. [39], who underlined that the multi-time resolution model is more sensitive to missing values than the original PMF model.

In order to avoid double counting, in this study S was chosen as the input variable instead of SO_4^{2-} as it was determined on both low time and high time resolution samples (by XRF and PIXE analysis, respectively [108]). However, SO_4^{2-} and S concentrations showed a high correlation (correlation coefficient R = 0.98) and the Deming regression gave a slope of 2.68 ± 0.13 (sulphate vs. sulphur) with an intercept of -198 ± 82 $ng\ m^{-3}$, i.e. compatible with zero within 3 standard deviations. The slight difference (of the order of 10%) between the estimated slope and the SO_4^{2-}-to-S stoichiometric coefficient (i.e. 3) can be ascribed to either a small fraction of insoluble sulphate or to the use of different analytical techniques.

As already mentioned in Sect. 2.5.1, PM10 mass concentrations were included in the model with uncertainties set at 4 times their values in order to lower their weight in the analysis. In the end, 22 low time resolution variables (PM10 mass, Na, Mg, Al, Si, S, K, Ca, Cr, Mn, Fe, Cu, Zn, Pb, EC, OC, levoglucosan, NO_3^-, b_{ap} 405 nm, b_{ap} 532 nm, b_{ap} 635 nm, b_{ap} 780 nm) and 17 hourly variables (Na, Mg, Al, Si, S, K, Ca, Cr, Mn, Fe, Cu, Zn, Pb, b_{ap} 405 nm, b_{ap} 532 nm, b_{ap} 635 nm, b_{ap} 780 nm) were considered.

The input matrix $\mathbf{X}$ consisted of 386 samples and the total number of time units (see Sect. 2.6.1 for the definition of time unit in the multi-time model) was 1117. The analysis was performed in the robust mode; the lower limit for $\mathbf{G}$ contribution was set to -0.2 [30] and the error model em=-14 was used for the main equation with C1 = input error, C2 = 0.0, and C3 = 0.1 (see Sect. 2.5.2) for both chemical and optical absorption data.

Sensitivity tests on the uncertainty of absorption data were performed starting from a minimum experimental uncertainty of 10%. Lower uncertainties were considered not physically meaningful from an experimental point of view. ME-2 analyses performed with 10% experimental uncertainty on absorption data gave very similar results to the base-case solution that will be presented in the following, with no differences in mass apportionment and a maximum variation in the concentrations of chemical and optical absorption profiles (matrix $\mathbf{F}$) of 7% when considering significant variables in each profile (i.e. explained variation for matrix $\mathbf{F}$ EVF - see Sect. 2.5.4 for EVF definition - higher or near 0.30). In contrast, considering an experimental uncertainty of 20% on absorption data, the solution significantly differed from the base-case one and showed less physical meaning (e.g. a couple of factors got mixed, or an additional unique factor appeared giving a null mass contribution). Thus, the estimated relative experimental uncertainty of 15% was here considered appropriate for optical variables.

Base-case solution

Different solutions (from 5 to 10 factors) were explored; after 30 convergent runs, the eight-factor base-case solution corresponding to the lowest Q value (2086.88) was firstly selected (Fig. 3.6). It is important to note that the model was run using all variables (chemical + optical). A lower or higher number of factors caused ambiguous chemical profiles and the physical interpretation suggested clearly mixed sources for a lower number of factors or unique factors in case of more factors (i.e. Pb for nine factors); moreover, inconsistent mass closure was detected by increasing the number of factors (e.g. the sum of species contribution was up to 25% higher than the mass for the 10-factor solution). In the eight-factor base-case solution, the mass was well reconstructed by the model ($R^2 = 0.98$), with a slope of 0.98 ± 0.02 and a negligible intercept (0.51 ± 0.89 $\mu g\ m^{-3}$).

The factor-to-source assignment was based on both EVF values – which are typically higher for chemical tracers [43] [44] – and the physical consistency of factor chemical profiles. In the chosen solution, the unexplained variation was lower than 0.25 for all variables. The uncertainty-scaled residuals (see Sect. 2.5.3 for the definition) showed a random distribution of negative and positive values in the ±3 range, with a Gaussian shape for most of the variables (Fig. 3.7).

Using EVF and chemical profiles reported in Fig. 3.6, the eight factors were tentatively assigned to nitrate, sulfate, resuspended dust, biomass burning, construction works, traffic, industry, and aged sea salt. In Table 3.2 absolute and relative average source contributions to PM10 mass are reported.

Although the above-mentioned base-case solution was a satisfactory representation of the main sources active in the area (as reported in previous works [33] [85] [89] [109] [110] [111]), the chemical profile of some factors was improved by exploring rotated solutions.
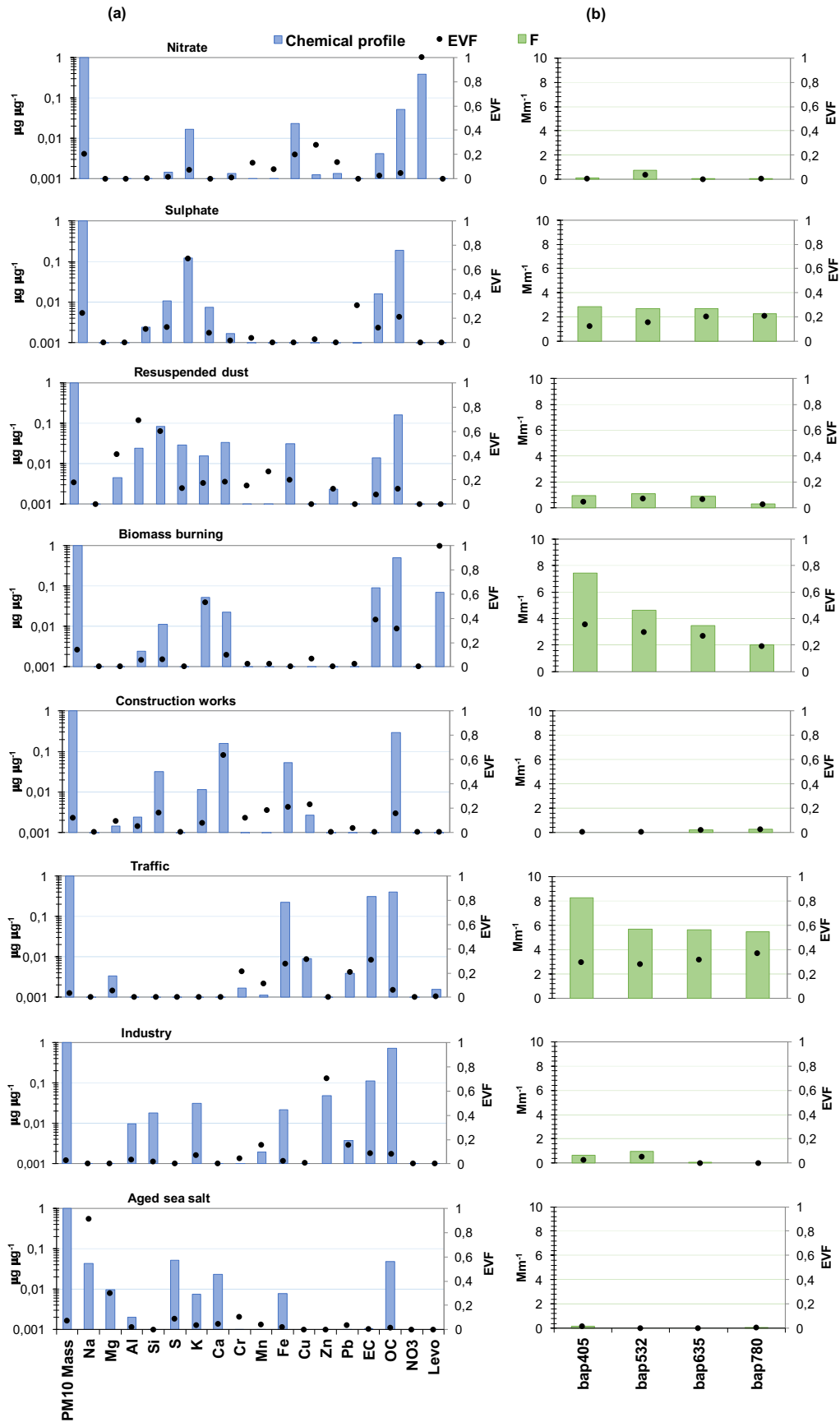
Figure 3.6: (a) Chemical profiles of the 8-factor base-case solution (b) b_{ap} apportionment of the 8-factor base-case solution. The blue bars represent the chemical profile (output of the matrix **F** normalised on mass), the green bars the output values of the matrix **F**, and the black dots the EVF [31].
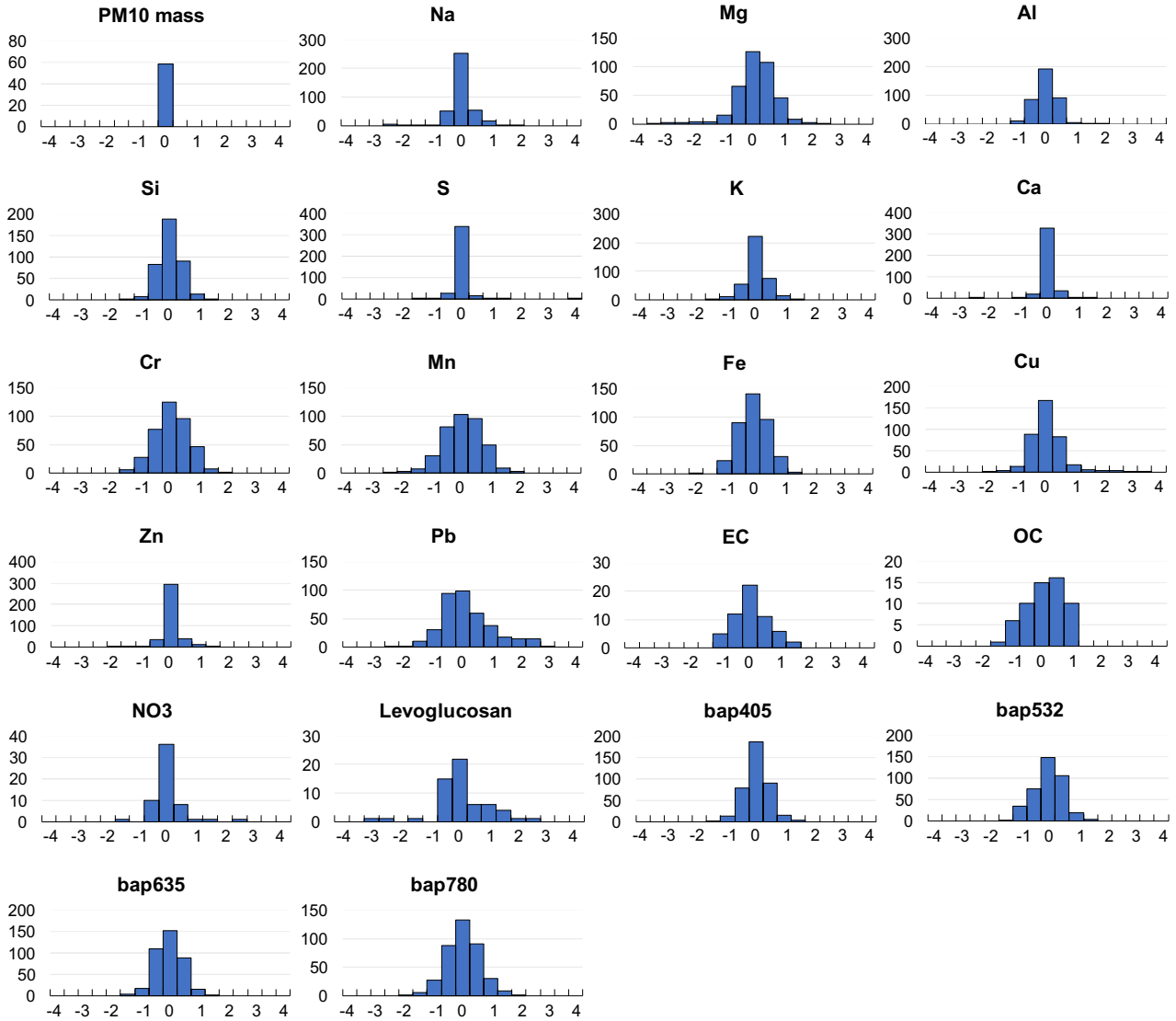
Figure 3.7: Uncertainty-scaled residuals (residuals divided by input data uncertainties) of the 8-factor base case solution. The x-axis represents the scaled residuals values, while the y-axis represents the frequencies [31].

Table 3.2: Absolute and relative average source contributions to PM10 mass in the 8-factor base case solution [31].

Factors	Source contributions $[\mu g\ m^{-3}]$
Nitrate	10.4 (31%)
Sulphate	6.2 (19%)
Resuspended dust	5.5 (16%)
Biomass burning	3.5 (11%)
Construction works	3.6 (11%)
Traffic	1.7 (5%)
Industry	1.1 (3%)
Aged sea salt	1.3 (4%)

Final constrained solution and bootstrap analysis

The multi-time resolution model implemented by the research group of Environmental Physics of the University of Milan [56] allows the imposition of constraints on modelled factors in order to explore rotated solutions, and to perform a bootstrap analysis to estimate solution uncertainties.

The aged sea salt factor was investigated further, because typical diagnostic ratios such as Mg/Na and Ca/Na (in bulk seawater equal to 0.12 and 0.04, respectively; see Table 1.1) were not well reproduced in the base-case solution and the chemical profile was too much impacted by the presence of Fe compared to bulk seawater composition. Therefore, the above-mentioned diagnostic ratios were here used as constraints and Fe was maximally pulled down in the chemical profile. The effective increase in Q was about 61 units ($Q =$ 2147), with a percentage increase of about 3%; as a rule of thumb, an increase in the Q value of a few tens is generally considered acceptable [25]. It is noteworthy that the constrained solution led to an improvement in the chemical profile of the aged sea salt, and negligible differences in all other relevant features of the solution (i.e. EVF, residuals, mass reconstruction, source apportionment) were found compared to the base-case solution. Therefore, the eight-factor constrained solution was considered the most physically reliable; results are presented in Table 3.3 and Fig. 3.8 and discussed in detail in the following.

Table 3.3: Absolute and relative average source contributions to PM10 mass in the 8-factor constrained solution [31].

Factors	Summer [$\mu g\ m^{-3}$]	Winter [$\mu g\ m^{-3}$]	Total [$\mu g\ m^{-3}$]
Nitrate	3.6 (15%)	21.1 (44%)	10.2 (31%)
Sulphate	6.3 (26%)	8.1 (17%)	7.0 (21%)
Resuspended dust	4.6 (19%)	1.7 (4%)	3.5 (11%)
Biomass burning	0.32 (1%)	8.3 (17%)	3.3 (10%)
Construction works	5.9 (24%)	3.4 (7%)	4.9 (15%)
Traffic	1.4 (6%)	2.2 (5%)	1.7 (5%)
Industry	0.86 (4%)	1.2 (3%)	1.0 (3%)
Aged sea salt	1.4 (6%)	1.8 (4%)	1.6 (5%)

The factor interpreted as nitrate fully accounted for the explained variation of NO_3^-. This factor contained a significant fraction of nitrate in the chemical profile (39%) and all nitrate was present only in this factor. This source was by far the most significant one at the investigated site, explaining about 31% of the PM10 mass over the whole campaign (a similar estimate – 26%– was reported by Amato et al. [33] during the AIRUSE campaign in Milan in 2013) increasing up to 44% during wintertime (comparable to the 37% reported by Vecchi et al. [110]). Indeed, the Po Valley is well known for experiencing very high nitrate concentrations during wintertime ([110] and references therein) because of large emissions of gaseous precursors related to urban and industrial activities, residential heating, high ammonia levels due to agricultural field manure and poor atmospheric dispersion conditions.

The factor associated with sulphate showed EVF = 0.47 for S and much lower EVF for all the other variables in the factor. Considering the sulfur contribution in the chemical profile in terms of ammonium sulphate, the relative contribution of sulphur components in the profile increased from 11% (S) up to 45% (ammonium sulphate). The latter is the main sulphur compound detected in the Po Valley as reported in previous papers (e.g. [112]) and was by far the highest contributor in the chemical profile. The other important contributor was OC (19%), whose impact on PM mass increased up to 30% when reported as organic matter using 1.6 as the organic carbon-to-organic matter conversion factor for this site [87]. Due to the secondary origin of the aerosol associated with this factor, it was not surprising to also find a significant OC contribution; indeed, aerosol chemical composition in Milan is impacted by highly oxygenated components due to aging processes favoured by strong atmospheric stability [81] [110]. In this factor, EC contributed about 1%. Considering the total EC concentration reconstructed by the model, the EC fraction related to the sulphate factor was about 6%. In contrast to sulphates, EC has a primary origin; however, its presence with a very similar percentage (4%–5%) in a sulphate chemical profile was previously pointed out in Milan, indicating a more complex mixing between primary and secondary sources [33], e.g. with sulphate condensation on primary emitted particles. The sulphate factor accounted for 21% of the PM10 mass.

The factor identified as resuspended dust was mainly characterised by high EVF and contributions coming from Al, Si and Mg, i.e. crustal elements. The Al/Si ratio was 0.31, very similar to the literature value for average crustal composition [106]; the relatively high OC contribution in the chemical profile (15%) together with the presence of EC (about 2.6%) was suggestive of a mixing with road dust [107]. This source explained for about 11% of the PM10 mass.

The factor identified as biomass burning was characterised by high EVF for levoglucosan (0.98), a known
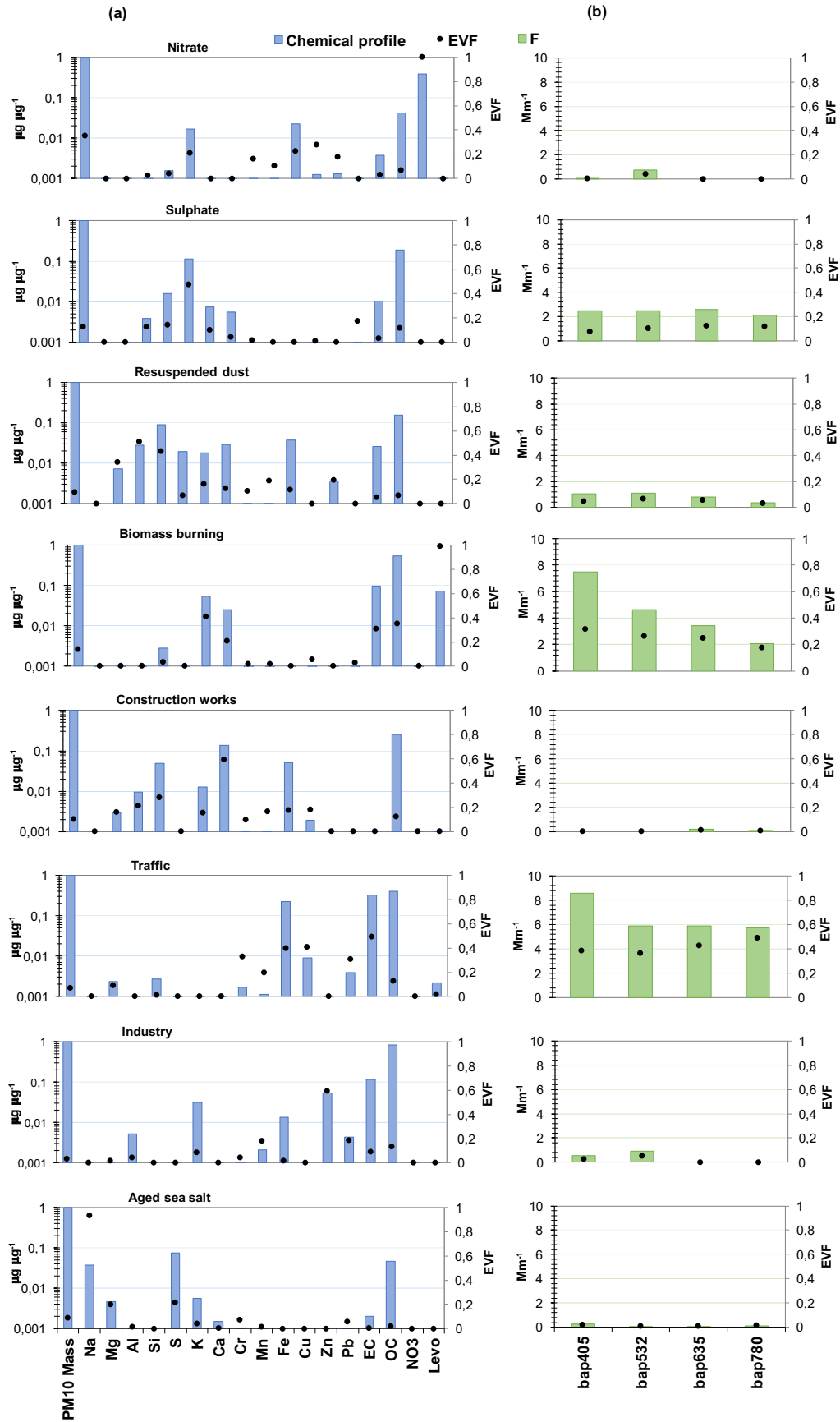
Figure 3.8: (a) Chemical profiles of the eight-factor constrained solution; (b) b_{ap} apportionment of the eight-factor constrained solution. The blue bars represent the chemical profile (output of the matrix **F** normalized on mass), the green bars the output values of the matrix **F** for the optical absorption variables, and the black dots the EVF [31].

tracer for this source as it is generated by cellulose pyrolysis; EVF higher than 0.3 was also found for K, OC, and EC. In the source chemical profile, OC contributed 54%, EC 10%, levoglucosan 7%, and K 5%. The average biomass burning contribution during this campaign was 10% (up to 17% in wintertime). Anticipating the discussion presented in detail in the next paragraph, it is worth noting that the second largest contribution to the aerosol absorption coefficient after traffic was detected in this factor.

The factor with high EVF (0.60) for Ca was associated with construction works, following literature works ([85] [89] [111] [113] [114] and references therein). Major contributors to the chemical profile were Ca (13%), OC (26%), Fe, and Si (5% each). This factor accounted on average for 15% to PM10 mass. As already mentioned, during the campaign a non-negligible contribution from this source was expected, due to the presence of a building site nearby the monitoring location.

In the factor assigned to traffic (primary contribution), EVF larger than 0.3 characterised EC, Cu, Fe, Cr, and Pb. The highest relative mass contributions in the chemical profile were given by OC (41%), EC (32%), Fe (23%), and Cu (1%). The lack of relevant crustal elements such as Ca and Al in the chemical profile suggested a negligible impact of road dust in this factor. As reported above, at our sampling site the road dust contribution was very likely mixed to resuspended dust and further separation of these contributions was not possible. This traffic (primary) contribution over the whole dataset accounted for 5% of the PM10 mass, with a slightly lower absolute contribution in summer (see Table 3.3). This contribution is comparable to the percentage (7%) reported by Amato et al. [33] for exhaust traffic emissions, but it is lower than our previous estimates [89] [110], i.e. 15% in 2006 in PM10 and 12% in PM1 recorded in winter 2012. However, the current estimate seems to still be reasonable when considering the efforts made in recent years to reduce vehicles exhaust particle emissions and the fraction of secondary nitrate due to high nitrogen oxides and ammonia emissions in the region [115], which has to be added to account for the overall traffic impact. Unfortunately, the non-linearity of the emission-to-ambient concentration level relationship and the high uncertainties in emission inventories still prevent a robust estimate of this secondary contribution to total traffic exhaust emissions. As shown in the next paragraph, traffic is the largest contributor to the aerosol absorption coefficient, thus strengthening the interpretation of this factor as a traffic emission source.

The industry factor showed high EVF for Zn (0.59) and the second highest EVF was related to Mn (0.13). Previous studies at the same sampling site identified these elements as tracers for industrial emissions (e.g. [110] and references therein). The chemical profile was enriched by heavy metals and, after traffic, it was the profile with the highest share of Cr, Mn, Fe, Cu, Zn, and Pb (explaining about 8% of the total PM10 mass in the profile). The industry contribution was not very high in the urban area of Milan, accounting for 3% on average.

The factor interpreted as aged sea salt was characterised by high EVF of Na (0.93) and this element was – as a matter of fact – present only in this factor chemical profile. To check the physical consistency of this assignment and considering that Milan is about 120 km away from the nearest sea coast, back-trajectories were calculated through the NOAA HYSPLIT trajectory model [116] [117] [118], classified on the basis of the aged sea salt concentration, and represented using R package Openair [51] [119] [120]. As an example, results from a very short event (13 July 16:00–18:00 LT) singled out by the model and representing the highest sea salt contribution during summer are reported in Fig. 3.9. Before and during the event, south-western air masses coming from the Ligurian Sea were observed, while soon after the event, there was a rapid change in wind direction. High wind speeds were recorded during the episode ($4.8 \pm 1.7\ m\ s^{-1}$ with a maximum peak of $9.5\ m\ s^{-1}$) compared to the $1.9 \pm 1.0\ m\ s^{-1}$ average wind speed characterising the summer campaign. When marine air masses are transported to polluted sites, sea salt particles show a Cl deficit due to reactions with sulphuric and nitric acid (see Sect. 1.2) and the factor chemical profile is expected to be enriched in sulphate and nitrate. In this work, nitrate was not present in the aged sea salt chemical profile; a very rough estimate [44] gave a maximum expected contribution of 2% (about $82\ ng\ m^{-3}$) of the total nitrate mass in the atmosphere that can be considered negligible in terms of mass contribution of the sources. Temporal patterns of Cl concentrations (not inserted in the multi-time resolution analysis as being a weak variable) during marine aerosol episodes were exploited to further confirm the factor-to-source association. Cl concentration and aged sea salt pattern showed an evident temporal coincidence in peak occurrence during the short summer event, thus supporting the source identification; moreover, during this episode only the Cl coarse fraction increased (Fig. 3.10) and reached about 90% of the total PM10 Cl concentration. The Cl/Na ratio was 0.38 ± 0.05, consistent with an aging of marine air masses during advection showing the typical Cl depletion.
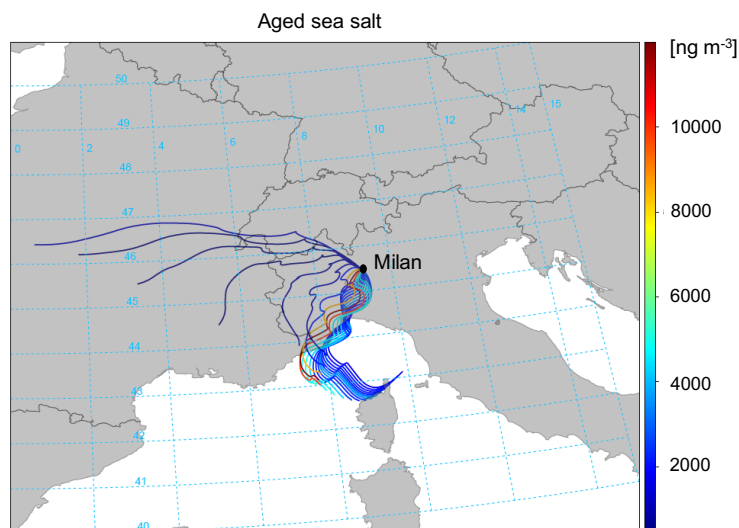
Figure 3.9: 72h Hysplit back-trajectories coloured by the aged sea salt concentration (in $ng\ m^{-3}$) on July 13th, 2016 [31].
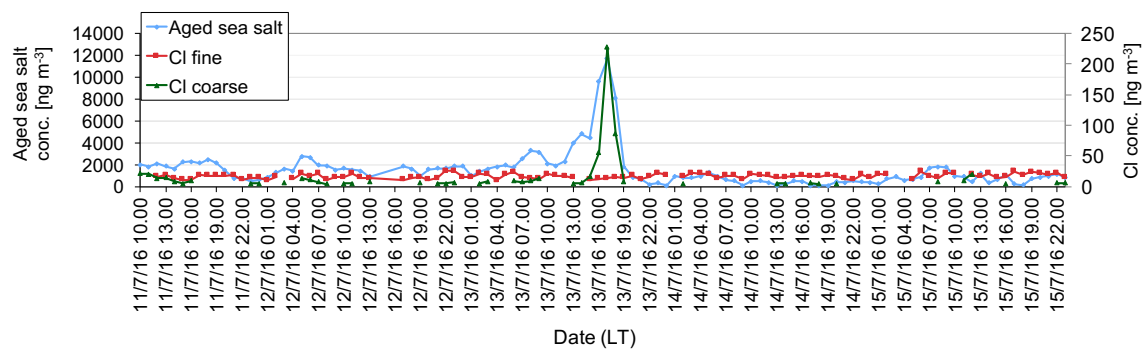
Figure 3.10: Temporal patterns of aged sea salt source retrieved from the multi-time resolution model and Cl concentrations measured in atmosphere in the fine and coarse fractions [31].

Bootstrap analysis was performed to evaluate the uncertainties associated with source profiles. One-hundred runs were carried out: factors were well mapped, with mapping always higher than 97% considering the Pearson coefficient (Table 3.4). In Fig. 3.11 a high interquartile range is associated to variables that are not significant for the considered factor (i.e. variables characterised by low EVF vaues), while main tracers for each source showed a small interquartile range, supporting the goodness of the final solution.

Table 3.4: Mapping of factors on the eight-factor constrained solution considering the Pearson coefficient (F1: Nitrate, F2: Sulphate, F3: Resuspended dust, F4: Biomass burning, F5: Construction works, F6: Traffic, F7: Industry, F8: Aged sea salt).

	F1	F2	F3	F4	F5	F6	F7	F8	Not explained
F1	99								
F2		99							
F3			99						
F4				99					
F5					97				2
F6						99			
F7							99		
F8								99	
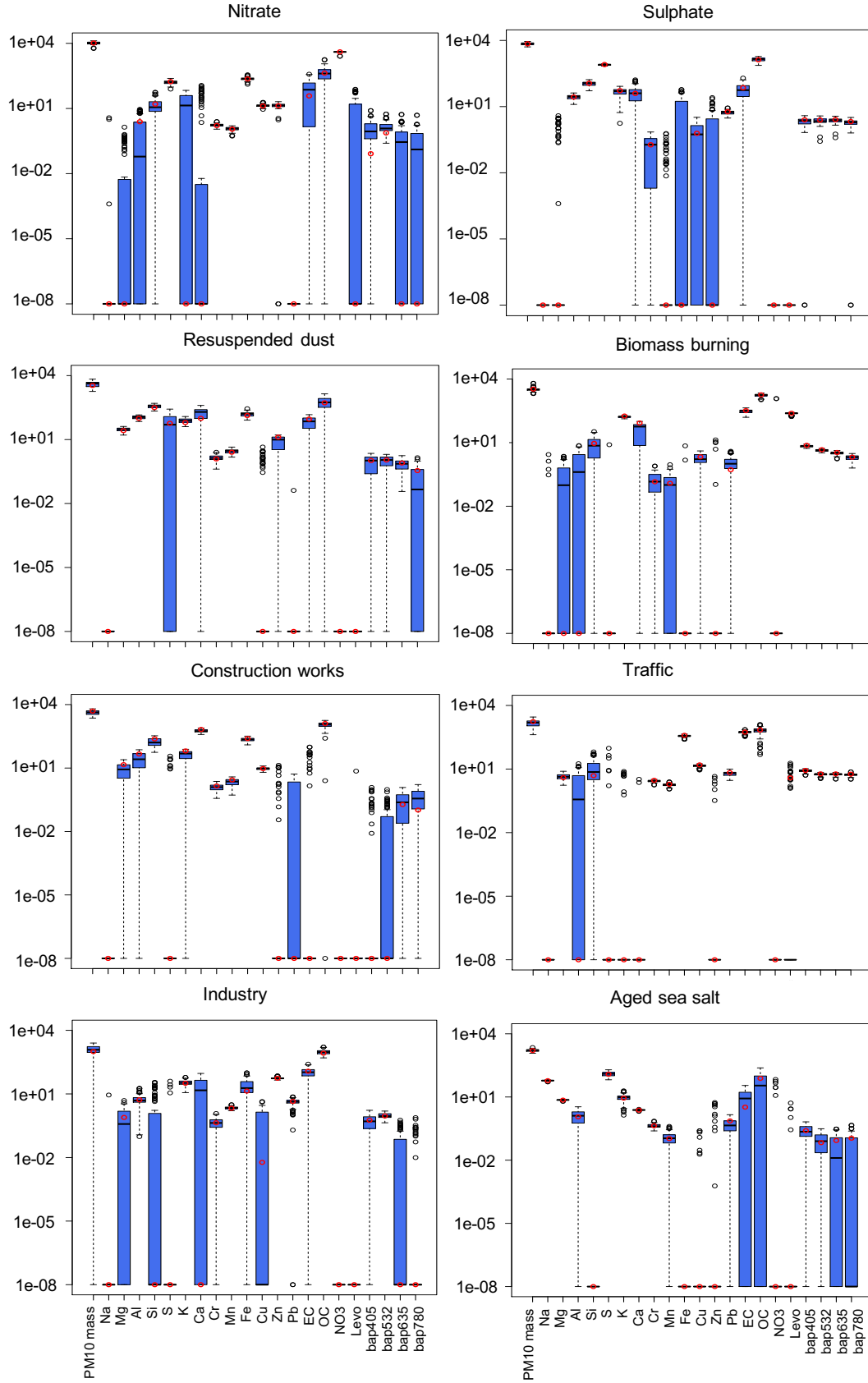
Figure 3.11: Box plot of the bootstrap analysis on the eight-factor constrained solution (values expressed in $ng\ m^{-3}$ for chemical variables and Mm^{-1} for optical absorption variables, on a logarithmic scale). The red dots represent the output values of the solution of the model; the black lines the medians from the bootstrap analysis; the blue bars the 25th and 75th percentiles; the dotted lines the interval equal to 1.5 times the interquartile range; and the black dots the outliers from this interval [31].

Improving source apportionment making use of aerosol light absorption coefficients

First of all, the use of the light absorption coefficients measured at different wavelengths as input variables in the multi-time resolution model strengthened the identification of sources, suggesting that they can be exploited when specific chemical tracers are not available (e.g. levoglucosan for biomass burning). To prove that, a separate source apportionment study was performed with EPA PMF 5.0 [27] using only hourly elemental concentrations from samples collected by the streaker sampler and hourly b_{ap} at different wavelengths measured by PP_UniMI on the same filters as input variables. Streaker samples typically lack a complete chemical characterisation; in particular, important chemical tracers such as levoglucosan and EC detected on filter samples collected in parallel are not available. In this analysis, b_{ap} assessed at different wavelengths was proved effective in identifying the biomass burning factor that explained a significant percentage (from 25% to 35% depending on λ) of the b_{ap} itself (Fig. 3.12); without the optical variables, the factor-to-source assignment would otherwise be based only on the presence of elemental potassium, although it is well known that K cannot be considered an unambiguous tracer as it is emitted by a variety of sources (see e.g. [121] and references therein). Furthermore, results showed that b_{ap} contribution was higher than 45% in the factor labelled as traffic, highlighting the importance of exhaust emissions in a factor that would be differently characterised by elements related to non-exhaust emissions (Cu, Fe, Cr).

In the following, results obtained from the introduction of the aerosol absorption coefficients at different wavelenghts in the multi-time resolution model will be thoroughly outlined. The two factors identified as biomass burning and traffic were the main contributors to aerosol absorption in the atmosphere and showed significant EVF values. At $\lambda = 780$ nm and $\lambda = 405$ nm, traffic contributions to b_{ap} were 55% and 42%, respectively; at the same wavelengths biomass burning accounted for 20% and 36%, respectively. The EVF of b_{ap} has the maximum value at $\lambda = 405$ nm for biomass burning (0.32) and at $\lambda = 780$ nm for traffic (0.49), showing the tendency to decrease and increase with the wavelength, respectively.

The third contributor to aerosol absorption in the atmosphere was the sulphate factor, with a contribution comparable to the biomass burning one at $\lambda = 780$ nm (about 20% of the total reconstructed b_{ap} at this wavelength). The sulphate factor contained a small fraction of EC, as previously discussed (see previous paragraph). This might be explained considering that non light-absorbing or weakly light-absorbing material can form a coating able to enhance particle absorption [122] [123] within a few days after emission. Laboratory experiments and simulations from in situ measurements highlighted absorption amplification for absorbing particles coated with secondary organic aerosol [124] [125]. Particle aging is a significant process in the Po Valley due to low atmospheric dispersion conditions and it might explain the relatively high contribution of the sulphate factor to the absorption coefficient in respect to the other sources (apart from traffic and biomass burning).

Among the remaining factors, resuspended dust was the main contributor at all wavelengths (between 3% and 7% of the total reconstructed b_{ap}, depending on the wavelength), likely due to the role of iron minerals. The other sources were less relevant in terms of EVF values and overall contributed less than 11%.

In contrast to the approach used in source apportionment models based on optical data like the widespread Aethalometer model [75] and MWAA (Multi-Wavelength Absorption Analyzer) model [76] [77], it is noteworthy that no a priori information about α values of the fossil fuel and biomass burning sources was introduced in the multi-time resolution model and an estimate for these values was directly retrieved from the model. In order to compare the multi-time resolution model and models based on optical absorption data, contributions due to traffic and industry (i.e. emissions most likely connected to fossil fuel usage) were added up and labelled as "fossil fuel emissions". In accordance with the two-source approach used in the Aethalometer model, the discussion about optical properties will be hereafter focused on the biomass burning and fossil fuel sources considering that sulphate and resuspended dust factors were less significant also in terms of EVF for optical absorption variables, ranging from 0.08 to 0.12 and from 0.03 to 0.06, respectively, depending on the wavelength.

In Fig. 3.13 the wavelength dependence of b_{ap} for the biomass burning and the fossil fuel profiles obtained with the multitime resolution model is shown; as α values can show significant differences when calculated using different pairs of λs [126], here we performed a fitting procedure considering $b_{ap} \propto \lambda^{-\alpha}$. Results were α_{BB} (α biomass burning) = 1.83 and α_{FF} (α fossil fuels) = 0.80; the range of variability of α values was estimated with the bootstrap analysis obtaining 0.78–0.88 for α_{FF} and 1.65–1.88 for α_{BB} (as the 25th and 75th percentiles, respectively).

Zotter et al. [78] reported a possible combination of $\alpha_{FF} = 0.8$ and $\alpha_{BB} = 1.8$ when EC concentration from fossil fuel combustion (estimated with radiocarbon measurements) is between 40% and 85% of the total
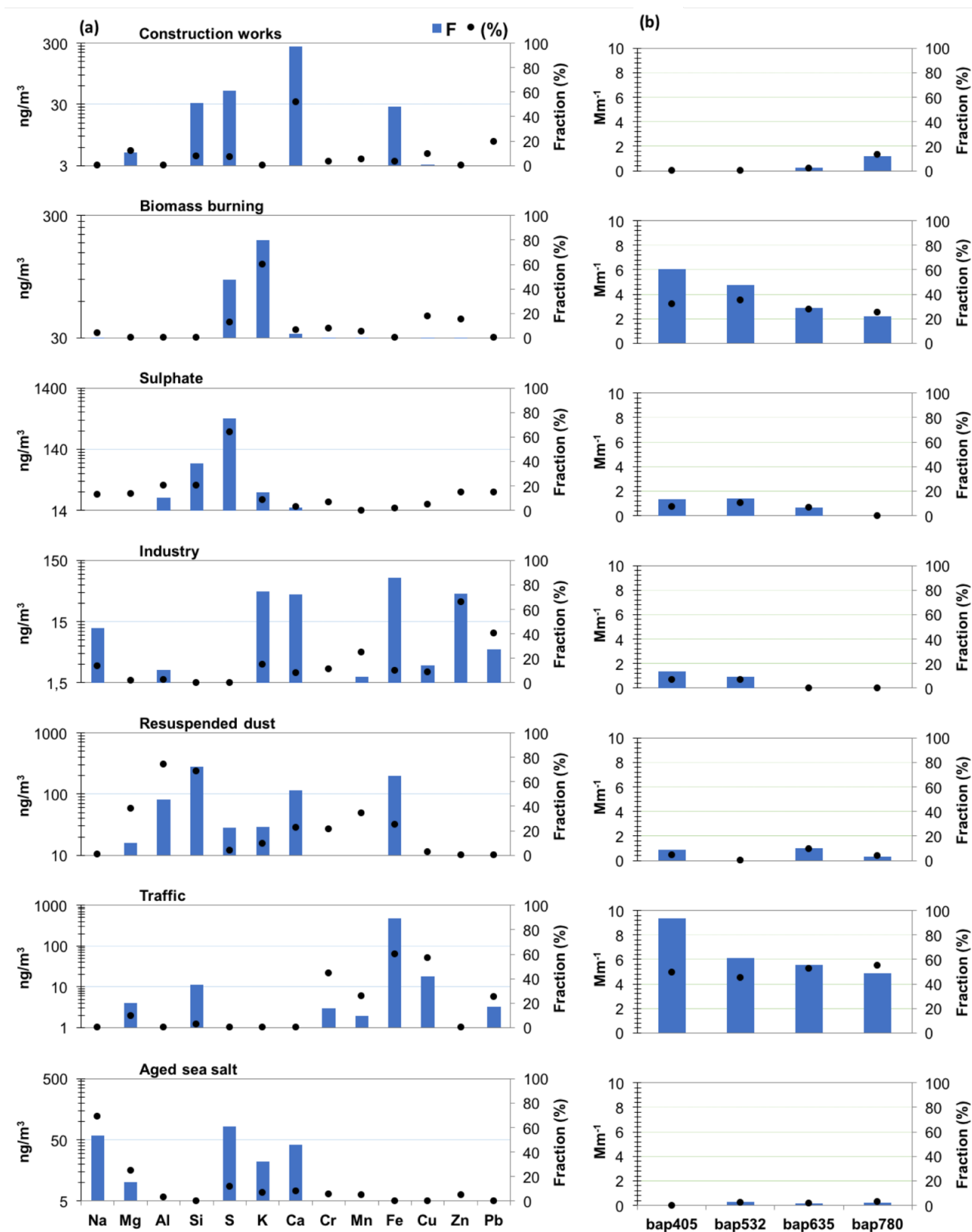
Figure 3.12: Source apportionment study performed with EPA PMF 5.0 on elemental concentrations and absorption coefficients at four wavelengths, both measured on high-time resolution samples collected by streaker sampler [31].
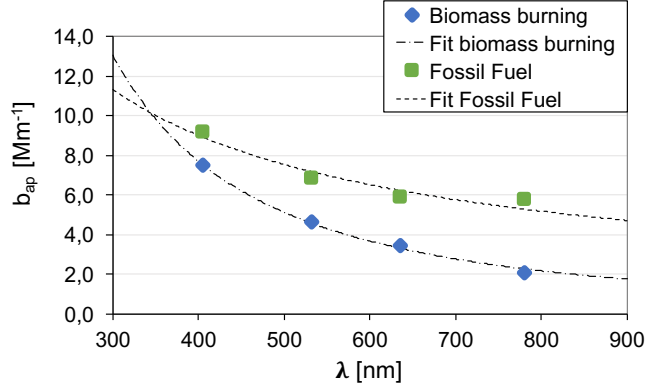
Figure 3.13: Aerosol absorptio coefficient (b_{ap}) dependence on the wavelength (λ) for biomass burning and fossil fuel emissions [31].

EC concentration; in this work, the fraction of EC ascribed by the multi-time model to fossil fuel sources was 56%. The combination 0.9 and 1.68 for α_{FF} and α_{BB}, respectively, was also suggested when in the study there is no or only limited additional information (e.g. from ^{14}C measurements). From the wide range of possible combinations reported in the literature it is clear that the assessment of α_{BC} (assumed to be equal to α_{FF} in source apportionment models based on optical data) is still an issue, and both experimental and simulation studies are in progress to reduce uncertainties and give a better evaluation of this key parameter.

The α_{FF} value resulted in the range 0.8–1.1 typically reported in source apportionment studies based on optical data ([77] [78] and references therein). Indeed, the sampling site was an urban background station in Milan and our samples were hardly impacted by fresh traffic emissions. Considering the aged nature of Milan aerosol, the average α_{FF} was comparable to estimates for BC-coated particles reported in the literature (approximately 0.6–1.3; see e.g. [127]) and obtained by both ambient measurement (e.g. [128] and references therein) and numerical simulations (e.g. [127] [129] and references therein).

The α_{BB} value retrieved by the model was very similar to values reported by Zotter et al. [78] and also comparable to 1.86 found for biomass burning by Sandradewi et al. [75] and 1.8 obtained by Massabò et al. [76], who also used independent ^{14}C measurements for checking.

Results here reported also allow us to study the relationship between the absorption coefficient and the mass of BC, i.e. the so-called mass absorption cross section (MAC), at different wavelengths. The $MAC(\lambda) = b_{ap}(\lambda)/BC$ relationship assumes that BC is the only light-absorbing species present; however, this assumption is not always valid since the transport of mineral dust from desert areas and brown carbon can significantly contribute to aerosol absorption. During our monitoring campaign, no contribution from Saharan dust was observed; in contrast, biomass burning was proved to be an important source, so that BrC was certainly a significant contributor [1], as also suggested by $\alpha_{BB} = 1.83$ in the biomass burning factor. The possible overestimation of BC when total b_{ap} is ascribed to BC only is usually minimised by choosing a wavelength longer than 600 nm, exploiting the spectral dependence of absorption from different aerosol compounds [130].

EC concentration retrieved from the chemical profiles (see Fig. 3.8) was used as a proxy for BC to estimate a source-dependent b_{ap}-to-BC ratio. Results are represented in Fig. 3.14.
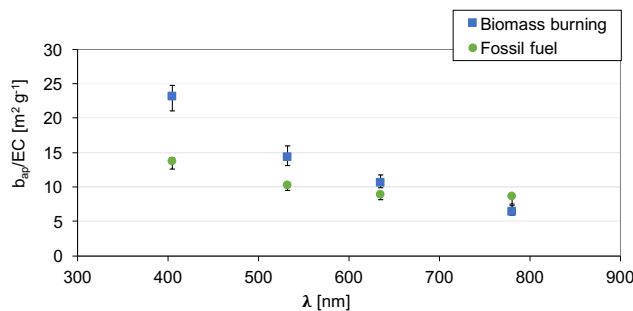
Figure 3.14: b_{ap}-to-EC ratio dependence on wavelenght (λ) for biomass burning and fossil fuel emissions. Error bars represent the 25th and 75th percentiles retrieved from the bootstrap analysis [31].

49

It is noteworthy that here this ratio is intentionally not indicated as MAC, since overestimation of the BC absorption especially at shorter wavelenghts might occur (see the previous discussion). BrC is expected to give a small contribution in the fossil fuel source. Therefore, the best approximations for $MAC(\lambda)$ values are likely the b_{ap}-to-BC ratios observed in the fossil fuel source at our monitoring site; values of these ratios are reported in Table 3.5.

Table 3.5: b_{ap}-to-EC ratios (in $m^2\ g^{-1}$) for biomass burning and fossil fuel emission source assessed by the multi-time resolution model. Results are retrieved considering b_{ap} and EC apportioned in each source; the 25th and 75th percentile is estimated by the bootstrap analysis [31].

λ	Biomass burning		Fossil fuel	
	b_{ap}/EC	$25^{th}-75^{th}perc.$	b_{ap}/EC	$25^{th}-75^{th}perc.$
405 nm	23.1	21.1 - 24.8	13.7	12.7 - 14.2
532 nm	14.3	13.2 - 16.0	10.2	9.6 - 10.4
635 nm	10.6	9.9 – 11.7	8.8	8.2 - 9.1
780 nm	6.4	6.0 – 7.3	8.6	7.6 - 8.9

For $\lambda = 550$ nm, Bond and Bergstrom [122] reported a MAC value of $7.5 \pm 1.2 m^2\ g^{-1}$ for uncoated fresh emitted particles and MAC values in polluted regions ranging from 9 to 12 $m^2\ g^{-1}$, attributable to absorption enhancement due to particle coating. The MAC estimate obtained in this work from the multi-time resolution model for $\lambda = 532$ nm is comparable to literature values reported above, thus confirming the importance of aging processes in the atmosphere for the optical properties of particles. Ratios represented in Fig. 3.14 are less comparable at $\lambda = 405$ nm (see also Table 3.5) due to the significant contribution of BrC to b_{ap} at this wavelength in the biomass burning factor.

From the outputs of the modelling approach here proposed, the apportionment of the biomass burning and fossil fuel contributions to b_{ap} at different wavelengths was also obtained (Table 3.6). As expected, the relative contribution to the total reconstructed b_{ap} ascribed to the biomass burning factor decreased with increasing λ, in contrast to the contribution from fossil fuel combustion which gave the highest contribution at $\lambda = 780$ nm; in addition, the fossil fuels contribution prevailed at all wavelengths at the investigated site.

Table 3.6: Average contribution to total reconstructed b_{ap} for the biomass burning and fossil fuel factors; in parentheses the 25th and 75th percentiles are reported [31].

λ	Biomass burning	Fossil fuel
405 nm	36% (31% − 36%)	45% (41% − 46%)
532 nm	29% (25% − 30%)	43% (39% − 44%)
635 nm	26% (23% − 27%)	45% (41% − 47%)
780 nm	20% (16% − 22%)	55% (48% − 55%)

3.3   Gaining knowledge on organics contribution through receptor modelling

This SA study was performed on data collected during February 2017 in Rome (Italy), in the frame of the CARE (Carbonaceous Aerosol in Rome and Environs) experiment [131]. A complete chemical characterisation was carried out by online and offline instrumentation, and the aerosol absorption coefficients $b_{ap}(\lambda)$ at 7 wavelengths were retrieved by an Aethalometer AE33 (Magee Scientific). All these variables (chemical + optical) were organised in a joint matrix (see Sect. 2.6.2) following the approach proposed in Forello et al. [31], and used as input to the multi-time resolution model (see Sect. 2.6.1). Thanks to the coupling of optical and chemical variables as input to the receptor model, information such as the impact of a mineral dust transport episode to the aerosol absorption in the atmosphere and estimates for the absorption Ångström exponent (α) of the sources were retrieved in addition to the source apportionment.

Moreover, an original way to analyse results from different SA studies was implemented to relate primary and secondary organic aerosol (OA) contributions to their emission sources, exploiting results from the source apportionment study carried out in this PhD work [74]. Total OA concentration measured by ACSM (Aerosol

Chemical Speciation Monitor) was apportioned to different sources by our receptor model, and afterwards compared with HOA (hydrocarbon-like organic aerosol), BBOA (biomass burning-like organic aerosol), and OOA (oxygenated organic aerosol) concentrations obtained from an independent source apportionment study previously performed on ACSM data only [131]. Insights on OA contributions were thus retrieved and the noteworthy result indicating that OOA apportionment made on ACSM data likely includes a secondary OA contribution due to biomass burning was obtained.

3.3.1   Site description

The CARE (Carbonaceous Aerosol in Rome and Environs) measurement campaign was carried out in Rome (Italy; latitude: 41.88°, longitude: 12.49°), in the middle of the Mediterranean sea, at an urban background site from 1st to 28th February 2017. Besides contribution from local sources, due to its position and meteorological conditions the site can be affected by air masses coming from the sea as Rome is about 30 km from the nearest coast, and by long-range transport from the Sahara desert [132].

3.3.2   Online and offline measurements

Mass

Hourly PM2.5 mass concentration was reconstructed from particle number size distribution (PNSD) data measured combining a scanning mobility particle sizer and an aerodynamic particle sizer [133]. Size distributions from these instruments were merged following the methodology reported in Khlystov et al. [134] and a size-dependent effective particle density was used to obtain the mass (details on the mass retrieval procedure and validation can be found in [131] and [135]).

Elemental composition

Hourly PM2.5 samples were collected by a streaker sampler [86]. Briefly, the streaker sampler collects with hourly resolution aerosol particles in the coarse (PM10 - 2.5) and fine (PM2.5) fractions on an impaction stage and a filter, respectively. For the aim of this campaign, only the fine fraction was analysed by Particle Induced X-ray Emission (PIXE) technique to obtain the elemental composition. More details about the technique and the set-up can be found e.g. in [95] and [96]. PIXE Minimum detection limits (MDLs) were in the range 1-10 $ng\ m^{-3}$ (depending on the element) and average experimental uncertainties for different species ranged from about 10% to about 40% (the latter refers to those elements measured with concentrations near MDL).

Non-refractory chemical components

Major non-refractory at 600°C components in PM1 were measured by an Aerodyne Aerosol Chemical Speciation Monitor (ACSM, see e.g. [136]) with a temporal resolution of 30 minutes. Shortly, in the ACSM particles are focused inside the instrument by a system of aerodynamic lenses, then thermally vaporised, and finally ionized by electron impact. Starting from the acquired mass spectrum, organic matter (OA), sulphate (SO_4^{2-}), ammonium (NH_4^+), nitrate (NO_3^-), and chloride (Cl^-) concentrations can be assessed. In a previous work [131], from ACSM data three factors for OA were singled out: HOA (hydrocarbon-like organic aerosol), BBOA (biomass burning-like organic aerosol), and OOA (oxygenated organic aerosol); as for the latter, it is generally associated to secondary compounds in the literature. MDLs were estimated following Ng et al. [136] as 0.105 $\mu g\ m^{-3}$, 0.201 $\mu g\ m^{-3}$, 0.017 $\mu g\ m^{-3}$, 0.008 $\mu g\ m^{-3}$, and 0.008 $\mu g\ m^{-3}$ for OA, NH_4^+, SO_4^{2-}, NO_3^-, and Cl^-, respectively. Sensitivity tests using different ranges of trial uncertainties as input to the model were performed; in the end, average uncertainties for ACSM measurements were set to 19% for OA, 36% for NH_4^+, 28% for SO_4^{2-}, and 15% for NO_3^-, in accordance with the reproducibility of relative uncertainties observed in ACSM intercomparison exercises [137] [138].

Carbonaceous components

EC and OC concentrations with two hour resolution were obtained by a Sunset Field Thermal-Optical Analyser (Sunset Laboratory Inc.). Briefly, this instrument collects particles on a quartz fibre filter; at the end of each sampling period (105 minutes of sampling and 15 minutes of analysis) the collected sample is analysed with the NIOSH-like temperature protocol [139]. The inlet was equipped with a cyclone with a cut point of 2.5 μm and a denuder for organics. MDL was 0.240 $\mu g\ m^{-3}$ for OC and EC concentrations. Average uncertainties used as

input to the model were 15% and 10% for EC and OC concentrations, respectively.

Light absorption coefficients

Optical properties were retrieved by online instrumentation with a time resolution of 1 minute. The aerosol absorption coefficient $b_{ap}(\lambda)$ at 7 wavelengths (370, 470, 520, 590, 660, 880, and 950 nm) was retrieved in PM10 by a dual-spot Aethalometer (AE33, Magee Scientific) [140] using the instrument specific mass absorption cross-sections (MACs) [141] and the measured equivalent black carbon (eBC) concentration. $b_{ap}(\lambda)$ values are calculated by the AE33 internal software considering attenuation measurements corrected for loading (k parameter) and multiple scattering (C factor) effects. It is noteworthy that recent literature studies (e.g. [142]) evidenced that the fixed C factor equal to 1.57 typically used in AE33 can lead to a significant $b_{ap}(\lambda)$ overestimation. Therefore, in this work a C factor of 2.66 was used at all wavelengths, as previously estimated by Valentini et al. [132] for the CARE campaign. MDLs were estimated in the range $0.36 - 0.92\ Mm^{-1}$ depending on the wavelength and average experimental uncertainty on $b_{ap}(\lambda)$ was 15% [143].

3.3.3   Implementation to the multi-time resolution model

The multi-time resolution model implemented by Crespi et al. [56] was used as a basis as it allows the estimation of uncertainties by bootstrap analysis (see Sect. 2.4.2 for details about this method). In this work, a physical constraint based on a mass balance equation was added and solved together with main equation (2.18) and regularisation equation (2.19): in each factor, the sum of concentrations of the species must be equal or smaller than the total mass concentration. The equation for each factor was introduced in the general form [25]:

$$0 = \sum_{j=1}^{N} c_j f_{kj} + r_v \tag{3.1}$$

where the index j represent the chemical species, N is the total number of chemical species, c_j are numerical coefficients set at $c_{mass} = +1$ for mass while $c_j = -1$ for species other than mass, f_{kj} is an element of the matrix $\mathbf{F}$ (i.e. chemical profiles), and r_v represents the residual of this auxiliary equation. Since contributions from not measured element oxides and water can be present, the equation was implemented in order to allow negative values of the residual r_v using the error model code $em = -17$ (see Sect. 2.5.2).

3.3.4   Source apportionment coupling online and offline measurements

Input data

One hour was chosen as the basic time unit in the model to study high time resolution changes in source contributions. Since no replicated species were present in the dataset after input data selection (selection criteria are explained hereafter), adjustment factors η_{jm} in the main equation (2.18) of the multi-time resolution model were set to one (see Sect. 2.6.1 for more information about adjustment factors).

Mass concentrations were included in the model with uncertainties set at 4 times their values [50] in order to reduce their relevance in the modelling process. All other variables were classified according to their signal-to-noise ratio (S/N) (see Sect. 2.5.2). All strong variables (S/N $\geq$ 1.2) and only some weak variables (i.e. Ti, V, Rb, and Pb) were used as input to the model. In the literature, Ti, V, Rb, and Pb are often indicated as tracers of specific sources (Saharan dust advection for Ti, residual oil combustion for V, biomass burning for Rb, and industry for Pb); for this reason, they were taken into account although strongly underweighted multiplying their uncertainties by a factor 3.

Ranges of uncertainties and MDLs for measured variables have been already reported in Sect. 3.3.2; in the input dataset, uncertainties and data below minimum detection limits were pre-treated according to Polissar et al. [40]. Missing values were substituted by linear interpolation, with uncertainties set as three times the interpolated concentration value. Among strong variables, Si showed a slightly higher percentage (26%) of missing data due to blank filter contamination. Linear interpolation was not possible in this case, since missing data were consecutive over time; therefore, in order to avoid artificial high values in modelled time contributions as already reported in literature works [31] [39], missing values were substituted by the median value calculated over the whole campaign, with uncertainties set at four times the median value. To avoid double counting for sulphur/sulphate, organic aerosol/organic carbon, and chlorine/chloride the selection of input data was performed as explained in the following.

Atmospheric concentrations of SO_4^{2-} (measured online by ACSM) and S (measured offline by PIXE analysis on streaker samples) display very similar temporal patterns. The linear regression shows a slope of 2.54 ± 0.02 ($R^2 = 0.89$) (Fig. 3.15). The difference of about 15% from the sulphate-to-sulphur stoichiometric value (equal to 3) is within average uncertainties (i.e. 12% for S and 28% for SO_4^{2-}), and can be mainly ascribed to different sampling and analytical techniques. In fact, the effect of different PM fractions collected by ACSM and streaker sampler (PM1 and PM2.5, respectively) can be considered less relevant, since clear discrepancies from the average ratio of the campaign are present only when S concentration is higher than about 1.1 $\mu g\ m^{-3}$ as can be seen from Fig. 3.15 (17 over 619 samples, about 3% of data). The largest discrepancy is registered on the 3rd of February, when there is a clear change in the SO_4^{2-}-to-S ratio (Fig. 3.16) and a polluted marine air mass transport episode was detected [132]. In Fig. 3.16, it can be noted that the S peak corresponds to a peak in Mg concentration, suggesting the presence of compounds (e.g. $MgSO_4$) [144] that are not measured by ACSM due to particle size and refractory properties. Tests performed using S or SO_4^{2-} as input to the model did not give any significant change in the solution in terms of number of factors or source contributions. In order to avoid double counting, SO_4^{2-} measured by ACSM was chosen as input variable, because sulphate is very often in the form of ammonium sulphate and NH_4^+ was measured by ACSM as well.
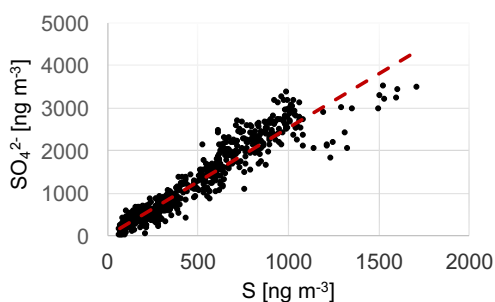
Figure 3.15: Linear regression of atmospheric concentrations of SO_4^{2-} (measured online by ACSM) and S (measured offline by PIXE analysis on streaker samples). Slope: 2.54 ± 0.02 ($R^2 = 0.89$). Figure from [74].
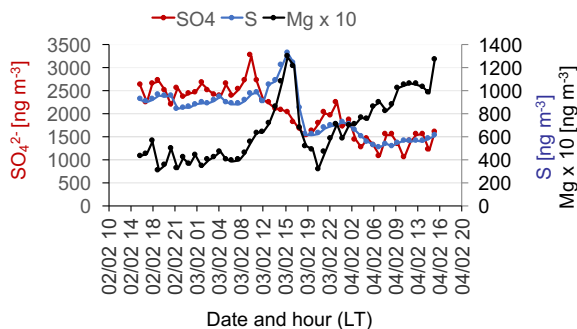
Figure 3.16: Temporal pattern of atmospheric concentrations of SO_4^{2-}, S and Mg on the 3rd of February 2017. Figure from [74].

Atmospheric concentrations of organic aerosol OA (measured online by ACSM) and organic carbon OC (measured online by Sunset Field Thermal-Optical Analyser) show very similar temporal patterns (Fig. 3.17). The two-hour median value of OA-to-OC ratio is 1.3 (1.1 and 1.5 are the limits of the interquartile range) that is lower than 1.6 used in previous literature studies performed in Rome [145] [146] [147]; also in this case it is likely due to different sampling and analytical techniques. Finally, OA was selected as input variable since it carries a larger fraction of the total mass.

As for Cl (measured by PIXE analysis) and Cl$^-$(given by ACSM) concentrations, the former was used as input variable to the model as it showed a much more reliable temporal pattern.

$b_{ap}(\lambda)$ values measured at 7 wavelengths in PM10 were inserted in the model together with chemical variables assessed in PM2.5 (and PM1 for ACSM data). The main issue in considering different size fractions for chemical and optical variables was the presence of a desert dust transport episode [132] lasting less than two days (24th -
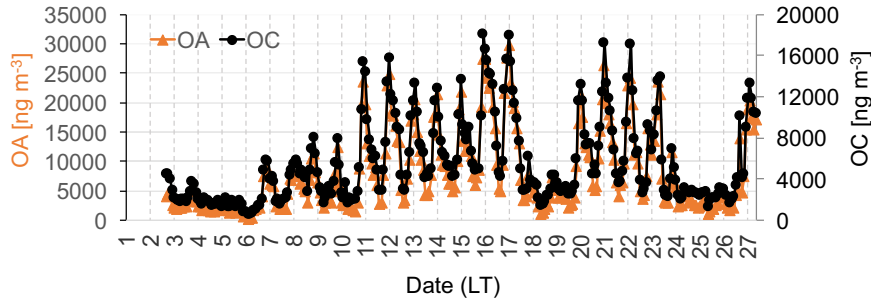
Figure 3.17: Temporal pattern of atmospheric concentrations of organic aerosol OA (measured online by ACSM) and organic carbon OC (measured online by Sunset Field Thermal-Optical Analyser) for February 2017. Vertical lines show midnight for each day. Figure from [74].

25th February) during the CARE campaign. In this work, samples impacted by desert dust were included in the input dataset in order to estimate optical absorption properties of the mineral dust source (see the following). It is interesting to note that in the simulation chamber study by Caponi et al. [12], desert dust samples in the PM10 and PM2.5 fractions showed very small differences in elemental composition and the absorption Ångström exponent (α) of dust in that work did not seem to be related to differences in particle size. As for the MAC of mineral dust, in Caponi et al. [12] it was defined considering the total mass concentration of dust and, opposite to the α value, the MAC seemed to be dependent on particle size. For this reason, the MAC values at different wavelenghts retrieved as a result of the model in this work for mineral dust have to be considered as an upper limit.

Finally, 30 variables with 1-hour resolution (PM2.5 mass, Na, Mg, Al, Si, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Br, Rb, Pb, OA, NH_4^+, SO_4^{2-}, NO_3^-, b_{ap}(370 nm), b_{ap}(470 nm), b_{ap}(520 nm), b_{ap}(590 nm), b_{ap}(660 nm), b_{ap}(880 nm), and b_{ap}(950 nm)) and 2-hour EC concentrations were inserted as input data in the multi-time resolution model. The input matrix $\mathbf{X}$ consisted of 916 samples distributed over 619 time units. The analysis was performed in the robust mode (see Sect. 2.5.2). The error model $em = $ -14 (see Sect. 2.5.2 for its definition) was used for the main equation with $C_1 = $ input error, $C_2 = 0.0$, and $C_3 = 0.1$ for both chemical and optical absorption data.

Base-case solution

Solutions from 5 to 10 factors were explored. In this analysis 30 convergent runs were obtained and a nine-factor base case solution corresponding to the minimum Q value was selected. A lower or higher number of factors gave mixing or artificial separation of sources, respectively, and a not satisfactory reconstruction of some variables during aerosol transport episodes (e.g. marine aerosol advection). In the selected base-case solution, all variables were well reconstructed by the model ($R^2 > 0.70$) with the exception of V ($R^2 = 0.52$) and Ni ($R^2 = 0.51$), that were characterised by concentration values near MDLs. Uncertainty-scaled residuals (see Sect. 2.5.3) were randomly distributed in the $\pm$ 3 range and their distribution was mostly symmetrical.

The explained variation for matrix $\mathbf{F}$ (EVF) (see Sect. 2.5.4) was firstly exploited for factor-to-source assignment; indeed, high EVF values are typically indicators for chemical species which are source tracers. The unexplained variation for matrix F was lower than 0.15 for all variables. In Fig. 3.18, EVF and chemical profiles for the base-case solution are reported. According to EVF, chemical profile, and temporal pattern, the nine factors were tentatively assigned to biomass burning, nitrate and aged aerosol, traffic exhaust, sulphate, mineral dust, marine aerosol, traffic non-exhaust, local source, and polluted marine aerosol (see also afterwards for details on the factor-to-source assignment motivation).

In Table 3.7, average source contributions to atmospheric PM2.5 mass are reported both in absolute and percentage values. Even if the base-case solution was largely satisfactory, constrained solutions were explored.

54
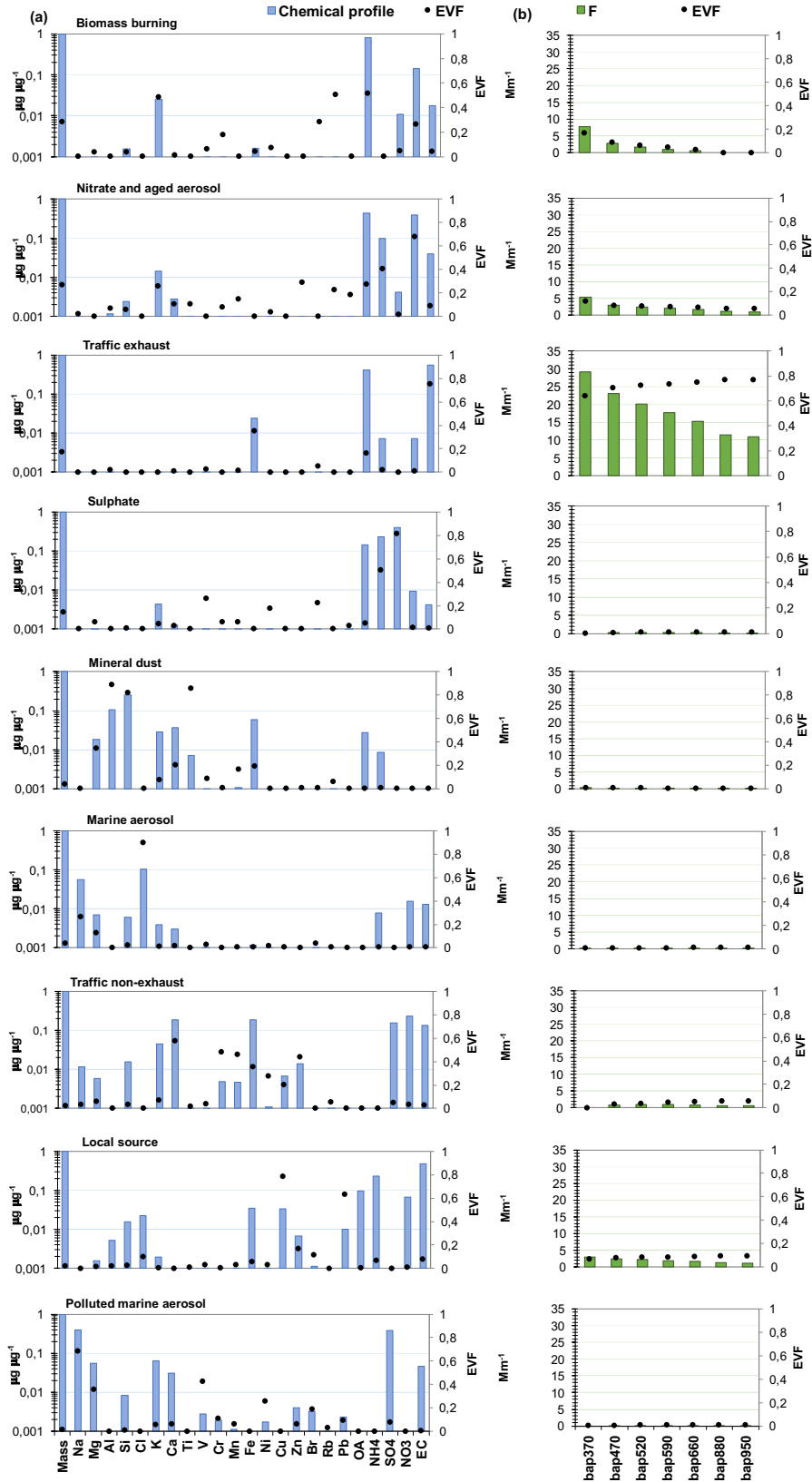
Figure 3.18: (a) Chemical profiles and (b) b_{ap} apportionment of the nine-factor base-case solution. The blue bars represent the chemical profile (output of the matrix **F** for chemical variables normalised on mass), the green bars the output of the matrix **F** for optical absorption variables, and the black dots the EVF. Figure from [74].

Table 3.7: Absolute and relative average source apportionment in the nine-factor base-case solution. Table adapted from [74].

Factors - sources	$\mu g\ m^{-3}$	%
F1 - Biomass burning	5.1	29%
F2 - Nitrate and aged aerosol	4.7	27%
F3 - Traffic exhaust	2.9	17%
F4 - Sulphate	2.5	14%
F5 - Mineral dust	0.67	3.9%
F6 - Marine aerosol	0.66	3.8%
F7 - Traffic non-exhaust	0.37	2.1%
F8 - Local source	0.35	2.0%
F9 - Polluted marine aerosol	0.24	1.4%

Final constrained solution and bootstrap analysis

Constraints were applied to the factor interpreted as marine aerosol. Indeed, the marine aerosol factor in the base-case solution was characterised by values of the typical diagnostic ratios Mg/Na and Cl/Na very similar to literature ones for bulk sea salt emissions (see Table 1.1), suggesting advection of fresh marine aerosol. However, contaminations appeared in the chemical profile due to EC, together with NH_4^+ and NO_3^-, which are often found in chemical profiles of aged marine emissions. From source temporal patterns, it was noted that the polluted marine aerosol episode was interrupted for a few hours by the advection of fresh marine aerosol; the former was characterised also by ship emissions so that some mixing between the two chemical profiles can be present. Therefore, in the constrained solution EC, NO_3^-, and NH_4^+ were pulled down maximally in the chemical profile of marine aerosol (fresh); as a consequence, also b_{ap} at all wavelengths decreased in agreement with the lack of light absorbing components in the profile. It is noteworthy that NO_3^- contribution appeared in the polluted marine aerosol chemical profile, as expected when compounds present in marine air masses react with polluted air masses during the transport, leading also to chloride deficit (see Sect. 1.2). Constraints led to an effective increase in Q of about 25 units with a 0.6% increase, which can be considered acceptable [25]. The constrained solution improved the chemical profiles of factors impacted by sea salt, with negligible differences in all other relevant features (i.e. EVF, residuals, source apportionment) respect to the base-case one. Thus, the constrained solution was considered the most reliable one from a physical point of view; results are presented in Fig. 3.19 and Fig. 3.20 and discussed in the following. The average apportionment during the CARE experiment is reported in Table 3.8.

Table 3.8: Absolute and relative average source apportionment in the nine-factor constrained solution; in parentheses, the 10th and 90th percentiles from the bootstrap analysis are reported. Table adapted from [74].

Factors - sources	$\mu g\ m^{-3}$	%
F1 - Biomass burning	5.5 (4.5 - 5.8)	32 (26 - 34)%
F2 - Nitrate and aged aerosol	4.4 (3.7 − 5.2)	25 (22 − 30)%
F3 - Traffic exhaust	2.8 (2.6 − 3.2)	16 (15 − 18)%
F4 - Sulphate	2.5 (2.1 − 2.8)	14 (12 − 16)%
F5 - Mineral dust	0.66 (0.57 − 0.71)	3.8 (3.3 − 4.1)%
F6 - Marine aerosol	0.63 (0.50 − 0.74)	3.6 (2.9 − 4.2)%
F7 - Traffic non-exhaust	0.38 (0.26 − 0.51)	2.2 (1.5 − 2.9)%
F8 - Local source	0.33 (0.25 − 0.63)	1.9 (1.4 − 3.7)%
F9 - Polluted marine aerosol	0.28 (0.20 − 0.81)	1.6 (1.1 − 4.6)%
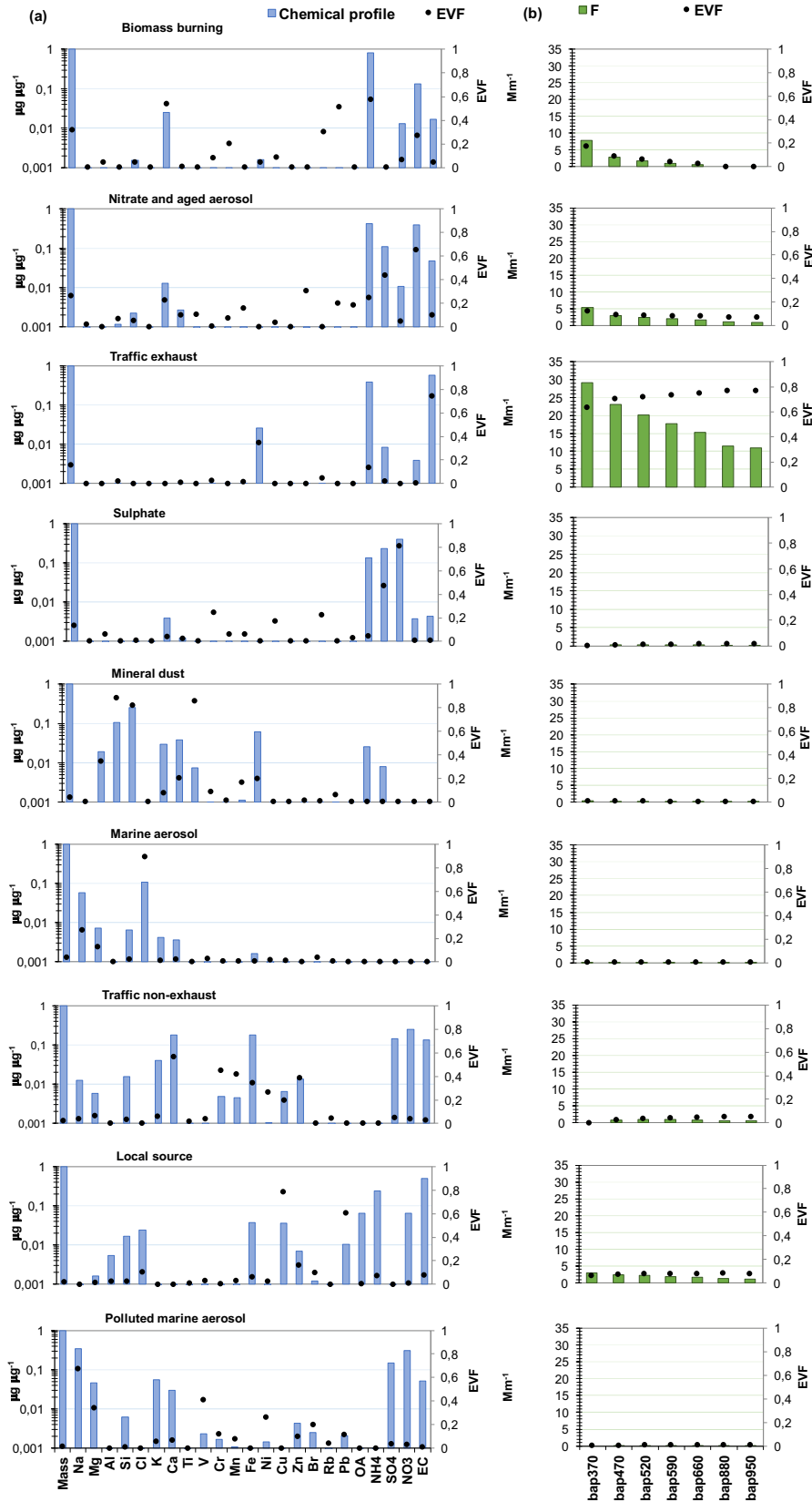
Figure 3.19: (a) Chemical profiles and (b) b_{ap} apportionment of the nine-factor constrained solution. The blue bars represent the chemical profile (output of the matrix **F** for chemical variables normalised on mass), the green bars the output of the matrix **F** for optical variables, and the black dots the EVF. Figure from [74].
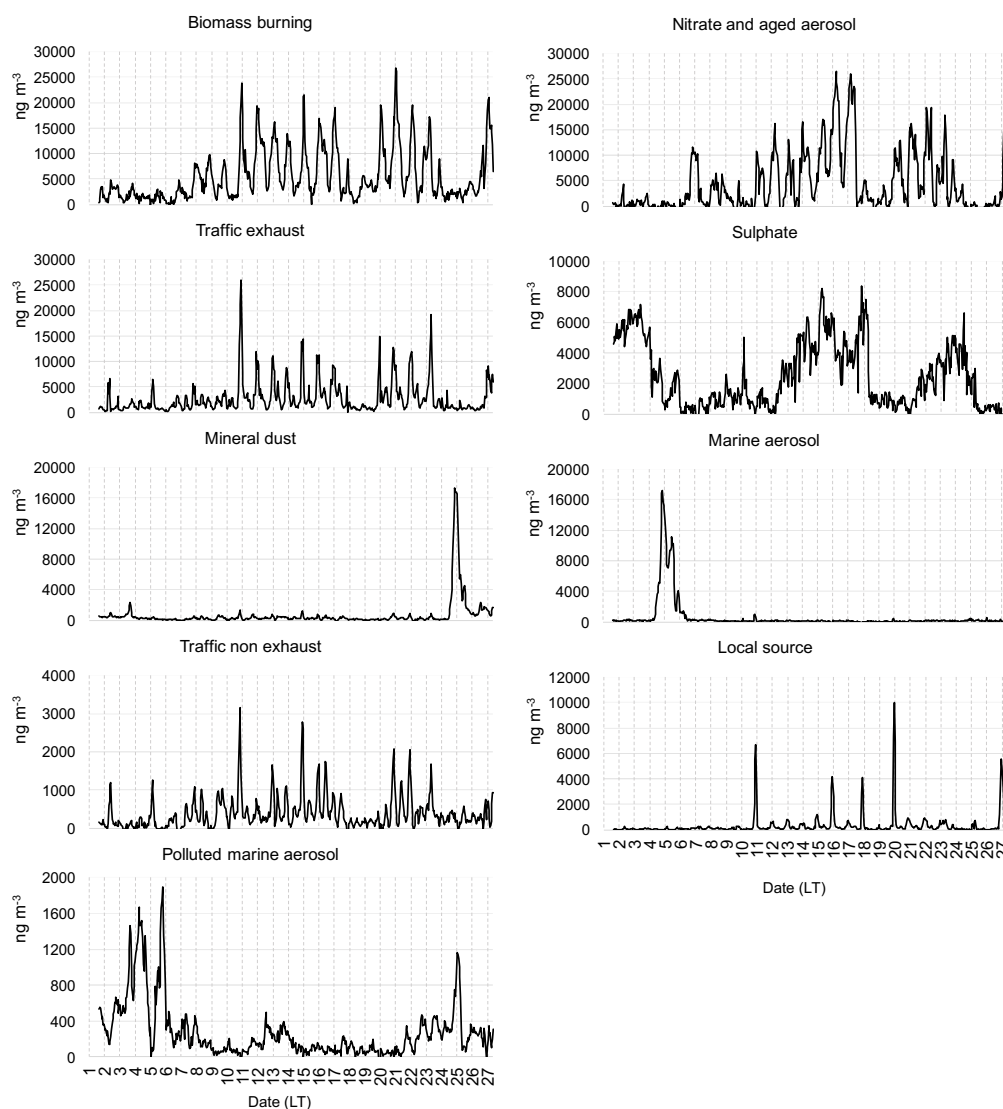
Figure 3.20: Hourly temporal patterns of the nine-factor constrained solution for February 2017. Vertical lines show midnight in each day. Figure from [74].

Factor 1 was identified as biomass burning because it was characterised by high EVF for OA (0.57), Rb (0.51), and K (0.54) [33] [148]. In the chemical profile, OA concentration contributed for 81% of the total mass apportioned to the source; the second highest contribution was 13% given by NO_3^-, followed by K (2.5%), EC (1.7%), and SO_4^{2-} (1.3%). Rb was less relevant in terms of mass contribution (about 0.01%). The biomass burning source gave a dominant contribution during the night, with highest values in the time interval 23:00 - 02:00 LT (temporal pattern in Fig. 3.20). Perrino et al. [149] already highlighted a similar temporal behaviour in levoglucosan concentrations (tracer of biomass burning emissions) in wintertime in the city centre of Rome; it was likely related to biomass burning products originated in the peri-urban area and then transported towards the city centre. The biomass burning primary contribution to PM2.5 estimated in that work was in the range 7.2% – 23.3% during 2013 – 2016 winter months. In this work, the biomass burning source explained 32% of the PM2.5 mass, a bit higher than the previous estimate as also aged organic aerosol contribution was accounted for as explained in Sect. 3.3.5. Influence of aerosol ageing in a chemical profile of biomass burning from PMF was already reported in the literature [150].

Factor 2 was related to a source called nitrate and aged aerosol. In fact, NO_3^- and NH_4^+ showed EVF of 0.65 and 0.44, respectively, and non negligible EVF values were also found for K, Zn, Rb, and OA. This observation suggested that, as already found in factor 1, the chemical profile of factor 2 showed some mixed contributions from biomass burning and nitrate. However, nitrate formation at urban sites is expected mainly from NOx traffic emissions thus justifying the share of EC in the chemical profile (4.7% of the apportioned mass, higher than in the biomass burning one) and the optical absorption contribution (see afterwards). The average mass

contribution of this factor was 25%.

Factor 3 was characterised by very high EVF (0.74) for EC and the only other significant chemical component in terms of EVF was Fe (0.35). The mass contribution of this source was ascribed to two major contributors, i.e. EC and OA accounting together for about 96% of the apportioned mass. The factor was thus identified as traffic (exhaust emissions) and impacted, on average, for 16% to the PM2.5 mass. Peaks in concentration values appeared in the evening approximately at 22:00 LT (Fig. 3.20 for the temporal pattern). Similar traffic emission concentration patterns were previously observed in Rome, independently of the season, and they were explained by boundary layer dynamics [151]. The observed modulation was also confirmed by the temporal pattern of natural radioactivity due to Radon progeny detected in the atmosphere during the CARE campaign (Fig. 3.21; more details in [131]); as well known, measurements of natural radionuclides can be used to trace the temporal evolution of atmospheric dispersion in the boundary layer and to estimate the mixing layer height (see e.g. [81] [152] and references therein). The traffic (exhaust) source is the main contributor to aerosol light absorption in the atmosphere, confirming the factor-to-source assignment (see the following).
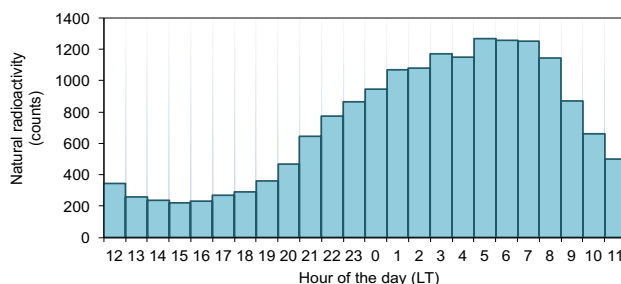
Figure 3.21: Mean diurnal cycle of natural radioactivity during the CARE campaign. Figure from [74].

Factor 4 was assigned to sulphate, since SO_4^{2-} and NH_4^+ showed high EVF. V and Ni were characterised by EVF slightly higher than the other elements, suggesting some mixing with ship emissions. In terms of mass, the most significant contribution in the chemical profile was given by OA (14%) after SO_4^{2-} and NH_4^+ (63% when added together). During the campaign, the average mass apportionment of this source was 14%.

Factor 5 was associated to the mineral dust source because of high EVF for Al (0.88), Ti (0.86), Si (0.82), and Mg (0.34). These variables are all crustal elements and tracers for mineral dust; it is noteworthy that the diagnostic ratios between these elements apportioned in the chemical profile are consistent with literature values [33]. During the CARE campaign, a desert dust transport episode lasting less than two days (24th - 25th February) was clearly identified exploiting optical properties [132]. Even if the impact of desert dust was dominant in this factor - with concentration values as high as 25 times the average over the whole campaign (see Fig. 3.20) - the source retrieved by the model probably included minor contributions also from local soil resuspension. The mass contribution of this source over the whole campaign was 3.8%, but during the mineral dust advection it accounted for a relevant fraction (49% on average) of the PM2.5 mass concentration.

Factor 6 was identified as a marine aerosol source being characterised by EVF = 0.89 for Cl, with the second highest EVF being 0.27 for Na. Typical diagnostic ratios for this source, i.e. Mg/Na and Cl/Na, were respectively 0.13 and 1.9, very similar to what expected for bulk sea salt aerosol (0.12 and 1.8, respectively; see Table 1.1). It is noteworthy that local atmospheric circulation in the area under investigation allows the inland penetration of weak sea breezes, even if Rome is about 30 km from the nearest sea cost. This episodic source contributed on average for about 3.6% of the total PM2.5 mass, up to 47% on average during the advection episode (Fig. 3.20).

Factor 7 was assigned to traffic non-exhaust emissions (including road dust resuspension), since high EVF were associated to main tracers for this source: Ca (0.57), Zn (0.38), Fe (0.34), Mn (0.42), and Cr (0.45) [107] [153]. EVF for Cu was a bit lower (0.20), because this element was found with higher concentrations (11.8 $ng\ m^{-3}$ compared to 2.5 $ng\ m^{-3}$) in the chemical profile of a factor that was associated to local emissions (see afterwards). Connection with traffic emissions was also confirmed by the presence of EC in the chemical profile (14%), likely due to road dust resuspension. The average contribution of traffic non-exhaust emissions to PM2.5 over the CARE campaign was quite low (2.2%), as already found for the same fraction by e.g. [33].

Factor 8 showed a strong episodic character (see Fig. 3.20) and presented a high EVF for Cu (0.78) and Pb (0.61). The high EC contribution in the chemical profile was likely associated to combustion emissions and the

optical absorption profile of this factor (see afterwards) suggested an influence of fossil fuel combustion ($\alpha \simeq 1$). This is the first time that a similar factor has been detected in the urban area of Rome and – as far as we know - it was not reported in previous literature works; therefore, this factor was tentatively assigned to local emissions but further investigation is needed in the future to identify the specific source. The local feature of the source is evidenced in Fig. 3.22 - realised using the Openair R package [51] [119] - which shows variation in source contributions by wind speed and wind direction. The episodic and late evening contribution of this source (Fig. 3.20) is also likely influenced by boundary layer dynamics (Fig. 3.21). The average mass contribution of this source was very low (1.9%).
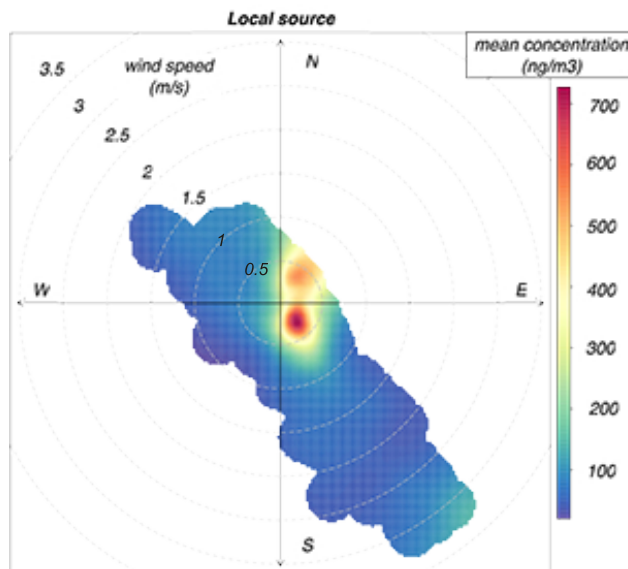
Figure 3.22: Polar Plot for the local source. Figure from [74].

Factor 9 was associated with polluted marine aerosol. Indeed, main tracers of aged sea salt aerosol are Na and Mg which showed EVF values of 0.67 and 0.34, respectively; moreover, EVF for V (0.41) and Ni (0.26) were also of interest as they are elemental markers for heavy oil combustion here likely related to ship emissions, as already highlighted for the CARE campaign [132]. Mg-to-Na ratio in the chemical profile was 0.14 (i.e. in fair agreement with 0.12 reported in the literature) and the chemical profile did not contain Cl; opposite, the chemical profile was clearly enriched in SO_4^{2-} and NO_3^{-}, highlighting the ageing of sea salt aerosol (see Sect. 1.2). Moreover, the presence of EC in the profile suggested the influence of ship emissions and the contamination due to air mass transport from the coast. The average mass contribution of this source was 1.6% at the receptor site.

A bootstrap analysis with 100 convergent runs was performed to evaluate the uncertainties associated with source profiles [56]; results are shown in Fig. 3.23. Main tracers of each source were characterised by small interquartile ranges (blue bars, with values expressed in $ng\ m^{-3}$ or Mm^{-1} on a logarithmic scale). Mapping of factors was always 99%, supporting the goodness of the solution presented in this work.
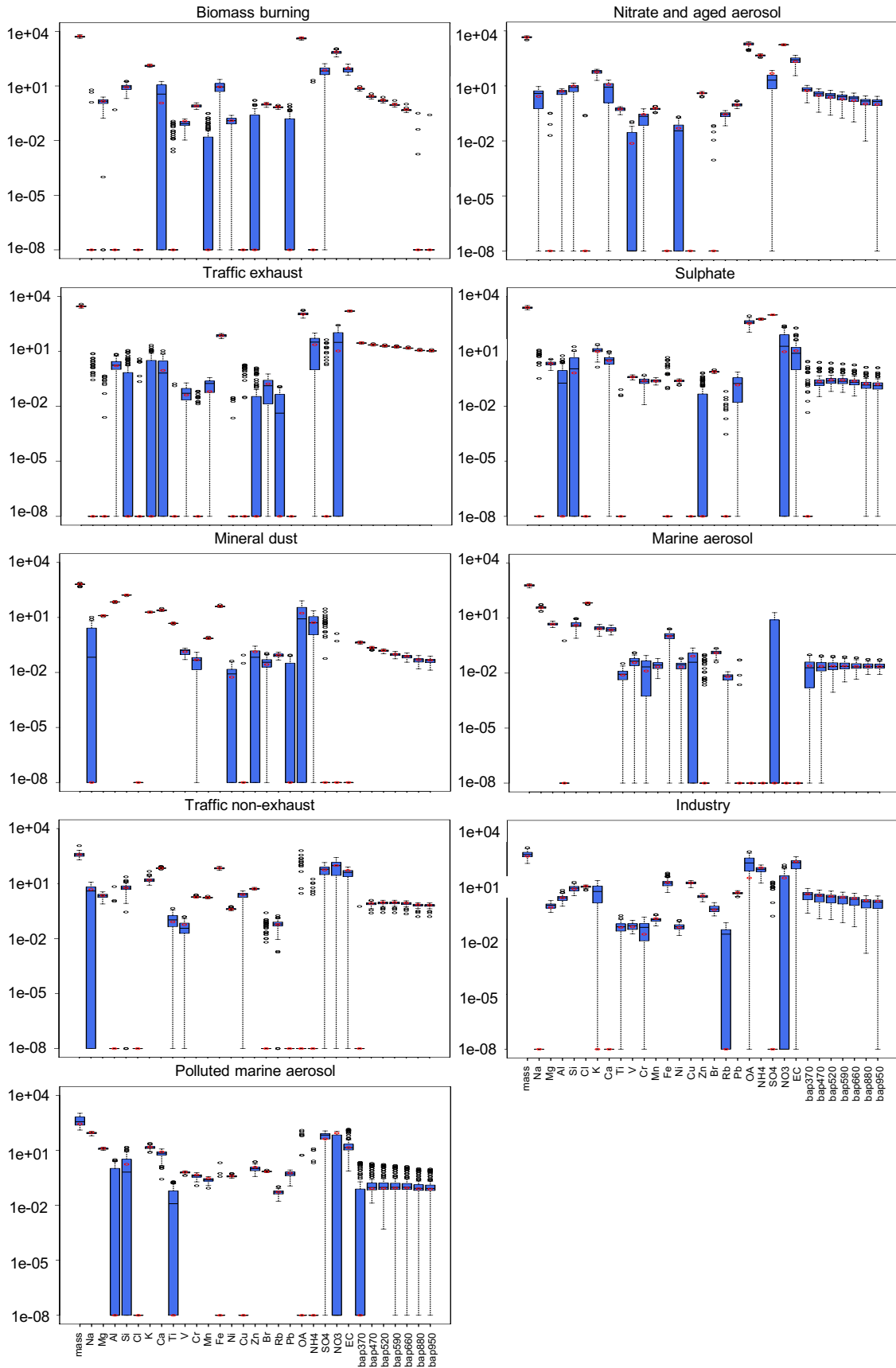
Figure 3.23: Box plot of the bootstrap analysis on the nine-factor constrained solution. The red dots represent the output values of the solution, the black lines the medians from the bootstrap analysis, the blue bars the 25th and 75th percentiles, the dotted lines the interval equal to 1.5 times the interquartile range, and the black dots the outliers from this interval. Figure adapted from [74].

Model results exploiting optical absorption variables

As already mentioned, the source apportionment of the light absorption coefficients b_{ap} at different wavelengths (see Fig. 3.19) strengthens the identification of sources giving additional information about their contribution to light absorption in the atmosphere. In addition, the multi-variable modelling approach introduced by Forello et al. [31] allows the retrieval of relevant source-dependent optical parameters – such as the absorption Ångström exponent and the mass absorption cross section - without any a-priori assumption. In Table 3.9 the b_{ap} apportionment at different wavelengths is shown; traffic exhaust and local source emissions are added together to consider total fossil fuel emissions.

Table 3.9: Average contribution to total b_{ap} reconstructed by the model at different wavelengths; in parentheses, the 10th and 90th percentiles are reported. Table adapted from [74].

	370nm	470nm	520nm	590nm	660nm	880nm	950nm
Biomass burning	17 %	8.5 %	6.0 %	4.1 %	2.5 %	0 %	0 %
	(14-18)	(6.8-9.0)	(4.9-6.4)	(3.4-4.4)	(2.0-2.7)	(0-0)	(0-0)
Nitrate and aged aerosol	12 %	9.2 %	8.5 %	8.1 %	7.8 %	7.1 %	6.9 %
	(8.9-22)	(5.9-19)	(5.2-18)	(4.6-18)	(4.3-17)	(3.6-17)	(3.4-17)
Mineral dust	0.9 %	0.7 %	0.6 %	0.4 %	0.4 %	0.3 %	0.3 %
	(0.8-1.1)	(0.6-0.8)	(0.5-0.6)	(0.3-0.5)	(0.2-0.4)	(0.2-0.4)	(0.2-0.4)
Fossil fuel	70 %	78 %	80 %	82 %	83 %	86 %	86 %
	(62-73)	(69-81)	(71-83)	(72-84)	(73-86)	(75-89)	(76-89)

As expected, the relative contribution to the total reconstructed b_{ap} ascribed to the factors related to biomass burning and mineral dust decreases with increasing λ, in contrast to the contribution from fossil fuel combustion. The most significant contribution to b_{ap} at all wavelengths is given by the traffic exhaust emission source (significant also in terms of EVF, ranging from 0.63 to 0.77 and increasing with increasing wavelength), followed by the factor assigned to the local source. These two main contributors to optical absorption in the atmosphere are related to fossil fuel combustion (traffic exhaust + local source emissions) as highlighted by the value of the absorption Ångström exponent (α) that is 1.1 (1.0 - 1.1 as 10th - 90th percentile from the bootstrap analysis); in fact, α values near 1 are typically associated to light absorption contribution dominated by fresh BC emissions. In Fig. 3.24, the wavelength dependence of b_{ap} for fossil fuel emissions together with the other sources is reported; the line corresponds to the data fitting considering $b_{ap} \propto \lambda^{-\alpha}$.
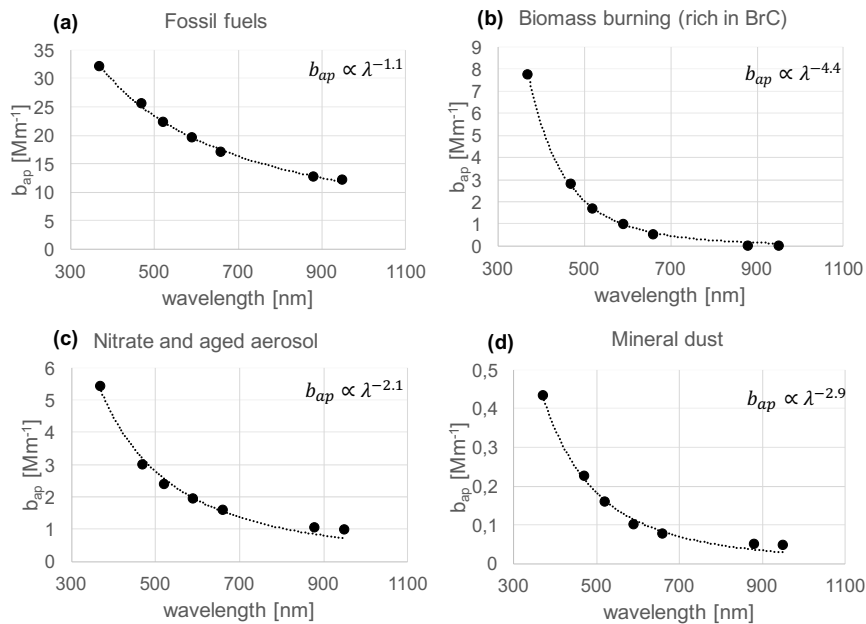
Figure 3.24: b_{ap} dependence on λ for (a) fossil fuels, (b) biomass burning, (c) nitrate and aged aerosol, and (d) mineral dust. Figure from [74].

Even if the other sources do not contribute as much as fossil fuel emissions to the light absorption in the atmosphere, useful information can be retrieved considering source-dependent optical parameters.

The wavelength dependence of b_{ap} apportioned to the biomass burning source in Fig. 3.24 is characterised by α value of 4.4 (4.4 - 4.5 as 10th - 90th percentile); it is higher than typical literature α values for biomass burning (e.g. [75] and references therein) but the significant role played by BrC in this source can account for it [154]. In the literature, BrC was already found in particles enriched in nitrate (that is the second highest contributor in the source chemical profile, after OA) and poor in BC, with a BC-to-OA ratio below 0.05 ± 0.03 [155]; considering EC as a proxy for BC, the same ratio in the biomass burning chemical profile was 0.02.

The wavelength dependence of b_{ap} for the nitrate and aged aerosol source in Fig. 3.24 has α value of 2.1 ($1.6 - 2.6$ as 10th - 90th percentile from bootstrap analysis), consistent with a mixed contribution from both BC and BrC.

Even if the mineral dust source is characterised by very low values of b_{ap} it has a clear wavelength dependence (Fig. 3.24), in contrast to the other remaining sources giving negligible contributions to light absorption. For this source, α is 2.9 ($2.6 - 3.5$ as 10th - 90th percentile), i.e. comprised in the typical range for desert dust reported in the literature (e.g. [12] and references therein). This result is noteworthy because values for the absorption Ångström exponent of mineral dust are still relatively scarce in the literature. Absolute b_{ap} values apportioned to the mineral dust source are much lower (ranging from 0.9% to 0.3% of the total reconstructed b_{ap} - depending on the wavelength - see Table 3.9) than the ones from fossil fuels combustion and biomass burning; this result can be expected since the transport episode of mineral dust is very short (lasting less than two days over the whole campaign). The picture is totally different when considering the time interval covering the transport event (from 24th February 12:00 until 25th February 15:00 LT, estimated considering the temporal pattern in Fig. 3.20); indeed, even if the dominant contribution is still given by fossil fuels combustion (from 59% to 75% of the total reconstructed b_{ap}, increasing with increasing wavelength), the mineral dust impact on light absorption ranges from 25% at $\lambda = 370$ nm to 10% at $\lambda = 950$ nm.

Another relevant result from this modelling approach is the estimate of the ratio between $b_{ap}(\lambda)$ and EC - here considered as a proxy for BC concentrations - for each source. It is noteworthy that when BC is the only absorbing component, $b_{ap}(\lambda)$-to-EC ratio provides the mass absorption cross-section of BC (MAC_{BC}) at different wavelengths; this assumption can be considered valid for fossil fuel emissions (for which $\alpha = 1.1$). Calculations of $b_{ap}(\lambda)/EC$ for biomass burning, fossil fuel, and nitrate and aged aerosol sources are reported in Fig. 3.25 and Table 3.10. The average MAC value for BC – not related to the specific sources – was estimated by Costabile et al. [131] during the same campaign as 8.7 $m^2 g^{-1}$ at $\lambda = 637$ nm.
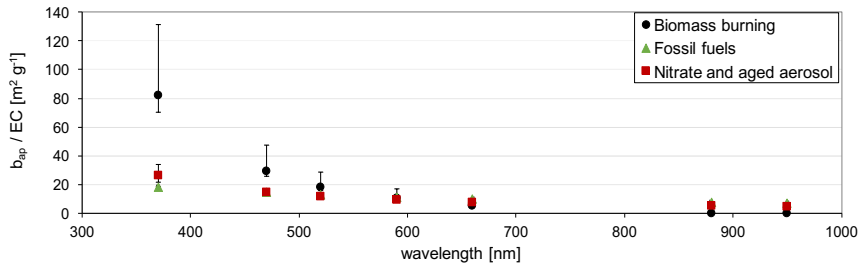
Figure 3.25: b_{ap}-to-EC ratio dependence on λ for biomass burning, fossil fuels, and nitrate and aged aerosol. Error bars represent the 10th and 90th percentiles from the bootstrap analysis. Figure from [74].

It is noteworthy the large difference at shorter wavelengths among the various sources, which is due to the contribution of BrC. The difference respect to fossil fuels - where BC can be considered as the only absorbing component - is clearly more significant for biomass burning, where BrC resulted to be dominant ($\alpha = 4.4$), than for nitrate and aged aerosol. For the latter, significant differences are present only at $\lambda = 370$ nm (α for this source was 2.1) .

Table 3.10: Ratio between $b_{ap}(\lambda)$ and EC for biomass burning, fossil fuel, and nitrate and aged aerosol sources; in parentheses, the 10th and 90th percentiles are reported. Table adapted from [74].

b_{ap}/EC $[m^2 g^{-1}]$	370nm	470nm	520nm	590nm	660nm	880nm	950nm
Biomass burning	81.9	29.5	17.8	10.4	5.4	0.0	0.0
	(70.4-131.4)	(25.5-47.4)	(15.5-28.8)	(9.1-17.0)	(4.7-8.9)	(0.0-0.0)	(0.0-0.0)
Nitrate and aged aerosol	26.2	14.4	11.4	9.4	7.7	5.0	4.7
	(21.7-33.8)	(13.6-16.6)	(11.0-12.3)	(8.8-10.3)	(6.9-8.6)	(4.1-6.0)	(3.8-5.7)
Fossil fuel	18.3	14.5	12.7	11.1	9.7	7.2	6.9
	(17.6-18.6)	(13.9-14.7)	(12.2-12.9)	(10.7-11.3)	(9.3-9.9)	(6.9-7.4)	(6.6-7.0)

3.3.5   Comparison between ME-2 modelling and ACSM results on organics

In order to obtain more insights on the organic aerosol (OA) apportionment, results from the modelling approach presented in this work coupling chemical and optical variables (ME-2_{all}, in the following) were compared with an independent source apportionment study previously performed [131] on OA measured by the ACSM (ME-2_{org}, in the following). Using the latter approach, three factors were recognised: HOA (hydrocarbon-like organic aerosol), BBOA (biomass burning-like organic aerosol), and OOA (oxygenated organic aerosol); HOA and BBOA (i.e. primary OA components) accounted for about 12% of the OA mass each, while OOA was the main component accounting for the remaining apportioned mass fraction.

Results from the application of ME-2_{all} showed that the main contributors to organic aerosol concentrations in the atmosphere (see also Fig. 3.19) were biomass burning (accounting nearly for 58% of the total OA concentrations reconstructed by the model), nitrate and aged aerosol (almost 24%), and traffic exhaust emissions (almost 14%). As an original contribution of this work, in Figure 3.26 a comparison between temporal patterns related to OA apportioned by ME-2_{all} (hereafter referred to as $OA_{biomass\ burning}$, $OA_{nitrate\ and\ aged\ aerosol}$, $OA_{traffic\ exhaust}$) vs. BBOA, OOA, and HOA obtained by ME-2_{org} is reported.
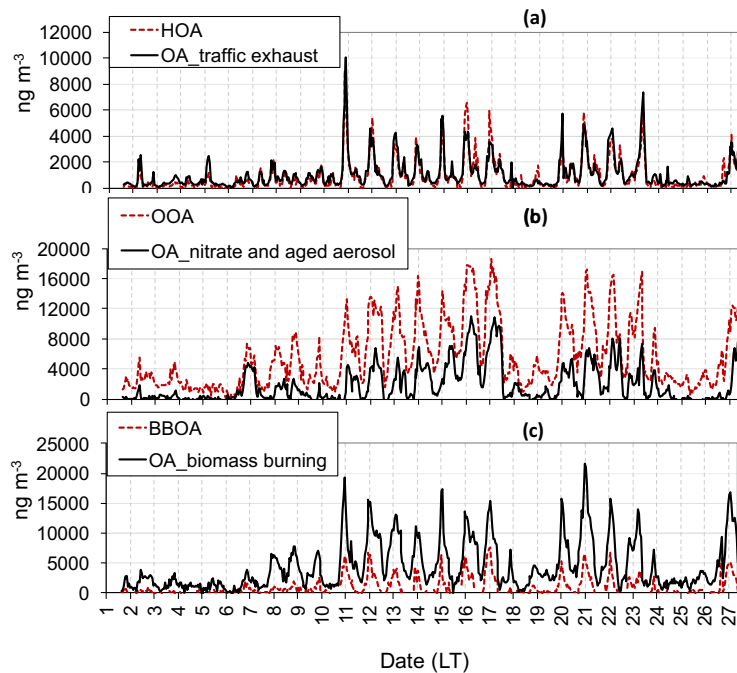
Figure 3.26: Hourly temporal patterns of (a) HOA from ME-2_{org} and OA apportioned to traffic exhaust by ME-2_{all}, (b) OOA from ME-2_{org} and OA apportioned to nitrate and aged aerosol by ME-2_{all}, (c) BBOA from ME-2_{org} and OA apportioned to biomass burning by ME-2_{all} for February 2017. Vertical lines show midnight for each day. Figure from [74].

The first noteworthy result is that HOA and $OA_{traffic\ exhaust}$ retrieved by the two different approaches are similar in temporal patterns ($R^2 = 0.85$) but – more importantly – fairly comparable in terms of absolute values (within 10% difference on average) (Fig. 3.26).

Also $OA_{nitrate\ and\ aged\ aerosol}$ shows similar features to OOA ($R^2 = 0.74$) in terms of temporal behaviour thus confirming that secondary aerosol and ageing processes impact on the source identified by ME-2_{all} as nitrate and aged aerosol. Correlation between the fraction of semi-volatile OOA (SV-OOA) and NO_3^- was already observed in Rome [151], in agreement with other literature studies [156]. However, $OA_{nitrate\ and\ aged\ aerosol}$ absolute values are much lower than OOA from ME-2_{org} (Fig. 3.26), suggesting that part of the OOA is apportioned to other sources by ME-2_{all}.

The biomass burning source retrieved by ME-2_{all} is characterised by a more complex mixture of organics showing a significant correlation with both BBOA ($R^2 = 0.74$) and OOA ($R^2 = 0.75$) from ME-2_{org}. However, one relevant difference is related to BBOA absolute concentration values, which do not account for all the OA apportioned by ME-2_{all} to the biomass burning source. In addition, the decrease of BBOA concentration values steeply reaches zero (typically during the time interval from 11 to 17 LT) while the $OA_{biomass\ burning}$ has higher concentration values (Fig. 3.26), especially during the period characterised by atmospheric stability (from about 10/02 until 24/02, excluding 18/02 and 19/02).

The discrepancies in organic aerosol absolute values mentioned for the latter two cases are very interesting and deserve a further discussion as they were never reported in previous works. Indeed, this observation can be explained considering that a consistent part of the OOA – generically ascribed to aged aerosol in literature works (see e.g. [156]) – is likely linked to the biomass burning source as shown by ME-2_{all} results and better described in the following. As can be seen in Fig. 3.27, the temporal pattern of the difference between $OA_{biomass\ burning}$ and BBOA is substantially overlapped with the difference between OOA from ME-2_{org} and $OA_{nitrate\ and\ aged\ aerosol}$ from ME-2_{all} (in the following, OOA-$OA_{nitrate\ and\ aged\ aerosol}$). Consistently, adding the contribution from OOA-$OA_{nitrate\ and\ aged\ aerosol}$ to BBOA apportioned by ME-2_{org}, the correlation with the biomass burning source from ME-2_{all} significantly increases ($R^2 = 0.92$ vs. 0.74) and also absolute concentration values are very similar, within 4% on average. Therefore, OOA-$OA_{nitrate\ and\ aged\ aerosol}$ can be considered a rough minimum estimation of the biomass burning contribution to OOA and on average it accounts for 60% of OOA concentrations, corresponding to 43% of total OA measured by ACSM. This is the second noteworthy result of this work, as it represents an estimate of the secondary contribution to OA due to biomass burning; therefore, it could be added to the 12% estimated as BBOA (typically associated only to primary aerosol content), evidencing the eminent role of biomass burning ($> 50\%$) - with its primary and secondary contributions - in explaining the total OA measured during the CARE campaign.

In contrast to the other two sources, the chemical profile of traffic exhaust from ME-2_{all} seems to be constituted mainly by primary emissions since $OA_{traffic\ exhaust}$ from ME-2_{all} corresponds to HOA from ME-2_{org}: thus, OOA contributions related to secondary organic components can be considered negligible in this source. Secondary organic compounds due to traffic emissions are likely mixed in the chemical profile of the nitrate and aged aerosol source from ME-2_{all}, so that minimum estimation of their contribution is not possible in this case.
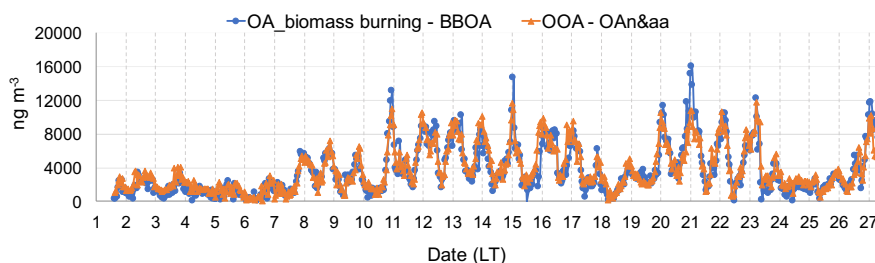
Figure 3.27: Hourly temporal patterns of the difference between OA apportioned by ME-2_{all} to the biomass burning source and BBOA from ME-2_{org} ($OA_{biomass\ burning}$ – BBOA) and the difference between OOA from ME-2_{org} and OA in the nitrate and aged aerosol source from ME-2_{all} (here denoted as OOA-$OA_{n\&aa}$) for February 2017. Vertical lines show midnight for each day. Figure from [74].

3.4 Brief description of my personal contribution

I contributed to the experimental work in the study described in 3.2 and reported in the publication [31]. In particular, I was in charge of the high time resolution samplings using the streaker sampler. I fitted the spectra resulting from PIXE analysis on streaker samples using the software GUPIXWIN; on these samples, I also performed the measurements of the aerosol light absorption coefficients at four wavelengths using the in-house polar photometer PP_UniMI.

Moreover, I carried out all the source apportionment studies presented in this Chapter, starting from the analysis and preparation of data with different time resolutions to be inserted into the model, including the aerosol light absorption coefficients. The latter were never inserted in a modelling procedure before (at least not in the same way); it is noteworthy that the results provided new insights for the aerosol science community using optical source apportionment methods. I implemented the multi-time resolution model to be applied to the final datasets and I performed the modelling analysis, including the final processing and evaluation of the model outputs.

As for the study reported in Sect. 3.3 and in the publication [74], I proposed a new approach to make a joint analysis using also ACSM outputs to retrieve new information on secondary contributions which are a known issue when using receptor modelling approaches. Finally, I wrote the first draft of the papers.

Chapter 4

Experimental methodologies and applications: atmospheric aerosol sampling

4.1   Brief description of my personal contribution

During my PhD, I took part in the INFN (National Institute of Nuclear Physics) experiment TRACCIA (Time Resolved Aerosol Characterization Challenging Improvements and Ambitions), in collaboration with other Italian research groups (INFN-Florence, INFN-Genoa, and INFN-Lecce). This experiment is devoted to the development, realisation, and characterisation of the new high time resolution sampler STRAS (Size and Time Resolved Aerosol Sampler), with the aim of improving the performance and technical characteristics of streaker sampler, which is now out of production.

In the literature, examples of high time (i.e. hourly) resolution samplers able to perfom size-segregated collection of particles are e.g. streaker sampler [86], and DRUM sampler [157]. Samples produced by these instruments are characerised by small deposit areas (e.g. about 1x8 mm^2 for streaker sampler) and low aerosol loads; because of these characteristics, the implementation of suitable techniques through dedicated set-up is mandatory for sample analyses.

As a member of the research group of Milan, I collaborated to carry out calculations for the sampler cut-off diameter (i.e. the value of the aerodynamic diameter for which particles are collected with an efficiency of 50% on the impaction stage) to be used in the new design, I personally performed the theoretical calculation of collection efficiencies and experimental pressure drops needed to evaluate suitable filters, and I took part in preliminary experimental tests both on field and in the atmospheric simulation chamber ChAMBRe.
In this chapter STRAS characterisation is outlined, together with main experimental results.

4.2   Development of a new aerosol sampler: STRAS

Many atmospheric aerosol samplers (including streaker sampler and STRAS) exploit the phenomenon of inertial impact to separate particles with different dimensions. The basic scheme is the following: the sampled air passes through a nozzle in which particles are accelerated, and the air flow is directed towards an obstacle (referred to as impaction stage or impaction plate) at the end of the nozzle itself. Air flow streamlines undergo a sharp deviation due to the presence of the impaction stage; particles with sufficient inertia collide with the impaction stage, while the rest are able to follow the flow streamlines (Fig. 4.1). Theoretically, all particles characterised by an aerodynamic diameter - see Sect. 1.1 for the definition - larger than a threshold value (called cut-off diameter) are removed from the flow. A thorough description of the theory behind size-segregated aerosol sampling and filtration can be found e.g. in [3] [5] [158] [159] [160] [161] [162] [163] and is also reported in Appendix A.
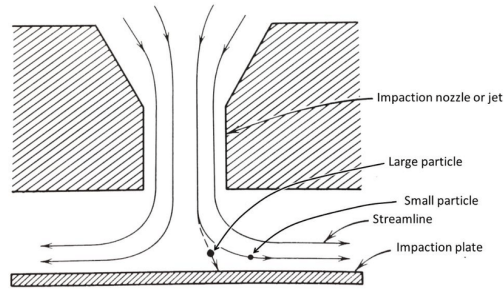
Figure 4.1: Basic scheme of the phenomenon of inertial impact. Figure modified from [3].

4.2.1 STRAS development and sizing

STRAS design

STRAS was developed aiming at improving streaker sampler performances; it is worth mentioning that the streaker sampler, widely used by the research groups of Milan, Florence, and Genoa (see e.g. [86]) is now out of production. STRAS was designed in order to obtain higher areal concentration of particle deposits (see the following) and operate with PM10 EPA inlet at a flow rate of 16.7 $l \ min^{-1}$ (i.e. $1 \ m^3 h^{-1}$) instead of $1 \ l \ min^{-1}$ as in the streaker sampler. For a robust and reliable particle sampling procedure, STRAS was integrated with a customised sampling unit (DADO LAB SRL) for flow regulation within $\pm$ 1%; this implementation lead to a stand-alone instrumentation for outdoor measurements (Fig. 4.2).
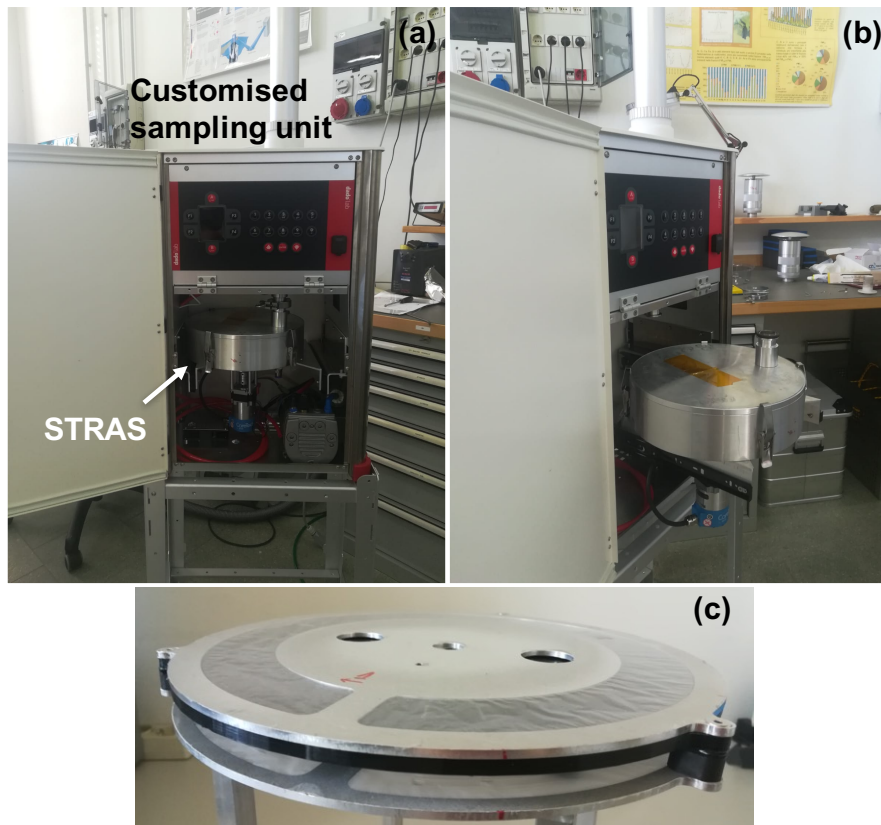
Figure 4.2: Configuration of STRAS integrated with the customised sampling unit (DADO LAB SRL). (a) STRAS is inserted in the unit as during sampling; (b) a retractable arm allows to easily handle STRAS (e.g. for support change); (c) the two rigidly coupled circular stages inside STRAS, consisting of the impaction stage (polypropilene foil) and the filter.

STRAS final configuration consists of two rigidly coupled circular stages (with diameters of about 25 cm): an impaction stage consisting in a polypropilene (PP) foil - typically coated with Apiezon to minimise particle bouncing - to collect particles in the coarse fraction, i.e. with an aerodynamic diameter d_{ae} between 2.5 and 10

μm, and a filter for the collection of the aerosol fine fraction, i.e. particles with $d_{ae} < 2.5$ μm (Fig. 4.2). The type of filter currently selected for STRAS is Nuclepore$^{\text{TM}}$, the same used by the streaker sampler. Particles are collected on the filter using a sucking orifice with a rectangular area of about 3x30 mm^2, for the sampling time of one hour (in case of high atmospheric aerosol concentrations, shorter sampling times or bigger deposit areas - 4x30 mm^2 or 5x30 mm^2 - can be also set to avoid filter clogging and the occurrence of high pressure drops); measurement of the deposit area on the PP foil is very critical due to difficulties in the correct estimate of the actual area of the sample produced by the impaction process. After each sampling time, the two stages are automatically moved to start a new sampling. In this way, 168 samplings can be automatically performed on both supports, corresponding to one week of hourly samples.

Proper sampler design is mandatory to obtain chosen cut-off diameters and efficiency curves close to ideality. STRAS design parameters are resumed in Table 4.1 and compared to literature guidelines, that are thoroughly described in Sect. A.2 in the Appendix. For STRAS, cut-off at 10 μm is operated by the PM10 EPA inlet; the separation between coarse and fine fraction at 2.5 μm is performed through a rectangular nozzle with dimensions width x length (WxL) of 1.24 ($\pm$ 0.01) x 12.33 ($\pm$ 0.01) mm. The jet-to-plate distance (S, see Fig. 4.3) is 1.90 ($\pm$ 0.01) mm, and the nozzle throath length (T, see Fig. 4.3) is 1.88 ($\pm$ 0.01) mm.
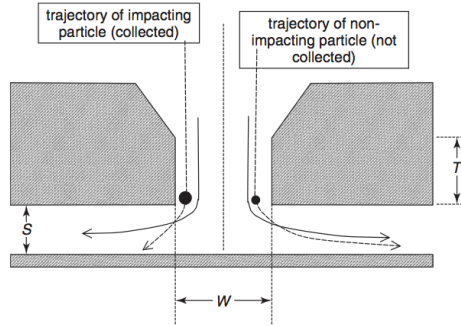
Figure 4.3: Crucial dimensional parameters of inertial classifiers. Figure from [164].

Table 4.1: STRAS design parameters compared to literature guidelines. The propagation of uncertainties related to S, W, T, and L was used to evaluate uncertainties reported in this Table.

	S/W	T/W	L/W	Re_{jet}
STRAS	1.53 $\pm$ 0.01	1.52 $\pm$ 0.01	9.94 $\pm$ 0.08	3000 $\pm$ 2
Literature	>1.5	1-2	$\geq$ 10	300-5000

Areal particulate matter concentration on filter

The dimension of the sucking orifice (3x30 mm^2) was chosen in order to increase the areal concentration of PM on the filter respect to streaker sampler; areal concentration is defined by dividing the total sampled PM mass ($C_{air} \cdot V_{sampled}$) by the sampling area:

$$C_{areal} = \frac{C_{air} \cdot V_{sampled}}{Area} = \frac{C_{air} \cdot Q \cdot t}{Area} = C_{air} \cdot t \cdot v_f \qquad (4.1)$$

where Q represents the volumetric flow, t the sampling time, and v_f the air face velocity on the filter. Considering STRAS characteristics ($Q = 16.7$ l min^{-1} and $Area = 3$x30 mm^2), we obtain $v_f \simeq 309$ cm s^{-1} that corresponds to an estimated theoretical increase of a factor of 1.9 in C_{areal} on the filter respect to streaker sampler. Experimental tests on C_{areal} were performed through parallel sampling by streaker sampler and STRAS with a deposit area of 5x30 mm^2 during test configuration, that corresponds to an estimated increase of a factor of 1.1 in C_{areal} on the filter. Samples were subsequently analysed by PIXE technique in order to obtain the elemental concentrations (see e.g. [96] for more details about PIXE analysis). Results for samples collected on Nuclepore$^{\text{TM}}$ filters (i.e. aerosol fine fraction) and on PP foils (i.e. aerosol coarse fraction) are reported in Fig. 4.4. Average experimental differences in C_{areal} on Nuclepore$^{\text{TM}}$ filters are fully consistent with the ones theoretically estimated above (a factor of $\simeq$ 1.1). Average experimental differences in C_{areal} on the PP

foil are around a factor of 2.6, but a theoretical value for comparison was not reported in this work due to the above-mentioned difficulties in the definition and measurement of areas of the samples produced by the impaction process.

Besides C_{areal}, v_f is also connected to the pressure drop and collection efficiency of the filter (see Sect. A.3 in the Appendix for more details), and these parameters were also investigated during STRAS development.
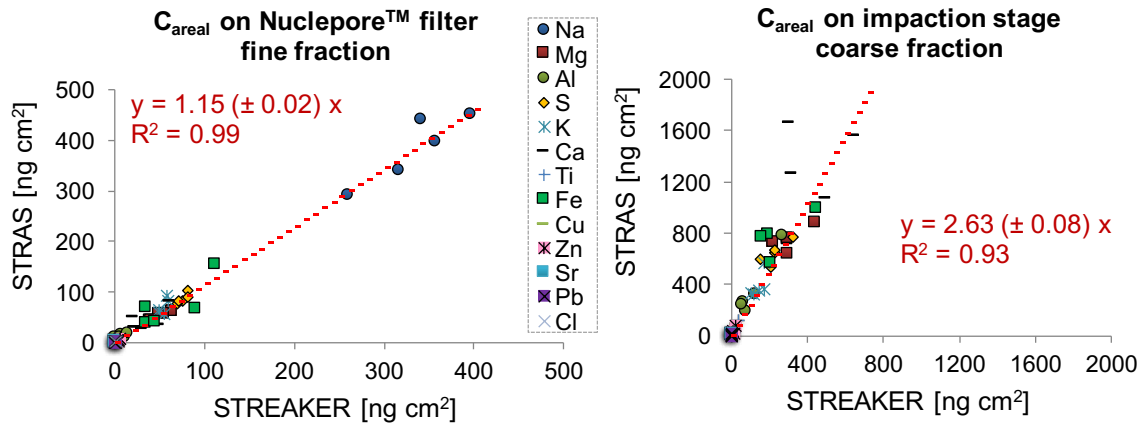
Figure 4.4: Comparison between areal concentration of different elements on samples collected by parallel sampling of STRAS and streaker. Left: Results for the fine fraction collected on Nuclepore[TM] filter; right: Results for the coarse fraction collected on the impaction stage (polypropilene foil). Measurements related to different elements are characterised by different colors and symbols, as reported in the legend.

Filter collection efficiency

The pressure drop on Nuclepore[TM] filters depends on pore size, and bigger pore diameters are generally associated to smaller pressure drops; this can be advantageous especially when a high sampling flow rate (e.g. 16.7 $l\ min^{-1}$ in the STRAS configuration, instead of 1 $l\ min^{-1}$ of the streaker sampler) is considered, but at the same time larger pores are also related to smaller collection efficiencies. As STRAS flow rate claims for Nuclepore[TM] filters with larger pores compared to streaker, the efficiency of filters with pore diameters of 0.4 μm and 0.8 μm were compared as explained in the following. Calculations are based on the theory reported in Sect. A.3 in the Appendix and results are presented in Fig. 4.5. The filters considered for STRAS were characterised by the same parameters - with the exception of the pore size - as reported in Table 4.2).
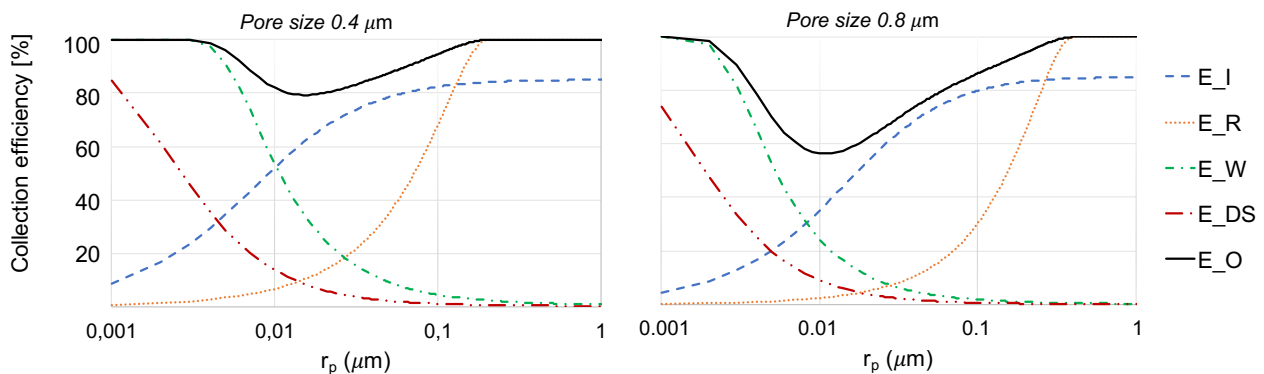
Figure 4.5: Partial and overall theoretical efficiencies for Nuclepore[TM] filters with pore diameters of 0.4 μm and 0.8 μm.

Table 4.2: Parameters used in the theoretical calculation of collection efficiencies of Nuclepore^{TM} filters. Air viscosity and mean free path were calculated at a temperature of 296.15 K and a pressure of 101.3 kPa.

Filter porosity P	0.15
Filter thickness L	9 μm
Face velocity v_f	309 $cm\ s^{-1}$
Particle density ρ_p	2.2 $g\ cm^{-3}$
Air viscosity η	1.83 $\cdot 10^{-4} g\ cm^{-1} s^{-1}$
Mean free path λ	0.067 μm

The minimum collection efficiency is at about the same particle radius for both types of filters ($\simeq$ 11 nm and 15 nm for filters with pore diameters of 0.8 μm and 0.4 μm, respectively) but with different values: $\simeq$ 56% for filters with pore size of 0.8 μm, and 80% for the ones with pore size of 0.4 μm. Considering typical aerosol mass size distributions for Milan urban area [165], the differences between the total aerosol mass collected in the two cases were estimated; the estimate was performed multiplying the atmospheric aerosol mass size distribution by the collection efficiency. The total PM2.5 mass collected on filters with pore diameters of 0.8 μm was estimated to be 1% smaller than the one collected on filters with pore diameters of 0.4 μm; calculations were also performed for e.g. sulphur - an element with most of the mass typically in the fine fraction - and the difference in the PM2.5 mass was negligible, i.e. around 2%.

A field test was also performed in the city of Florence (unfortunately typical aerosol mass size distributions were not available for theoretical calculations) mounting the two different types of filters on two streaker samplers (since two STRAS samplers were not available yet) and analysing the samples by PIXE technique to obtain the elemental concentrations. The difference in PM2.5 mass collected by the two samplers for elements such as S, K, Ca, Fe, Cu, and Zn was always in the range 10-30% (two examples are reported in Fig. 4.6). Therefore, different collecton efficiencies do not seem to be dominant on the difference in concentration values, since the range 10-30% is very similar to the one obtained in the past during parallel samplings performed by streaker samplers [166]; in that case, discrepancies were likely related to slight differences in the construction parameters of the different samplers, giving not negligible variations in the overall results. In the end, Nuclepore^{TM} filters with pore diameters of 0.8 μm were chosen for the final configuration of STRAS, and they will be considered in the following.
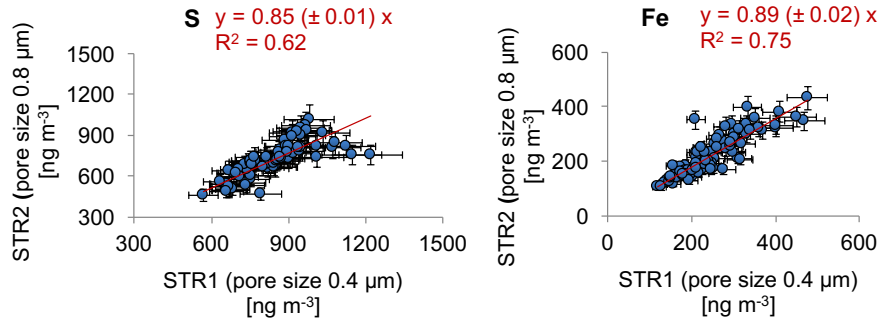
Figure 4.6: Comparison between concentrations of sulphur (left) and iron (right) on samples collected by two parallel streaker samplers (STR1 and STR2). STR1 was equipped with a Nuclepore^{TM} filter with pore diameters of 0.4 μm; STR2 with a Nuclepore^{TM} filter with pore diameters of 0.8 μm. Note the lower limit of the axes in the graph on the left.

Pressure drop on filter

The dependence of pressure drop of Nuclepore^{TM} filters on different face velocities v_f was experimentally tested. Tests were performed using an impactor prototype set to operate at 1 $m^3 h^{-1}$ (i.e. 16.7 $l\ min^{-1}$) and equipped with an EPA PM10 inlet, simulating STRAS impaction stages but suitable for aerosol sampling on 47 mm Nuclepore^{TM} filters. Different face velocities were obtained gradually decreasing the filter sampling area through the use of aluminum masks with different diameters (some examples are reported in Fig. 4.7). Pressure drop was measured by a sensor placed downstream the filter; experimental pressure drops are reported in Table 4.3.
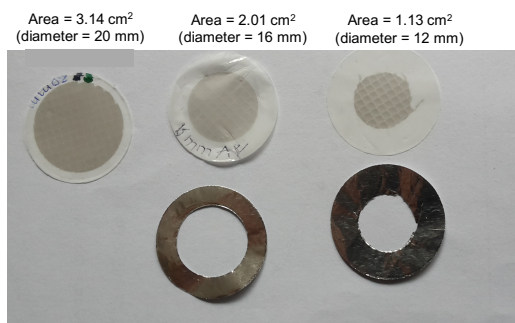
Figure 4.7: Different Nuclepore^{TM} filter sampling areas obtained through the use of aluminum masks with different diameters.

Table 4.3: Experimental relation between face velocity and pressure drop on Nuclepore^{TM} filters.

Sampling area cm^2	Average flow $l\ min^{-1}$	Face velocity $cm\ s^{-1}$	Pressure drop $kPa\ (cmHg)$
3.14	16.7	89	20 (15)
2.01	16.7	138	32 (24)
1.20	16.7	232	45 (34)
1.13	16.7	246	53 (40)
0.90	16.7	309	60 (45)

Values in Table 4.3 are in agreement with literature ones available for Nuclepore^{TM} filters, as shown in Fig. 4.8 and considering filter characteristics reported in Table 4.4 [167]. For the face velocity reached inside STRAS during normal sampling ($v_f \simeq 309\ cm\ s^{-1}$), the pressure drop on filter is around 60 kPa: this parameter was taken into account in the choice of the pump to be used for sampling. In the end, the dry vacuum pump Becker VT 4.8 was considered as the best choice in terms of performance and manageability. The possibility to use other types of filter producing lower pressure drops - and, at the same time, suitable for subsequent analyses such e.g. PIXE - is currently under investigation.
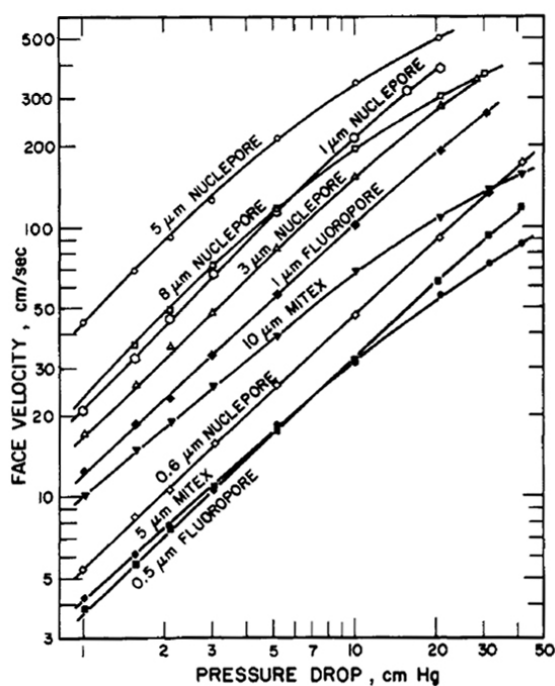
Figure 4.8: Experimental relation between face velocity and pressure drop for different types of membrane filters. Figure from [167].

Table 4.4: Parameters for Nuclepore^{TM}filters reported in Fig. 4.8. Table adapted from [167].

Nominal pore diameter μm	Pore density $pores\ cm^{-2}$	Porosity $\%$	Thickness μm
0.6	$3 \cdot 10^7$	8.4	10
0.8 (this work)	$3 \cdot 10^7$	15	9
1	$2 \cdot 10^7$	15.6	10
3	$2 \cdot 10^6$	14.1	10
5	$4 \cdot 10^5$	7.8	10
8	$1 \cdot 10^7$	5	10

4.2.2 Preliminary field tests
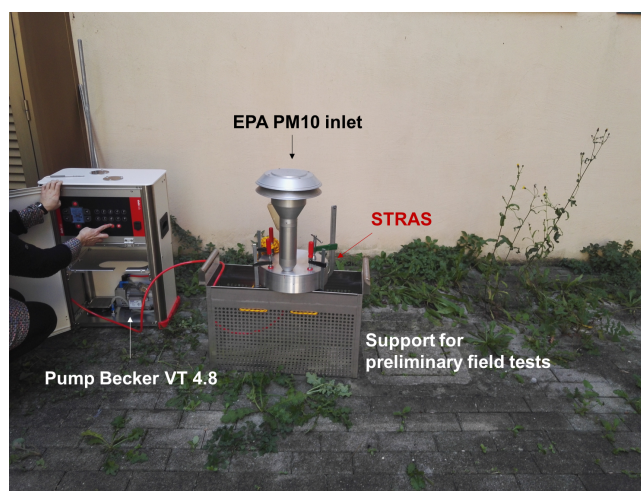
Figure 4.9: Experimental set-up for preliminary field tests of STRAS.

Preliminary field samplings were performed to test STRAS flow conditions (Fig. 4.9), and the target flow rate of 16.7 $l\ min^{-1}$ was reached during STRAS sampling. In order to verify the shape of the aerosol deposit on both the impaction stage (aerosol coarse fraction) and the filter (aerosol fine fraction), some icing sugar (expected in the coarse fraction) was resuspended and a cigarette was smoked near the sampler to produce fine aerosol. Examples of obtained samples are reported in Fig. 4.10.
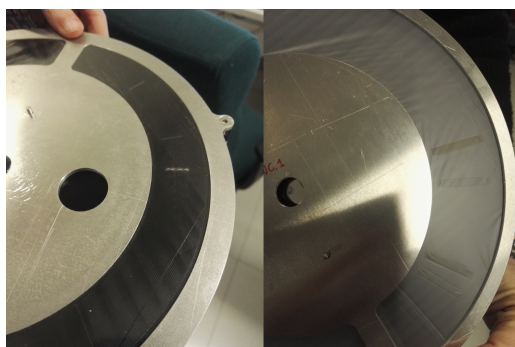
Figure 4.10: Samples produced during preliminary field tests of STRAS. On the left, samples produced on the polypropilene foil of the impaction stage (aerosol coarse fraction); on the right, samples produced on Nuclepore^{TM} filter (aerosol fine fraction).

4.2.3 Preliminary tests in the atmospheric simulation chamber ChAMBRe

The atmospheric simulation chamber ChAMBRe (Chamber for Aerosol Modelling and Bio-aerosol Research) was exploited to resuspend particles with known size in order to perform a preliminary characterisation of
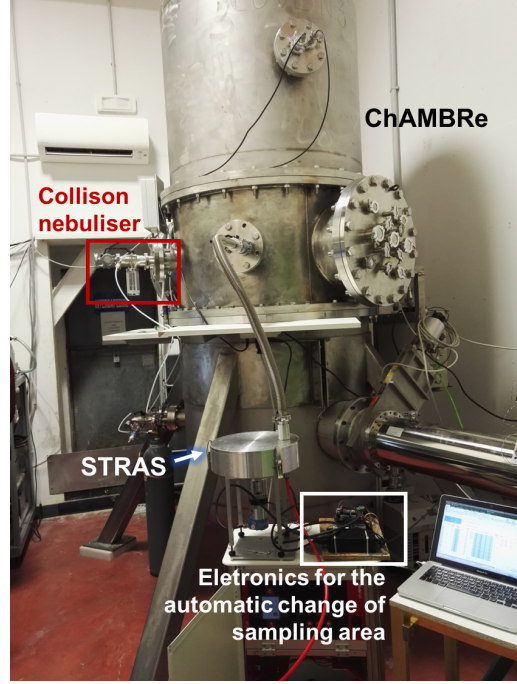
Figure 4.11: Experimental set-up for the preliminary characterisation of STRAS cut-off diameter in 2019. A very similar configuration was used in 2018 too.

STRAS cut-off diameter. ChAMBRe is installed at the National Institute of Nuclear Physics in Genoa (INFN-Genoa) and developed in collaboration with the Environmental Physics Laboratory at the Physics Department of the University of Genoa. The chamber is a node of the EUROCHAMP-2020 consortium since 2017; a detailed description and characterisation can be found in [168]. Briefly, ChAMBRe is made of stainless steel, with a total volume of about 2.2 m^3; it was designed to work at atmospheric pressure, even if a composite pumping system allows to reach a vacuum level of about $5 \cdot 10^{-2}$ $mbar$ in about 15 minutes. Aerosol injection can be performed through nebulisation, and particles lifetime inside the chamber varies from about 2 to 10 hours depending on their dimensions. After chamber cleaning, the background particle concentration is not significant (0.5 $particle$ cm^{-3} as measured by an Optical Particle Counter - OPC). Aerosol samplers can be easily connected through flanges and maintained in operation for times depending on their flow rate and the needs of the particular experiment.

STRAS sampling line was connected to ChAMBRe, as shown in Fig. 4.11. SiO_2 particles (density ρ_p: 2.65 g cm^{-3}) with known size (Corpuscolar Inc.) were diluted in milliQ water and the solution was mixed by centrifugation. Unfortunately, particles used for these tests were found to be contaminated with Al and organic components likely resulting from the mechanical production process (see e.g. Fig. 4.14 and the discussion in the following); contamination was so high in SiO_2 particles with geometric diameter $d_{geo} \geq 2.5$ μm making impossible the conversion from certified geometric diameters into aerodynamic diameters, so that only particles with d_{geo} from 0.2 μm to 2 μm, corresponding to 0.33 - 3.3 μm as aerodynamic diameters d_{ae} (since $d_{ae} = d_{geo}\sqrt{\frac{\rho_p}{\rho_0}}$, where ρ_0 is the unit density), are considered in this work. SiO_2 particles were chosen for these tests because Si concentration on both the polypropilene (PP) foil (i.e. the impaction stage) and the Nuclepore$^{\text{TM}}$ filter can be easily assessed by PIXE analysis (see e.g. [96] for details about this technique). Future tests of STRAS cut-off diameter exploiting different techniques with different types of particles (e.g. analysing polystirene particle fluorescence by optical microscope) are currently under investigation.

To check STRAS cut-off diameter, two tests were performed in 2018 and 2019 at different steps of STRAS development; anyway, the configuration of the experimental set-up was similar (Fig. 4.11) and main differences are outlined in the following.

Particles were injected in the chamber one size at a time, through a Blaustein Atomiser (BLAM, single-jet model, CH Technologies) at about 2 l min^{-1} in 2018, and through a Collison nebuliser (CH Technologies) at 4-5 l min^{-1} in 2019; an OPC was used to monitor particle concentration inside the chamber. When particle concentration was around 10 μg m^{-3} or higher (in order to collect enough mass to allow detection by PIXE,

considering a maximum sampling time of about 30 minutes), STRAS started the sampling at a flow rate of 16.7 $l\ min^{-1}$. Technical problems in the test performed in 2018 did not allow to reach the target flow rate of 16.7 $l\ min^{-1}$, since STRAS development was still at a preliminary stage; the flow rate was between 12.4 $l\ min^{-1}$ and 14.4 $l\ min^{-1}$ (see Table 4.5), and this issue will be taken into account in data analysis presented in the following.

After each sampling, the chamber was cleaned by making the vacuum, and STRAS impaction stage and filter were rotated in order to perform the new sampling on clean support areas; this rotation was performed manually in the test carried out in 2018, while in 2019 a motor placed inside the sampler allowed the automatic change of sampling area. The entire procedure of injection, sampling, and cleaning was repeated for each particle size.

For the characterisation of particles deposited on both the filter and the impaction stage, two different analyses were performed: PIXE (Particle Induced X-ray Emission) analysis at the LABEC-INFN in Sesto Fiorentino (Florence) [96] and SEM-EDS (Scanning Electron Microscopy - Energy Dispersive X-Ray Spectroscopy) at the Chemistry Department of the University of Genoa [169]. For each particle size, at least one sample devoted to SEM-EDS analysis was collected by STRAS with sampling times of few minutes to avoid overloading of particle concentration on the surface; opposite, a longer sampling time (about 30 min) was carried out to deal with PIXE minimum detection limit.

The first step was the evaluation of the homogeneity of sample spots on the impaction stage (see Fig. 4.10 for an example of particle deposit on the PP foil). Samples were irradiated in different points by PIXE technique, and Si concentration level was stable; an example is reported in Fig. 4.12.
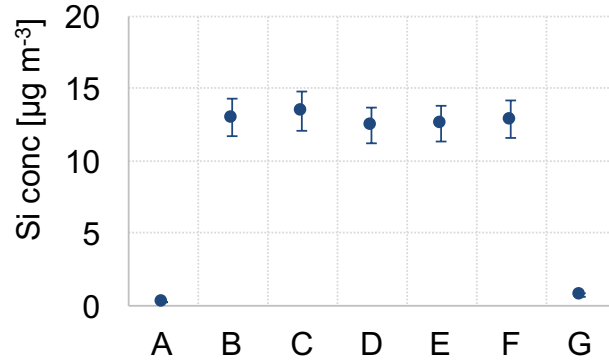
Figure 4.12: Si concentration in different points of a sample collected on the impaction stage (polypropilene foil); A and G are points outside the visible spot.

The experimental collection efficiency on the impaction stage was evaluated dividing the concentration of Si obtained by PIXE technique on the PP foil (Si_{imp}) by the total one collected on both the impaction stage and the Nuclepore™ filter (Si_{filter}):

$$Collection\ efficiency\ (\%) = \frac{Si_{imp}}{Si_{imp} + Si_{filter}} \cdot 100 \qquad (4.2)$$

Results of these preliminary tests are reported in Table 4.5 and in Fig. 4.13. Figure 4.13 shows also the calculated STRAS cut-off diameter corresponding to a flow rate of 16.7 $l\ min^{-1}$ instead of 14.2-14.4 $l\ min^{-1}$. This estimation was made according to the following equation (see Sect. A.2 in the Appendix for more details about its derivation):

$$d_{50} = \sqrt{\frac{9\eta W}{U\rho_0 C_c}} \sqrt{Stk_{50}} \qquad (4.3)$$

where d_{50} is the cut-off diameter; W is the nozzle width; U is the mean air velocity at the exit of the nozzle, corresponding to $\frac{Q}{LW}$ where Q is the volumetric flow through the nozzle and L is the second dimension (length) of the nozzle; η indicates the fluid viscosity; C_c the slip correction factor; ρ_0 the particle unit density; $\sqrt{Stk_{50}}$ is set to 0.77 [3].

From these analyses, STRAS aerodynamic cut-off diameter was estimated to be around 2.8 μm for tests done in 2018, and around 3.3 μm for the ones performed in 2019. Anyway, these results are still preliminary since more tests with certified particle diameters larger than 2 μm should be performed in order to obtain a

reliable collection efficiency curve. Since no commercial SiO_2 particles with these characteristics are available, and issues related to particle properties were highlighted especially by the SEM-EDS analysis (see the discussion in the following), future tests will exploit different types of particles such as fluorescent polystyrene particles, and different techniques for their analysis e.g. optical microscopy. Unfortunately, these tests have been delayed due to the long lock-down period caused by the COVID-19 pandemic.

Table 4.5: Preliminary results from the experimental tests on STRAS cut-off diameter carried out in 2018 and 2019. In 2019, two samplings were performed for some particle diameters (equal to or bigger than $d_{geo} = 0.7 \ \mu m$); the average is reported in Fig. 4.13. An uncertainty of about 10% is associated to each concentration value reported in this Table.

	d_{geom} μm	d_{ae} μm	STRAS flow rate $l \ min^{-1}$	Si_{imp} $\mu g \ m^{-3}$	Si_{filter} $\mu g \ m^{-3}$	Collection efficiency %
2018	0.6	1.0	13.0	1.6	110.2	1.4
	0.8	1.3	13.4	1.4	86.1	1.6
	1.0	1.6	12.4	1.8	99.4	1.8
	1.5	2.4	14.4	11.4	38.6	23
	2.0	3.3	14.2	18.2	11.3	62
2019	0.2	0.3	16.7	0.103	68.3	0.15
	0.5	0.8	16.7	0.117	47.2	0.25
	0.6	1.0	16.7	0.186	38.5	0.48
	0.7	1.1	16.7	0.085	40.0	0.21
	0.7	1.1	16.7	0.116	38.3	0.30
	0.8	1.3	16.7	0.63	76.7	0.81
	0.8	1.3	16.7	0.55	72.7	0.75
	0.9	1.5	16.7	0.84	75.9	1.1
	0.9	1.5	16.7	0.66	67.8	0.96
	1.0	1.6	16.7	0.89	56.6	1.5
	1.0	1.6	16.7	1.24	62.1	2.0
	1.5	2.4	16.7	9.9	31.6	24
	1.5	2.4	16.7	7.6	26.3	22
	2.0	3.3	16.7	16.4	15.0	52
	2.0	3.3	16.7	9.9	12.6	44
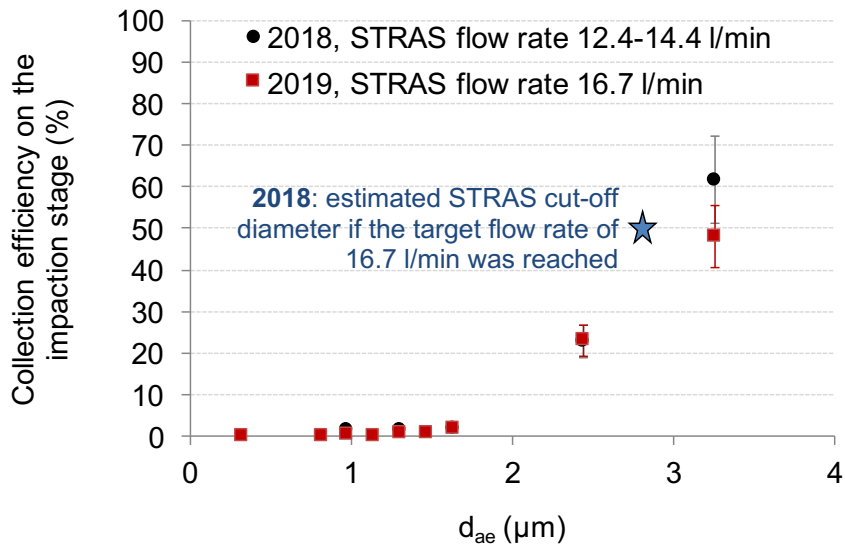
Figure 4.13: Preliminary collection efficiency of STRAS calculated with Eq. (4.2) for tests performed in 2018 and 2019. Error bars represent the propagation of uncertainties due to PIXE analysis on the impaction stage and the filter.

SEM-EDS analyses were performed only on samples collected during the tests in 2018 (analyses on samples collected during tests in 2019 will be carried out as soon as possible). These analyses pointed out problems likely related to the production phase of SiO_2 particles, such as Al contaminations (confirmed also by PIXE technique) and particles with notable differences as for size and composition respect to the ones declared by the producer (see e.g. Fig. 4.14). Anyway, information complementary to the one obtained by PIXE analysis could be derived in some cases. For example, the majority of SiO_2 particles with declared $d_{geo} = 1.0 \ \mu m$ ($d_{ae} = 1.6 \ \mu m$) were collected on the Nuclepore$^{\text{TM}}$ filter, in contrast to what happened for SiO_2 particles with declared $d_{geo} = 2.0 \ \mu m$ ($d_{ae} = 3.3 \ \mu m$) that were mainly present on the PP foil. Images with examples of measurements, and comparison between the Nuclepore$^{\text{TM}}$ filter and the PP foil are reported in Fig. 4.15 and Fig. 4.16, respectively.

As already said, these analyses are still preliminary and they highlighted the need to consider alternative methods for a complete characterisation of STRAS cut-off diameter, even if the latter was found to be qualitatively around the expected values of 2-3 μm.
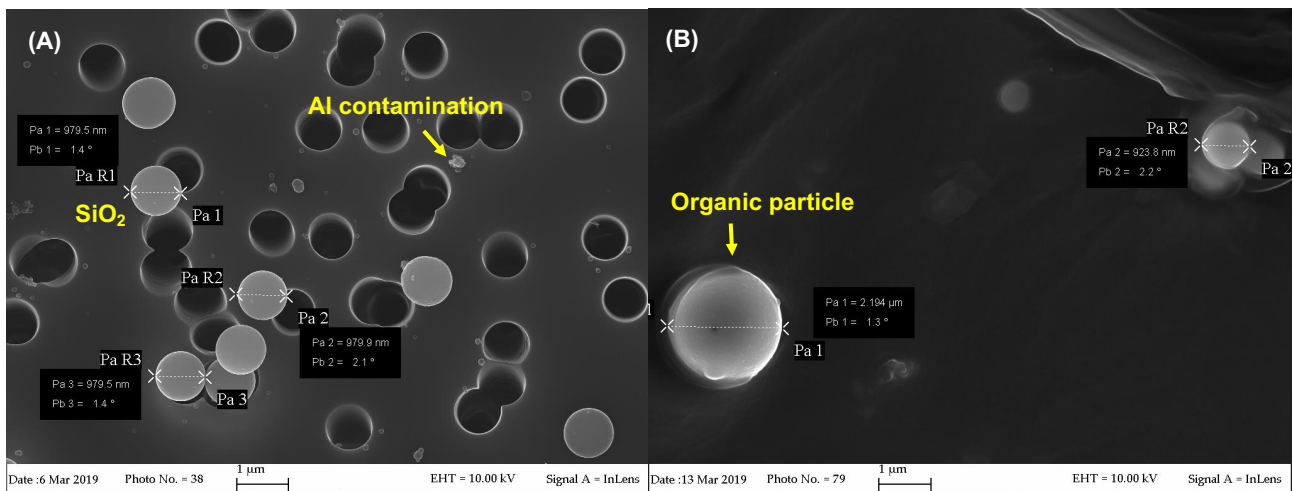
Figure 4.14: Example of images obtained by SEM-EDS analysis on (A) a Nuclepore$^{\text{TM}}$ filter and (B) a polypropilene foil. The sample in (A) corresponds to the collection of SiO_2 particles with declared $d_{geo} = 0.8 \ \mu m$ ($d_{ae} = 1.3 \ \mu m$); actual particle dimension is around 0.98 μm, and Al contamination is visible. The sample in (B) corresponds to the collection of SiO_2 particles with declared $d_{geo} = 2.5 \ \mu m$ ($d_{ae} = 4.1 \ \mu m$); particles have a d_{geo} of 2.2 μm and are constituted of organic material instead of Si.
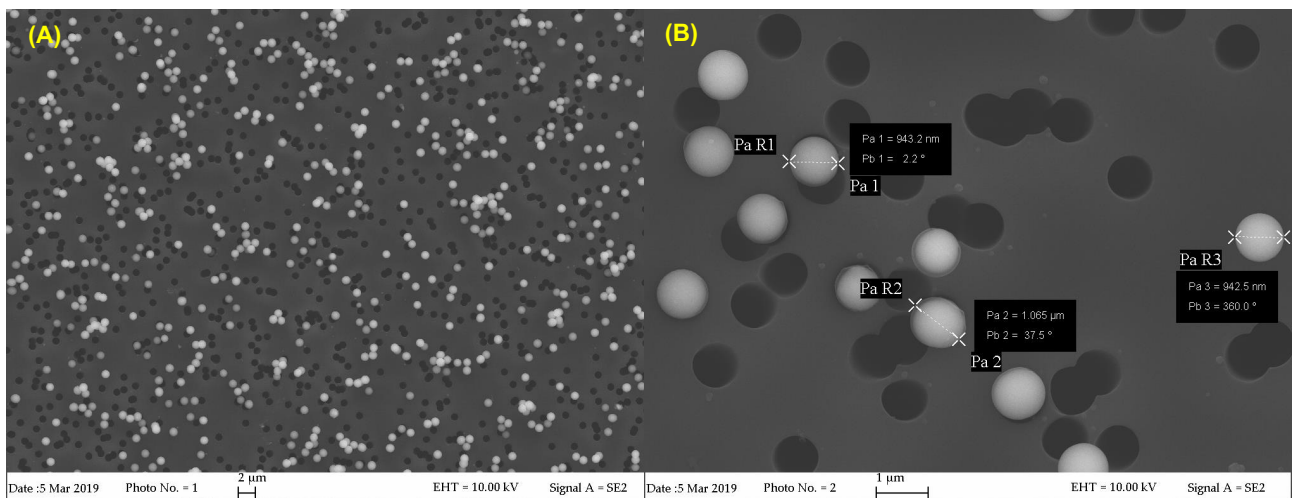
Figure 4.15: Example of images obtained by SEM-EDS analysis on a Nuclepore$^{\text{TM}}$ filter for the sample corresponding to the collection of SiO_2 particles with declared $d_{geo} = 1.0 \ \mu m$ ($d_{ae} = 1.6 \ \mu m$). (A) Particles appear to be collected homogeneously on the Nuclepore$^{\text{TM}}$ filter; (B) particles size is in agreement with the one declared by the producer.
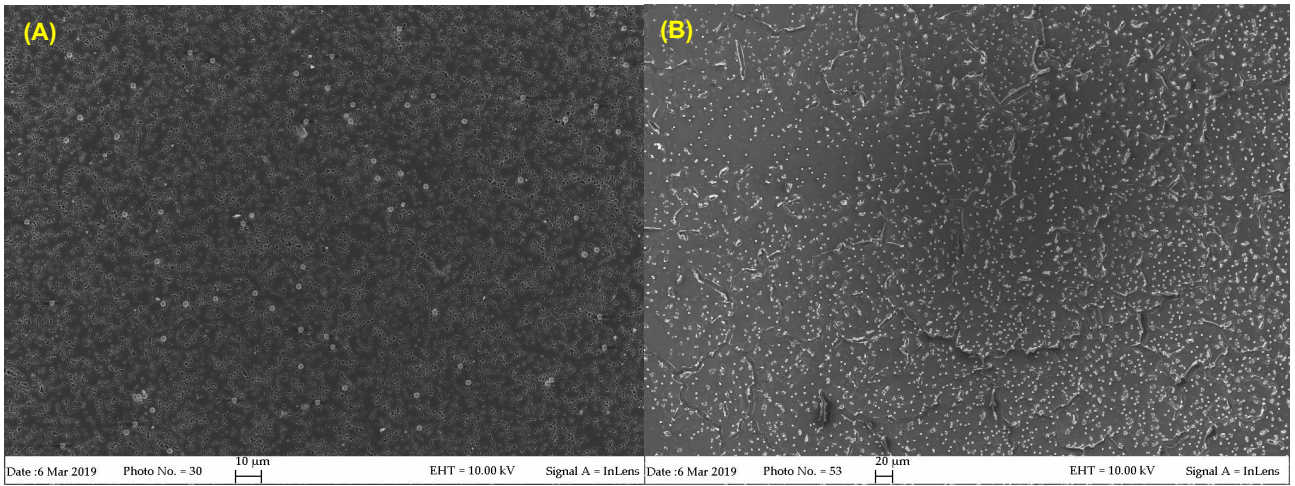
Figure 4.16: Example of images obtained by SEM-EDS analysis on (A) a Nuclepore$^{\text{TM}}$ filter and (B) a polypropilene foil, for the sample corresponding to the collection of SiO_2 particles with declared $d_{geo} = 2.0$ μm ($d_{ae} = 3.3$ μm). Note that different scales are used for the two images, and the presence of the Apiezon coating on the polypropilene foil.

Conclusions and perspectives

In this thesis, experimental and modelling approaches were developed and optimised in order to improve the knowledge on atmospheric aerosol and its emission sources. Atmospheric aerosol causes adverse effects on human health, impacts on environment, visibility, and climate; for these reasons, source identification and quantification is mandatory to establish mitigation strategies on atmospheric aerosol concentration.

This thesis work provides contribution to the state of the art of atmospheric aerosol science and research, through develpoment in both modelling and experimental methods. Major advancements addressed in this PhD thesis are reported and briefly described hereafter:

- In the last few years, part of the aerosol science community has focused on the study of optical properties of aerosol due to its crucial role in the Earth radiative budget and on visibility. Optical source apportionment methods (e.g. the Aethalometer model) have been developed and implemented in order to retrieve the contribution of different sources to the light absorption in the atmosphere; this piece of information can be very useful for formulating strategies for pollution abatement in order to improve air quality and face climate challenges. Assumptions at the basis of these modelling approaches, i.e. number of sources (two) and values for source-specific absorption Ångström exponent (α) fixed by the modeller, limit their applicability and cause large part of the uncertainty in the model results. On the other hand, at the state of the art receptor modelling approaches - which are increasingly used in source apportionment studies - are mostly based only on chemical composition data.
  In this PhD thesis, an original approach was proposed in order to exploit aerosol light absorption properties in receptor models. Aerosol chemical composition data collected during 2016 in Milan were coupled with the aerosol light absorption coefficient measured at different wavelengths, and used together as input to receptor modelling. This approach was successfully applied to experimental data with different time resolutions, exploiting the detailed chemical speciation at low time resolution and the temporal information given by high time resolution samples as input to the multi-time resolution model. Results on the light absorption characteristics of sources allowed to overcome the main issues related to the application of the Aethalometer model, leading to: (1) a more robust identification of sources themselves, which in this case are no more limited to only two sources (fossil fuels combustion and biomass burning); (2) estimation of source-related α values and mass absorption cross sections at different wavelengths without any a priori assumption. Besides the traditional source apportionment analysis, the impact of different sources on the aerosol light absorption in the atmosphere was thus assessed, providing hints for abatement measures to be implemented in the future. In particular, at the investigated site secondary compounds constituted the highest contribution in terms of PM10 mass (52% on average), while the two factors identified as biomass burning and traffic were found to be the most significant contributors to aerosol light absorption in the atmosphere.
  Furthermore, it is important to remark that the original approach here described can be used in source apportionment studies using any suitable dataset, not necessarily with multi-time resolution; an application to data collected with daily resolution is now under investigation in the frame of a collaboration with the Air Quality research group of the Department of Engineering and Nuclear Science of the Instituto Superior Técnico of Lisbon (Portugal).

- The assessment of the origin of secondary aerosol is an open issue for receptor modelling due to the underlying hypotheses of the models. Secondary aerosol comprises both inorganic (SIA) and organic (SOA) components; in particular, production mechanisms of SOA are not completely understood in comparison e.g. to the formation of secondary sulphate and nitrate (stemming from gaseous precursors SO_2

and NO_x). For Chemical Transport Models the estimation of organic aerosol (OA) and in particular of SOA is still challenging since dynamic processes in which low-volatility material evaporates, oxidises, and recondenses over time, are not fully understood yet thus limiting the modelling performances. In the last decades, advancements in aerosol mass spectrometry techniques have given new insights on the atmospheric OA fraction. Mass spectra have been exploited in receptor modelling and data collected at a large number of sites showed that the Oxygenated Organic Aerosol (OOA) factor, which is generally associated to secondary contributions in the literature, is dominant.

In this PhD work, a fully chemically characterised and high time resolved dataset was used to test an original approach to data analysis (results obtained by ME-2 application) aiming at relating primary and secondary OA contributions to their emission sources. Data were collected with high time resolution during February 2017 in Rome, in the frame of the CARE (Carbonaceous Aerosol in Rome and Environs) international project. A complete chemical characterisation to assess inorganic and organic aerosol contributions was carried out, and aerosol light absorption coefficients $b_{ap}(\lambda)$ at seven wavelenghts in the range 370-950 nm were retrieved by an Aethalometer AE33. Total OA concentration was measured by an ACSM (Aerosol Chemical Speciation Monitor), apportioned by our ME-2 receptor model, and afterwards compared with HOA (Hydrocarbon-like Organic Aerosol), BBOA (Biomass Burning-like Organic Aerosol), and OOA concentrations obtained as results from an independent source apportionment study previously performed. Additional information on OA contributions were thus obtained, e.g. indicating that (1) the contribution of OA assigned by ME-2 to the traffic exhaust source was fully comparable to HOA assessed by ACSM data analysis; (2) the OOA apportionment made on ACSM data likely includes a secondary contribution due to biomass burning emissions corresponding to 43% of total OA, that is dominant respect to the primary one (12%). It is important to remark that this result is of particular interest for the receptor modelling community, since identification of emission sources of secondary aerosol is still an open issue for this type of models.

In addition, the joint use of chemical and light absorption data allowed to perform a further test on the approach previously described in a case-study impacted by episodic sources, such as e.g. a transport of mineral dust that gave not negligible contribution to aerosol absorption in the atmosphere during a short time interval (impacting on $b_{ap}(\lambda)$ apportionment from 25% to 10%, decreasing with increasing wavelength).

- Routine sampling of atmospheric aerosol on filters is carried out with a time resolution of 24 hours, in order to collect enough material to perform full aerosol chemical characterisation. However aerosol emission, transformation, and removal processes are related to shorter time scales ($\sim$hour); reducing sampling duration was proved effective to improve the ability of receptor models to resolve sources, and allow the identification of episodic sources that might not negligibly impact on a short time interval. The availability of reliable and well characterised instrumentation has a crucial role for advancements in atmospheric aerosol science, and high time resolution measurements are needed to study in detail properties of atmospheric aerosol and its emission sources.

  In this PhD thesis, a contribution to the development and characterisation of the new high time resolution aerosol sampler STRAS (Size and Time Resolved Aerosol Sampler) was given in the frame of the INFN (National Institute of Nuclear Physics) experiment TRACCIA (Time Resolved Aerosol Characterization Challenging Improvements and Ambitions) in collaboration with other Italian research groups. Main efforts of this PhD thesis concerned contributions to sampler design and experimental sampler characterisation in the atmospheric simulation chamber ChAMBRe (Chamber for Aerosol Modelling and Bio-aerosol Research). In order to investigate STRAS cut-off diameter, SiO_2 particles with known size were injected in ChAMBRe and sampled; then, PIXE (Particle Induced X-ray Emission) and SEM-EDS (Scanning Electron Microscopy - Energy Dispersive X-Ray Spectroscopy) analyses were performed on collected samples to derive information about STRAS collection efficiency. Preliminary tests showed that the cut-off diameter is qualitatively around the expected range of 2-3 μm.

In perspective, main contributions from this PhD thesis pave the way to fully explore the potential of receptor models and to implement analyses able to integrate data which are usually considered separately. STRAS has the chance to become a sampler of interest for the aerosol science community, replacing the streaker sampler that has been used in many international campaigns.

Appendix A

A.1 Physical principles for size-segregated aerosol sampling

Equations for air motion can be derived from the application of the Newton's second law to an incompressible elementary air volume, and considering the continuity equation in order to ensure mass conservation. In this way, the Navier-Stokes equations for incompressible fluids (i.e. with $div \ \vec{u} = 0$) are obtained [164]:

$$\rho \frac{D\vec{u}}{Dt} = -grad \ p + \eta \nabla^2 \vec{u} \tag{A.1}$$

where ρ is the fluid density, $\vec{u}$ is the fluid velocity, p indicates the pressure, and η is the fluid viscosity. The mathematical D-operator represents the material derivative. In Eq. (A.1), inertial forces (left side of the equation) are balanced by pressure gradient forces and viscous shearing forces (first and second term on the right side, respectively).

Fluid dynamic problems are solved considering dimensionless quantities through the introduction of variables that are characteristic of the physical problem under investigation. For example, we can introduce the variables U and L, that are the characteristic velocity and length of the problem here of interest, respectively; in the study of the air flow in a nozzle, U and L represent the fluid velocity in the nozzle, and the nozzle dimension, respectively. We can define the dimensionless quantities:

$$\vec{u'} = \frac{\vec{u}}{U} \quad ; \quad p' = \frac{p}{\rho U^2} \quad ; \quad t' = \frac{tU}{L} \quad ; \quad \vec{x'} = \frac{\vec{x}}{L} \tag{A.2}$$

to be substituted in Eq. (A.1):

$$\frac{D\vec{u'}}{Dt'} = -grad \ p' + \frac{1}{Re} \nabla'^2 \vec{u'} \tag{A.3}$$

where Re represents the dimensionless Reynolds number, defined as:

$$Re = \frac{\rho U L}{\eta} \tag{A.4}$$

Re is one of the fundamental dimensionless numbers in fluid dynamics; it allows the evaluation of the flow regime, since it is related to the ratio between inertial and viscous forces:

$$Re \equiv \frac{inertial forces}{viscous forces} \propto \frac{\rho u \frac{\partial u}{\partial x}}{\eta \frac{\partial^2 u}{\partial x^2}} \propto \frac{\rho \frac{U^2}{L}}{\eta \frac{U}{L^2}} = \frac{\rho U L}{\eta} \tag{A.5}$$

Besides Re of the fluid flow, the Reynolds number can be also defined for the motion of a particle in a fluid (see Fig. A.1): $Re_p = \frac{\rho V d}{\eta}$, where ρ is the fluid density, V is the relative velocity between the particle and the fluid, d is the particle diameter, and η is the fluid viscosity. When Re_p is lower than 1, the Stokes regime yields; motion of atmospheric particles in the air is generally characterised by small Re_p (see Tab. A.1), because of small particle dimensions and low velocities [3] [5].

Table A.1: Reynolds number for spherical particles falling in the air at their terminal velocities at T=298 K and p=1 atm. Table adapted from [5].

Diameter $[\mu m]$	Re_p
0.1	$7 \cdot 10^{-9}$
1	$2.8 \cdot 10^{-6}$
10	$2.5 \cdot 10^{-3}$
20	0.02
60	0.4
100	2
300	20
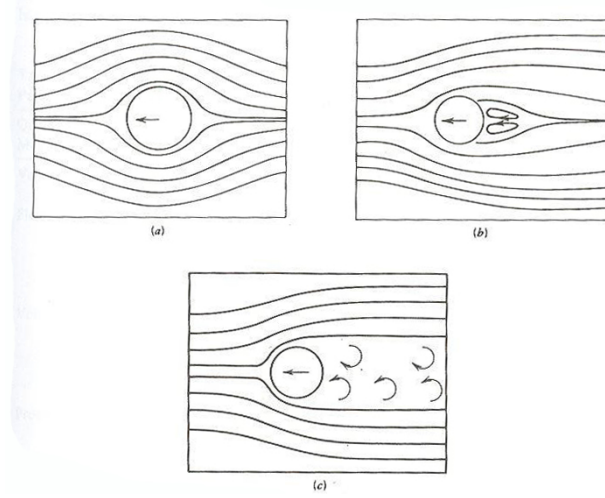
Figure A.1: Basic scheme of the flow around a sphere for different particle Reynolds numbers. (a) $Re_p = 0.1$, laminar flow; (b) $Re_p \simeq 2$, turbolent flow; (c) $Re_p \simeq 250$, turbolent flow. Figure from [3].

In the Stokes regime, the drag force on the particle can be approximated by the Stokes law:

$$F_{drag} = 3\pi\eta V d \tag{A.6}$$

where η is the fluid viscosity, V is the relative velocity between the particle and the fluid, and d is the particle diameter. Equation (A.6) is obtained solving the Navier-Stokes equations assuming that inertial forces are negligible compared to viscous forces; at $Re_p = 1$, this assumption leads to a drag force predicted by the Stokes law that is 13% lower than the real one. To account for terms related to viscosity, an empirical drag coefficient can be taken into account in the equation (see [3] [5] for more details). Other approximations present in Eq. (A.6) are that the fluid is incompressible, the particle is a rigid sphere, no other particles or obstacles are present nearby, and the velocity of the fluid on the particle surface is zero. The latter assumption is not satisfied by small particles with dimensions (d) comparable to the mean free path (λ) of the molecules of the fluid (Fig. A.2); e.g. for air at T=298 K and p=1 atm, $\lambda \simeq 0.065~\mu m$ [5] . The fundamental dimensionless number relating particle dimensions and fluid mean free path is the Knudsen number:

$$Kn = \frac{2\lambda}{d} \tag{A.7}$$

The Stokes law in Eq. (A.6) is valid in the limit $Kn \to 0$, and it is not applicable to high Kn values; in the latter case, the real drag force is smaller than the one predicted by the Stokes law. To account for this effect, the slip correction Cunningham factor $C_c(Kn)$ is introduced:

$$F_{drag} = \frac{3\pi\eta V d}{C_c(Kn)} \tag{A.8}$$

$C_c(Kn)$ can be estimated through an empirical equation obtained by experimental measurements of slip [3]:

$$C_c = 1 + \frac{Kn}{2}[2.34 + 1.05 \cdot exp(-0.39\frac{2}{Kn})] \tag{A.9}$$
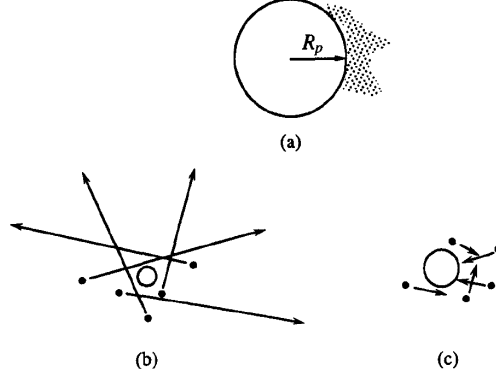
Figure A.2: Basic scheme of the interaction between the particle and the fluid in which it is suspended. (a) $Kn \to 0$, continuum regime; (b) $Kn \to \infty$, free molecule (kinetic) regime; (c) $Kn \simeq 1$, Transition regime. Figure from [5].

Examples of slip correction factor calculations are reported in Table A.2.

Table A.2: Slip correction factor C_c for spherical particles in air at T=298 K and p=1 atm. Table adapted from [5].

Diameter [μm]	C_c
0.001	216
0.002	108
0.005	43.6
0.01	22.2
0.02	11.4
0.05	4.95
0.1	2.85
0.2	1.865
0.5	1.326
1.0	1.164
2.0	1.082
5.0	1.032
10.0	1.016
20.0	1.008
50.0	1.003
100.0	1.0016

The drag force has to be considered in the equation for the motion of a particle in a fluid; in three dimensions:

$$m\frac{d\vec{V}(t)}{dt} = \vec{F_{ext}} + \vec{F_{drag}} \tag{A.10}$$

where $\vec{V}(t)$ is the three-dimensional relative velocity between the particle and the fluid as a function of time, m is the particle mass, and $\vec{F_{ext}}$ is the resulting from external forces. If the only external force is the gravity force $\vec{F_g}$ and we assume that it has the same direction of $\vec{F_{drag}}$, the problem can be treated as an unidimensional one. In this case, Eq. (A.10) in the Stokes regime can be written as:

$$m\frac{dV}{dt} = F_g - F_{drag} = mg - \frac{3\pi\eta V d}{C_c} \tag{A.11}$$

where g is the acceleration of gravity. Dividing Eq. (A.11) by m, that can be expressed as the product of particle density ρ_p and volume ($m = \rho_p \cdot \frac{4}{3}\pi(\frac{d}{2})^3$), the equation becomes:

$$\frac{dV}{dt} + \frac{V(t)}{\tau} - g = 0 \tag{A.12}$$

83

where $\tau = \frac{\rho_p d^2 C_c}{18\eta}$ (or $\tau = \frac{\rho_0 d_{ae}^2 C_c}{18\eta}$ if the aerodynamic diameter d_{ae} is considered in place of d) has the units of a time. Solution to Eq. (A.12) (with the initial condition $V(0) = 0$) is:

$$V(t) = \tau g [1 - exp(-\frac{t}{\tau})] \tag{A.13}$$

For $t \to \infty$ particle velocity reaches its limit value, called terminal velocity; τ is called relaxation time. From an experimental point of view, particles reach their terminal velocities when $t \simeq 3\tau$ as it corresponds to 95% of its velocity limit value. This time interval is very short: at T=298 K and p=1 atm, it is less than 1 ms for particles with an aerodynamic diameter $d_{ae} < 10 \ \mu m$.

Equation (A.10) can be also solved considering $\vec{F_{ext}}$ as negligible, and $V(0) = V_0$. The solution in this case is:

$$V(t) = V_0 \ exp(-\frac{t}{\tau}) \tag{A.14}$$

and can be integrated over time to obtain $x(t)$; the limit of $x(t)$ for $t \to \infty$ gives the stopping distance $S = V_0 \ \tau$. This quantity is important to study the motion of particles in curved flow streamlines, e.g. when the flow encounters an obstacle. After calculation of the fluid flow near the obstacle, particle trajectories can be determined (it is noteworthy that only very simple geometries can be resolved analitically). To study this type of problems, the dimensionless Stokes number is introduced: it is defined as the ratio between the stopping distance S and a dimension that is characteristic of the obstacle D_{obs}:

$$Stk = \frac{S}{D_{obs}} = \frac{V_0 \tau}{D_{obs}} = \frac{V_0 \rho_0 d_{ae}^2 C_c}{18\eta D_{obs}} \tag{A.15}$$

As can be seen from Eq. (A.15), Stk represents the ratio of the particle relaxation time over the time $\frac{D_{obs}}{V_0}$ necessary to overcome the obstacle. Particles characterised by $Stk >> 1$ will continue their rectilinear motion; opposite, if $Stk << 1$ the particle will follow flow streamlines (Fig. A.3).

The Stokes number is a fundamental parameter to characterise aerosol samplers that exploit the phenomenon of inertial impact to collect particles; these samplers can be also called inertial classifiers or impactors.
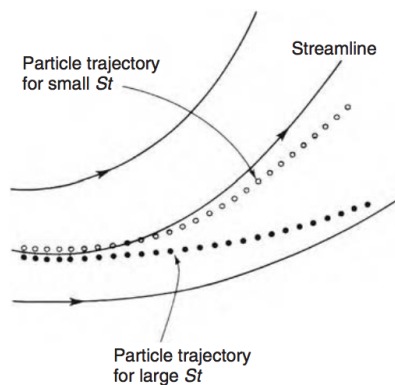
Figure A.3: Basic scheme of the effect of the Stokes number (Stk) on particle trajectories in the fluid flow. Figure from [164].

A.2   Inertial classifiers

The basic scheme of inertial classifiers (or impactors, in the following) consists in a nozzle to accelerate particles, and in an impaction stage on which - theoretically - all particles having an aerodynamic diameter larger than a threshold value are collected. Multistage impactors are constituted by more than one impaction stage in sequence.

From an experimental point of view, the typical dependance of particle collection efficiency on particle aerodynamic diameter (d_{ae}) is shown in Fig. A.4; sometimes, $\sqrt{Stk}$ is equivalently used instead of d_{ae} (since they are related according to Eq. (A.15)). For the characterisation of inertial classifiers, it is important to define the characteristic dimension D_{obs} (see Eq. (A.15)), that is related in this case to nozzle characteristics. Nozzles can
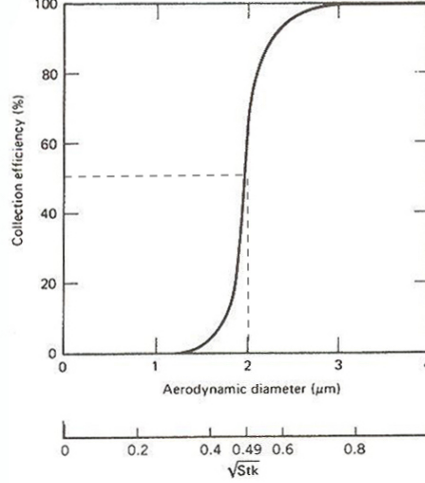
Figure A.4: Typical particles collection efficiency curve. Figure adapted from [3].

be circular or rectangular; D_{obs} is represented by the nozzle radius and the nozzle halfwidth for round impactors and rectangular impactors, respectively [158]. Equation (A.15) can be written as:

$$Stk = \frac{U\rho_0 d_{ae}^2 C_c}{9\eta W} \tag{A.1}$$

where $W = 2 \cdot D_{obs}$; U is the mean air velocity at the exit of the nozzle, corresponding to $\frac{Q}{\pi(\frac{W}{2})^2}$ for round impactors, and to $\frac{Q}{LW}$ for rectangular impactors where Q is the volumetric flow through the nozzle and L is the second dimension (length) of the rectangular impactor. η indicates the fluid viscosity, C_c the slip correction factor (see the empirical definition in Eq. (A.9)), d_{ae} the particle aerodynamic diameter, and ρ_0 the particle unit density.

d_{50} (also called ECD, i.e. effective cut-off diameter) is the impactor cut-off diameter, defined as the value of d_{ae} for which particles are collected with an efficiency of 50% on the impaction stage (e.g. $d_{50} = 2~\mu m$ in Fig. A.4). The critical parameter $\sqrt{Stk_{50}}$ is the value of $\sqrt{Stk}$ corresponding to d_{50} (e.g. $\sqrt{Stk_{50}} = 0.49$ in Fig. A.4); for impactors meeting recommended design criteria (see afterwards for more details), literature values for this parameter are 0.49 and 0.77 for round and rectangular impactors, respectively [3]. Starting from Eq. (A.1), the relation between d_{50} and $\sqrt{Stk_{50}}$ can be made explicit:

$$d_{50} = \sqrt{\frac{9\eta W}{U\rho_0 C_c}}\sqrt{Stk_{50}} \tag{A.2}$$

Inertial impactors have been extensively studied from both a theoretical and an experimental point of view, with comparable results [158]. Construction parameters of impactors must satisfy specific design criteria, obtained in the literature through numerical solutions of the Navier-Stokes equations for the determination of the flow field, and subsequent numerical integration of particle trajectories. The use of theoretical guidelines has led to the construction of impactors that present experimental collection efficiency curves very close to ideality. Design criteria are discussed in the following, and resumed in Table A.3. The most important parameters for the sizing of inertial classifiers are (see Fig. A.5):

- the ratio between the nozzle throat length (T) and the nozzle dimension (W): $\frac{T}{W}$;

- the ratio between the jet-to-plate distance (S) and the nozzle dimension (W): $\frac{S}{W}$;

- the Reynolds number of the flow Re_{jet}, defined as $\frac{\rho U W}{\eta}$ for round impactors and $\frac{2\rho U W}{\eta}$ for rectangular impactors (with ρ representing the air density) [159].

Moreover, for rectangular impactors:

- the ratio between the length of the nozzle (L) and the nozzle width (W): $\frac{L}{W}$

85
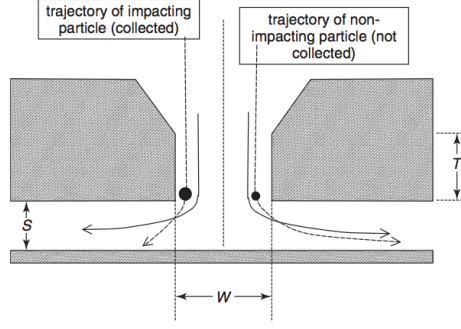
Figure A.5: Crucial dimensional parameters of inertial classifiers. Figure from [164].

$\frac{T}{W}$ ratio has a small influence on cut-off characteristics, especially when the impactor presents a tapered or conical inlet section, as shown in Fig. A.5. In this case, particles have enough time to adjust their motion to the fluid velocity in the nozzle throat, and particle losses due to sharp inlet entrance can be minimised. Moreover, $\frac{T}{W}$ ratio must be small enough in order to avoid the development of a parabolic velocity profile in the flow at the exit of the nozzle. The general criterion is $\frac{T}{W} \geq 1$ [159] and $\frac{T}{W} < 5$; in the literature, the conservative suggested range is $1 < \frac{T}{W} < 2$ [164].

Small variations in the $\frac{S}{W}$ ratio might impact on the $\sqrt{Stk_{50}}$ value, as can be noted from Fig. A.6. $\sqrt{Stk_{50}}$ values are quite constant for $\frac{S}{W} > 0.5$ and $\frac{S}{W} > 1$ for round and rectangular impactors, respectively; for values outside this range, small variations in the $\frac{S}{W}$ ratio cause significant variations in $\sqrt{Stk_{50}}$ (and consequently in d_{50}, see Eq. (A.2)). Therefore, suggested conservative values are $\frac{S}{W} > 1$ and $\frac{S}{W} > 1.5$ for round and rectangular impactors, respectively. The upper limit of the $\frac{S}{W}$ ratio is not clear, but values as high as 10 can be found in the literature.

The value of Re_{jet} has an influence especially in the sharpness of the collection efficiency curve (Fig. A.7). In order to have negligible impact on $\sqrt{Stk_{50}}$ values, the range suggested in the literature for Re_{jet} is 500-3000 for both round and rectangular impactors.

Finally, for rectangular impactors $\frac{L}{W} \geq 10$ is suggested [164].

Table A.3: Literature guidelines on construction parameters for impactor design.

Impactor	S/W	T/W	L/W	Re_{jet}
round	>1	1-2		300-5000
rectangular	>1.5	1-2	≥ 10	300-5000
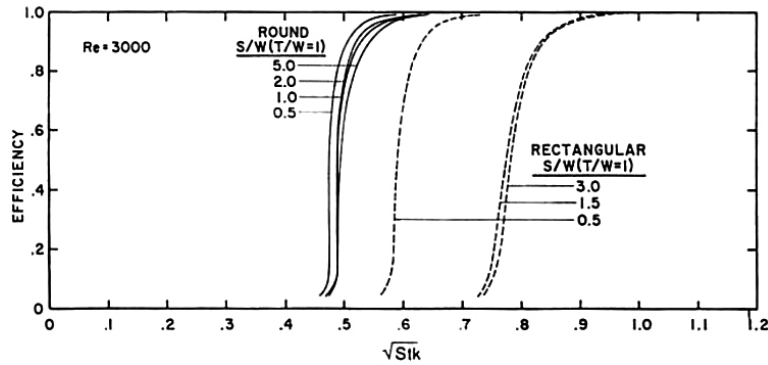
Figure A.6: Example of theoretical impactor efficiency curves for rectangular and round impactors showing the effect of S/W ratio. Figure from [158].
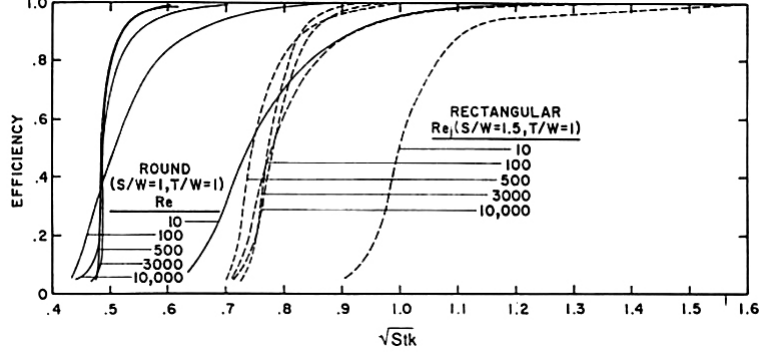
Figure A.7: Example of theoretical impactor efficiency curves for rectangular and round impactors showing the effect of the parameter Re_{jet}. Figure from [158].

A.3 Physical principles of aerosol filtration: collection efficiency

Streaker sampler and STRAS exploit aerosol filtration to collect particles that do not previously deposit on an impaction stage. Filter types can be divided in depth filters (e.g. quartz fibre filters) and membrane filters such as Nuclepore$^{\text{TM}}$ (see e.g. [170] for more detailed information about filters); the latters are considered in this work.

The theory of particle collection efficiency of Nuclepore$^{\text{TM}}$ filters was developed in the 1960-1970s ([160] [161] [171] [172] and references therein). This type of filters can be parametrised through a physical model for which filters are constituted by parallel capillaries; in this way, filter parameters (pore size, filter thickness, and porosity i.e. ratio of open space over the total filter volume), particle parameters (size and density), and filtration conditions (e.g. the air face velocity on the filter) can be used for the calculation of particle collection efficiency. Particle collection by Nuclepore$^{\text{TM}}$ filters can be described as the combination of different processes: (1) inertial impact on the filter surface, (2) interception at the pore opening, (3) Brownian diffusion to the pore walls, and (4) Brownian diffusion to the filter surface [162]. For very low Reynolds number (this is not the case of STRAS configuration), impaction and interception can be modelled together [171]. In the following, all equations are based on the more recent work of Ogura et al. [163], where slight modifications to the 1960-1970s theory were introduced in order to take into account the Cunningham slip correction factor (C_c, see Eq. (A.9)).

The collection efficiency due to impaction (E_I) can be calculated as:

$$E_I = \frac{2\epsilon_i}{1+\zeta} - (\frac{\epsilon_i}{1+\zeta})^2, \quad where \quad \epsilon_i = 2Stk\sqrt{\zeta} + 2Stk^2\zeta exp(-\frac{1}{Stk\sqrt{\zeta}}) - 2Stk^2\zeta \tag{A.1}$$

$$\zeta = \frac{\sqrt{P}}{1-\sqrt{P}}$$

$$Stk = \frac{2\rho_p r_p^2 v_f C_c}{9\eta r_0}$$

Stk is the Stokes number, P is the filter porosity, ρ_p is the particle density, r_p is the particle radius, v_f is the air face velocity on the filter, η is the fluid viscosity, and r_0 is the pore radius (which is considered as the characteristic dimension for the problem under investigation).

The collection efficiency due to interception (E_R) can be obtained as:

$$E_R = \frac{4N_R^2}{1+4N_G}(1 + 2\frac{N_G}{N_R}), \quad with \quad N_R = N_r(1 - \frac{N_r}{2}) \tag{A.2}$$

$$N_G = N_g(1 + \frac{N_g}{2})$$

$$N_r \leq 1$$

where $N_r = r_p/r_0$, N_g is the slip parameter defined as $N_g = 1.126 \cdot \lambda/r_0$, where λ is the air mean free path. When $N_r > 1$, E_R is equal to 1.

The collection efficiency due to diffusion to the pore walls (E_W) can be calculated as:

$$E_W = 2.56 N_D^{2/3} - 1.2 N_D - 0.177 N_D^{4/3}, \qquad\qquad with \quad N_D < 0.01, \quad or \qquad (A.3)$$

$$E_W = 1 - 0.81904 \cdot exp(-3.6568 N_D) - 0.09752 \cdot exp(-22.3045 N_D)$$
$$- 0.03248 \cdot exp(-56.95 N_D) - 0.0157 \cdot exp(-107.6 N_D), \qquad with \quad N_D > 0.01$$

N_D is the coefficient of diffusive collection $N_D = \frac{L D_p P}{r_0^2 v_f}$, where L is the filter thickness, D_p is the particle diffusion coefficient, and the other quantities have been already defined above.

The collection efficiency due to diffusion to the filter surface (E_{DS}) can be retrieved as:

$$E_{DS} = 1 - exp\{ \frac{-\alpha_1 D_n^{2/3}}{1 + (\alpha_1/\alpha_2) D_n^{7/15}} \}, \quad with \quad \alpha_1 = 4.57 - 6.46 P + 4.59 P^2 \qquad (A.4)$$

$$\alpha_2 = 4.5$$

D_n is the normalised particle diffusion coefficient $\frac{D_p}{D_0 v_f}$, where D_0 is the hypothetical radius of the cylindrical flow approaching a pore that is introduced for modelling purposes (see Fig. A.8); D_0 is related to the filter porosity since $P = (\frac{\pi r_0^2}{\pi D_0^2})$. The fitting parameters α_1 and α_2 in Eq. (A.4) are valid within the limits $0.05 \leq P \leq 0.64$ (that is the case of Nuclepore^{TM} filters that will be likely used in STRAS).

The overall collection efficiency is:

$$E_O = 1 - (1 - E_I)(1 - E_R)(1 - E_W)(1 - E_{DS}) \qquad (A.5)$$
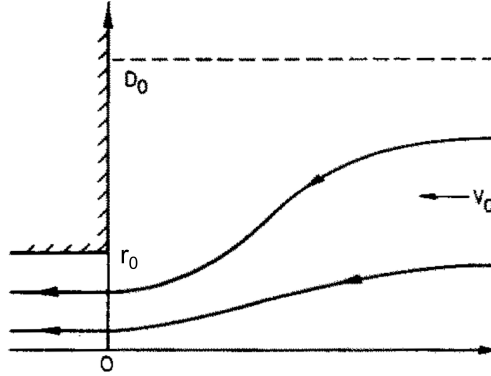
Figure A.8: Basic scheme of the axisymmetric cylindrical flow approaching a pore of Nuclepore^{TM} filter. D_0 represents the flow radius, r_0 the pore radius, and V_0 the face velocity. Figure modified from [163].

Bibliography

[1] Fuzzi S., Baltensperger U., Carslaw K., Decesari S., Denier van der Gon H., Facchini M.C., Fowler D., Koren I., Langford B., Lohmann U., Nemitz E., Pandis S., Riipinen I., Rudich Y., Schaap M., Slowik J.G., Spracklen D.V., Vignati E., Wild M., Williams M., Gilardoni S.: *Particulate matter, air quality and climate: lessons learned and future needs*, Atmospheric Chemistry and Physics, 15, 8217-8299, **2015**, doi: 10.5194/acp-15-8217-2015.

[2] Pöschl U.: *Atmospheric Aerosols: Composition, Transformation, Climate and Health Effects*, Angewandte Chemie International Edition, 44, 7520 – 7540, **2005**, doi: 10.1002/anie.200501122.

[3] Hinds W.C.: *Aerosol technology: properties, behavior, and measurement of airborne particles*, John Wiley and Sons, **1999**.

[4] Colbeck I., and Lazaridis M.: *Aerosol Science: Technology and Applications*, John Wiley & Sons, Chichester, West Sussex, United Kingdom, **2014**.

[5] Seinfeld J. H. and Pandis S. N.: *Atmospheric chemistry and physics: from air pollution to climate change*, 2nd edition, John Wiley & Sons, INC, Hoboken, New Jersey, **2006**.

[6] U.S. EPA. Integrated Science Assessment (ISA) For Particulate Matter (Final Report, Dec 2009). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/139F, **2009**.

[7] Calvo A.I., Alves C., Castro A., Pont V., Vicente M., and Fraile R.: *Research on aerosol sources and chemical composition: Past, current and emerging issues*, Atmospheric Research, 120-121, 1-28, **2013**, doi: 10.1016/j.atmosres.2012.09.021.

[8] Knippertz P., and Stuut J.-B.W.: *Mineral dust: A Key Player in the Earth System*, Springer, **2014**.

[9] Bohren C.F., and Huffman D.R.: *Absorption and Scattering of Light by Small Particles*, Wiley, **1983**.

[10] Moosmüller H., Chakrabarty R.K., and Arnott W.P.: *Aerosol light absorption and its measurement: A review*, Journal of Quantitative Spectroscopy & Radiative Transfer, 110, 844–878, **2009**, doi: 10.1016/j.jqsrt.2009.02.035.

[11] Horvath H.: *Size segregated light absorption coefficient of the atmospheric aerosol*, Atmospheric Environment, 29, 8, 875-883, **1995**.

[12] Caponi L., Formenti P., Massabò D., Di Biagio C., Cazaunau M., Pangui E., Chevaillier S., Landrot G., Andreae M. O., Kandler K., Piketh S., Saeed T., Seibert D., Williams E., Balkanski Y., Prati P., and Doussin J.-F.: *Spectral- and size-resolved mass absorption efficiency of mineral dust aerosols in the shortwave spectrum: a simulation chamber study*, Atmospheric Chemistry and Physics, 17, 7175–7191, **2017**, doi: 10.5194/acp-17-7175-2017.

[13] Belis C.A., Favez O., Mircea M., Diapouli E., Manousakas M.-I., Vratolis S., Gilardoni S., Paglione M., Decesari S., Mocnik G., Mooibroek D., Salvador P., Takahama S., Vecchi R., and Paatero P.: *European guide on air pollution source apportionment with receptor models - Revised version 2019*, Publications Office of the European Union, Luxembourg, **2019**, doi: 10.2760/439106.

[14] Mircea M., Calori G., Pirovano G., and Belis C.A.: *European guide on air pollution source apportionment for particulate matter with source oriented models and their combined use with receptor models*, Publications Office of the European Union, Luxembourg, **2020**, doi: 10.2760/470628

[15] Watson J.G., Zhu T., Chow J.C., Engelbrecht J., Fujita E.M., and Wilson W.E.: *Receptor modeling application framework for particle source apportionment*, Chemosphere, 49, 1093-1136, **2002**, doi: 10.1016/S0045-6535(02)00243-6.

[16] Viana M., Kuhlbusch T.A.J., Querol X., Alastuey A., Harrison R.M., Hopke P.K., Winiwarter W., Vallius M., Szidat S., Prévôt A.S.H., Hueglin C., Bloemen H., Wåhlin P., Vecchi R., Miranda A.I., Kasper-Giebl A., Maenhaut W., and Hitzenberger R.: *Source apportionment of particulate matter in Europe: A review of methouds and results*, Aerosol Science, 39, 827-849, **2008**, doi: 10.1016/j.jaerosci.2008.05.007.

[17] Schauer J.J., Lough G.C., Shafer M.M., Christensen W.F., Arndt M.F., DeMinter J.T., Park J.-S., *Characterization of Metals Emitted from Motor Vehicles*, Research Report Health Effect Institute, 133, 1-76, **2006**.

[18] Hopke P.K.: *Review of receptor modeling methods for source apportionment*, Journal of the Air and Waste Management Association, 66:3, 237-259, **2016**, doi: 10.1080/10962247.2016.1140693.

[19] Paatero P.: *The Multilinear Engine: A Table-Driven, Least Squares Program for Solving Multilinear Problems, including the n-way Parallel Factor Analysis Model*, Journal of Computational and Graphical Statistics, 8, 4, 854-888, **1999**, doi: 10.2307/1390831.

[20] Paatero P., and Tapper U.: *Analysis of different modes of factor analysis as least squares fit problems*, Chemometrics and Intelligent Laboratory Systems, 18, 183-194, **1993**, doi: 10.1016/0169-7439(93)80055-M.

[21] Paatero P., and Tapper U.: *Positive Matrix Factorization: a non-negative factor model with optimal utilization of error estimates of data values*, Environmetrics, 5, 111–126, **1994**, doi: 10.1002/(ISSN)1099-095X.

[22] Paatero P.: *Least square formulation of robust non-negative factor analysis*, Chemometrics and Intelligent Laboratory Systems, 37, 23–35, **1997**, doi: 10.1016/S0169-7439(96)00044-5.

[23] Amato F., Pandolfi M., Escrig A., Querol X., Alastuey A., Pey J., Perez N., and Hopke P.K.: *Quantifying road dust resuspension in urban environment by Multilinear Engine: A comparison with PMF2*, Atmospheric Environment 43, 2770-2780, **2009**, doi: 10.1016/j.atmosenv.2009.02.039

[24] Paatero P.: *User's guide for the Multilinear Engine program "ME2" for fitting multilinear and quasi-multilinear models*, University of Helsinki, Department of Physics, Finland, **2000**.

[25] Paatero P., and Hopke P.K.: *Rotational Tools for Factor Analytic Models*, Journal of Chemometrics, 23, 91-100, **2009**, doi: 10.1002/cem.1197

[26] Paatero, P.: *The Multilinear Engine (ME-2) script language (v. 1.352)*, available with the program ME-2 (me2scrip.txt), **2012**.

[27] Norris G., Duvall R., Brown S., and Bai S.: *EPA Positive Matrix Factorization (PMF) 5.0. Fundamentals and User Guide*, U.S. Environmental Protection Agency, Washington, DC, **2014**.

[28] Paatero P., Hopke P.K., Song X.-H., and Ramadan Z.: *Understanding and controlling rotations in factor analytic problems*, Chemometrics and Intelligent Laboratory Systems, 60, 253– 264, **2002**, doi: 10.1016/S0169-7439(01)00200-3.

[29] Paatero P., Eberly S., Brown S.G., and Norris G.A.: *Methods for estimating unvertainty in factor analytic solutions*, Atmospheric Measurement Techniques, 7, 781-797, **2014**, doi: 10.5194/amt-7-781-2014.

[30] Brown S.G., Eberly S., Paatero P., and Norris G.A.: *Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results*, Science of the Total Environment, 518-519, 626-635, **2015**, doi: 10.1016/j.scitotenv.2015.01.022.

[31] Forello A.C., Bernardoni V., Calzolai G., Lucarelli F., Massabò D., Nava S., Pileci R.E., Prati P., Valentini S., Valli G., and Vecchi R.: *Exploiting multi-wavelength aerosol absorption coefficients in a multi-time resolution source apportionment study to retrieve source-dependent absorption parameters*, Atmospheric Chemistry and Physics, 19, 11235–11252, **2019**, doi: 10.5194/acp-19-11235-2019.

[32] Reff A., Eberly S.I., and Bhave P.V.: *Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods*, Journal of the Air & Waste Management Association, 57, 2, 146-154, **2007**, doi: 10.1080/10473289.2007.10465319.

[33] Amato F., Alastuey A., Karanasiou A., Lucarelli F., Nava S., Calzolai G., Severi M., Becagli S., Gianelle V.L., Colombi C., Alves C., Custódio D., Nunes T., Cerqueira M., Pio C., Eleftheriadis K., Diapouli E., Reche C., Minguillón M., Manousakas M.-I., Maggos T., Vratolis S., Harrison R.M., and Querol X.: *AIRUSE-LIFE+: a harmonized PM speciation and source apportionment in five southern European cities*, Atmospheric Chemistry and Physics, 16, 3289-3309, **2016**, doi: 10.5194/acp-16-3289-2016.

[34] Kim E., Hopke P.K., Kenski D.M. and Koerber M.: *Sources of Fine Particles in a Rural Midwestern U.S. Area*, Environmental Science and Technology, 39, 4953-4960, textbf2005, doi: 10.1021/es0490774.

[35] Paatero P.: *User's guide for positive matrix factorization programs PMF2 and PMF3, part 1: Tutorial*, available at: https://www.helsinki.fi/~paatero/PMF/pmf2.zip (last access: 21 June 2018), **2015**.

[36] Vecchi R., Bernardoni V., Cricchio D., D'Alessandro A., Fermo P., Lucarelli F., Nava S., Piazzalunga A., and Valli G.: *The impact of fireworks on airborne particles*, Atmospheric Environment, 42, 6, 1121-1132, **2008**, doi: 10.1016/j.atmosenv.2007.10.047.

[37] Paatero P., and Hopke P.K.: *Discarding or downweighting high-noise variables in factor analytic models*, Analytica Chimica Acta, 490, 277-289, **2003**, doi: 10.1016/S0003-2670(02)01643-4.

[38] Van Grieken R.E., and Markowicz A.A.: *Handbook of X-Ray Spectrometry: Methods and Techniques*, Marcel Dekker Inc., New York, **1993**.

[39] Zhou L., Hopke P. K., Paatero P., Ondov J. M., Pancras J. P., Pekney N. J., and Davidson C. I.: *Advanced factor analysis for multiple time resolution aerosol composition data*, Atmospheric Environment, 38, 4909–4920, **2004**, doi: 10.1016/j.atmosenv.2004.05.040.

[40] Polissar A.V., Hopke P.K., Paatero P., Malm W.C., Sisler J.F.: *Atmospheric aerosol over Alaska 2. Elemental composition and sources*, Journal of Geophysical Research, 103, D15, 19045-19057, **1998**, doi: 10.1029/98JD01212.

[41] U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards: *Air Toxics Data Analysis Workbook*, Research Triangle Park, North Carolina, available at: https://www3.epa.gov/ttnamti1/files/ambient/airtox/workbook/AirToxicsWorkbook6-09.pdf (last access: 03 April 2020), **2009**.

[42] Paatero P.: *End User's Guide to Multilinear Engine Applications*, available with the program ME-2 (ME2_EndUsrGuid.pdf), **2016**.

[43] Paatero P.: *User's guide for positive matrix factorization programs PMF2 and PMF3, Part 2: reference*, available at: https://www.helsinki.fi/~paatero/PMF/pmf2.zip (last access: 21 June 2018), **2010**.

[44] Lee E., Chan C.K., and Paatero P.: *Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong*, Atmospheric Environment, 33, 3201-3212, **1999**, doi: 10.1016/S1352-2310(99)00113-2.

[45] Simon H., Beck L., Bhave P.V., Divita F., Hsu Y., Luecken D., Mobley J.D., Pouliot G.A., Reff A., Sarwar G., and Strum M.: *The development and uses of EPA's SPECIATE database*, Atmospheric Pollution Research, 1, 196-206, **2010**, doi: 10.5094/APR.2010.026.

[46] Bray C.D., Strum M., Simon H., Riddick L., Kosusko M., Menetrez M., Hays M.D., and Rao V.: *An assessment of important SPECIATE profiles in the EPA emissions modeling platform and current data gaps*, Atmospheric Environment, 207, 93-104, **2019**, doi: 10.1016/j.atmosenv.2019.03.013.

[47] Pernigotti D., Belis C.A., and Spanò L.: *SPECIEUROPE: The European data base for PM source profiles*, Atmospheric Pollution Research, 7, 2, 307-314, **2016**, doi: 10.1016/j.apr.2015.10.007.

[48] Pernigotti D., and Belis C.A.: *DeltaSA tool for source apportionment benchmarking, description and sensitivity analysis*, Atmospheric Environment, 180, 138-148, **2018**, doi: 10.1016/j.atmosenv.2018.02.046.

[49] Amato F., Pandolfi M., Viana M., Querol X., Alastuey A., and Moreno T.: *Spatial and chemical patterns of $PM_1 0$ in road dust deposited in urban environment*, Atmospheric Environment, 43, 1650-1659, **2009**, doi: 10.1016/j.atmosenv.2008.12.009.

[50] Kim E., Hopke P.K., and Edgerton E.S.: *Source Identification of Atlanta Aerosol by Positive Matrix Factorization*, Journal of the Air & Waste Management Association, 53, 6, 731-739, **2003**, doi: 10.1080/10473289.2003.10466209.

[51] Carslaw D.C.: *The openair manual - open-source tools for analysing air pollution data*, Manual for version 2.6-6, University of York, **2019**.

[52] Ogulei D., Hopke P. K., Zhou L., Paatero P., Park S. S., and Ondov J. M.: *Receptor modeling for multiple time resolved species: the Baltimore supersite*, Atmospheric Environment, 39, 3751–3762, **2005**, doi: 10.1016/j.atmosenv.2005.03.012.

[53] Liao H.-T., Kuo C.-P., Hopke P.K., and Wu C.-F.: *Evaluation of a Modified Receptor Model for Solving Multiple Time Resolution Equations: A Simulation Study*, Aerosol and Air Quality Research, 13, 1253–1262, **2013**, doi: 10.4209/aaqr.2012.11.0322.

[54] Kuo C.-P., Liao H.-T., Chou C.C.-K., and Wu C.-F.: *Source apportionment of particulate matter and selected volatile organic compounds with multiple time resolution data*, Science of the Total Environment, 472, 880–887, **2014**, doi: 10.1016/j.scitotenv.2013.11.114.

[55] Liao H.-T., Chou C.C.-K., Chow J.C., Watson J.G., Hopke P.K., and Wu C.-F.: *Source and risk apportionment of selected VOCs and $PM_{2.5}$ species using partially constrained receptor models with multiple time resolution data*, Environmental Pollution, 205, 121-130, **2015**, doi: 10.1016/j.envpol.2015.05.035.

[56] Crespi A., Bernardoni V., Calzolai G., Lucarelli F., Nava S., Valli G., and Vecchi R.: *Implementing constrained multi-time approach with bootstrap analysis in ME-2: An application to PM2.5 data from Florence (Italy)*, Science of the Total Environment, 541, 502–511, **2016**, doi: 10.1016/j.scitotenv.2015.08.159.

[57] Sofowote U.M., Healy R.M., Su Y., Debosz J., Noble M., Munoz A., Jeong C.-H., Wang J.M., Hilker N., Evans G.J., and Hopke P.K.: *Understanding the $PM_{2.5}$ imbalance between a far and near-road location: Results of high temporal frequency source apportionment and parameterization of black carbon*, Atmospheric Environment, 173, 277–288 **2018**, doi: 10.1016/j.atmosenv.2017.10.063.

[58] Srivastava D., Favez O., Petit J.-E., Zhang Y., Sofowote U.M., Hopke P.K., Bonnaire N., Perraudin E., Gros V., Villenave E., and Albinet A.: *Speciation of organic fractions does matter for aerosol source apportionment. Part 3: Combining off-line and on-line measurements*, Science of the Total Environment, 690, 944–955, **2019**, doi: 10.1016/j.scitotenv.2019.06.378.

[59] Paatero P.: *Interactive comment on "Receptor modelling of both particle composition and size distribution from a background site in London, UK - the two step approach" by David C.S. Beddows and Roy M. Harrison*, Atmospheric Chemistry and Physics Discussions, **2018**, https://www.atmos-chem-phys-discuss.net/acp-2018-784/acp-2018-784-RC2.pdf.

[60] Zhou L., Hopke P.K., Stanier C.O., Pandis S.N., Ondov J.M., and Pancras J.P.: *Investigation of the relationship between chemical composition and size distribution of airborne particles by partial least squares and positive matrix factorization*, Journal of Geophysical Research, 110, D07S18, **2005**, doi: 10.1029/2004JD005050.

[61] Ogulei D., Hopke P.K., Zhou L., Pancras J.P., Nair N., and Ondov J.M.: *Source apportionment of Baltimore aerosol from combined size distribution and chemical composition data*, Atmospheric Environment, 40, S396–S410, **2006**, doi: 10.1016/j.atmosenv.2005.11.075.

[62] Gu J., Pitz M., Schnelle-Kreis J., Diemer J., Reller A., Zimmermann R., Soentgen J., Stoelzel M., Wichmann H.-E., Peters A., and Cyrys J.: *Source apportionment of ambient particles: Comparison of positive matrix factorization analysis applied to particle size distribution and chemical composition data*, Atmospheric Environment, 45, 1849-1857, **2011**, doi: 10.1016/j.atmosenv.2011.01.009.

[63] Masiol M., Hopke P.K., Felton H.D., Frank B.P., Rattigan O.V., Wurth M.J., and LaDuke G.H.: *Source apportionment of $PM_{2.5}$ chemically speciated mass and particle number concentrations in New York City*, Atmospheric Environment, 148, 215-229, **2017**, doi: 10.1016/j.atmosenv.2016.10.044.

[64] Leoni C., Pokorn P., Hovorka J., Masiol M., Topinka J., Zhao Y., Křůmal K., Cliff S., Mikuška P., and Hopke P.K.: *Source apportionment of aerosol particles at a European air pollution hot spot using particle number size distributions and chemical composition*, Environmental Pollution, 234, 145-154, **2018**, doi: 10.1016/j.envpol.2017.10.097.

[65] Beddows D.C.S. , Harrison R.M., Green D.C., and Fuller G.W.: *Receptor modelling of both particle composition and size distribution from a background site in London, UK*, Atmospheric Chemistry and Physics, 15, 10107–10125, **2015**, doi: 10.5194/acp-15-10107-2015.

[66] Beddows D.C.S., and Harrison R.M.: *Receptor modelling of both particle composition and size distribution from a background site in London, UK – a two-step approach*, Atmospheric Chemistry and Physics, 19, 4863–4876, **2019**, doi: 10.5194/acp-19-4863-2019.

[67] Paatero P., and Hopke P.K.: *Utilizing wind direction and wind speed as independent variables in multilinear receptor modeling studies*, Chemometrics and Intelligent Laboratory Systems, 60, 25– 41, **2002**, doi: 10.1016/S0169-7439(01)00183-6.

[68] Kima E., Hopke P.K., Paatero P., and Edgerton E.S.: *Incorporation of parametric factors into multilinear receptor model studies of Atlanta aerosol*, Atmospheric Environment, 37, 5009–5021, **2003**, doi: 10.1016/j.atmosenv.2003.08.035.

[69] Begum B.A., Hopke P.K., and Zhao W.: *Source Identification of Fine Particles in Washington, DC, by Expanded Factor Analysis Modeling*, Environmental Science and Technology, 39, 1129-1137, **2005**, doi: 10.1021/es049804v.

[70] Buset K.C., Evans G.J., Leaitch W.R., Brook J.R., and Toom-Sauntry D.: *Use of advanced receptor modelling for analysis of an intensive 5-week aerosol sampling campaign*, Atmospheric Environment, 40, S482–S499, **2006**, doi: 10.1016/j.atmosenv.2005.12.074.

[71] Zhou L., Hopke P.K., and Zhao W.: *Source Apportionment of Airborne Particulate Matter for the Speciation Trends Network Site in Cleveland, OH*, Journal of the Air & Waste Management Association, 59, 3, 321-331, doi: 10.3155/1047-3289.59.3.321.

[72] Crawford J., Cohen D.D., Chambers S., Williams A., and Stelcer E.: *Incorporation of Radon-222 as a parameter in ME-2 to improve apportionment of $PM_{2.5}$ sources in the Sydney region*, Atmospheric Environment, 80, 378-388, **2013**, doi: 10.1016/j.atmosenv.2013.08.015.

[73] Crawford J., Chambers S., Cohen D.D., Dyer L., Wang T., and Zahorowski W.: *Receptor modelling using Positive Matrix Factorisation, back trajectories and Radon-222*, Atmospheric Environment, 41, 6823–6837, **2007**, doi: 10.1016/j.atmosenv.2007.04.048.

[74] Forello A.C., Amato F., Bernardoni V., Calzolai G., Canepari S., Costabile F., Di Liberto L., Gualtieri M., Lucarelli F., Nava S., Perrino C., Petralia E., Valentini S., Valli G., and Vecchi R.: *Gaining knowledge on source contribution to aerosol optical absorption properties and organics by receptor modelling*, Atmospheric Environment, 243, 117873, **2020**, doi: 10.1016/j.atmosenv.2020.117873.

[75] Sandradewi J., Prévôt A.S.H., Szidat S., Perron N., Alfarra M.R., Lanz V.A., Weingartner E., and Balternsperger U.: *Using aerosol light absorption measurements for the quantitative determination of wood burning and traffic emission contributions to particulate matter*, Environmental Science and Technology, 42, 3316–3323, **2008**, doi: 10.1021/es702253m.

[76] Massabò D., Caponi L., Bernardoni V., Bove M.C., Brotto P., Calzolai G., Cassola F., Chiari M., Fedi M.E., Fermo P., Giannoni M., Lucarelli F., Nava S., Piazzalunga A., Valli G., Vecchi R., and Prati P.: *Multi-wavelength optical determination of black and brown carbon in atmospheric aerosols*, Atmospheric Environment, 108, 1–12, **2015**, doi: 10.1016/j.atmosenv.2015.02.058.

[77] Bernardoni V., Pileci R.E., Caponi L., and Massabò D.: *The Multi-Wavelength Absorption Analyzer (MWAA) model as a tool for source and component apportionment based on aerosol absorption properties: application to samples collected in different environments*, Atmosphere, 8, 218, **2017**, doi: 10.3390/atmos8110218.

[78] Zotter P., Herich H., Gysel M., El-Haddad I., Zhang Y., Mocnik G., Huglin C., Baltensperger U., Szidat S., and Prévôt A.S.H.: *Evaluation of the absorption Ångström exponents for traffic and wood burning in the Aethalometer-based source apportionment using radiocarbon measurements of ambient aerosol*, Atmospheric Chemistry and Physics, 17, 4229–4249, **2017**, doi: 10.5194/acp-17-4229-2017.

[79] Harrison R.M., Beddows D.C.S., Jones A.M., Calvo A., Alves C., and Pio C.: *An evaluation of some issues regarding the use of aethalometers to measure woodsmoke concentrations*, Atmospheric Environment, 80, 540-548, **2013**, doi: 10.1016/j.atmosenv.2013.08.026.

[80] Vecchi R., Marcazzan G., and Valli G.: *A study on nighttimedaytime PM_{10} concentration and elemental composition in relation to atmospheric dispersion in the urban area of Milan (Italy)*, Atmospheric Environment, 41, 2136–2144, **2007**, doi: 10.1016/j.atmosenv.2006.10.069.

[81] Vecchi R., Piziali F.A., Valli G., Favaron M., and Bernardoni V.: *Radon-based estimates of equivalent mixing layer heights: A long-term assessment*, Atmospheric Environment, 197, 150–158, **2019**, doi: 10.1016/j.atmosenv.2018.10.020.

[82] Perrone M.G., Larsen B.R., Ferrero L., Sangiorgi G., De Gennaro G., Udisti R., Zangrando R., Gambaro A., and Bolzacchini E.: *Sources of high $PM_{2.5}$ concentrations in Milan, Northern Italy: Molecular marker data and CMB modelling*, Science of the Total Environment, 414, 343–355, **2012**, doi: 10.1016/j.scitotenv.2011.11.026.

[83] Bigi A. and Ghermandi G.: *Long-term trend and variability of atmospheric PM_{10} concentration in the Po Valley*, Atmospheric Chemistry and Physics, 14, 4895–4907, **2014**, doi: 10.5194/acp-14-4895-2014.

[84] Perrino C., Catrambone M., Dalla Torre S., Rantica E., Sargolini T., and Canepari S.: *Seasonal variations in the chemical composition of particulate matter: a case study in the Po Valley. Part I: macrocomponents and mass closure*, Environmental Science and Pollution Research, 21, 3999–4009, **2014**, doi: 10.1007/s11356-013-2067-1.

[85] Vecchi R., Bernardoni V., Fermo P., Lucarelli F., Mazzei F., Nava S., Prati P., Piazzalunga A., and Valli G.: *4-hours resolution data to study PM_{10} in a "hot spot" area in Europe*, Environmental Monitoring and Assessment, 154, 283–300, **2009**, doi: 10.1007/s10661-008-0396-1.

[86] D'Alessandro A., Lucarelli F., Mandò P.A., Marcazzan G., Nava S., Prati P., Valli G., Vecchi R., and Zucchiatti A.: *Hourly elemental composition and sources identification of fine and coarse PM_{10} particulate matter in four Italian towns*, Journal of Aerosol Science, 34, 243–259, **2003**, doi: 10.1016/S0021-8502(02)00172-6.

[87] Vecchi R., Marcazzan G., Valli G., Ceriani M., and Antoniazzi C.: *The role of atmospheric dispersion in the seasonal variation of PM_1 and $PM_{2.5}$ concentration and composition in the urban area of Milan (Italy)*, Atmospheric Environment, 38, 4437–4446, **2004**, doi: 10.1016/j.atmosenv.2004.05.029.

[88] Piazzalunga A., Fermo P., Bernardoni V., Vecchi R., Valli G., and De Gregorio M.A.: *A simplified method for levoglucosan quantification in wintertime atmospheric particulate matter by high performance anion-exchange chromatography coupled with pulsed amperometric detection*, International Journal of Environmental Analytical Chemistry, 90, 934–947, **2010**, doi: 10.1080/03067310903023619.

[89] Bernardoni V., Vecchi R., Valli G., Piazzalunga A., and Fermo P.: *PM$_{10}$ source apportionment in Milan (Italy) using time-resolved data*, Science of the Total Environment, 409, 4788–4795, **2011**, doi: 10.1016/j.scitotenv.2011.07.048.

[90] Robinson A.L., Donahue N.M., and Rogge W.F.: *Photochemical oxidation and changes in molecular composition of organic aerosol in the regional context*, Journal of Geophysical Research, 111, D03302, **2006**, doi: 10.1029/2005JD006265.

[91] Hennigan C.J., Sullivan A.P., Collett Jr. J.L., and Robinson A.L.: *Levoglucosan stability in biomass burning particles exposed to hydroxyl radicals*, Geophysical Research Letters, 37, L09806, **2010**, doi: 10.1029/2010GL043088.

[92] Piazzalunga A., Bernardoni V., Fermo P., and Vecchi R.: *Optimisation of analytical procedures for the quantification of ionic and carbonaceous fractions in the atmospheric aerosol and application to ambient samples*, Analytical Bioanalytical Chemistry, 405, 1123–1132, **2013**, doi: 10.1007/s00216-012-6433-5.

[93] Piazzalunga A., Bernardoni V., Fermo P., Valli G., and Vecchi R.: *Technical Note: On the effect of water-soluble compounds removal on EC quantification by TOT analysis in urban aerosol samples*, Atmospheric Chemistry and Physics, 11, 10193–10203, **2011**, doi: 10.5194/acp-11-10193-2011.

[94] Campbell J.L., Boyd N.I., Grassi N., Bonnick P., and Maxwell J.A.: *The Guelph PIXE software package IV*, Nuclear Instruments and Methods in Physics Research Section B, 268, 3356-3363, **2010**, doi: 10.1016/j.nimb.2010.07.012.

[95] Calzolai G., Lucarelli F., Chiari M., Nava S., Giannoni M., Carraresi L., Prati P., and Vecchi R.: *Improvements in PIXE analysis of hourly particulate matter samples*, Nuclear Instruments and Methods B, 363, 99–104, **2015**, doi: 10.1016/j.nimb.2015.08.022.

[96] Lucarelli F., Calzolai G., Chiari M., Nava S., and Carraresi L.: *Study of atmospheric aerosols by IBA techniques: The LABEC experience*, Nuclear Instruments and Methods in Physics Research B, 417, 121–127, **2018**, doi: 10.1016/j.nimb.2017.07.034.

[97] Vecchi R., Bernardoni V., Paganelli C., and Valli G.: *A filter-based light absorption measurement with polar photometer: effects of sampling artefacts from organic carbon*, Journal of Aerosol Science, 70, 15–25, **2014**, doi:10.1016/j.jaerosci.2013.12.012.

[98] Bernardoni V., Valli G., and Vecchi R.: *Set-up of a multi-wavelength polar photometer for the off-line measurement of light absorption properties of atmospheric aerosol collected with high-temporal resolution*, Journal of Aerosol Science, 107, 84–93, **2017**, doi: 10.1016/j.jaerosci.2017.02.009.

[99] Hänel G.: *Radiation budget of the boundary layer. Part II: Simultaneous Measurement of Mean Solar Volume Absorption and Extinction Coefficients of Particles*, Beitraege zur Physik der Atmosphaere, 50, 2, 241-247, **1987**.

[100] Petzold A., and Schönlinner M.: *Multi-angle absorption photometry - a new method for the measurement of aerosol light absorption and atmospheric black carbon*, Journal of Aerosol Science, 35, 421-441, **2004**, doi: 10.1016/j.jaerosci.2003.09.005.

[101] Coakley J.A., and Chylek P.: *The Two-Stream Approximation in Radiative Transfer: Including the Angle of the Incident Radiation*, Journal of the Atmospheric Sciences, 32, 409-418, **1975**.

[102] Cappa C.D., Lack D.A., Burkholder J.B., and Ravishankara A.R.: *Bias in filter-based aerosol light absorption measurements due to organic aerosol loading: Evidence from laboratory measurements*, Aerosol Science and Technology, 42, 1022–1032, **2008**, doi: 10.1080/02786820802389285.

[103] Lack D.A., Cappa C.D., Covert D.S., Baynard T., Massoli P., Sierau B., Bates T.S., Quinn P.K., Lovejoy E.R., and Ravishankara A.R.: *Bias in filter-based aerosol light absorption measurements due to organic aerosol loading: Evidence from ambient measurements*, Aerosol Science and Technology, 42, 1033–1041, **2008**, doi: 10.1080/02786820802389277.

[104] Davies N.W., Fox C., Szpek K., Cotterell M.I., Taylor J.W., Allan J.D., Williams P.I., Trembath J., Haywood J.M., and Langridge J.M.: *Evaluating biases in filter-based aerosol absorption measurements using photoacoustic spectroscopy*, Atmospheric Measurement Techniques, 12, 3417–3434, **2019**, doi: 10.5194/amt-12-3417-2019.

[105] Simoneit B.R.T., Schauer J.J., Nolte C.G., Oros D.R., Elias V.O., Fraser M.P., Rogge D.F., and Cass G.R.: *Levoglucosan, a tracer for cellulose in biomass burning atmospheric particles*, Atmospheric Environment, 33, 173–182, **1999**, doi: 10.1016/S1352-2310(98)00145-9.

[106] Mason B.: *Principles of geochemistry*, 3rd Edition, John Wiley & Sons, New York, **1966**.

[107] Thorpe A. and Harrison R.M.: *Sources and properties of non-exhaust particulate matter from road traffic: A review*, Science of the Total Environment, 400, 270–282, **2008**, doi: 10.1016/j.scitotenv.2008.06.007.

[108] Calzolai G., Chiari M., Lucarelli F., Mazzei F., Nava S., Prati P., Valli G., and Vecchi R.: *PIXE and XRF analysis of particulate matter samples: an inter-laboratory comparison*, Nuclear Instruments and Methods B, 266, 2401–2404, **2008**, doi: 10.1016/j.nimb.2008.03.056.

[109] Marcazzan G.M., Ceriani M., Valli G., and Vecchi R.: *Source apportionment of PM_{10} and $PM_{2.5}$ in Milan (Italy) using receptor modelling*, Science of the Total Environment, 317, 137–147, **2003**, doi: 10.1016/S0048-9697(03)00368-1.

[110] Vecchi R., Bernardoni V., Valentini S., Piazzalunga A., Fermo P., and Valli G.: *Assessment of light extinction at a European polluted urban area during wintertime: Impact of PM1 composition and sources*, Environmental Pollution, 233, 679–689, **2018**, doi: 10.1016/j.envpol.2017.10.059.

[111] Bernardoni V., Elser M., Valli G., Valentini S., Bigi A., Fermo P., Piazzalunga A., and Vecchi R.: *Size-segregated aerosol in a hot-spot pollution urban area: Chemical composition and three-way source apportionment*, Environmental Pollution, 231, 601–611, **2017**, doi: 10.1016/j.envpol.2017.08.040.

[112] Marcazzan G M., Vaccaro S., Valli G., and Vecchi R.: *Characterisation of PM_{10} and $PM_{2.5}$ particulate matter in the ambient air of Milan (Italy)*, Atmospheric Environment, 35, 4639–4650, **2001**, doi: 10.1016/S1352-2310(01)00124-8.

[113] Dall'Osto M., Querol X., Amato F., Karanasiou A., Lucarelli F., Nava S., Calzolai G., and Chiari M.: *Hourly elemental concentrations in $PM_{2.5}$ aerosols sampled simultaneously at urban background and road site during SAPUSS – diurnal variations and PMF receptor modelling*, Atmospheric Chemistry and Physics, 13, 4375–4392, **2013**, doi: 10.5194/acp-13-4375-2013.

[114] Crilley L.R., Lucarelli F., Bloss W.J., Harrison R.M., Beddows D.C., Calzolai G., Nava S., Valli G., Bernardoni V., and Vecchi R.: *Source Apportionment of Fine and Coarse Particles at a Roadside and Urban Background Site in London during the Summer ClearfLo Campaign*, Environmental Pollution, 220, 766–778, **2017**, doi: 10.1016/j.envpol.2016.06.002.

[115] INEMAR – ARPA Lombardia: *INEMAR, Inventario Emissioni in Atmosfera: emissioni in Regione Lombardia nell'anno 2014 – dati finali*, ARPA Lombardia Settore Monitoraggi Ambientali, available at: http://www.inemar.eu/xwiki/bin/view/Inemar/HomeLombardia (last access: 14 January 2019), **2014**.

[116] Draxler R.R. and Hess G.D.: *An overview of the HYSPLIT_4 modelling system for trajectories, dispersion, and deposition*, Australian Meteorological Magazine, 47, 295–308, **1998**.

[117] Stein A.F., Draxler R.R., Rolph G.D., Stunder B.J.B., Cohen M.D., and Ngan F.: *NOAA's Hysplit atmospheric transport and dispersion modeling system*, Bulletin of the Americal Meteorological Society, 96, 2059– 2077, **2015**, doi: 10.1175/BAMS-D-14-00110.1.

[118] Rolph G., Stein A., and Stunder B.: *Real-time Environmental Application and Display sYstem: READY*, Environmental Modelling & Software, 95, 210–228, **2017**, doi: 10.1016/j.envsoft.2017.06.025.

[119] Carslaw D.C. and Ropkins K.: *Openair – an R package for air quality data analysis*, Environmental Modelling & Software, 27/28, 52–61, **2012**, doi: 10.1016/j.envsoft.2011.09.008.

[120] R Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, available at: http://www.R-project.org (last access: 7 January 2019), **2019**.

[121] Pachon J.E., Weber R.J., Zhang X., Mulholland J.A., and Russell, A.G.: *Revising the use of potassium (K) in the source apportionment of $PM_{2.5}$*, Atmospheric Pollution Research, 4, 14–21, **2013**, doi: 10.5094/APR.2013.002.

[122] Bond T.C. and Bergstrom R.W.: *Light absorption by carbonaceous particles: an investigative review*, Aerosol Science and Technology, 40, 27–67, **2006**, doi: 10.1080/02786820500421521.

[123] Fuller K.A., Malm W.C., and Kreidenweis S.M.: *Effects of mixing on extinction by carbonaceous particles*, Journal of Geophysical Research, 104, 15941–15954, **1999**, doi: 10.1029/1998JD100069.

[124] Schnaiter M., Horvath H., Möhler O., Naumann K.-H., Saathoff H., and Schöck O.W.: *UV-VIS-NIR spectral optical properties of soot and soot-containing aerosols*, Journal of Aerosol Science, 34, 1421–1444, **2003**, doi: 10.1016/S0021-8502(03)00361-6.

[125] Moffet R.C. and Prather K.A.: *In-situ measurements of the mixing state and optical properties of soot with implications for radiative forcing estimates*, Proceedings of the National Academy of Sciences of the United States of America, 106, 11872–11877, **2009**, doi: 10.1073/pnas.0900040106.

[126] Sandradewi J., Prévôt A.S.H., Weingartner E., Schmidhauser R., Gysel M., and Balternsperger U.: *A study of wood burning and traffic aerosols in an Alpine valley using a multi-wavelength Aethalometer*, Atmospheric Environment, 2, 101–112, **2008**, doi: 10.1016/j.atmosenv.2007.09.034.

[127] Liu C., Chung C.E., Yin Y., and Schnaiter M.: *The absorption Ångström exponent of black carbon: from numerical aspects*, Atmospheric Chemistry and Physics, 18, 6259–6273, **2018**, doi: 10.5194/acp-18-6259-2018.

[128] Fischer D.A. and Smith G.D.: *A portable, four wavelength, single-cell photoacoustic spectrometer for ambient aerosol absorption*, Aerosol Science and Technology, 52, 393–406, **2018**, doi: 10.1080/02786826.2017.1413231.

[129] Gyawali M., Arnott W.P., Lewis K., and Moosmüller H.: *In situ aerosol optics in Reno, NV, USA during and after the summer 2008 California wildfires and the influence of absorbing and nonabsorbing organic coatings on spectral light absorption*, Atmospheric Chemistry and Physics, 9, 8007–8015, **2009**, doi: 10.5194/acp-9-8007-2009.

[130] Petzold A., Ogren J.A., Fiebig M., Laj P., Li S.-M., Baltensperger U., Holzer-Popp T., Kinne S., Pappalardo G., Sugimoto N., Wehrli C., Wiedensohler A., and Zhang X.-Y.: *Recommendations for reporting "black carbon" measurements*, Atmospheric Chemistry and Physics, 13, 8365–8379, **2013**, doi: 10.5194/acp-13-8365-2013.

[131] Costabile F., Alas H., Aufderheide M., Avino P., Amato F., Argentini S., Barnaba F., Berico M., Bernardoni V., Biondi R., Calzolai G., Canepari S., Casasanta G., Ciampichetti S., Conidi A., Cordelli E., Di Ianni A., Di Liberto L., Facchini M.C., Facci A., Frasca D., Gilardoni S., Grollino M.G., Gualtieri M., Lucarelli F., Malaguti A., Manigrasso M., Montagnoli M., Nava S., Padoan E., Perrino C., Petralia E., Petenko I., Querol X., Simonetti G., Tranfo G., Ubertini S., Valli G., Valentini S., Vecchi R., Volpi F., Weinhold K., Wiedensohler A., Zanini G., and Gobbi G.P.: *First Results of the "Carbonaceous Aerosol in Rome and Environs (CARE)" Experiment: Beyond Current Standards for PM_{10}*, Atmosphere, 8, 249, **2017**, doi: 10.3390/atmos8120249.

[132] Valentini S., Barnaba F., Bernardoni V., Calzolai G., Costabile F., Di Liberto L., Forello A.C., Gobbi G.P., Gualtieri M., Lucarelli F., Nava S., Petralia E., Valli G., Wiedensohler A., Vecchi R.: *Classifying aerosol particles through the combination of optical and physical-chemical properties: Results from a wintertime campaign in Rome (Italy)*, Atmospheric Research, 235, 104799 , **2020**, doi: 10.1016/j.atmosres.2019.104799.

[133] Costabile F., Gualtieri M., Canepari S., Tranfo G., Consales C., Grollino M.G., Paci E., Petralia E., Pigini D., and Simonetti G.: *Evidence of association between aerosol properties and in-vitro cellular oxidative response to PM$_1$, oxidative potential of PM$_{2.5}$, a biomarker of RNA oxidation, and its dependency on combustion sources*, Atmospheric Environment, 213, 444-455, **2019**, doi: 10.1016/j.atmosenv.2019.06.023.

[134] Khlystov A., Stanier C., and Pandis S.N.: *An algorithm for combining electrical mobility and aerodynamic size distributions data when measuring ambient aerosol*, Aerosol Science and Technology, 38, 229–238, **2004**, doi: 10.1080/02786820390229543.

[135] Alas H.D.C., Weinhold K., Costabile F. Di Ianni A., Müller T., Pfeifer S., Di Liberto L., Turner J.R., and Wiedensohler A.: *Methodology for high-quality mobile measurement with focus on black carbon and particle mass concentrations*, Atmospheric Measurement Techniques, 12, 4697-4712, **2019**, doi: 10.5194/amt-12-4697-2019.

[136] Ng N.L., Herndon S.C., Trimborn A., Canagaratna M.R., Croteau P., Onasch T.M., Sueper D., Worsnop D.R., Zhang Q., Sun Y.L., and Jayne T.: *An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol*, Aerosol Science and Technology, 45, 770–784, **2011**, doi: 10.1080/02786826.2011.560211.

[137] Crenn V., Sciare J., Croteau P.L., Verlhac S., Fröhlich R., Belis C.A., Aas W., Äijälä M., Alastuey A., Artiñano B., Baisnée D., Bonnaire N., Bressi M., Canagaratna M., Canonaco F., Carbone C., Cavalli F., Coz E., Cubison M.J., Esser-Gietl J.K., Green D.C., Gros V., Heikkinen L., Herrmann H., Lunder C., Minguillón M.C., Močnik G., O'Dowd C.D., Ovadnevaite J., Petit J.E., Petralia E., Poulain L., Priestman M., Riffault V., Ripoll A., Sarda-Estève R., Slowik J.G., Setyan A., Wiedensohler A., Baltensperger U., Prévôt A.S.H., Jayne J.T., Favez O.: *ACTRIS ACSM intercomparison - Part 1: reproducibility of concentration and fragment results from 13 individual Quadrupole Aerosol Chemical Speciation Monitors (Q-ACSM) and consistency with co-located instruments*, Atmospheric Measurement Techniques, 8 (12), 5063–5087, **2015**, doi: 10.5194/amt-8-5063-2015.

[138] Belis C.A., Pikridas M., Lucarelli F., Petralia E., Cavalli F., Calzolai G., Berico M., and Sciare J.: *Source apportionment of fine PM by combining high time resolution organic and inorganic chemical composition datasets*, Atmospheric Environment X, 3, 100046, **2019**, doi: 10.1016/j.aeaoa.2019.100046.

[139] Sunset Laboratory Inc.: *Semi-continuous OCEC carbon aerosol analyzer, A guide to running and maintaining the sunset laboratory semi-continuous OCEC analyser*, **2005**.

[140] Drinovec L., Močnik G., Zotter P., Prévôt A.S.H., Ruckstuhl C., Coz E., Rupakheti M., Sciare J., Müller T., Wiedensohler E., and Hansen D.A.: *The "dual-spot" Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation*, Atmospheric Measurement Techniques 8, 1965-1979, **2015**, doi: 10.5194/amt-8-1965-2015.

[141] Magee Scientific Aethalometer® Model AE33 – User Manual, Version 1.54, **2016**.

[142] Goetz J.D., Giordano M.R., Stockwell C.E., Christian T.J., Maharian R., Adhikari S., Bhave P.V., Praveen P.S., Panday A.K., Jayarathne T., Stone E.A., Yokelson R.J., and DeCarlo P.F.: *Speciated online PM1 from South Asian combustion sources – Part1: Fuel-based emission factors and size distributions*, Atmospheric Chemistry and Physics, 18, 14653–14679, **2018**, doi: 10.5194/acp-18-14653-2018.

[143] U.S. EPA (Environmental Protection Agency), Office of Air Quality Planning and Standards, Air Quality Assessment Division, Ambient Air Monitoring Group: *Sunset Carbon Evaluation Project, Quality Assurance Project Plan (QAPP)*, Version 1, **2011**.

[144] Kester D.R., and Pytkowicx R.M.: *Sodium, magnesium, and calcium sulfate ion-pairs in seawater at 25C*, Limnology and Oceanography, 14, 5, 686-692, **1969**.

[145] Perrino C., Canepari S., Catrambone M., Dalla Torre S., Rantica E., and Sargolini T.: *Influence of natural events on the concentration and composition of atmospheric particulate matter*, Atmospheric Environment, 43, 4766-4779, **2009**, doi: 10.1016/j.atmosenv.2008.06.035.

[146] Tofful L., and Perrino C.: *Chemical Composition of Indoor and Outdoor PM2.5 in Three Schools in the City of Rome*, Atmosphere, 6, 1422-1443, **2015**, doi: 10.3390/atmos6101422.

[147] Perrino C., Catrambone M., Farao C., and Canepari S.: *Assessing the contribution of water to the mass closure of PM10*, Atmospheric Environment, 140, 555-564, **2016**, doi: 10.1016/j.atmosenv.2016.06.038.

[148] Reid J.S., Koppmann R., Eck T.F., and Eleuterio D.P.: *A review of biomass burning emissions part II: intensive physical properties of biomass burning particles*, Atmospheric Chemistry and Physics, 5, 799–825, **2005**, doi: 10.5194/acp-5-799-2005.

[149] Perrino C., Tofful L., Dalla Torre S., Sargolini T., and Canepari S.: *Biomass burning contribution to PM10 concentration in Rome (Italy): Seasonal, daily and two-hourly variations*, Chemosphere, 222, 839-848, **2019**, doi: 10.1016/j.chemosphere.2019.02.019.

[150] Piazzalunga A., Belis C., Bernardoni V., Cazzuli O., Fermo P., Valli G., and Vecchi R.: *Estimates of wood burning contribution to PM by the macro-tracer method using tailored emission factors*, Atmospheric Environment, 45, 6642-6649, **2011**, doi: 10.1016/j.atmosenv.2011.09.08.

[151] Struckmeier C., Drewnick F., Fachinger F., Gobbi G.P., and Borrmann S.: *Atmospheric aerosols in Rome, Italy: sources, dynamics and spatial variations during two seasons*, Atmospheric Chemistry and Physics, 16, 15277–15299, **2016**, doi: 10.5194/acp-16-15277-2016.

[152] Salzano R., Pasini A., Casasanta G., Cacciani M., and Perrino C.: *Quantitative Interpretation of Air Radon Progeny Fluctuations in Terms of Stability Conditions in the Atmospheric Boundary Layer*, Boundary-Layer Meteorology, 160, 529–550, **2016**, doi: 10.1007/s10546-016-0149-6.

[153] Jeong C.H., Wang J.M., Hilker N., Debosz J., Sofowote U., Su Y., Noble M., Healy R.M., Munoz T., Dabek-Zlotorzynska E., Celo V., White L., Audette C., Herod D., and Evans G.J.: *Temporal and spatial variability of traffic-related PM2.5 sources: Comparison of exhaust and non-exhaust emissions*, Atmospheric Environment, 198, 55-69, **2019**, doi: 10.1016/j.atmosenv.2018.10.038.

[154] Laskin A., Laskin J., and Nizkorodov S.A.: *Chemistry of Atmospheric Brown Carbon*, Chemical Reviews, 115, 4335-4382, **2015**, doi: 10.1021/cr5006167.

[155] Costabile F., Gilardoni S., Barnaba F., Di Ianni A., Di Liberto L., Dionisi D., Manigrasso M., Paglione M., Poluzzi V., Rinaldi M., Facchini M.C., Gobbi G.P.: *Characteristics of brown carbon in the urban Po Valley atmosphere*, Atmospheric Chemistry and Physics, 17, 313–326, **2017**, doi: 10.5194/acp-17-313-2017.

[156] DeCarlo P.F., Ulbrich I.M., Crounse J., de Foy B., Dunlea E.J., Aiken A.C., Knapp D., Weinheimer A.J., Campos T., Wennberg P.O., and Jimenez J.L.: *Investigation of the sources and processing of organic aerosol over the Central Mexican Plateau from aircraft measurements during MILAGRO*, Atmospheric Chemistry and Physics, 10, 5257–5280, **2010**, doi: 10.5194/acp-10-5257-2010.

[157] Raabe O.G., Braaten D.A., Axelbaum R.L., Teague S.V., and Cahill T.A.: *Calibration studies of the DRUM impactor*, Journal of Aerosol Science, 19, 2, 183-195, **1988**.

[158] Marple V.A., Olson B.A., and Rubow K.L.: *Aerosol measurement: principles, techniques, and applications*, John Wiley and Sons, **2001-2005**.

[159] Marple V.A., and Willeke K.: *Impactor design*, Atmospheric Environment, 10, 891-896, **1976**.

[160] Spurny K.R., Lodge J.P., Frank E.R., and Sheesley D.C.: *Aerosol Filtration byMeans of Nuclepore Filters: Structural and Filtration Properties*, Environmental Science and Technology, 3, 5, 453–464, **1969**.

[161] Spurny K.R., and Lodge Jr. J.P.: *Collection Efficiency Tables for Membrane Filters Used in the Sampling and Analysis of Aerosols and Hydrosols*, NCAR Technical Note NCAR/TN-77+STR, University Corporation for Atmospheric Research, **1972**, doi: 10.5065/D6F769JJ.

[162] Ogura I., Kotake M., Sakurai H., and Honda K.: *Surface-collection efficiency of Nuclepore filters for nanoparticles*, Aerosol Science and Technology, 50, 8, 846-856, **2016**, doi: 10.1080/02786826.2016.1200007.

[163] Ogura I., Hashimoto N., Kotake M., Sakurai H., Kishimoto A., and Honda K.: *Aerosol Particle Collection Efficiency of Holey Carbon Film-Coated TEM Grids*, Aerosol Science and Technology, 48, 7, 758-767, **2014**, doi: 10.1080/02786826.2014.924614.

[164] Vincent J.H.: *Aerosol sampling: science, standards, instrumentation and applications*, John Wiley and Sons, **2007)**.

[165] Elser M.: *Le particelle ultrafini in atmosfera: metodi di campionamento e caratterizzazione*, Masters degree Thesis in Physics, Department of Physics, University of Milan, **2011-2012**.

[166] Forello A.C.: *Metodologie sperimentali per la caratterizzazione ad alta risoluzione temporale dell'aerosol atmosferico*, Masters degree thesis in Physics, Department of Physics, University of Milan, **2016-2017**.

[167] Liu B.Y.H., and Lee K.W.: *Efficiency of Membrane and Nuclepore filters for Submicrometer Aerosols*, Environmental Science and Technology, 10, 4, **1976**.

[168] Massabò D., Danelli S.G., Brotto P., Comite A., Costa C., Di Cesare A., Doussin J.F., Ferraro F., Formenti P., Gatta E., Negretti L., Oliva M., Parodi F., Vezzulli L., and Prati P.: *ChAMBRe: a new atmospheric simulation chamber for aerosol modelling and bio-aerosol research*, Atmospheric Measurement Techniques, 11, 5885-5900, **2018**, doi: 10.5194/amt-11-5885-2018.

[169] Capannelli G., Castello E., Comite A., Costa C., and Mamolini G.: *Electron microscopy characterization of airborne micro- and nanoparticulate matter*, Journal of Electron Microscopy, 60, 2, 117-131, **2011**, doi: 10.1093/jmicro/dfr001.

[170] Chow J.C.: *Measurement Methods to Determine Compliance with Ambient Air Quality Standards for Suspended Particles*, Journal of the Air and Waste Management Association, 45, 5, 320-382, **1995**, doi: 10.1080/10473289.1995.10467369.

[171] Manton M.J.: *The impaction of aerosols on a nuclepore filter*, Atmospheric Environment, 12, 1669-1675, **1978**.

[172] Manton M.J.: *Brownian diffusion of aerosols to the face of a nuclepore filter*, Atmospheric Environment, 13, 525-531, **1979**.

List of Publications

Peer-reviewed papers

[**P1**] Valentini S., Bernardoni V., Bolzacchini E., Ciniglia D., Ferrero L., Forello A.C., Massabó D., Pandolfi M., Prati P., Soldan F., Valli G., Yus Díez J., and Vecchi R.: Applicability of Benchtop Multi-Wavelength Polar Photometers to Off-line Measurements of the Multi-Angle Absorption Photometer (MAAP) Samples, in press on Journal of Aerosol Science, 105701, 2020, doi: 10.1016/j.jaerosci.2020.105701.

[**P2**] Forello A.C., Amato F., Bernardoni V., Calzolai G., Canepari S., Costabile F., Di Liberto L., Gualtieri M., Lucarelli F., Nava S., Perrino C., Petralia E., Valentini S., Valli G., and Vecchi R.: Gaining knowledge on source contribution to aerosol optical absorption properties and organics by receptor modelling, Atmospheric Environment, 243, 117873, 2020, doi: 10.1016/j.atmosenv.2020.117873.

[**P3**] Valentini S., Barnaba F., Bernardoni V., Calzolai G., Costabile F., Di Liberto L., Forello A.C., Gobbi G.P., Gualtieri M., Lucarelli F., Nava S., Petralia E., Valli G., Wiedensohler A., Vecchi R.: Classifying aerosol particles through the combination of optical and physical-chemical properties: Results from a wintertime campaign in Rome (Italy), Atmospheric Research, 235, 104799 , 2020, doi: 10.1016/j.atmosres.2019.104799

[**P4**] Forello A.C., Bernardoni V., Calzolai G., Lucarelli F., Massabò D., Nava S., Pileci R.E., Prati P., Valentini S., Valli G., and Vecchi R.: Exploiting multi-wavelength aerosol absorption coefficients in a multi-time resolution source apportionment study to retrieve source-dependent absorption parameters, Atmospheric Chemistry and Physics, 19, 11235–11252, 2019, doi: 10.5194/acp-19-11235-2019.

Papers submitted to peer-reviewed journals

[**S1**] Crova F., Valli G., Bernardoni V., Forello A.C., Valentini S., and Vecchi R.: Effectiveness of airborne radon progeny assessment for atmospheric studies, Atmospheric Research, under revision.

[**S2**] Bernardoni V., Ferrero L., Bolzacchini E., Forello A.C., Gregorič A., Massabò D., Močnik G., Prati P., Rigler M., Santagostini L., Soldan F., Valentini S., Valli G., and Vecchi R.: Determination of Aethalometer multiple-scattering enhancement parameters and impact on source and component apportionment during the winter 2017-2018 EMEP/ACTRIS/COLOSSAL campaign in Milan, Atmospheric Measurement Techniques, under revision.

Contributions in volumes

[**V1**] Bernardoni V., Forello A.C., Mariani F., Paroli B., Potenza M.A.C., Pullia A., Riccobono F., Sanvito T., Valentini S., Valli G., and Vecchi R.: Innovative instrumentation for the study of atmospheric aerosol optical properties, Toward a Science Campus in Milan. A Snapshot of Current Research at the Physics Department Aldo Pontremoli, Springer, ISBN 978-3-030-01628-9 ISBN 978-3-030-01629-6 (eBook), doi: 10.1007/978-3-030-01629-6.

Oral presentations as invited speaker

[**A1**] Forello A.C.: Positive Matrix Factorization mediante Multilinear Engine: Analisi multi-time, Workshop IAS Source Apportionment: nuove tecniche e frontiere dell'applicazione, Milan, Italy (20-21/02/2019).

[A2] <u>Forello A.C.</u>: Metodologie sperimentali per la caratterizzazione ad alta risoluzione temporale dell'aerosol atmosferico, VIII Convegno Nazionale sul Particolato Atmosferico PM2018, Matera, Italy (23-25/05/2018).

Presentations to international conferences

Oral presentations (author in bold to indicate the presenter)

[IO1] **<u>Forello A.C.</u>**, Bernardoni V., Calzolai G., Lucarelli F., Massabò D., Nava S., Pileci R.E., Prati P., Valentini S., Valli G., and Vecchi R.: Multi-time source apportionment: exploiting multi-wavelength aerosol absorption coefficient to better identify sources, European Aerosol Conference – EAC 2019 in Gothenburg, Sweden (25-30/08/2019).

[IO2] Valentini S., Bernardoni V., Bundke U., <u>Forello A.C.</u>, Massabò D., Petzold A., Prati P., Soldan F., Valli G., and **Vecchi R.**: Multi-wavelength polar photometers to assess absorption optical properties of aerosol collected on filters, European Aerosol Conference – EAC 2019 in Gothenburg, Sweden (25-30/08/2019).

[IO3] **Vecchi R.**, Bernardoni V., <u>Forello A.C.</u>, Massabò D., Prati P., Soldan F., Valentini S., and Valli G.: Exploiting the features of multi-lambda polar photometers to retrieve optical properties of aerosols collected on filters, 12th International Conference on Carbonaceous Particles in the Atmosphere (ICCPA) 2019 in Vienna, Austria (03-06/04/2019).

[IO4] Bernardoni V., <u>Forello A.C.</u>, Pileci R.E., Valentini S., Valli G., **Vecchi R.**, Caponi L., Prati P., and Massabò D.: Multi-wavelength absorption coefficient measurements of aerosol collected on filters: instrumentation and modelling developments for brown carbon studies, Aerosol Technology 2018 (AT2018) in Bilbao, Spain (17-20/06/18).

Poster presentations (author in bold to indicate the presenter)

[IP1] **<u>Forello A.C.</u>**, Bernardoni V., Calzolai G., Lucarelli F., Massabò D., Nava S., Pileci R.E., Prati P., Valentini S., Valli G., and Vecchi R.: Retrieving information on black and brown carbon emission sources exploiting aerosol optical properties in an advanced receptor model, 12th International Conference on Carbonaceous Particles in the Atmosphere (ICCPA) 2019 in Vienna, Austria (03-06/04/2019).

[IP2] **Vecchi R.**, Bernardoni V., Bigi A., Calzolai G., Elser M., Fermo P., <u>Forello A.C.</u>, Lucarelli F., Massabò D., Nava S., Piazzalunga A., Pileci R.E., Prati P., Valentini S., and Valli G.: Advanced receptor models as a tool to improve the knowledge of aerosol emission sources at a hot-spot pollution site (Milan-Italy), International Aerosol Conference 2018 (IAC2018) in St Louis, Missouri, USA (02-07/09/18).

Presentations to national conferences

Oral presentations (author in bold to indicate the presenter)

[NO1] Calzolai G., Chiari M., Giardi F., **Lucarelli F.**, Nava S., Bernardoni V., <u>Forello A.C.</u>, Valli G., Vecchi R., Massabò D., Danelli S., Prati P., and Vernocchi V.: Un nuovo campionatore per studi di composizione delle frazioni fine e grossolana ad alta risoluzione temporale: STRAS (Size and Time Resolved Aerosol Sampler), IX Convegno Nazionale sul Particolato Atmosferico PM2020 in Lecce, Italy (14-16/10/2020).

[NO2] **Crova F.**, Bernardoni V., <u>Forello A.C.</u>, Valentini S., Valli G., and Vecchi R.: Caratterizzazione di emissioni non-exhaust da traffico in campioni di aerosol atmosferico separato dimensionalmente, IX Convegno Nazionale sul Particolato Atmosferico PM2020 in Lecce, Italy (14-16/10/2020).

[NO3] **Crova F.**, Bernardoni V., <u>Forello A.C.</u>, Valentini S., Valli G., and Vecchi R.: Caratterizzazione di emissioni non-exhaust da traffico in campioni di aerosol atmosferico separato dimensionalmente, 106° Congresso Nazionale - Società Italiana di Fisica, online, (14-18/09/2020).

[**NO4**] **Forello A.C.**, Bernardoni V., Calzolai G., Massabò D., Lucarelli F., Nava S., Pileci R.E., Prati P., Valentini S., Valli G., and Vecchi R.: Multi-time source apportionment. Un approccio avanzato per l'identificazione delle sorgenti di particolato atmosferico a Milano, VIII Convegno Nazionale sul Particolato Atmosferico PM2018 in Matera, Italy (23-25/05/2018).

[**NO5**] **Bernardoni V.**, Elser M., Valli G., Forello A.C., Valentini S., Bigi A., Fermo P., Piazzalunga A., and Vecchi R.: Source apportionment 3-D di aerosol urbano separato dimensionalmente mediante impattore multistadio, VIII Convegno Nazionale sul Particolato Atmosferico PM2018 in Matera, Italy (23-25/05/2018).

Acknowledgments

Ringraziamenti

Un grazie particolare a Roberta per aver reso possibile questo lavoro e per il sostegno che ho ricevuto durante tutto il dottorato. Grazie per spronarmi sempre quando sto per gettare la spugna. Grazie a Gianlu (il vero guru della ME-2) e a Vera per l'aiuto, la collaborazione, le discussioni illuminanti e il supporto. Mi spiace non aver goduto fino alla fine dei momenti di spensieratezza in ufficio, che permettono di ricaricare le batterie e imparare le citazioni dei film.

Grazie a Sara per tutti gli anni di chiacchierate, consigli e ginseng; sono contenta di averti avuto a fianco (in tutti i sensi) durante tutto il percorso. Grazie a Fede per la bella compagnia e la positività. È sempre un piacere lavorare e passare il tempo con voi.

Un grazie speciale alle mie amiche e ai miei amici, vicini e lontani (solo geograficamente), che mi hanno fatto sentire la loro presenza anche in questo periodo difficile per tutti, e con i quali condividerò in qualche modo questo traguardo. È difficile fare un elenco, ma mi riprometto di far sentire la mia gratitudine ad ognuno di voi. Grazie a chi c'è sempre e mi saluta dal balcone, a chi mi chiama Elis, a chi vive vicino al mare lontano dai cicianebbia, a chi festeggia con Magister e Fonzies. Grazie a Javier e Lucrezia per aver condiviso con me sei bellissimi mesi della mia vita e per essermi ancora vicini (Gracias a Javier y Lucrezia por compartir conmigo seis hermosos meses de mi vida y por seguir estando a mi lado).

Infine (non per importanza) grazie a tutta la mia grande famiglia. Grazie alla mia mamma e al mio papà per essere i miei punti di riferimento, per avermi insegnato a vedere sempre la luce in fondo al tunnel, per appoggiarmi in tutte le mie scelte. Se mi guardo indietro vedo una sola costante, e senza il vostro sostegno non ce l'avrei sicuramente fatta. Grazie ai miei fratelli Samuele e Gionata, fondamentalmente per rendere la mia vita più felice. Un pensiero speciale va ai miei nonni Rosario, Carmela, Nino e Virginia; in particolare, grazie a mio nonno Nino per avermi influenzato con il suo amore per lo studio e i libri, e a mia nonna Virginia per avermi insegnato a ridere della vita e semplicemente per essere stata così importante per me.