How to use efficiently airborne criteria pollutants and radon-222 in source apportionment: A self-organizing maps approach

Alessandro Zappi [a,*], Erika Brattich [b], Mariassunta Biondi [a], Laura Tositti [a,**]

[a] Department of Chemistry "G. Ciamician", University of Bologna, Via F. Selmi, 2, 40126, Bologna, Italy
[b] Department of Physics and Astronomy "Augusto Righi", University of Bologna, Viale Berti Pichat 6/2, 40127, Bologna, Italy

HIGHLIGHTS

- Hourly concentration data of eight gaseous pollutants and black-carbon were collected.
- Source profiling analysis based on self-organizing maps (SOM) was carried out.
- SOM Results were compared with meteorological parameters and radon-222 activity concentration.
- Six pollution sources were identified in the area of Savona municipality.

GRAPHICAL ABSTRACT
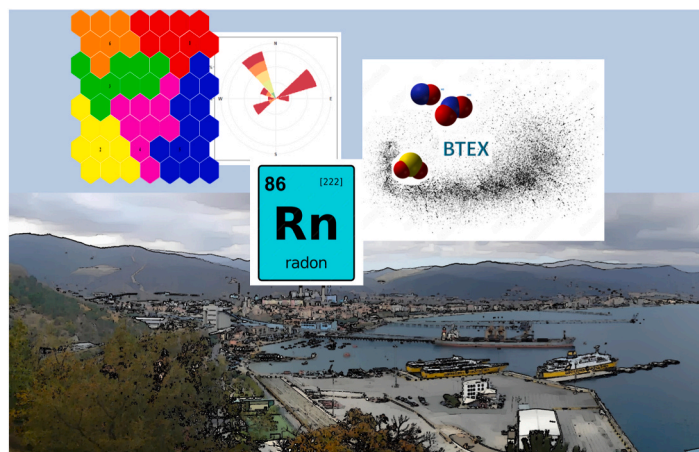
ABSTRACT

Pollutant source apportionment represents one of the fundamental activities in environmental science. Several efficient chemometric tools are available to the scope, mostly based on multivariate techniques and usually applied to aerosol chemical speciation data. In the present work, an alternative source profiling method is proposed, based on the self-organizing maps (SOM) algorithm. Moreover, the dataset used includes typical criteria pollutants and physical parameters related to airborne particulate matter widely used as a complement of aerosol source apportionment and largely available at a higher time resolution than bulk aerosol samplings, allowing the information on the dynamic behavior of the local airshed to be extended. In this work, data was collected at a coastal location in NW Italy, between January and July 2012. Hourly concentrations of typical gaseous pollutants (SO_2, NO, NO_2, benzene, toluene, (m-p)-xylene, o-xylene), black-carbon and particle number concentrations by an optical particle sizer (OPS) were collected. The dataset was integrated with radon-222 activity concentration and meteorological parameters to enrich and refine the information obtained by SOM computation as well as to improve the air pollution source localization. Despite the lower specificity of criteria

* Corresponding author. Via Selmi 2, 40126, Bologna, Italy.
** Corresponding author. Via Selmi 2, 40126, Bologna, Italy.
E-mail addresses: alessandro.zappi4@unibo.it (A. Zappi), laura.tositti@unibo.it (L. Tositti).

pollutants, the approach developed was capable of revealing distinct pollution sources such as the urban background traffic, the coal-fired power plant active at the time of the study, and the harbor, in agreement with previous PM-based source apportionment studies carried out locally, while enlightening peculiar dynamical conditions detectable at the sub-daily time scale. The application of the SOM algorithm, with the integration of meteorological parameters and atmospheric radon, proved to be very efficient in unveiling the air pollution sources.

1. Introduction

Source apportionment is one of the main objectives of environmental science, independent of the type of ecosystem or environmental matrix. Owing to environmental complexity and dynamicity together with pollution source multiplicity, several experimental approaches can be applied, covering either monitoring, usually within legally issued protocols as managed by Environmental Protection Agencies, or more detailed chemical speciation. Overall, source apportionment by receptor modeling represents a consolidated, though continuously evolving, approach in the case of airborne particulate matter, wherein highly accurate results are achieved at the cost of demanding and costly aerosol chemical speciation (Hopke et al., 2020; Mircea et al., 2020). In this framework, atmospheric monitoring mostly based on gaseous criteria pollutants usually play a complementary role on account of their parental association with secondary aerosol species, co-emission with aerosol primary species, and post-formation/emission aerosol processing in the complex multiphase troposphere (Corral et al., 2020; Perrone et al., 2022; Pöschl and Shiraiwa, 2015; Seinfeld and Pandis, 2016). The exploitation of gaseous criteria pollutants for source apportionment purposes can be, however, fairly ambiguous, due to their limited source specificity and variable reactivity, as expressed by individual atmospheric residence time (Hobbs, 2000). Moreover, daily, or even lower, sampling time resolution smooths or annihilates variability of criteria pollutants caused by short-scale circulation, masking transport processes useful for source identification/profiling and obliterating the associated chemical information.

Coastal sites represent an interesting case of overall environmental complexity. The land-sea discontinuity produces remarkable gradients in physicochemical properties, significantly affecting heat and matter exchange at the borders with considerable influence on local circulation and atmospheric composition. Depending on the latitude, coastal areas are historically heavily inhabited owing to favorable climatology and strategic trade conditions. The Mediterranean area is one of the most representative areas in this sense, being characterized by an extremely extended and topographically complex coastal line, heavily inhabited and economically developed since ancient times, with a major influence on air quality. Near-surface flow in complex terrain with coastal and urban influences is impacted by several different mechanisms including the large-scale pressure gradients and thermally driven flows generated by diurnal insolation (Leo et al., 2015; Millán et al., 2002). Thermally driven flows in such regions include sea-land breezes, slope and valley flows, and circulations induced by the urban heat island effects, mainly characterized by weak winds and conducive air mass highly localized recirculation. All of these flows have a significant impact on air pollution variability at the sub-daily time scale (Tan et al., 2021), and therefore on air mass aging (Millán et al., 2002). The understanding of near-surface dynamics in such conditions, however, is still limited, requires high time/space resolution data, and is still challenging to be accurately modeled due to the parameterized turbulence calculation in the planetary boundary layer (Song and Shao, 2023). In this framework, the use of tracers and radiotracers with consolidated and well-known sources offers reliable alternatives.

Radon-222 has a long and consolidated history as a valuable radiotracer, dating back to the early 1900s (Satterly, 1910; Wilkening, 1981, 2004; Wright and Smith, 1915). Due to its distinct emission source in the lithogenic materials such as soil and rocks, radon from parent ^{238}U–^{226}Ra in rocks and soil, constitutes a measurable constraint to transport, suggestive of the continental origin of air masses loaded in ^{222}Rn as compared to radon depleted areas, e.g., the upper troposphere or marine regions. As a result, the significant radon gradients among distinct areas make this species an excellent tracer of directionality in circulation patterns, thus highlighting their physical location. ^{222}Rn is also an efficient tracer of atmospheric stability (Chambers et al., 2015; Kikaj et al., 2019; Perrino et al., 2001; Williams et al., 2016) due to its physicochemical properties. It is, indeed, a radioactive noble gas exhaled from the ice-free terrestrial surface at an approximately uniform emission rate over the continental areas (0.72–1.2 atoms cm^{-2} s^{-1}), relatively water-soluble, and unreactive, whose concentration in the lower troposphere is controlled by the stability of the boundary layer as a result of local radiative budget in combination with the local wind field (Chambers et al., 2015). Its half-life of 3.83 d makes it a suitable descriptor for atmospheric processes at a scale between the local and the regional ones with a focus on low troposphere stability problems (Chambers et al., 2015, 2017, 2018).

Multivariate techniques are widely applied to many environmental problems with an emphasis on atmospheric science due to their ability in the identification and apportionment of pollutant emission sources. The so-called receptor models, indeed, are capable of extracting chemical fingerprints from complex chemical databases, conducive to emission sources, usually characterized by recognized species (tracers), wherein advanced techniques based on Positive Matrix Factorization (PMF), Chemical Mass Balance or others can support even the quantification of each emission source identified (Hopke et al., 2020; Mircea et al., 2020; Mooibroek et al., 2022; Tositti et al., 2022; Watson and Chow, 2015) or enable a comparison between several areas (Pietrodangelo et al., 2024; Veld et al., 2021). Moreover, large atmospheric datasets may contain latent information associated with physical variables, particularly those represented by specific meteo-climatic conditions and dynamics sensibly controlling the fluctuations in airborne compounds and particulate concentrations at the various time scales, from the sub-daily to the daily, weekly, seasonal ones, and more.

In this work, we propose an alternative strategy for data mining based on self-organizing maps (SOM) (Kohonen, 1998). SOM is a consolidated chemometric method allowing to cluster a series of observations based on an Artificial Neural Network (ANN) approach (Kohonen, 1998; Licen et al., 2020). Such a procedure aims at revealing statistically consistent sources of covariance among a complex database of atmospheric species (in this case, trace gases and airborne particulate) in connection with the controlling meteorological parameters. Source profiling analysis obtained by SOM was also evaluated considering radon-222 activity concentration. The use of ^{222}Rn in this framework is especially innovative due to its discriminant potential thus enriching and consolidating SOM outputs in treating complex and dynamic systems like coastal airsheds.

2. Materials and methods

2.1. Dataset

The dataset was collected in the small town of San Genesio, a fraction of Vado Ligure near Savona, a mid-size coastal town in NW Italy. Fig. 1 shows a map of the sampling area, while Fig. S1 shows a topographic map (Tarquini et al., 2023) of the surrounding region (Northern Italy).

This study area has been already presented in several previous articles by this group, all concerning airborne PM_{10}, source apportionment, as well as DNA and bacterial luminescence, and its environmental implications (Morozzi et al., 2021; Palladino et al., 2021; Tositti et al., 2018). Such papers showed that the main emission sources active in this coastal district were naval traffic due to the port activities, marine emissions (sea salt and biogenic secondary species), road traffic, a coal-fired power plant active in the period of data collection, and, to a lesser extent, biomass burning from wood stoves.

For the present study, hourly data of meteorological variables, gaseous pollutants, particulate matter number densities, and black carbon (BC) were collected together with ^{222}Rn atmospheric concentration activity. The sampling campaign was carried out from January 11 until July 23, 2012. The choice of the 6-months study period aimed to cover the typical seasonal oscillations and variability, in this case from winter to summer, in order to maximize meteorologically associated variance and its influence on physical and chemical variability of air pollutants. Data were collected as follows.

- Meteorological variables: air temperature, air relative humidity, barometric pressure, wind speed, wind direction, and accumulated rain were measured and recorded using a Davis Vantage Pro2 (Davis Instrument, Hayward, CA 94545, USA) meteorological station.
- The concentrations of eight gaseous species: SO_2, NO, NO_2, benzene, toluene, (m-p)-xylene, o-xylene, and ^{222}Rn were kindly provided by the Liguria regional Environmental Protection Agency (ARPAL). ^{222}Rn was measured with a RAD7 Radon Detector (Durridge, Billerica, MA, USA), while details on the air quality monitoring conducted by ARPAL are provided at https://www.arpal.liguria.it/tematiche/aria/monitoraggio-e-inquinanti.html (in Italian)
- Equivalent black carbon (EBC) was measured using a micro-aethalometer microAeth AE51 (AethLabs Inc., San Francisco, CA 94110, USA)
- Particle number density was collected by an optical particle sizer (OPS) Profiler Model 212 (MetOne Instruments, Inc., Grants Pass, OR, USA). Number density is determined in the following size bins: 0.3–0.5; 0.5–0.7; 0.7–1.0; 1.0–2.0; 2.0–3.0; 3.0–5.0; 5.0–10.0;

>10.0 μm (Brattich et al., 2020). For each channel, the number of particles of the corresponding size is recorded. This data was recorded from April 5 until the end of the field campaign.

All the instruments were installed on a mobile station provided and operated by the local Environmental Protection Agency.

2.2. Chemometric analyses

To reveal data behavior and properties and extract the correlations among variables, a chemometric approach based on the self-organizing maps (SOM) was applied. All computations were performed in R environment (R Core Team, Vienna, Austria), while descriptive analyses were performed with the "Openair" R package (Carslaw and Ropkins, 2012).

SOM (Kohonen, 1998; Licen et al., 2019) is a consolidated chemometric method, based on ANN, whose aim is to cluster observations in an unsupervised way (i.e. no *a priori* knowledge is needed regarding experimental data classification). Moreover, this approach can handle large datasets and even non-linear problems. The graphical output is a 2D-map of SOM-units (represented as circles or hexagons), whose overall number is a function of the number of experimental observations available. Each unit represents vectors, whose length is equal to the number of experimental variables, that are iteratively updated during the SOM training phase to better adapt to the data. The algorithm initializes the calculations from a random point, attributing random values to the vectors; subsequently, every experimental observation is presented to each unit and assigned to the best-matching unit (BMU), based on Euclidean distance. The units are then updated to the mean vector of the assigned observations and the computation starts again. This process of allocation and update is iterated for a given number of epochs (up to thousands), or until convergence is reached and no further changes are observed in the output. In the final SOM map, similar observations are allocated in the same BMU so that close units report similar information. The unit-vectors can then be further grouped by cluster analysis to efficiently reveal data behavior and properties. Clustering was carried out by K-means cluster analysis, using the minimum Davies-Bouldin
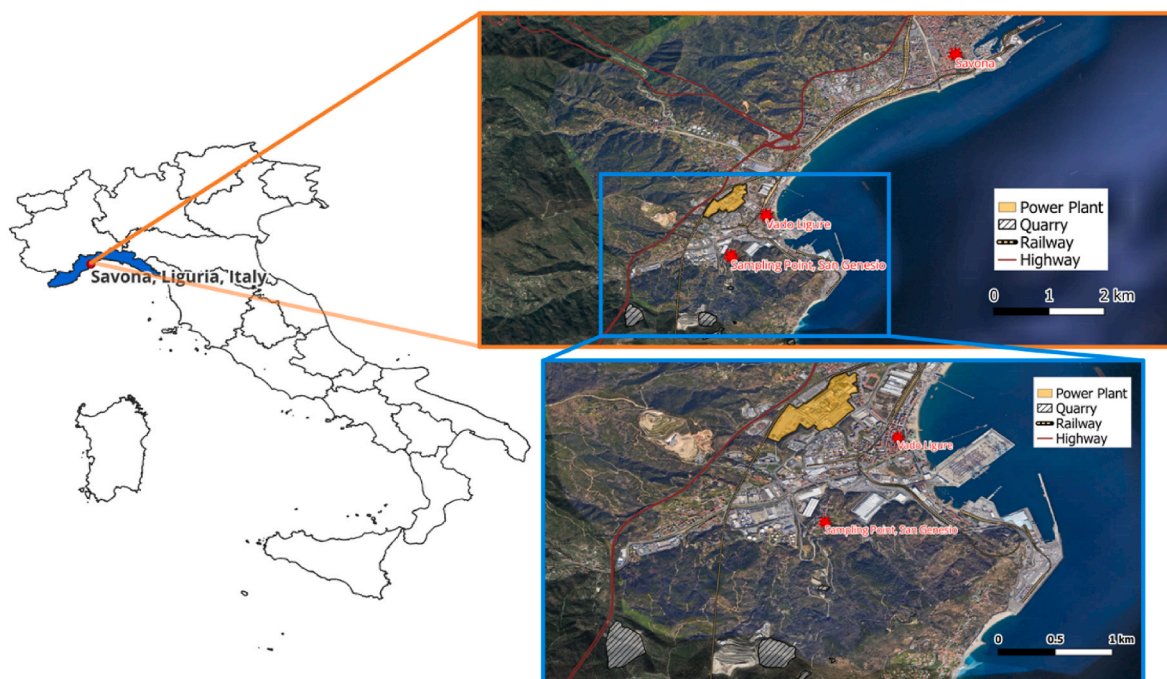
Fig. 1. Sampling area, San Genesio near Savona town, in Northwestern Italy. The maps on the right show the location of the sampling site and some major pollution sources in its vicinity. Maps were created with QGIS v.3.34.1 software (https://qgis.org/en/site/).

(DB) index (Davies and Bouldin, 1979) to choose the optimal number of clusters. The "map" computed by SOM, therefore, reports the temporal trend of data, and must not be confused with a geographical map.

The SOM training phase requires defining the map size at the beginning of the computation. Several decision criteria have been proposed to fulfill this task; in the present work, we used the one proposed by Nakagawa et al. (2020). It calculates the best number of units (m) as a heuristic function of the number n of observations (eq. (1)):

$$m = 5\sqrt{n} \tag{1}$$

The actual number of units is chosen as the integer closest to m. It has been observed that a rectangular, rather than square, SOM can better describe the training data, allowing a distribution along a dominant axis (May et al., 2010). Therefore, both side lengths of the map are set as the proportion of the first two eigenvalues obtained by a principal component analysis (PCA) previously computed on training data that best fits m (Hentati et al., 2010).

SOM maps were calculated with the R package SOMEnv (Licen et al., 2021), which also provides a graphical user interface (GUI) running on all browsers.

3. Results and discussion

All the meteorological variables and the basic statistics (Tables S1 and S2) for the analyzed period (11 January – July 23, 2012) are reported in the Supplementary Information (SI). During the sampling campaign, the temperature (Table S1) ranged between −3.5 and 33.1 °C (mean: 15.0 °C, σ: 7.3 °C), with the monthly absolute minimum registered in February and the absolute maximum in July (Fig. S2). The overall daily variation reflects the typical Mediterranean climate, with daily maxima registered in the afternoon (between 13:00 and 16:00) and minima early in the morning (3:00 to 6:00). Relative humidity (RH) (Table S1 and Fig. S3) was in the range 20–91% (mean 63%, standard deviation 18%), with monthly absolute minimum in February and absolute maximum in June. No deviations from the Mediterranean climate were observed in the daily trends, with maxima during the night (23:00 to 6:00) and minima during the afternoon (13:00 to 16:00). Atmospheric pressure was always in the range 1000–1039 hPa (Table S1 and Fig. S4) with an average daily trend presenting two maxima at 12:00 and 0:00. Despite an anomalously huge snowfall in February, the sampling period was characterized by limited precipitation, associated with a severe weather event early in April with 100 mm rainfall and several minor events scattered throughout the campaign. The wind rose reported in Fig. S5 shows a prevalence of weak to moderate NW winds. The Liguria region is characterized by a distinct seasonal pattern in wind circulation, with prevailing lower wind speeds in the warm season and more intense winds in the cold one. The wind rose presented in Fig. S5 is in perfect agreement with that shown in a previous publication (Burlando et al., 2017) that analyzed the wind climatology based on observations from a network of 15 (ultra)sonic anemometers in the Liguria region, including one in the port of Savona Vado Ligure in the vicinity of the study site. Liguria, and in particular the main valleys, is characterized by a special kind of downslope winds called gap winds, occurring when the air to the North of the Maritime Alps and the Apennines gets to the top of the mountains and, after becoming cooler and denser than the maritime air above the warmer Mediterranean Sea, flows down the sloping surfaces towards the sea. Under peculiar meteorological conditions, these winds can be locally reinforced by larger-scale low-pressure systems, as during secondary cyclogenesis events in the Gulf of Genoa (Trigo et al., 2002). According to the local wind intensity, this means that the topographic flow is mostly confined to come from the Po Valley and alpine regions, but eventually can be extended to strong winds from France and Swiss regions. However, based on the wind rose presented in Fig. S4, as well as from the wind climatology by Burlando et al., such events are rare as the wind velocity is mostly in the range 0.5–10 m s^{-1}.

Radon-222 (Fig. 2) shows the typical diurnal variation highlighted in previous works (Chambers et al., 2015; Kikaj et al., 2020), with maxima registered early in the morning, around 7:00, and minima during the late afternoon, around 19:00, in association respectively with maximum and minimum stability conditions independently of the season (Fig. 2b). No differences were observed in ^{222}Rn diurnal trend between cold (Fig. 2c) and warm (Fig. 2d) seasons. It is well known, indeed, that ^{222}Rn, which is exhaled by soil and rocks, accumulates near the ground during the night to be subsequently redistributed and dispersed during the day by thermal convection and turbulence i.e. by an unstable boundary layer (Chambers et al., 2015, 2017, 2018, 2019; Kataoka et al., 2001; Kikaj et al., 2020; Sesana et al., 2003; Tositti et al., 2002; Turekian et al., 1977). Considering the whole sampling period (Fig. 2a), the radon activity showed a maximum in January and a minimum in April, with a secondary maximum in June. The polar plot reported in Fig. S6 shows two main sources for ^{222}Rn, one close to the sampling point connected to weak and local circulation patterns, and one associated with weak-to-moderate NW circulation from the inland. This second source might be ascribed to gap winds previously described, which may have transported radon to the sampling point from the Alps and the Po Valley.

SOM was thereafter computed using gas and BC concentrations with and without PM (Particulate Matter) number densities, due to the shorter sampling period of OPS data (from April 5 to July 23). Meteorological variables and ^{222}Rn data were excluded at this stage to use them *a posteriori* to gain a deeper insight into air pollution phenomenology in a cascade mode. Therefore, two distinct SOMs were computed, the first one considering only chemical variables (gases and BC), covering the whole sampling period (3549 valid observations, without empty cells in the data matrix), and the second one with all the formerly mentioned variables plus number densities covering only the warm season (1954 valid observations).

3.1. SOM with plain chemical variables

The first SOM was calculated on observations of air pollutants (gases and BC) recorded from January 11 to July 23. In this case an 11 × 7 map was trained. The number of epochs for the training phase was set to 100, based on Gaussian neighborhood. At the end of the computation, all units were populated by at least one observation. Fig. 3 shows the heatmaps representing the data distribution on the map for each modeled variable (Licen et al., 2021). The content of the hexagons represents the basic statistics of the observations populating the units (i. e., white: lower outliers; grayscale: quartiles; black: upper outliers). From a preliminary visual evaluation of heatmaps in Fig. 3, based on similar distribution patterns some correlations between nitrogen oxides, and among the xylenes (m,p-xylene and o-xylene) can be assumed as likely. The other variables do not reveal apparent correlations, as shown by different quartile profiles on the heatmap. However, these visual considerations represent only the starting point of SOM results elaboration and will be treated in more detail in the following sections.

Based on the frequency distribution of each parameter, SOM units can be further rationalized with suitable grouping techniques. The clustering computation was performed in the range from two to eight clusters, and the best solution resulted with six clusters. The right portion of Fig. 3 shows the SOM map with units distributed into the six calculated clusters.

The obtained cluster distribution enables us to obtain more details on the behavior of the various parameters. For each cluster, boxplots of the original variables were thus calculated, allowing highlighting the peculiar conditions characterizing each cluster. Such boxplots are reported in Fig. 4. Variables in Fig. 4 are autoscaled: the zero line corresponds to the average value of each variable in the overall dataset. Therefore, boxplots whose median is significantly higher than zero indicate a higher median concentration for the corresponding species, while medians significantly lower than zero indicate a depletion of that species.
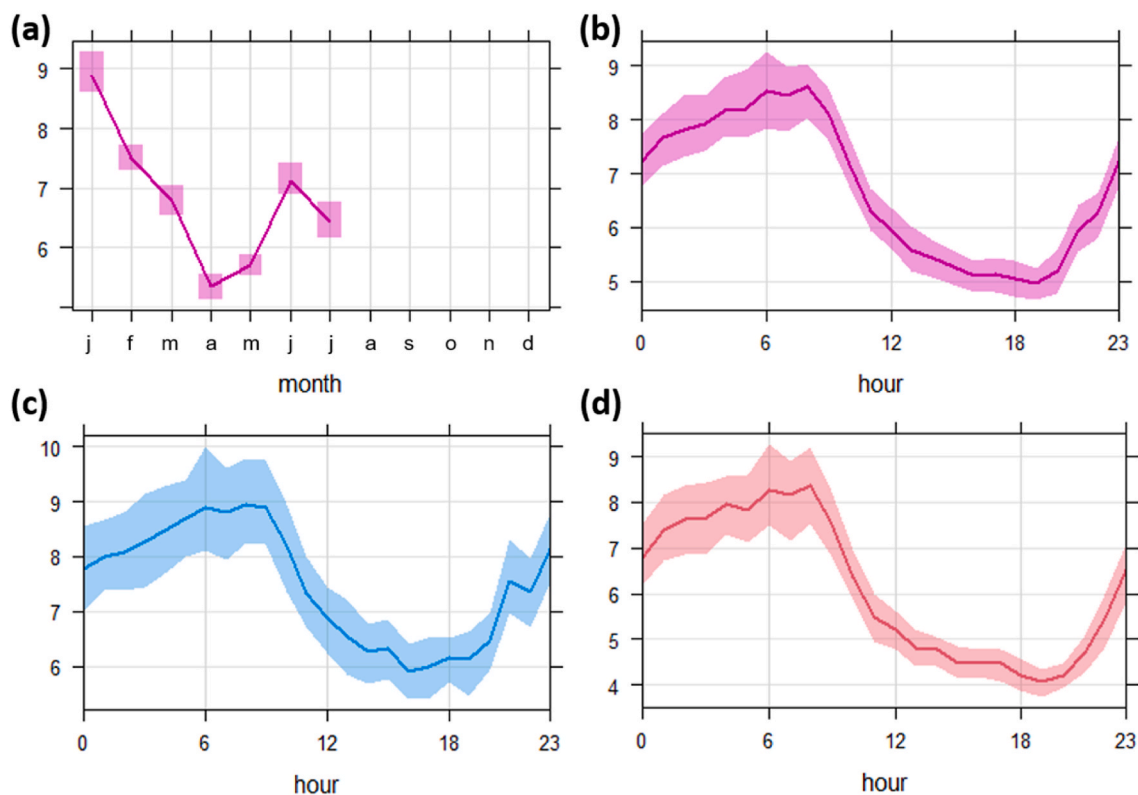
Fig. 2. Average time trends of radon activity (in Bq m^{-3}): a) monthly trend; b) hourly trend over the whole sampling period; c) hourly trend in the cold season (between January 11 and April 5); d) hourly trend in the warm season (between April 5 and July 23). The line represents the average while the shadowed area represents the 95% confidence interval.

Based on Fig. 4, Cl6 (in orange, in the top-right corner of the SOM) is the one with the lowest concentrations of all species. It can be considered as a "local-background" cluster representing the local airshed not significantly affected by particular sources in the immediate vicinity of the receptor site. Cl1 (in red, in the bottom-left corner) represents the highest air pollution conditions observed during the campaign in terms of nitrogen oxides, toluene, and xylenes. However, intermediate situations highlight higher values for single pollutants, which suggest multiple emission sources, including some with simpler emissive spectra. Indeed, the monitored species are criteria pollutants, typically shared by a range of emission sources, thus requiring further complementary information to allow source identification. Given the limited diagnostic potential of criteria pollutants, what affects the degree of population of such clusters might be due to wind direction and speed, location of the source or their physicochemical behavior. Overall, the SOM pattern shows how the mean concentration of the pollutants considered increases starting from the top-right corner down to the bottom-left. The Cluster in the top right corner (Cl6) is characterized by a lower concentration of all analytes, connected with the periods of lower pollution, while the one in the bottom left portion of the map (Cl1) has the highest pollution level, with almost all species above the campaign overall average concentration value, except for benzene and (m,p)-xylene. The other clusters show intermediate situations, with only one or a few pollutants having higher concentrations than Cl1 at the same time. Cl2 (in yellow), indeed, is characterized by slight pollution levels, showing only a minor peak of concentration of SO$_2$. Cl3 (in green) reveals a significant contribution of BC, without other significant contributions. Additionally, on account of its peculiar quartile behavior (Fig. 3), it is confirmed that BC is weakly correlated with the other pollutants. SO$_2$ and BC are generally considered, together, as marker of fuel and coal burning. However, they do not easily show a net correlation: firstly, gaseous and particulate states are not equivalent though both species are

primary. This means that the way they are measured might affect correlation. Moreover, their concentration in the atmosphere is influenced by way to the open air from the emission points (e.g., from the furnace to the stack through different mitigation stages as in coal fired power plants) or even by the presence, usually neglected, of fugitive emission, i.e. not the stack emissions only, usually very high, but also the leaks at low level from a plant, as a hole. Cl4 (in purple) and Cl5 (in blue) reveal more complex air pollution patterns. Indeed, Cl4 shows higher concentrations of the aromatic compounds (except toluene), with a contribution of NO$_2$. Cl5, instead, shows higher levels of NO$_2$ and SO$_2$, with a negligible difference of NO from the overall mean, suggesting an association with aged air masses. The simultaneous occurrence in Cl4 and lower relevance in Cl1 confirm the correlation between benzene and (m-p)-xylene, from distinct sources.

A further step in SOM processing allows the temporal behavior of the analyzed species as function of meteorological fluctuations to be better defined. To the scope, Fig. 5a shows the hourly observations count for each cluster, and Fig. 5b shows the daily distribution of the clusters along the whole study period. The combination of Fig. 5a and b reveals therefore the daily distribution of clusters computed by the SOM procedure. However, to better understand the variability within each SOM cluster, boxplots of the meteorological parameters and a wind rose were computed, with the aim of better reconstructing the origin of the pollutants (Fig. 6).

Considering the daily distribution of clusters (Fig. 5) and the meteorological (Fig. 6a) and anemological (Fig. 6b) influence on air pollutants, several considerations about the SOM clusters and ^{222}Rn behavior can be drawn.

Cl1 (387 observations), showing the highest pollution levels, concerns the cold period (January to March, Fig. 5b). It is characterized by the lowest median air temperature, compared to the other clusters, the highest barometric pressure and highest relative humidity. Winds are
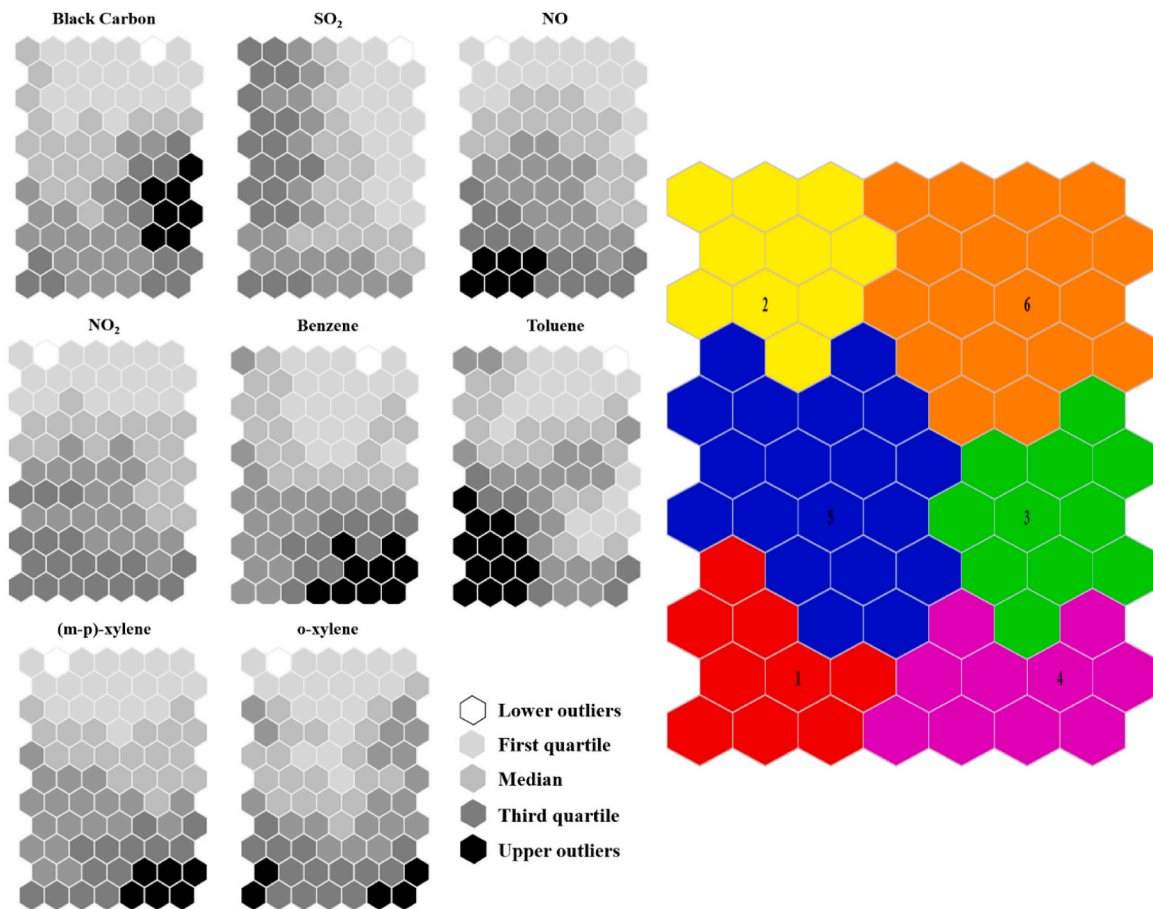
Fig. 3. Left portion: heatmap of the quartile distribution of chemical variables; right portion: SOM map for chemical data distributed across six clusters. Cluster numbers are placed into the cluster-centroid unit.
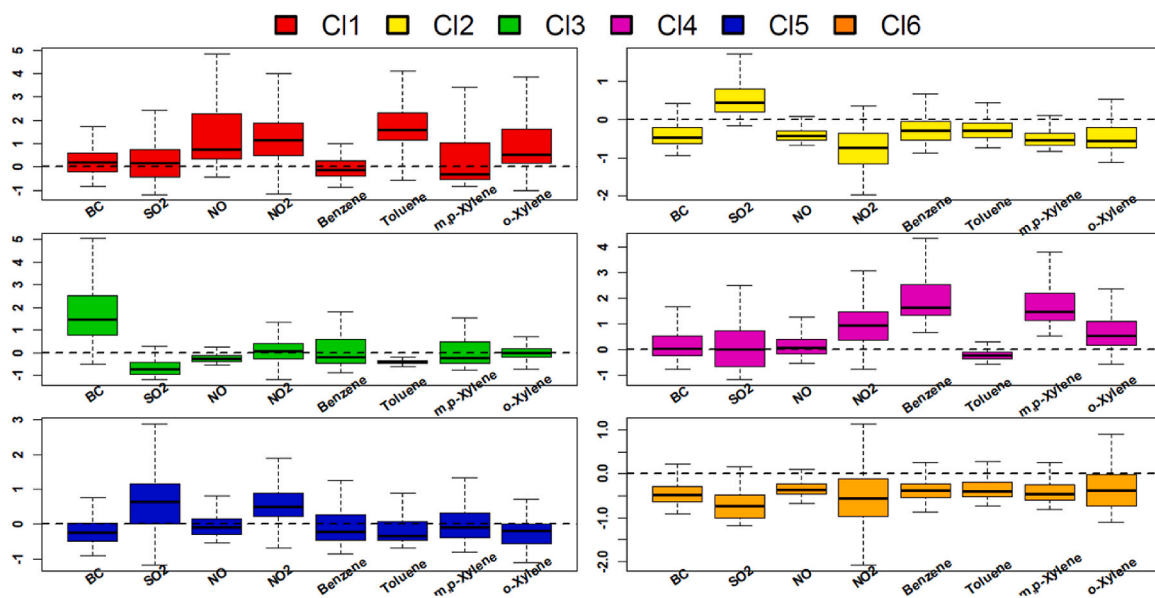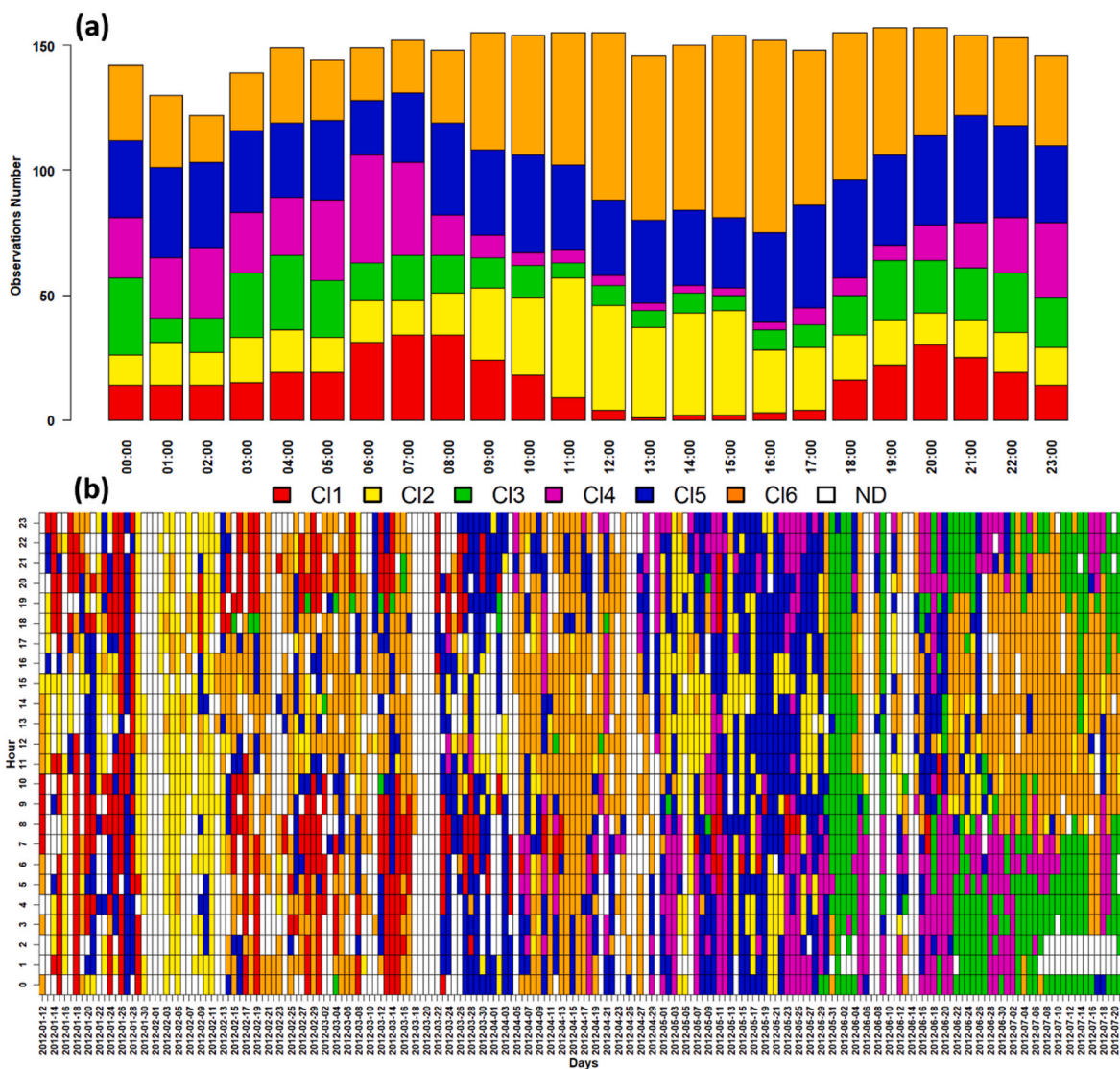

Fig. 4. Boxplots of chemical variables based on the cluster division. Colors are based on the clusters calculated for the SOM as depicted in Fig. 2. The black dashed line indicates the average values of the auto-scaled parameters in the overall dataset.

Fig. 5. a) Hourly clusters distribution with the corresponding observations counts; b) daily clusters distribution along the whole sampling period: each hour (observation) of each day is assigned to a specific SOM cluster. White cells (ND in the legend) represent hours for which the observation was not available.

mainly from W and NW, i.e., from the continental side of the district investigated, lying on a coastal area characterized by a complex topography. Moreover, most of the observations of this cluster (270) are nocturnal, from 19:00 to 8:00 a.m. (Fig. 5a). Therefore, westerly wind observed for this cluster can be related to local circulation pattern influenced by the nocturnal boundary layer and the behavior of the continental landscape. The more intense winds from NW are most likely associated with larger-scale katabatic flushing of the Po Valley region. This cluster describes the air pollution accumulation in winter, particularly NOx, toluene, and o-xylene, occurring during the night and at least partially, deriving from inland. The nighttime character of Cl1 and the wind direction from NW is supported by the high concentration of ^{222}Rn observed for this cluster and might be assimilated to "nocturnal air drainage" mentioned in a previous work (Wilkening, 1981). Such drainage may also include influences from the nearby freeway tunnel entrance/exit.

Cl2 (553 observations) is independent of the season; its observations are, indeed, distributed both in the cold and in the warm season. Moreover, the atmospheric temperatures show a wide range, and the median barometric pressure is close to that of the other clusters. It is characterized, however, by lower median relative humidity compared to the other clusters, and by winds from NW directions also up to 5.8 m s^{-1},

the highest values recorded in the whole period. These conditions could represent air circulating anticlockwise around the Po Valley topography, or foehn wind events from the Alps. Foehn wind events might favor the low relative humidity, and lower than average pollution (for most species). The wind rose of this cluster is also very similar to that of Cl1, but Cl2 is, on average, characterized by much lower pollutant concentrations. Such observation might be explained by the distinct diurnal pattern of the two clusters and the distinct mixing height during daytime and nighttime hours. Indeed, Cl1 appears mainly associated with the nocturnal hours when the mixing height is very low while Cl2 is instead linked with the well-mixed boundary layer typical of the daytime hours. Most observations for this cluster are concentrated in two periods, from 3 to 11 February and in the entire month of May (Fig. 5b). This cluster is also characterized by higher SO$_2$ concentration suggesting that air masses featured by this cluster are affected by man-made industrial sources shortly upwind the receptor site, in particular a coal-fired power plant active at the time of the sampling campaign (Langner and Rodhe, 1991; Morozzi et al., 2021). The lower concentration of ^{222}Rn revealed by this cluster is associated with a prevalence of morning observations (Fig. 5a), when radon tends to dissipate due to increasing instability, or to the observed stronger wind.

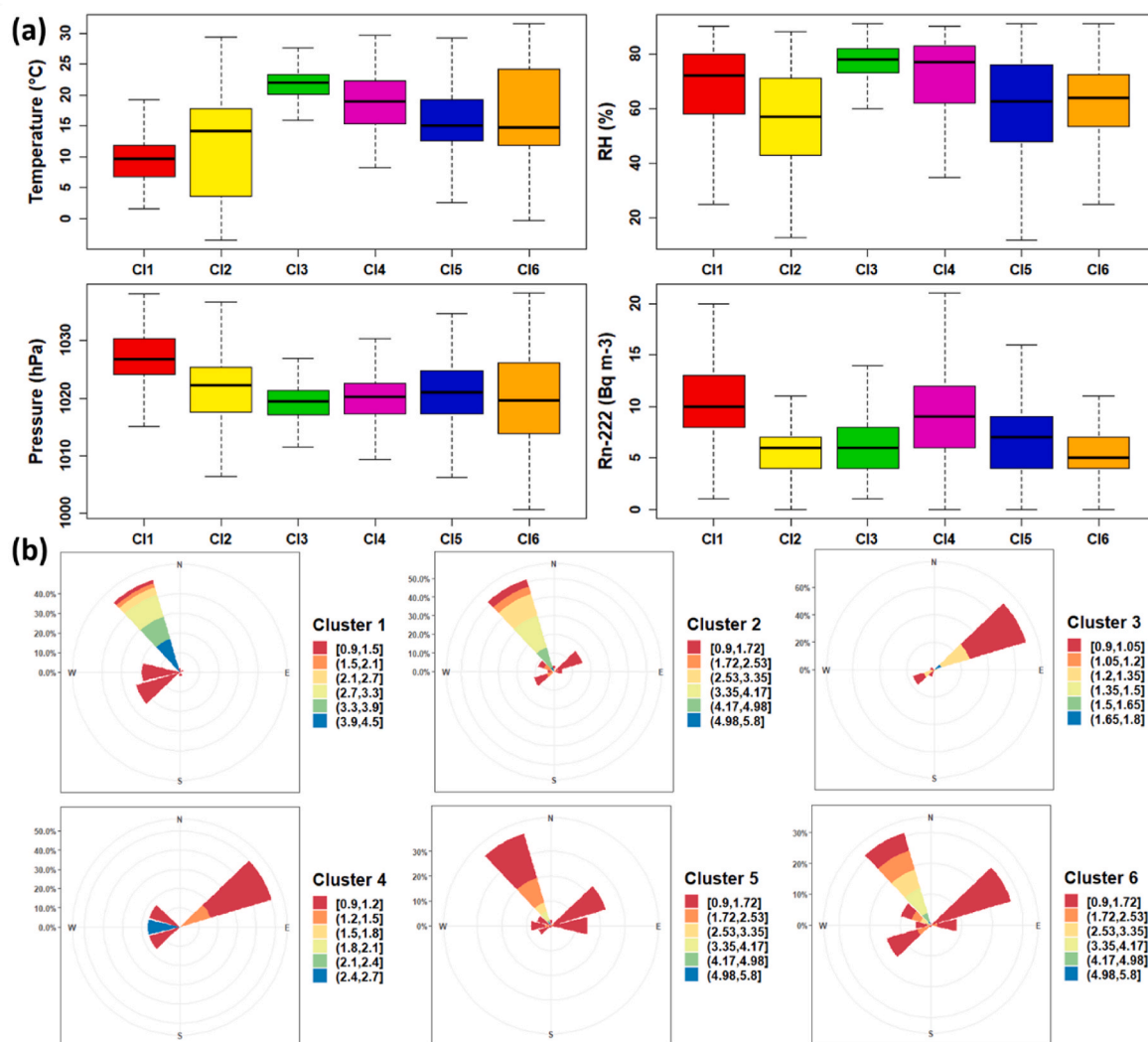Cl3 (385 observations) is a purely warm-season cluster. It is indeed

Fig. 6. a) Boxplots of the meteorological variables (temperature, relative humidity, pressure) and activity concentration of ^{222}Rn; b) wind roses for each cluster.

characterized by higher air temperature and relative humidity and lower barometric pressure, within very narrow variability ranges, as compared to the other clusters. It features low ^{222}Rn concentrations and high BC levels. This cluster is characterized by slow NE winds, addressing to the sea (higher relative humidity, negligible ^{222}Rn outgassing), excluding its association with strong synoptic forcing. Thus, although its nighttime prevalence excludes summer sea-breeze circulation, we cannot rule out its association with other breeze regimes typical of the night conditions. NE wind direction also intercepts Savona harbor emissions. Therefore, the significant occurrence of BC can be ascribed to ship emissions, (Chambers et al., 2019; Turekian et al., 1977). It is interesting to note the occurrence of a single episode, from 31 May to 4 June, almost completely assigned to Cl3 (Fig. 5b). Such an event might be due to the contingency of a north-easterly flow and the prolonged presence of several ships at berth in the harbor.

Cl4 (390 observations) is typical of the warm season, the median temperature, humidity, and pressure are like Cl3, although with wider ranges. However, ^{222}Rn activity concentration is higher than for most of the other clusters, except Cl1, while the most concentrated stable species in this cluster are benzene, (m,p)-xylene, o-xylene, and NO$_2$. All these species are characteristic of road traffic (Tan et al., 2021), and their presence in this cluster can be again ascribed to local circulation. Indeed, besides a weak NE wind component conducive of a local origin (0.9–1.5 m s^{-1}), Cl4 is characterized by westerlies (up to 2.7 m s^{-1}), which

transport air masses from the highway and inland in agreement with the higher concentrations of ^{222}Rn in the cluster. The low wind speeds with occasional faster flows down valley from the west suggest that this cluster is dominated by katabatic flow conditions (stable nocturnal shallow mixing conditions). This cluster therefore suggests the influence of road traffic. The higher correlation in this cluster between aromatic species and NO$_2$, rather than with NO, suggests a relatively aged air mass where primary NO tends to be converted to secondary NO$_2$, especially in the low wind speeds conditions characterizing this cluster.

Cl5 (810 observations) is independent of the season with observations scattered throughout the day (Fig. 5), no peculiar meteo-climatic conditions are observed, while a slight increase of ^{222}Rn is present. The winds are weak from all directions (to the utmost, they reach 3.4 m s^{-1} from NW), while the most significant pollutants are NO$_2$ and SO$_2$. These species can be ascribed to fossil fuel combustion (Langner and Rodhe, 1991; Prospero et al., 1983), therefore this cluster represents the local recirculation of relatively aged pollutants.

Cl6 (1041 observations), finally, represents periods characterized by lower pollution levels. It is independent of the season, features no distinct atmospheric conditions with highly variable wind direction, nor a high concentration of ^{222}Rn. This cluster therefore represents efficient redistribution and dispersal of pollutants resulting in better-than-average air quality. The wind rose of this cluster suggests the prevalence of low wind speeds from both NW and NE directions. The NE

direction from the Ligurian Sea is also shared by Cl3 and Cl4, which however are characterized also by stronger wind speeds from this direction. Conversely, Cl6 appears mostly connected to reduced wind speeds, pointing to a very local origin (order of 1–2 km or less) which possibly explains, at least partially, the low ^{222}Rn contribution. In addition, this cluster's prevailing daytime contribution suggests the connection with intense mixing and dilution of the emissions. The NW direction, instead, where most of the pollution sources are located, is shared with Cl1 and Cl2. These clusters are mostly associated with higher wind speeds from this direction. In addition, as previously described, the diurnal patterns of these clusters are distinct, with Cl1 and Cl2 mostly present at night and Cl6 mostly present at daytime when mixing is much deeper.

This strong variability in air pollutant concentration at the sub-daily scale reflects the combination of meteorological variability, complex topography of the Ligurian coast, strongly affecting circulation patterns from the local to the regional scale, emission sources modulation, and different reactivity (atmospheric lifetime) of the single chemical species due to intrinsic molecular properties and to photochemical conditions. Such a circumstance is roughly conducive to a coastal breeze regime wherein morning flows from land seaward promote pollution seaward dispersal from local sources (industrial settlements, fossil fuel power plant, traffic and shipping, dense conurbation), leading to air quality transitory improvement over the urban airshed, while after midday air masses rich in ship emissions, secondary and aged pollutants from the sea with typical shipping markers and/or enriched in secondary species such as ozone for example, are returned to the coast (Fortezza et al., 1993; Millán et al., 2002).

The reported results, however, are in good agreement with those reported in the previous work from our group (Palladino et al., 2021), carried out on daily PM$_{10}$ chemical speciation data. The PMF model based on those data revealed the presence of seven source factors, that can be compared with the clusters obtained here by SOM i.e. with a completely independent methodology and dataset. The present SOM did not find "crustal" and "sea spray" sources due to the absence of specific markers in the dataset. However, both approaches recognized "traffic" (Factor 2 in the previous PMF study and Cl4 in the present one), "naval-maritime" (Factor 5 and Cl3), and "coal burning" (Factor 6 and Cl2) sources. The previous PMF study also identified two factors (3 and 4) described by secondary NO_3^- (Factor 3), SO_4^{2-} and NH_4^+ (Factor 4). The secondary nature of these factors can be compared with what reported for Cl5, representing locally aged air masses. Cl1 and Cl6, instead, cannot be compared with any factor of the previous PMF study. Both approaches agree in terms of emission source diagnostics, due to co-emissions of gaseous and particulate components in all the emerged sources. It has to be noted that PMF data and SOM data were collected during the same campaign and the same time lag. Nevertheless, chemical speciation data and sampling time resolution used in the two independent elaborations we performed, address different aspects of the same emissive phenomenology in the studied airshed. The complementarity of the two studies enriches the understanding of chemical (and physical) phenomenology. Noticeably, while PMF focuses on the specificity of chemical tracers with robust multivariate methods, SOM seems able to shed light on the dynamics of emissions, enabling a reasonable capacity of disentangling emission source contributions of criteria pollutants.

3.2. SOM with chemicals and OPS data

As already mentioned, the OPS data were available only in the second part of the campaign, consistently limiting SOM computation and preventing the comprehensive evaluation of air quality behavior. This paragraph thus is exclusively devoted to warm-season data discussion.

In this case, an 8 × 7 map was trained, and the number of epochs for the training phase was set to 100, using Gaussian neighborhood as in the previous elaboration. At the end of the computation, all units were
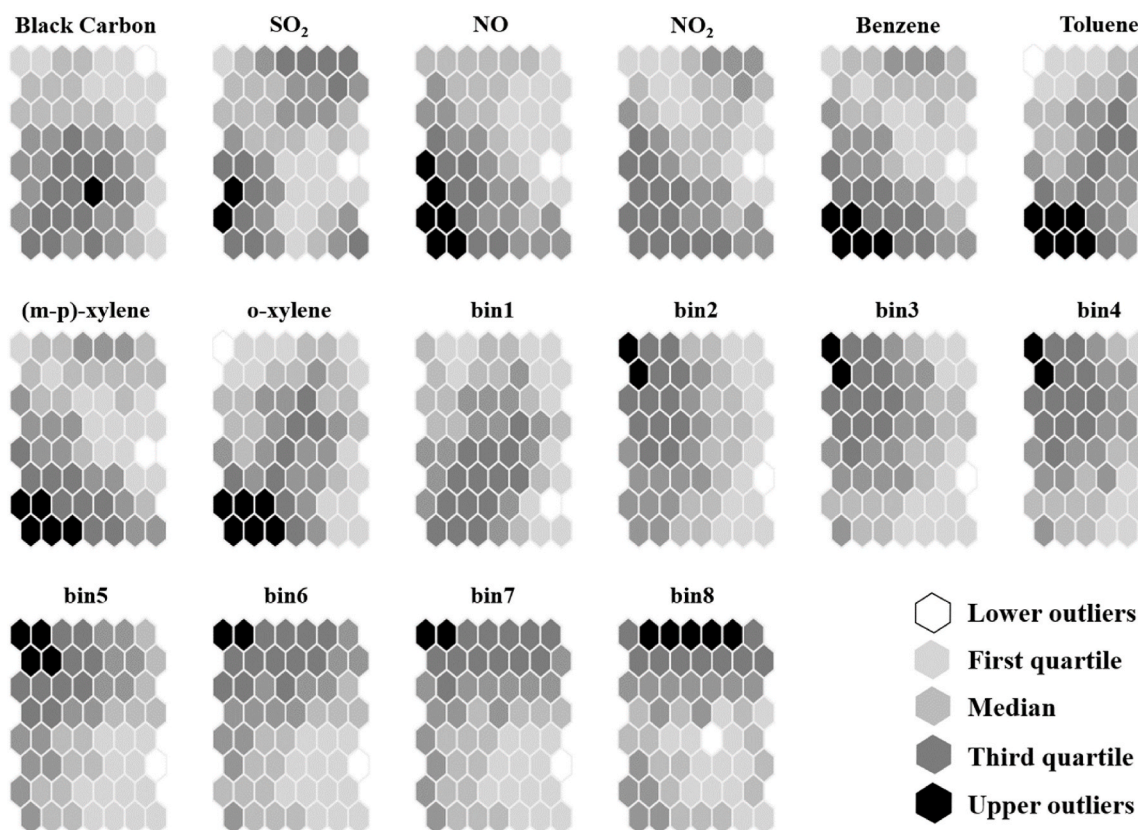

Fig. 7. Heatmap of the quartile distribution of chemical and OPS variables.

populated by at least one observation. The heatmaps in Fig. 7 show a different behavior of air pollutants compared to that of Fig. 3, which can be ascribed both to the specificity of the season covered (previously not resolved from the cold one) and to the inclusion of aerosol information from the OPS to the dataset. In this case, the correlation between compositional parameters is enhanced, in particular between two couples of organic species: benzene and (m,p)-xylene, and toluene and o-xylene; the profiles of both couples are, indeed, almost identical. The inorganic gaseous species reveal a limited correlation, with the highest correlation between NO and NO_2, due to their "genetic" connection. NO_2 seems partially correlated also to benzene and (m,p)-xylene which suggests a local source due to the relatively short lifetime of the species involved in the temperate Mediterranean conditions. NO_2, indeed, though a secondary species (only a few percent fractions being initially released at any combustion/high-temperature source compared to the highly reactive NO), can be as reactive as the aromatic species to which it is correlated (Friedrich et al., 2021). BC, instead, as already discussed in paragraph 3.1, does not show any apparent correlation with any other parameters as it can be attributed to a series of combustion sources active across the Savona airshed, unless using suitable constraints such as directionality at the sub-daily scale or seasonality (e.g., absence of

domestic wood heating in the warm season or modulation/trends in maritime traffic) and/or ^{222}Rn.

OPS data, finally, shows good to excellent correlation among the coarse fraction bins, likely associated with their buoyancy (distinct from sub-micron particles), though with an even more limited correlation with bin8 (>10 μm), reasonably due to the low population of this fraction. Bin1 (0.3–0.5 μm), instead, largely representing the accumulation mode, seems uncorrelated with all the others, because of its significant involvement in secondary aerosol formation (Raes et al., 2000) and in aerosol aging (Morozzi et al., 2021; Palladino et al., 2021; Tositti et al., 2018). Such lack of cross-correlations among macro-pollutants can be ascribed to the absence of specificity typical of criteria pollutants at the source, i.e., they are emitted simultaneously by multiple sources and recirculated across the same airshed. Nevertheless, their different emission sources might be distinguished based on wind direction and speed, distinctively affecting detection at the receptor site in terms of location/distance from the source and of their chemical kinetics (i.e. reaction rates, typical of each chemical species) all affecting their atmospheric residence time, but also the buildup of secondary species, often including a change of phase as in the case of submicron particles (e.g. the accumulation mode among the OPS fractions described above).
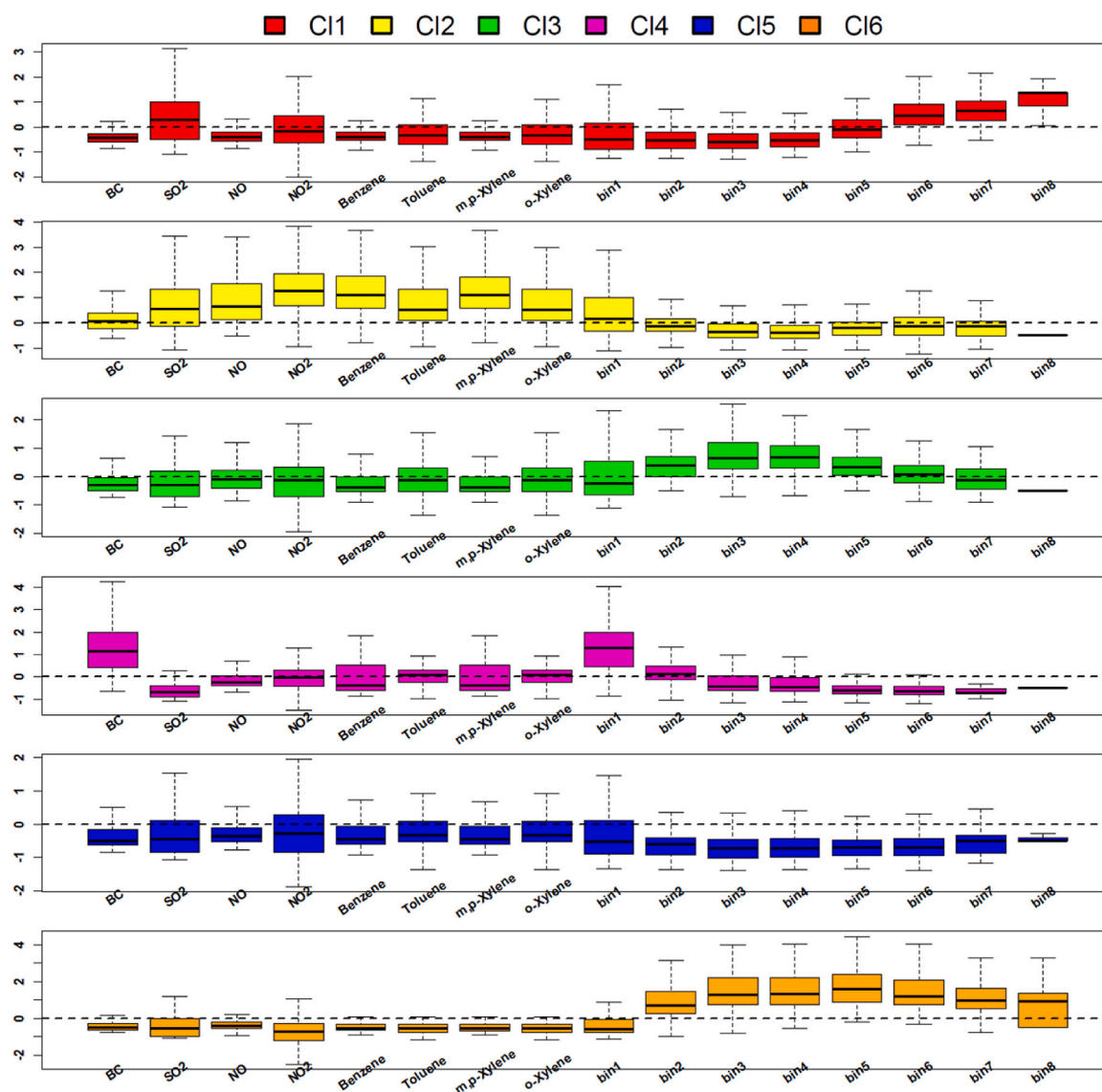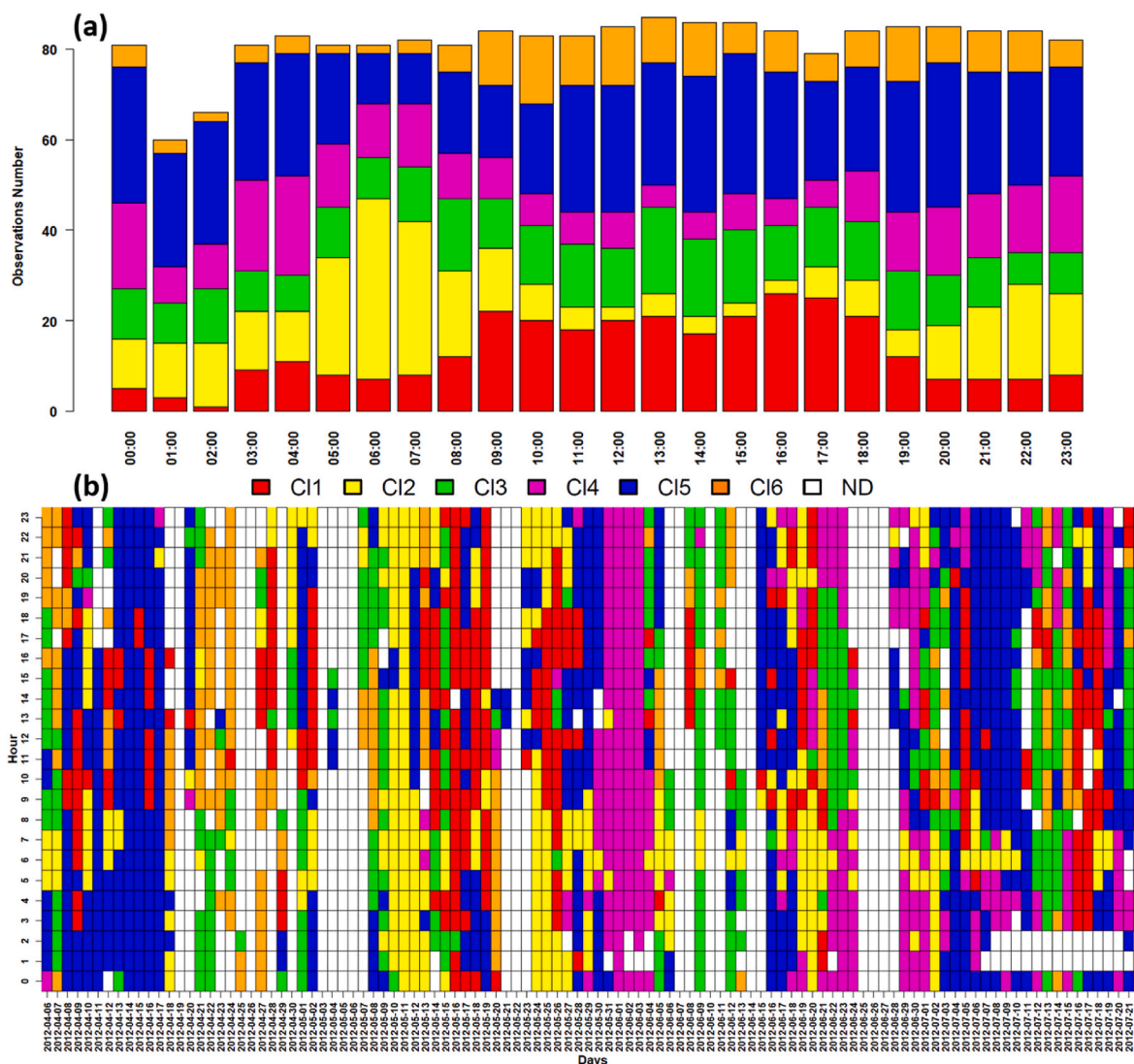

Fig. 8. Boxplots of chemical and optical particle counting variables based on the cluster division.

Fig. 9. a) Hourly cluster distribution with the corresponding observations counts; b) daily cluster distribution along the whole sampling period: each hour (observation) of each day is assigned to a specific SOM cluster. White cells (ND in the legend) represent hours for which the observation was not available.

Again, clustering based on DB-Index divided the dataset into six clusters (Fig. 8). Similarly to section 3.1, the behavior of the variables in the six clusters, and the daily distribution of observations are reported in Fig. 9. Meteorological variables, ^{222}Rn concentrations, and wind roses for each cluster are shown in Fig. 10.

As previously reported, this analysis was conducted only on warm season data, thus no seasonal behavior can be deduced from SOM results. Indeed, no specific meteorological conditions characterize any of the clusters, except for some peculiar humidity levels (Fig. 10a). However, such conditions allow drawing the attention to distinct pollution events that can be argued from the analysis of the six clusters.

Cl1 (316 observations) is mainly a diurnal cluster (observations are concentrated between 9:00 to 18:00) as confirmed by the lower daytime concentration of ^{222}Rn (Fig. 9a), due to enhanced instability conditions (Chambers et al., 2015), and scattered along all days. Winds are mainly from W-NW (Fig. 10b), with the lowest median RH among these clusters. It is also characterized by a significant contribution of coarse particles, OPS-bins 5 to 7 (3.0 to > 10 μm). This cluster can be considered analogous to Cl2 observed in the previous SOM (par. 3.1), with the contribution of SO$_2$ due to the coal-fired power plant in the northwesterly direction from the sampling point, addressing the opencast coal storage facility.

Cl2 (313 observations) is characterized by higher concentration

levels of all the chemical pollutants including fine particulate matter (i. e., bin1, 0.3–0.5 μm) and ^{222}Rn. The most interesting meteorological feature of this cluster is the wind, which is typically very weak (<1.3 m s^{-1}) addressing to low efficiency of dispersion processes. Except for the data collected between 9 and 11 May, all the observations cover nighttime data, collected between the hours 21:00 and 9:00 (mostly between 5:00 and 7:00). This is in connection with the higher boundary layer stability developing as soon as temperature decreases over the ground, but also the marked local and continental character reflected by the higher ^{222}Rn levels. The days 9–11 May are characterized by high atmospheric stability, with a complete absence of wind and high pressure (up to 1030 hPa), that enhance the accumulation of atmospheric pollutants.

Cl3 (289 observations) is characterized by higher-than-average concentrations of fine particles and the lighter among the coarse ones (bins 2 to 5, in the range 0.5–3.0 μm) and by southwesterly winds that also reach 4.5 m s^{-1}, suggesting a limestone quarry SW of the receptor site (Fig. 1) as a likely source (Ekpa et al., 2023; Khazins et al., 2022), compatible with the average wind speed and coarse particle size, typical of mechanically produced particles.

Cl4 (276 observations) is well characterized and confirms Cluster 3 of the previous SOM (Figs. 4 and 5). It is characterized by higher concentrations of BC, in this case in association with bin1. Indeed, this
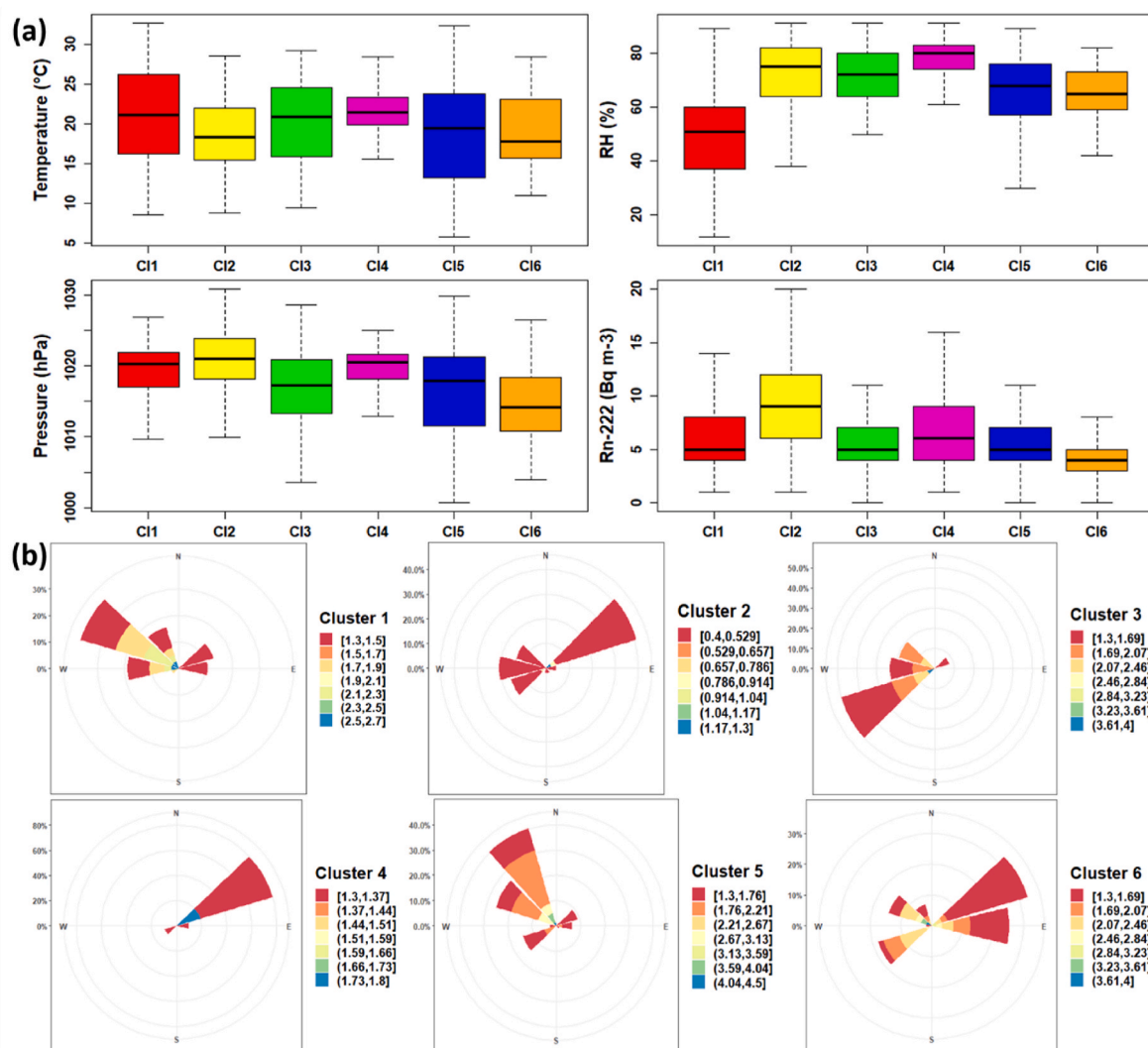
Fig. 10. a) Boxplots of the meteorological variables (temperature, relative humidity, pressure) and activity concentration of ^{222}Rn; b) wind roses for each cluster.

agrees with the typical fine/ultra-fine nature of elemental carbon, while the NE wind direction emerging from the clustering confirms the high-temperature fossil fuel source identified in shipping activities at the nearby commercial harbor. The days between 31 May and 4 June are univocally assigned to this cluster, confirming the previous attribution in paragraph 3.1.

Cl5 (585 observations) is characterized by lower pollution conditions similar to Cluster 6 of the previous SOM though now extended to aerosol fractions. This cluster is characterized by northerly winds from the polluted Po Valley airshed. Such winds were sufficiently intense to improve the original air quality by efficient pollutant dispersion as well as by wet removal by orographic clouds during the transit across the Alpine–Apennine range.

Cl6 (178 observations) is characterized by higher-than-average counts of all OPS bins, except for bin1. This cluster shares some characteristics with Cl3 (it is also present on the same days), particularly the influence of W-SW winds. The position of this cluster in the SOM, top-left corner (Fig. S7), between Cl1 and Cl3, and the lower number of observations included indicate that this cluster contains unresolved observations with coarse particulate matter, of lower environmental and health concerns, but involving multiple sources such as the highway, the cave, and others upwind the receptor site.

4. Conclusions

A six-month sampling campaign was carried out near Savona, a medium-sized city in northern Italy. The concentrations of eight chemical species and the number of particles at eight OPS bin dimensions were collected on an hourly basis. A source profiling method using Self-Organizing Maps was computed. Two SOMs were trained with chemical variables only over the whole sampling period and using the full set of variables over the hot season only. Meteorological parameters and the concentration of ^{222}Rn were finally used to evaluate the clustering obtained by SOM and to highlight pollution sources for each cluster. In both cases, six clusters were obtained and, for most of them, a well-defined air pollution source has been described, considering both the pollution profile and the meteorological characteristics. The SOM method was also able to highlight correlations between the variables, and the presence of OPS counts made it also possible to determine a dimension range for some kind of pollution, such as the fine particulate due to BC. These results were obtained with a data-analysis methodology requiring lower computational effort than PMF, showing its capability for source profiling applications, in particular when dealing with small datasets, not suitable for PMF. Moreover, although the chemical species used in the present work were criteria pollutants, generally considered unsuitable for source profiling due to the multiple co-emissive sources, the use of SOM and the integration with meteorological parameters, above all wind direction and speed, enabled emission

sources by spatial association to be revealed, rather than by time consuming and costly chemical speciation based methods. Indeed, SOM allowed the identification of groups of pollutants (or single species) described by the SOM clusters. Then, their origin was traced thanks to the complementary use of meteorological data, that was independent of SOM computation, considering wind speed and direction (evaluating local or distant origins) and the circadian (day/night) behavior of clusters, useful to evaluate the reactivity of the identified pollutants.

Financial support

CRediT authorship contribution statement

Alessandro Zappi: Writing – original draft, Software, Methodology, Investigation, Data curation. **Erika Brattich:** Writing – review & editing, Validation, Supervision, Software, Investigation, Formal analysis, Data curation. **Mariassunta Biondi:** Writing – review & editing, Visualization, Data curation. **Laura Tositti:** Writing – review & editing, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2024.143619.

Data availability

Data will be made available on request.

References

Brattich, E., Bracci, A., Zappi, A., Morozzi, P., Sabatino, S.D., Porcù, F., Nicola, F.D., Tositti, L., 2020. How to get the best from low-cost particulate matter sensors: guidelines and practical recommendations. Sensors 20, 1–33. https://doi.org/10.3390/s20113073.

Burlando, M., Tizzi, M., Solari, G., 2017. Characteristics of downslope winds in the Liguria region. Wind Struct. 24, 613–635. https://doi.org/10.12989/was.2017.24.6.613.

Carslaw, D.C., Ropkins, K., 2012. Openair - an r package for air quality data analysis. Environ. Model. Software 27 (28), 52–61. https://doi.org/10.1016/j.envsoft.2011.09.008.

Chambers, S.D., Choi, T., Park, S.J., Williams, A.G., Hong, S.B., Tositti, L., Griffiths, A.D., Crawford, J., Pereira, E., 2017. Investigating local and remote terrestrial influence on air masses at contrasting antarctic sites using radon-222 and back trajectories. J. Geophys. Res. Atmos. 122 (13). https://doi.org/10.1002/2017JD026833, 525-13,544.

Chambers, S.D., Podstawczyńska, A., Pawlak, W., Fortuniak, K., Williams, A.G., Griffiths, A.D., 2019. Characterizing the state of the urban surface layer using radon-222. J. Geophys. Res. Atmos. 124, 770–788. https://doi.org/10.1029/2018JD029507.

Chambers, S.D., Preunkert, S., Weller, R., Hong, S.B., Humphries, R.S., Tositti, L., Angot, H., Legrand, M., Williams, A.G., Griffiths, A.D., Crawford, J., Simmons, J., Choi, T.J., Krummel, P.B., Molloy, S., Loh, Z., Galbally, I., Wilson, S., Magand, O., Sprovieri, F., Pirrone, N., Dommergue, A., 2018. Characterizing atmospheric

transport pathways to Antarctica and the remote southern ocean using radon-222. Front. Earth Sci. 6. https://doi.org/10.3389/FEART.2018.00190.

Chambers, S.D., Williams, A.G., Crawford, J., Griffiths, A.D., 2015. On the use of radon for quantifying the effects of atmospheric stability on urban emissions. Atmos. Chem. Phys. 15, 1175–1190. https://doi.org/10.5194/ACP-15-1175-2015.

Corral, A.F., Dadashazar, H., Stahl, C., Edwards, E.L., Zuidema, P., Sorooshian, A., 2020. Source apportionment of aerosol at a coastal site and relationships with precipitation chemistry: a case study over the southeast United States. Atmosphere 11. https://doi.org/10.3390/ATMOS11111212.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909.

Ekpa, I.D., Laniyan, D.G., Agbor, C.N., Ben, U.C., Okon, J.E., 2023. Effect of particulate matter from quarry activities on crops and plant biodiversity in South-Eastern Nigeria. Environ. Monit. Assess. 195, 837. https://doi.org/10.1007/s10661-023-11445-w.

Fortezza, F., Strocchi, V., Giovanelli, G., Bonasoni, P., Georgiadis, T., 1993. Transport of photochemical oxidants along the northwestern adriatic coast. Atmospheric Environment. Part A. General Topics 27, 2393–2402. https://doi.org/10.1016/0960-1686(93)90407-P.

Friedrich, N., Eger, P., Shenolikar, J., Sobanski, N., Schuladen, J., Dienhart, D., Hottmann, B., Tadic, I., Fischer, H., Martinez, M., Rohloff, R., Tauer, S., Harder, H., Pfannerstill, E.Y., Wang, N., Williams, J., Brooks, J., Drewnick, F., Su, H., Li, G., Cheng, Y., Lelieveld, J., Crowley, J.N., 2021. Reactive nitrogen around the arabian peninsula and in the Mediterranean Sea during the 2017 AQABA ship campaign. Atmos. Chem. Phys. 21, 7473–7498. https://doi.org/10.5194/acp-21-7473-2021.

Hentati, A., Kawamura, A., Amaguchi, H., Iseri, Y., 2010. Evaluation of sedimentation vulnerability at small hillside reservoirs in the semi-arid region of Tunisia using the Self-Organizing Map. Geomorphology 122, 56–64. https://doi.org/10.1016/J.GEOMORPH.2010.05.013.

Hobbs, P.V., 2000. Introduction to Atmospheric Chemistry. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511808913.

Hopke, P.K., Dai, Q., Li, L., Feng, Y., 2020. Global review of recent source apportionments for airborne particulate matter. Sci. Total Environ. 740, 140091. https://doi.org/10.1016/j.scitotenv.2020.140091.

Kataoka, T., Yunoki, E., Shimizu, M., Mori, T., Tsukamoto, O., Ohashi, Y., Sahashi, K., Maitani, T., Miyashita, K., Iwata, T., Fujikawa, Y., Kudo, A., Shaw, R.H., 2001. A study of the atmospheric boundary layer using radon and air pollutants as tracers. Boundary-Layer Meteorol. 101, 131–156. https://doi.org/10.1023/A:1019219708361.

Khazins, V.M., Solov'ev, S.P., Loktev, D.N., Krasheninnikov, A.V., Shuvalov, V.V., 2022. Nearsurface air layer pollution with micronic dust particles in large-scale blasting in open pit mining. J. Min. Sci. 58, 676–689. https://doi.org/10.1134/S1062739122040160.

Kikaj, D., Chambers, S.D., Kobal, M., Crawford, J., Vaupotič, J., 2020. Characterizing atmospheric controls on winter urban pollution in a topographic basin setting using Radon-222. Atmos. Res. 237, 104838. https://doi.org/10.1016/J.ATMOSRES.2019.104838.

Kikaj, D., Vaupotič, J., Chambers, S.D., 2019. Identifying persistent temperature inversion events in a subalpine basin using radon-222. Atmos. Meas. Tech. 12, 4455–4477. https://doi.org/10.5194/amt-12-4455-2019.

Kohonen, T., 1998. The self-organizing map. Neurocomputing 21, 1–6. https://doi.org/10.1016/S0925-2312(98)00030-7.

Langner, J., Rodhe, H., 1991. A global three-dimensional model of the tropospheric sulfur cycle. J. Atmos. Chem. 13, 225–263. https://doi.org/10.1007/BF00058134.

Leo, L.S., Fernando, H.J.S., Di Sabatino, S., 2015. Near-surface flow in complex terrain with coastal and urban influence. Environ. Fluid Mech. 15, 349–372. https://doi.org/10.1007/s10652-013-9327-y.

Licen, S., Cozzutto, S., Barbieri, G., Crosera, M., Adami, G., Barbieri, P., 2019. Characterization of variability of air particulate matter size profiles recorded by optical particle counters near a complex emissive source by use of Self-Organizing Map algorithm. Chemometr. Intell. Lab. Syst. 190, 48–54. https://doi.org/10.1016/J.CHEMOLAB.2019.05.008.

Licen, S., Cozzutto, S., Barbieri, P., 2020. Assessment and comparison of multi-annual size profiles of particulate matter monitored at an urban-industrial site by an optical particle counter with a chemometric approach. Aerosol Air Qual. Res. 20, 800–809. https://doi.org/10.4209/aaqr.2019.08.0414.

Licen, S., Franzon, M., Rodani, T., Barbieri, P., 2021. SOMEnv: an R package for mining environmental monitoring datasets by Self-Organizing Map and k-means algorithms with a graphical user interface. Microchem. J. 165, 106181. https://doi.org/10.1016/J.MICROC.2021.106181.

May, R.J., Maier, H.R., Dandy, G.C., 2010. Data splitting for artificial neural networks using SOM-based stratified sampling. Neural Network. 23, 283–294. https://doi.org/10.1016/J.NEUNET.2009.11.009.

Millán, M.M., José Sanz, M., Salvador, R., Mantilla, E., 2002. Atmospheric dynamics and ozone cycles related to nitrogen deposition in the western Mediterranean. Environ. Pollut. 118, 167–186. https://doi.org/10.1016/S0269-7491(01)00311-6.

Mircea, M., Calori, G., Pirovano, G., Belis, C., 2020. European Guide on Air Pollution Source Apportionment for Particulate Matter with Source Oriented Models and Their Combined Use with Receptor Models. https://doi.org/10.2760/470628 (online).

Mooibroek, D., Sofowote, U.M., Hopke, P.K., 2022. Source apportionment of ambient PM10 collected at three sites in an urban-industrial area with multi-time resolution factor analyses. Sci. Total Environ. 850. https://doi.org/10.1016/J.SCITOTENV.2022.157981.

Morozzi, P., Bolelli, L., Brattich, E., Ferri, E.N., Girotti, S., Sangiorgi, S., Orza, J.A.G., Piñero-García, F., Tositti, L., 2021. Chemiluminescent fingerprints from airborne

particulate matter: a luminol-based assay for the characterization of oxidative potential with kinetic implications. Sci. Total Environ. 789, 148005. https://doi.org/10.1016/J.SCITOTENV.2021.148005.

Nakagawa, K., Yu, Z.Q., Berndtsson, R., Hosono, T., 2020. Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps. J. Hydrol. 582, 124519. https://doi.org/10.1016/J.JHYDROL.2019.124519.

Palladino, G., Morozzi, P., Biagi, E., Brattich, E., Turroni, S., Rampelli, S., Tositti, L., Candela, M., 2021. Particulate matter emission sources and meteorological parameters combine to shape the airborne bacteria communities in the Ligurian coast, Italy. Sci. Rep. 11 (1), 1–12. https://doi.org/10.1038/s41598-020-80642-1, 2021 11.

Perrino, C., Pietrodangelo, A., Febo, A., 2001. An atmospheric stability index based on radon progeny measurements for the evaluation of primary urban pollution. Atmos. Environ. 35, 5235–5244. https://doi.org/10.1016/S1352-2310(01)00349-1.

Perrone, M.R., Paladini, F., Becagli, S., Amore, A., Romano, S., 2022. Daytime and nighttime chemical and optical properties of fine and coarse particles at a central Mediterranean coastal site. Environ. Sci. Pollut. Control Ser. 29, 43401–43420. https://doi.org/10.1007/S11356-021-18173-Z.

Pietrodangelo, A., Bove, M.C., Forello, A.C., Crova, F., Bigi, A., Brattich, E., Riccio, A., Becagli, S., Bertinetti, S., Calzolai, G., Canepari, S., Cappelletti, D., Catrambone, M., Cesari, D., Colombi, C., Contini, D., Cuccia, E., De Gennaro, G., Genga, A., Ielpo, P., Lucarelli, F., Malandrino, M., Masiol, M., Massabò, D., Perrino, C., Prati, P., Siciliano, T., Tositti, L., Venturini, E., Vecchi, R., 2024. A PM10 chemically characterized nation-wide dataset for Italy. Geographical influence on urban air pollution and source apportionment. Sci. Total Environ. 908, 167891. https://doi.org/10.1016/j.scitotenv.2023.167891.

Pöschl, U., Shiraiwa, M., 2015. Multiphase chemistry at the atmosphere-biosphere interface influencing climate and public health in the anthropocene. Chem. Rev. 115, 4440–4475. https://doi.org/10.1021/cr500487s.

Prospero, J.M., Charlson, R.J., Mohnen, V., Jaenicke, R., Delany, A.C., Moyers, J., Zoller, W., Rahn, K., 1983. The atmospheric aerosol system: an overview. Rev. Geophys. 21, 1607–1629. https://doi.org/10.1029/RG021I007P01607.

Raes, F., Van Dingenen, R., Vignati, E., Wilson, J., Putaud, J.-P., Seinfeld, J.H., Adams, P., 2000. Formation and cycling of aerosols in the global troposphere. Atmos. Environ. 34, 4215–4240. https://doi.org/10.1016/S1352-2310(00)00239-9.

Satterly, J., 1910. I. On the amount of radium emanation in the lower regions of the atmosphere and its variation with the weather. London, Edinburgh Dublin Phil. Mag. J. Sci. 20, 1–36. https://doi.org/10.1080/14786441008636877.

Seinfeld, J.H., Pandis, S.N., 2016. Atmospheric chemistry and physics: from air pollution to climate change. Environment. https://doi.org/10.1080/00139157.1999.10544295.

Sesana, L., Caprioli, E., Marcazzan, G.M., 2003. Long period study of outdoor radon concentration in Milan and correlation between its temporal variations and dispersion properties of atmosphere. J. Environ. Radioact. 65, 147–160. https://doi.org/10.1016/S0265-931X(02)00093-0.

Song, Y., Shao, M., 2023. Impacts of complex terrain features on local wind field and PM2.5 concentration. Atmosphere 14. https://doi.org/10.3390/atmos14050761.

Tan, Y., Han, S., Chen, Y., Zhang, Z., Li, H., Li, W., Yuan, Q., Li, X., Wang, T., Lee, S.C., 2021. Characteristics and source apportionment of volatile organic compounds (VOCs) at a coastal site in Hong Kong. Sci. Total Environ. 777. https://doi.org/10.1016/J.SCITOTENV.2021.146241.

Tarquini, S., Isola, I., Favalli, M., Battistini, A., Dotta, G., 2023. TINITALY, a Digital Elevation Model of Italy with a 10 Meters Cell Size (Version 1.1).

Tositti, L., Brattich, E., Parmeggiani, S., Bolelli, L., Ferri, E., Girotti, S., 2018. Airborne particulate matter biotoxicity estimated by chemometric analysis on bacterial luminescence data. Sci. Total Environ. 640–641, 1512–1520. https://doi.org/10.1016/J.SCITOTENV.2018.06.024.

Tositti, L., Morozzi, P., Brattich, E., Zappi, A., Calvello, M., Esposito, F., Lettino, A., Pavese, G., Sabia, S., Speranza, A., Summa, V., Caggiano, R., 2022. Apportioning PM1 in a contrasting receptor site in the Mediterranean region: aerosol sources with an updated sulfur speciation. Sci. Total Environ. 851, 158127. https://doi.org/10.1016/j.scitotenv.2022.158127.

Tositti, L., Pereira, E.B., Sandrini, S., Capra, D., Tubertini, O., Bettoli, M.G., 2002. Assessment of summer trends of tropospheric radon isotopes in a coastal antarctic station (Terra Nova Bay). Int. J. Environ. Anal. Chem. 82, 259–274. https://doi.org/10.1080/03067310290027767.

Trigo, I.F., Bigg, G.R., Davies, T.D., 2002. Climatology of cyclogenesis mechanisms in the mediterranean. Mon. Weather Rev. 130, 549–569. https://doi.org/10.1175/1520-0493(2002)130<0549:COCMIT>2.0.CO;2.

Turekian, K.K., Nozaki, Y., Benninger, L.K., 1977. Geochemistry of atmospheric radon and radon products. Annu. Rev. Earth Planet Sci. 5, 227–255. https://doi.org/10.1146/annurev.ea.05.050177.001303.

Veld, M. in t., Alastuey, A., Pandolfi, M., Amato, F., Pérez, N., Reche, C., Via, M., Minguillón, M.C., Escudero, M., Querol, X., 2021. Compositional changes of PM2.5 in NE Spain during 2009–2018: a trend analysis of the chemical composition and source apportionment. Sci. Total Environ. 795, 148728. https://doi.org/10.1016/J.SCITOTENV.2021.148728.

Watson, J.G., Chow, J.C., 2015. Chapter 20 - receptor models and measurements for identifying and quantifying air pollution sources. In: Murphy, B.L., Morrison, R.D. (Eds.), B.T.-I. To E.F., Third E. Academic Press, San Diego, pp. 677–706. https://doi.org/10.1016/B978-0-12-404696-2.00020-5.

Wilkening, M., 2004. Natural radioactivity as a tracer in the sorting of aerosols according to mobility. Rev. Sci. Instrum. 23, 13–16. https://doi.org/10.1063/1.1746056.

Wilkening, M., 1981. Radon in Atmospheric Studies: a Review. United States.

Williams, A.G., Chambers, S.D., Conen, F., Reimann, S., Hill, M., Griffiths, A.D., Crawford, J., 2016. Radon as a tracer of atmospheric influences on traffic-related air pollution in a small inland city. Tellus B. https://doi.org/10.3402/tellusb.v68.30967.

Wright, J.R., Smith, O.F., 1915. The variation with meteorological conditions of the amount of radium emanation in the atmosphere, in the soil gas, and in the air exhaled from the surface of the ground, at Manila. Phys. Rev. 5, 459–482. https://doi.org/10.1103/PhysRev.5.459.