atmosphere

Supplementary Materials:

Iasonas Stavroulas [1,2], Georgios Grivas [2,*], Eleni Liakakou [2], Panayiotis Kalkavouras [1,2], Aikaterini Bougiatioti [2], Dimitris G. Kaskaoutis [1,2], Maria Lianou [2], Kyriaki Papoutsidaki [1], Maria Tsagkaraki [1], Pavlos Zarmpas [1], Evangelos Gerasopoulos [2] and Nikolaos Mihalopoulos [1,2,*]

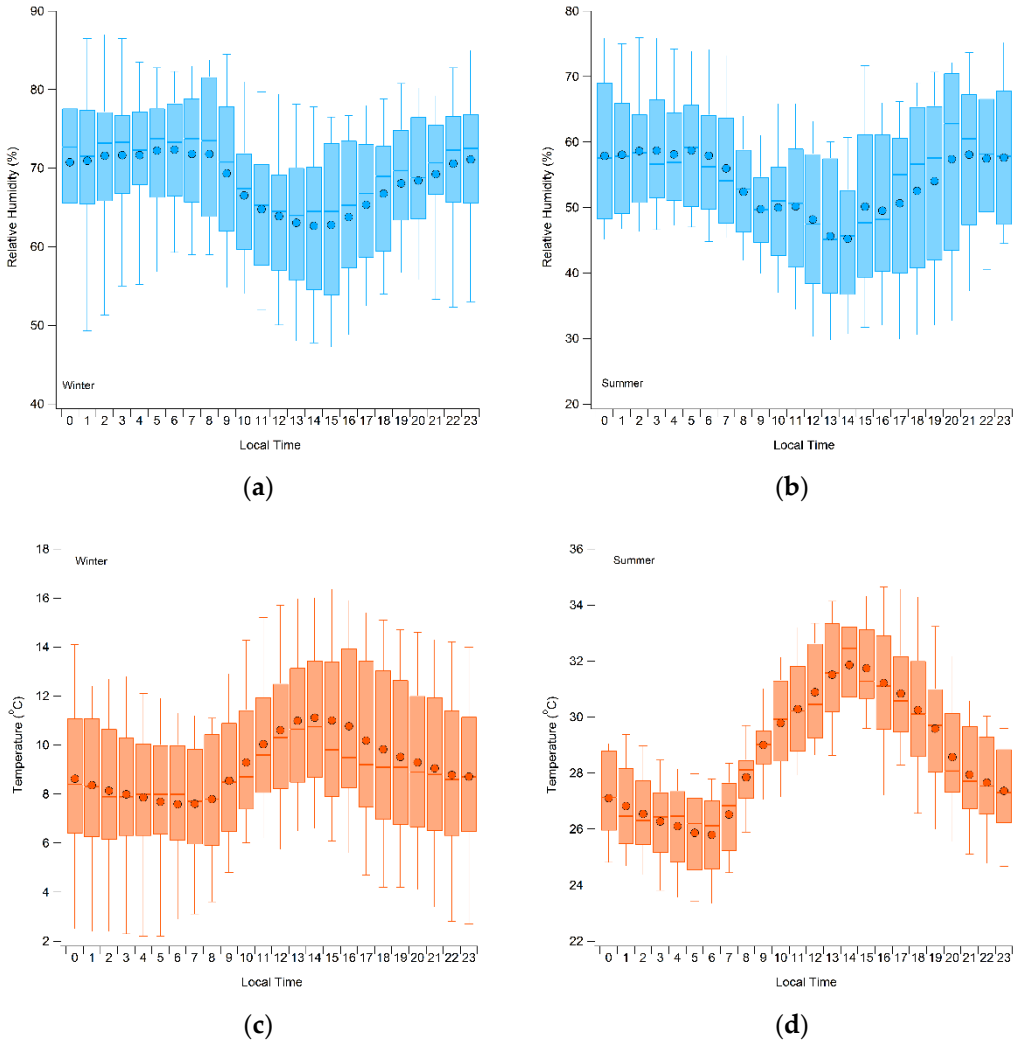


Figure S1. Diurnal box-whisker plots of relative humidity (**a**,**b**), and temperature (**c**,**d**) for the winter and summer periods, respectively. Boxes are the 25th, 50th and 75th percentiles, whiskers are the 10th and 90th percentiles, while dot markers are average values.

S1. PMF Overview

S1.1 Theory

Source apportionment of the obtained mass spectra makes use of a bi-linear model, representing the obtained OA mass spectra matrix *X*, in which data in each column *j* are measured intensities for a given *m*/*z*, and data in each row *i* represent a single mass spectrum obtained at a given time. The bi-linear model analyses the matrix *X* to the linear combination of two static matrices *G* and *F*:

$$X = GF + E$$

where E is the model residual matrix. In this case PMF was applied [1], meaning that the resulting matrices are required to include non-negative values only, which makes sense when dealing with concentrations in the atmosphere. The number (*p*) of columns in *G*, or for that matter, the number of rows in *F*, represent the predefined number of factors, while each column in G represents the time-series of a factor and each row in F is the factor's profile, effectively the factor's mass spectrum.

A least squares algorithm is implemented, which minimizes the quantity *Q*:

$$Q^m = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{e_{ij}}{\sigma_{ij}}\right)^2$$

where e_{ij} are the model residuals and σ_{ij} are the corresponding measurement uncertainties of the model input matrix X of size $m \times n$. When assessing model performance, the relative change in the normalized *Q* over its expected value Q_{exp} is monitored, where:

$$Q_{exp} = n \cdot m - p \cdot (m + n)$$

Given that model solutions exhibit rotational ambiguity [2], the ME-2 solver, as implemented in the SoFi software package [3] provides tools to address the issue. Introducing a priori information, e.g. predefined factor profiles, or predefined factor time-series, significantly reduces this rotational ambiguity [4]. In this study we are using the a-value approach to introduce these a priori information, thus constraining the resulting matrices *F* and/or *G*. The a-value is a measure of the extent to which the resulting profile or time series is allowed to vary from the ones introduced as parts of the constrained matrices *G* or *F*:

$$g_{i,solution} = g_i \pm a \cdot g_i$$

$$f_{j,solution} = f_j \pm a \cdot f_j$$

where *g* and *f* are a column of *G* or a row of *F* respectively. The a-value can vary between 0 and 1 representing a fully constrained and a fully unconstrained result respectively.

S1.2 Applied Strategy

For both seasons, PMF was initialized with no factor profile constrains, exploring solutions ranging from 3 to 8 factors. For the winter dataset, the 3-factor unconstrained PMF revealed 2 primary factors– one associated with fossil fuel combustion, the other with biomass burning – and a secondary OA factor (Figure S1a). The OOA factor showed significant diurnal variability peaking in the evening, thus exhibiting important influence from RWB emissions. While the BBOA factor had high correlations with BC_{bb} (r^2 above 0.8), the HOA factor's correlation to BC_{ff} was rather low with r^2 values around 0.3. The later led to the exploration of the 4-factor solution. The 4-factor solution revealed another HOA-like factor, with limited variability between seeds and runs, consistently exhibiting highly similar relative intensity patterns in the mass spectra, with pronounced contributions to *m*/*z* 41 and 43 and lower peaks at the rest of the alkane and alkene related fragments (e.g. 55, 57 etc.) as illustrated in the spectrum of factor four in Figure S2. This factor

appeared as slightly more oxidized, with higher contributions at *m*/*z* 44 implying atmospheric processing and was also present in runs within the 5 and 6 factor solutions.

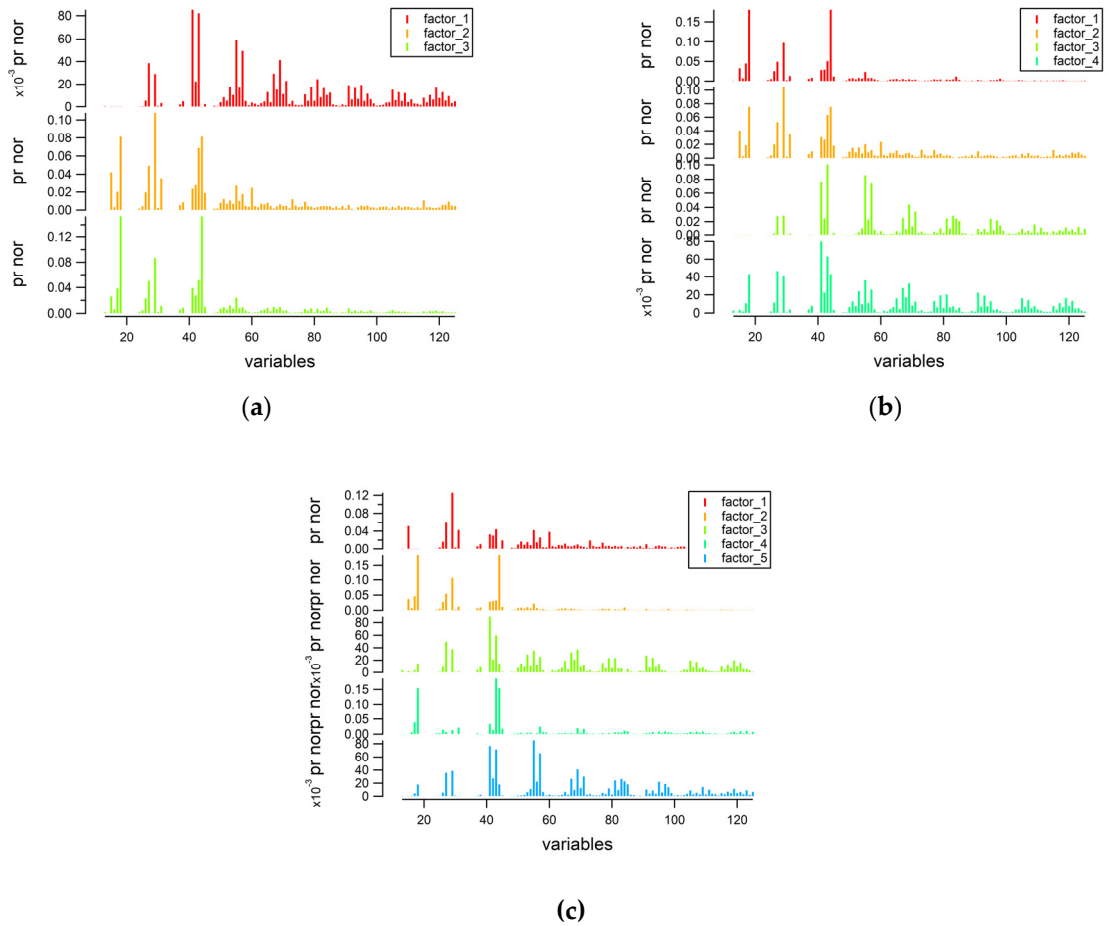


Figure S2. Factor profiles of wintertime unconstrained PMF solutions for (**a**) 3, (**b**) 4, and (**c**) 5 factors respectively.

Inspecting factors 3 and 4 in Figures S2c and S2b, respectively, the factor's small difference in intensities between the *m*/*z* 55 and 57 ions, points away from a cooking source, a conclusion further corroborated by plotting the measured f_{55} over f_{57} scatter plot [5] . When coloring the data points according to the time of day there is no evident shift of data points towards the COA PMF and "cooking source" lines during typical meal hours (Figure S3). This observation, also present in summertime data, lead to the adjustment of the implemented PMF strategy, thus not applying any COA factor constrains in subsequent analysis.

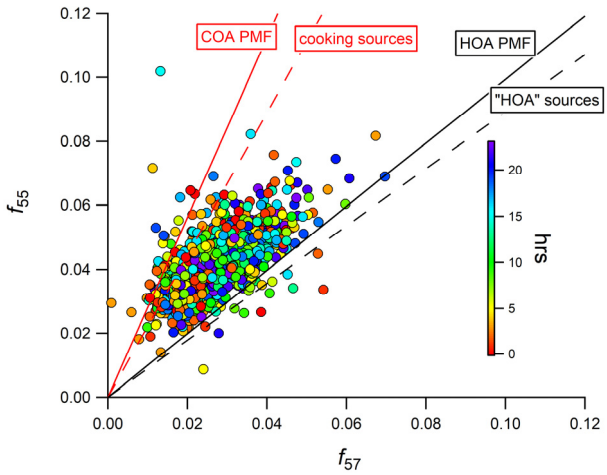
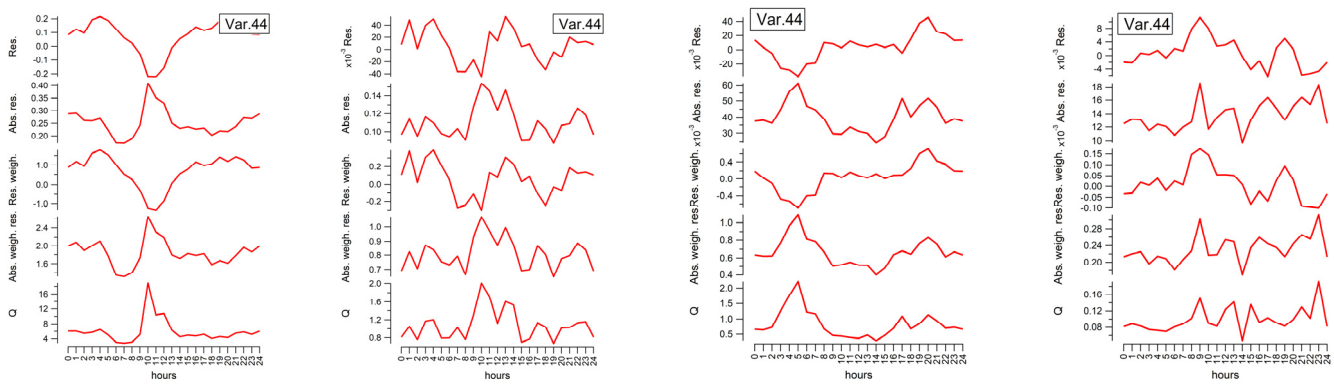


Figure S3. Scatter plot of f_{55} vs f_{57} for wintertime data in Piraeus according to Mohr et al. [5]. Data points are color coded according to time of day. Red and black lines correspond to linear fits of f_{55} and f_{57} extracted from HOA COA factors as well as spectra from cooking and traffic emission studies grouped accordingly.

Furthermore, in winter, while exploring the unconstrained runs and moving to more than 5 factors, factor splitting behavior started to emerge, while the stability of the solutions was hindered with large variability observed for resulting profiles between different seed runs.

As a next step, constrains were introduced to the primary OA factors by implementing the a-value approach [6]. Initially, only the HOA traffic-related factor was constrained using the factor profile acquired by Crippa et al. during wintertime in Paris [7]. The a-values applied ranged from 0.02 to 0.2. The applied constrain improved substantially the observed correlation of the HOA traffic-related factor to BC_{ff}, bringing r^2 values close to 0.5 while all 4 and 5 factor runs incorporated the second HOA-like factor, which retained its distinct profile among different seed runs. It should be mentioned here that the 3-factor solution, while constraining the HOA traffic factor, other than exhibiting larger Q/Qexp values, showed a clear structure in the diurnal variation of both the Q value as well as the residuals of key variables (e.g. *m*/*z* 44 or *m*/*z* 60) indicating an inferior performance of the PMF model (Figure S4a and S4b). When exploring the diurnal variability of residuals for the summer dataset, a similar behavior was observed (Fig. S4c and S4d). Consequently, we placed the focus on solutions with at least 4 factors for both measuring periods.



(**a**) (**b**) (c) (d)

Figure S4. Diurnal variability of the PMF model residuals and the value Q for the variable *m*/*z* 44 in the case of (**a**) a 3-factor solution and (**b**) a 4-factor solution when constraining an HOA factor for wintertime and in the case of 3- factor (**c**) and 4-factor (**d**) solution for summertime respectively.

In winter, given the fact that when constraining the HOA factor only, the resulting unconstrained BBOA factor was lacking stability among different seeds runs, we introduced constrains using the BBOA factor profile derived in Ng et al. [8]. The a-values used, ranged from 0.2 to 0.6. Different combinations of a-values for the HOA and BBOA factors were tested and the resulting HOA and BBOA time-series were assessed in terms of their correlation with BC_{ff} and BC_{bb}, respectively. Exploring 4-factor solutions with both BBOA and HOA constrained, while the second HOA-like factor persisted, appearing in most of the examined seed runs, the remaining factor, largely resembling an OOA factor in terms of mass spectra, seemed to be highly affected by primary sources, exhibiting large peaks during evening hours and tracking the BBOA factor's nighttime variability related to residential heating. In order to tackle this behavior, 5-factor solutions were consequently explored. While keeping the two aforementioned constrains, in the majority of the 5-factor runs, together with than the second HOA-like factor and a highly oxygenated OOA factor, a seemingly less oxygenated factor was revealed, with considerable contributions at the *m*/*z* 60 fragment. Most of the nighttime elevated concentrations were attributed to this less oxidized OOA factor. Other than the Q/Qexp values, behavior of the residual diurnal cycles improved, while the 3 unconstrained factors of the 5-factor solution showed low variability among different seed runs, supporting the choice of the 5-factor solution in winter. Eventually, the selected final wintertime 5-factor solution includes two HOA factors (HOA-1 & HOA-2) a BBOA, a less oxidized and a more oxidized (LO-OOA and a MO-OOA) factor.

In summer, after constraining the traffic-related HOA factor with the same anchor profile and going through the 4-factor solutions, a HOA-like factor similar to the one found in the winter dataset (Fig S4), could be observed in the majority of the different seed runs. This HOA-like factor was characterized by limited stability, mostly in terms of the *m*/*z* 44 contribution to its profile, which for some of the runs was as high as the respective contributions of *m*/*z* 43 and 41. When *m*/*z* 44 contributions were rather low, the factor showed similarities with the HOA-2 factor obtained for wintertime (r^2~0.8). In order to extract this HOA-like factor more effectively, we used the wintertime HOA-2 factor profile as an anchor, applying rather loose a-values (0.1 – 0.3). No BBOA-resembling factor could be identified for the summer season, since no substantial contribution from the *m*/*z* 60 and *m*/*z* 73 fragments was observed. In fact, an average f_{60} value of 0.27% was recorded for the summer data, in contrast to the 0.7% average f_{60} measured in winter. Therefore, a 4-factor solution was chosen, incorporating two HOA factors, a less oxidized and a more oxidized OOA factor

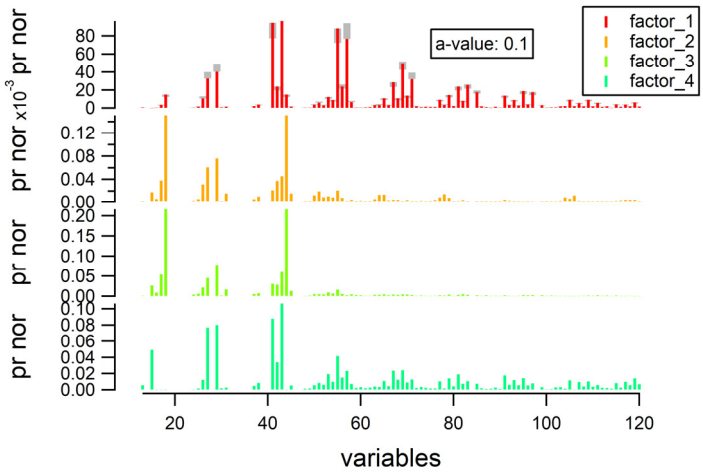


Figure S5. PMF run constraining an HOA factor (a-value=0.1) for the summertime period. A second HOA like factor with major contributions at *m*/*z* 41 and 43, can be spotted in position 4.

Table S1. Brief overview of the constrains and model performance for the selected PMF solutions in the two measurement periods.

	Winter	Summer
Number of Factors	5	4
Modeled vs. Measured OA r^2	0.99	0.99
Modeled vs. Measured OA Slope	0.99	1.03
HOA-1 Anchor Profile	Crippa et al., 2013	Crippa et al., 2013
HOA-1 a-value	0.02	0.1
HOA-2 Anchor Profile	-	This Study
HOA-2 a-value	-	0.2
BBOA Anchor Profile	Ng et al., 2011	-
BBOA a-value	0.6	-

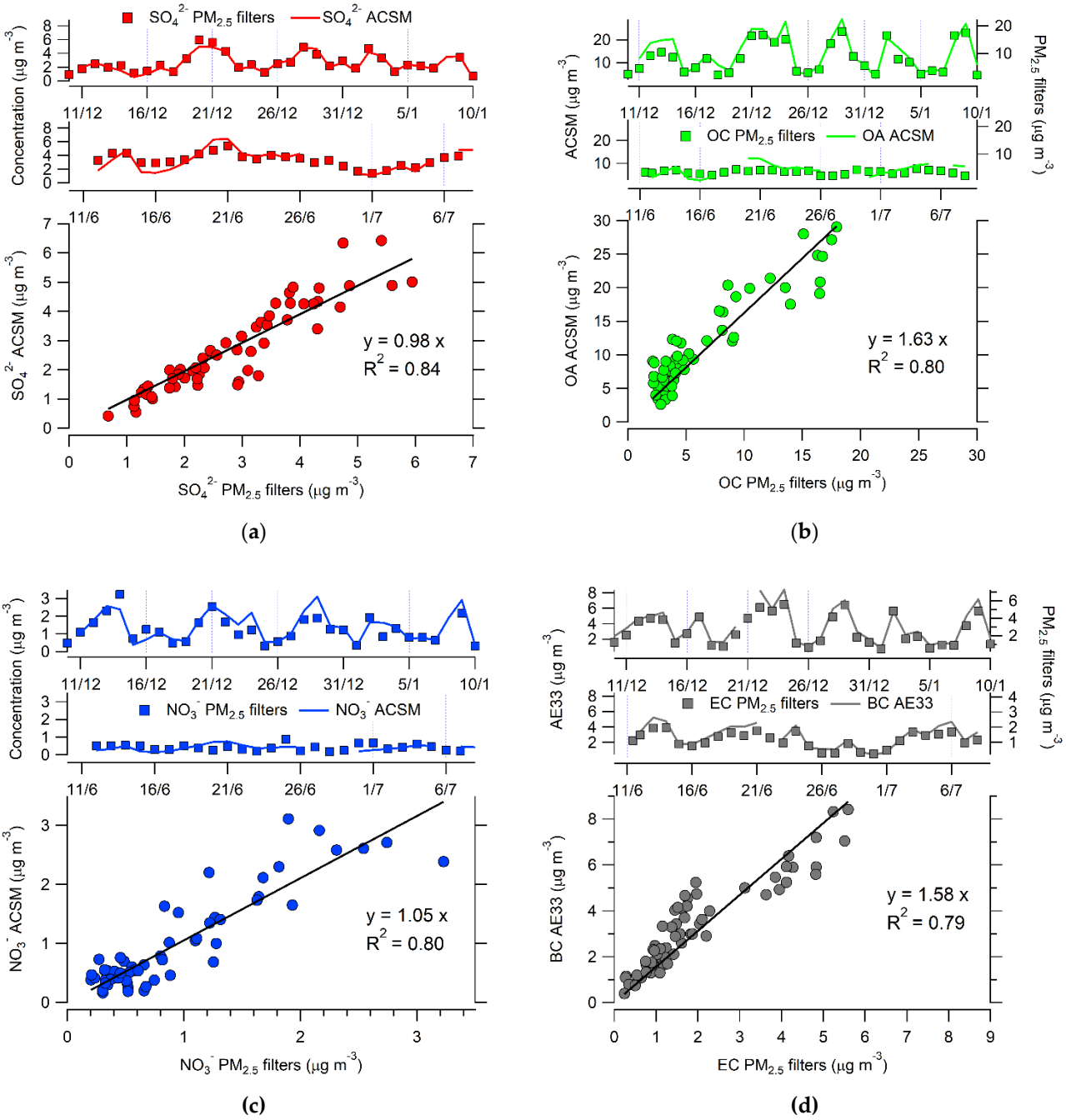


Figure S6. Comparison of 24 h averaged concentrations determined with the on-line instruments and from filter-based analyses, for major fine aerosol species in Piraeus, for the two campaigns combined. Daily averaged time-series for both seasons are also included. (**a**) SO_4^{2-} $PM_{2.5}$ filters, (**b**) OC $PM_{2.5}$ filters, (**c**) NO_3 $PM_{2.5}$ filters, (**d**) EC $PM_{2.5}$ filters.

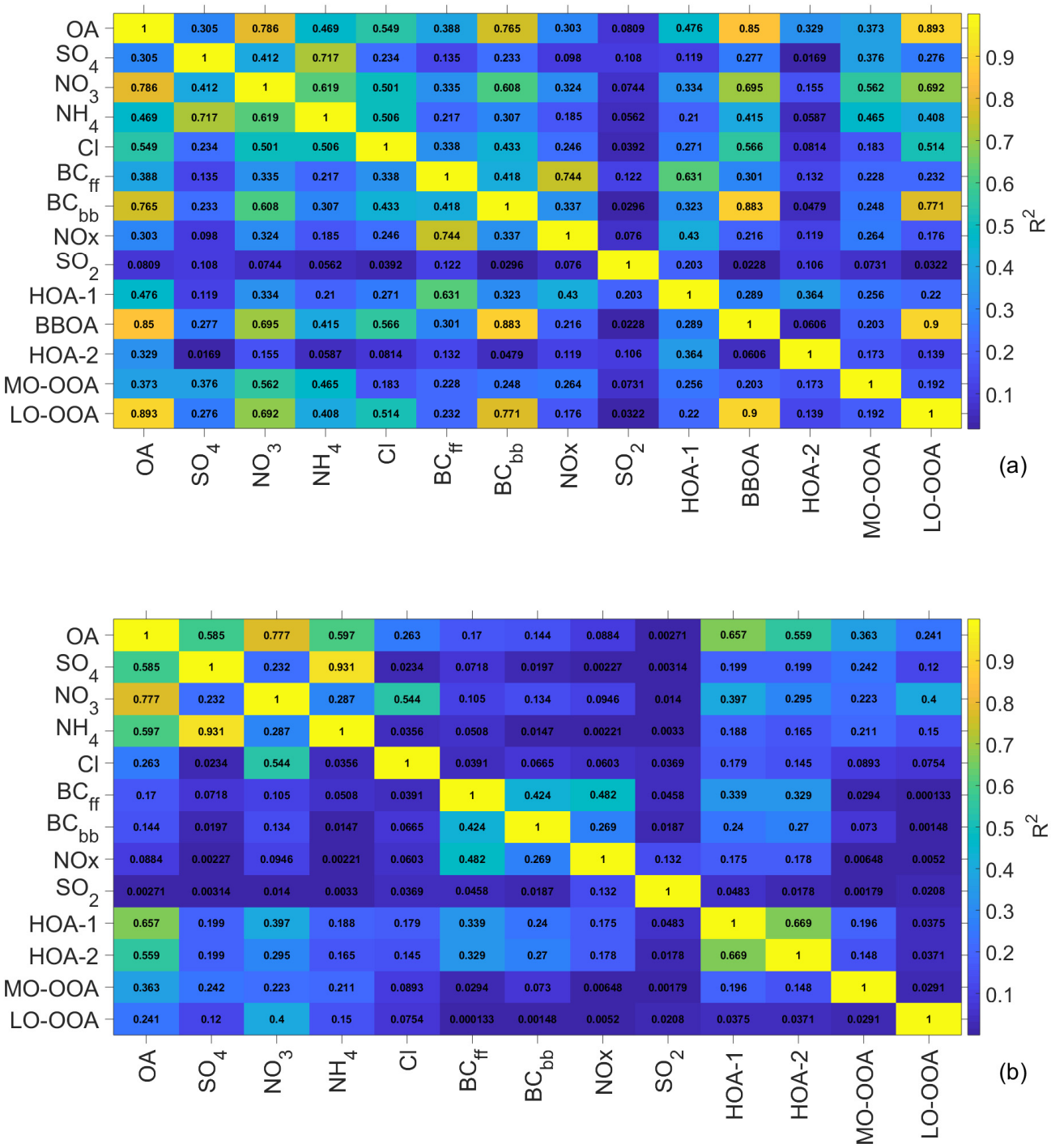


Figure S7. Correlation plot for major species, source factors, BC constituents, NO_x and SO_2 for wintertime (**a**) and summertime (**b**) measurements.

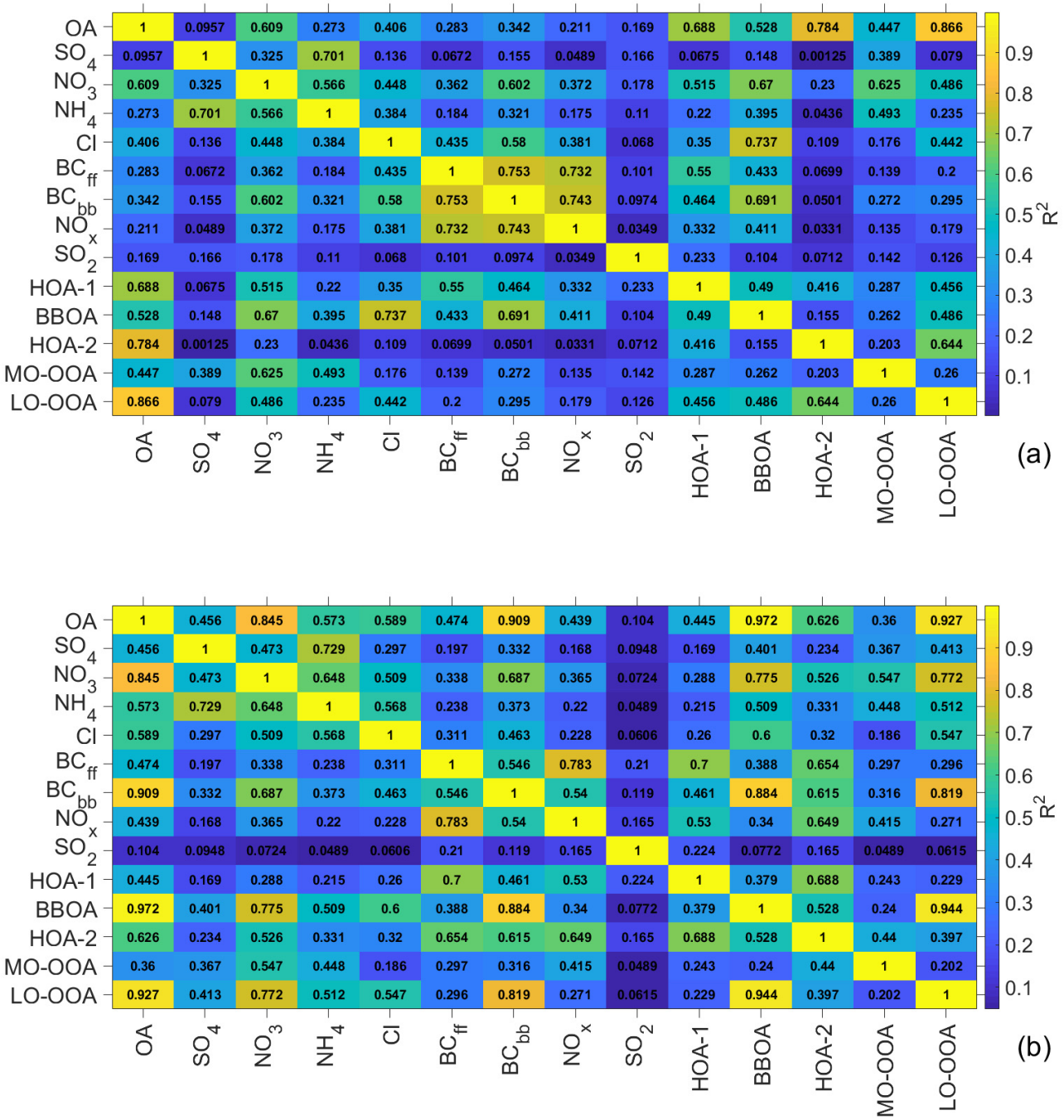


Figure S8. Correlation plot for major species, source factors, BC constituents, NO_x and SO_2 during the winter period for daytime (**a**) and nighttime data (**b**).

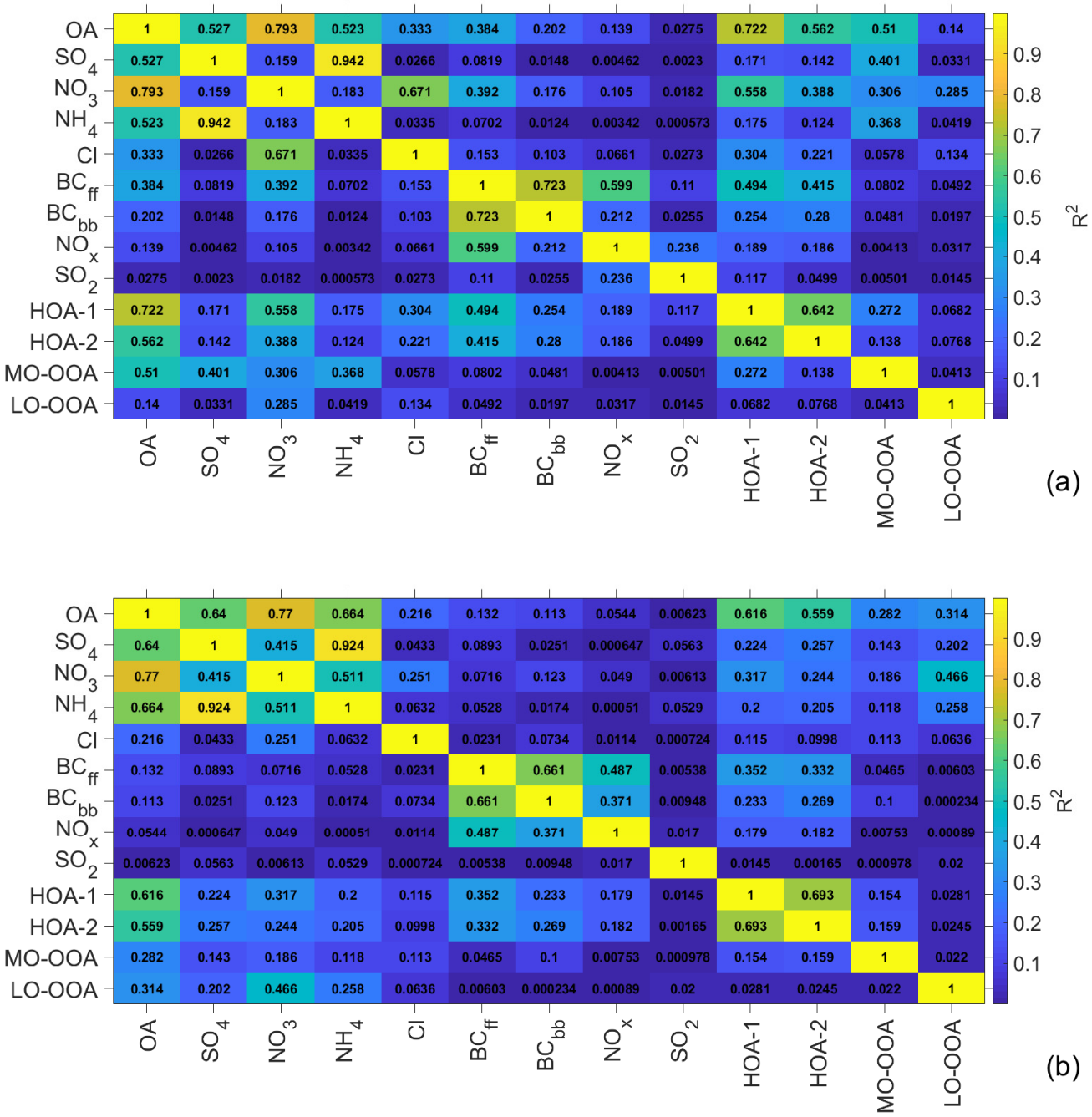


Figure S9. Correlation plot for major species, source factors, BC constituents, NO_x and SO_2 during the summer period for daytime (**a**) and nighttime data (**b**).

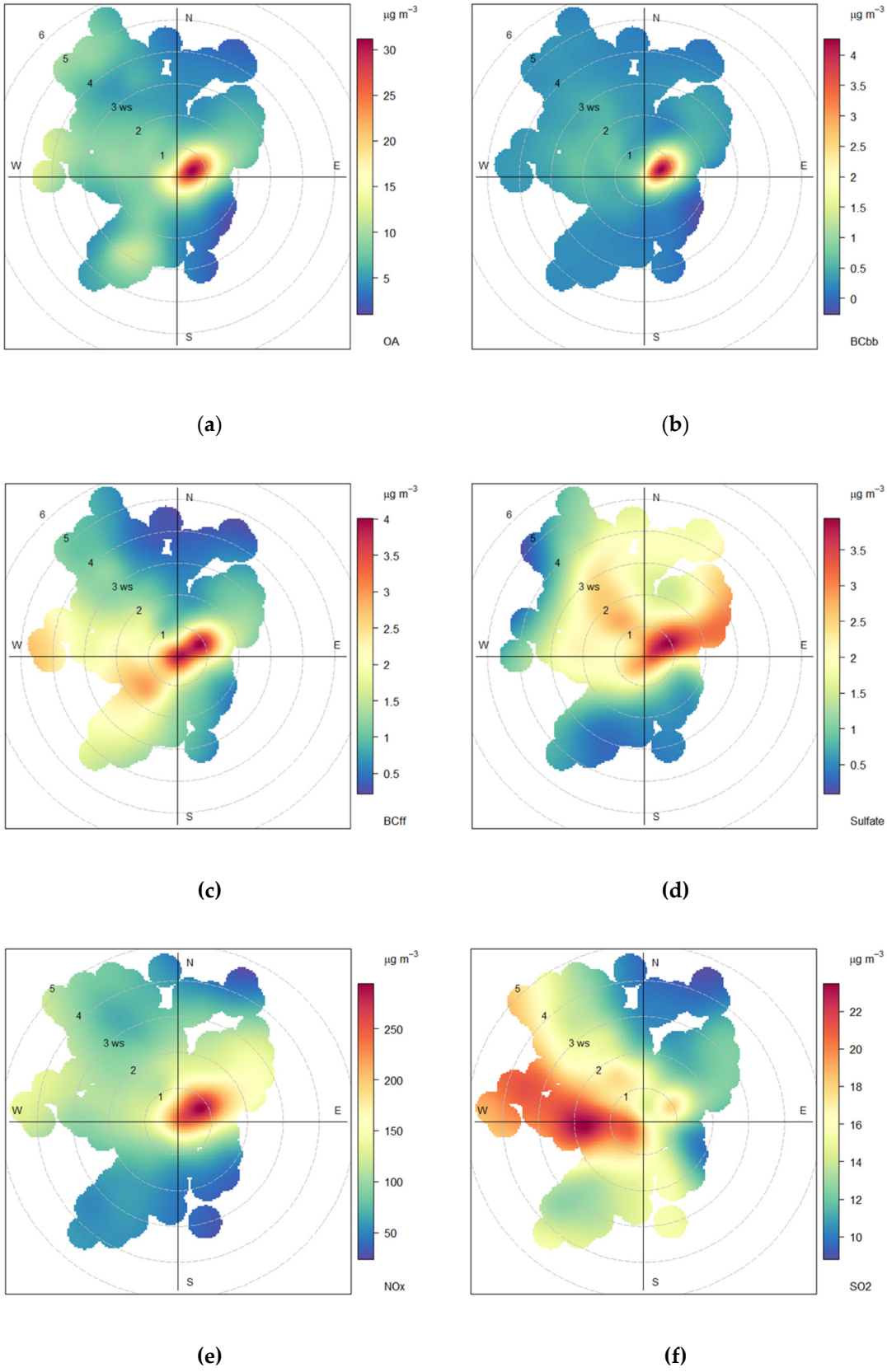


Figure S10. Bivariate polar wind plots for the effects of wind direction and wind speed (radial axis in m s^{-1}) to concentrations of OA (**a**), BC_{bb} (**b**), BC_{ff} (**c**), sulfate (**d**) as measured at P1 and NO_x (**e**) and SO_2 (**f**) as measured at PEI-1, during the winter period.

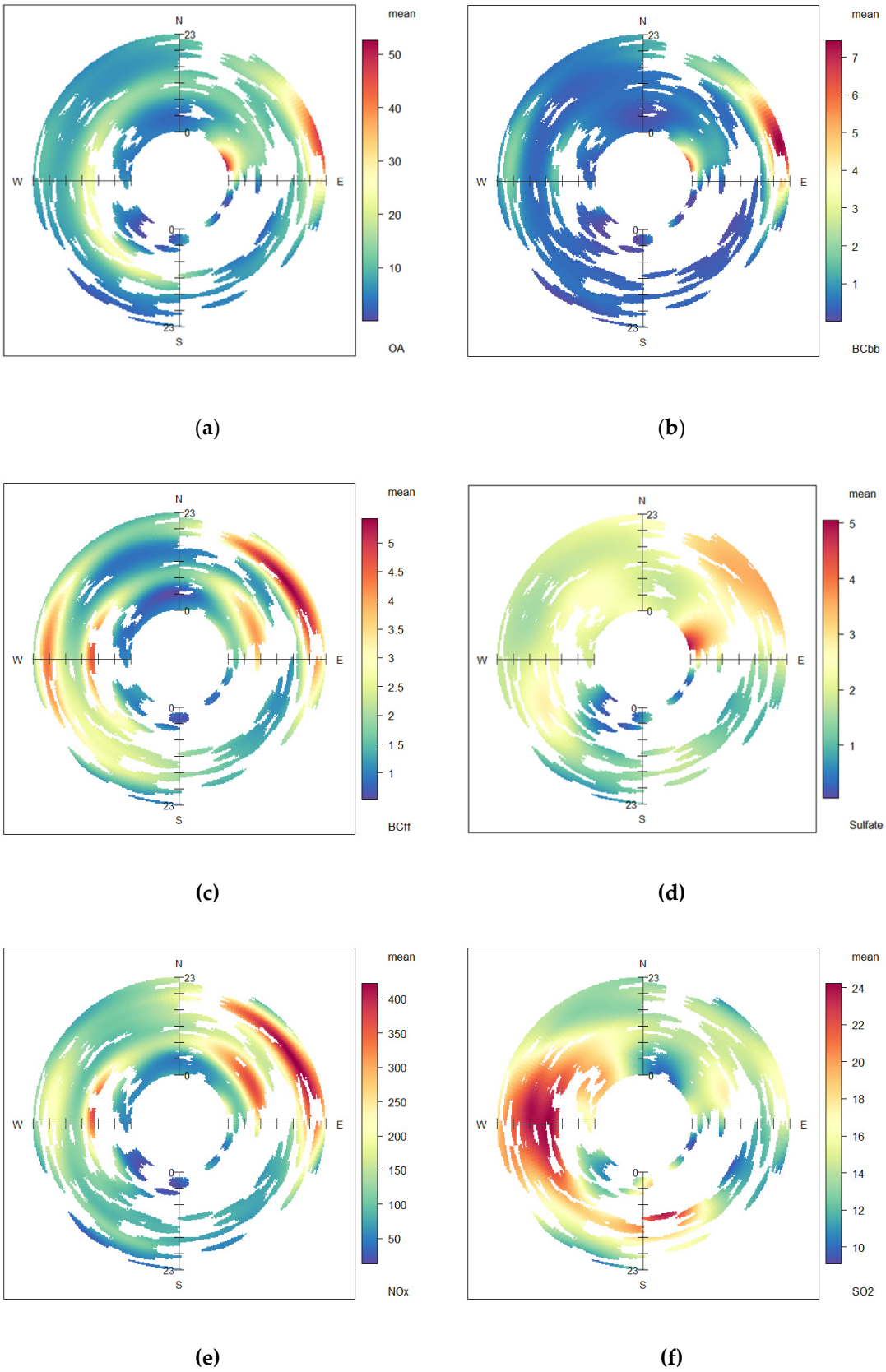


Figure S11. Annular wind plots for the effects of wind direction per hour of the day (radial axis, increments by 4 hr) to concentrations of OA (**a**), BC_{bb} (**b**), BC_{ff} (**c**), sulfate (**d**) as measured at P1 and NO_x (**e**) and SO_2 (**f**) as measured at PEI-1, during the winter period.

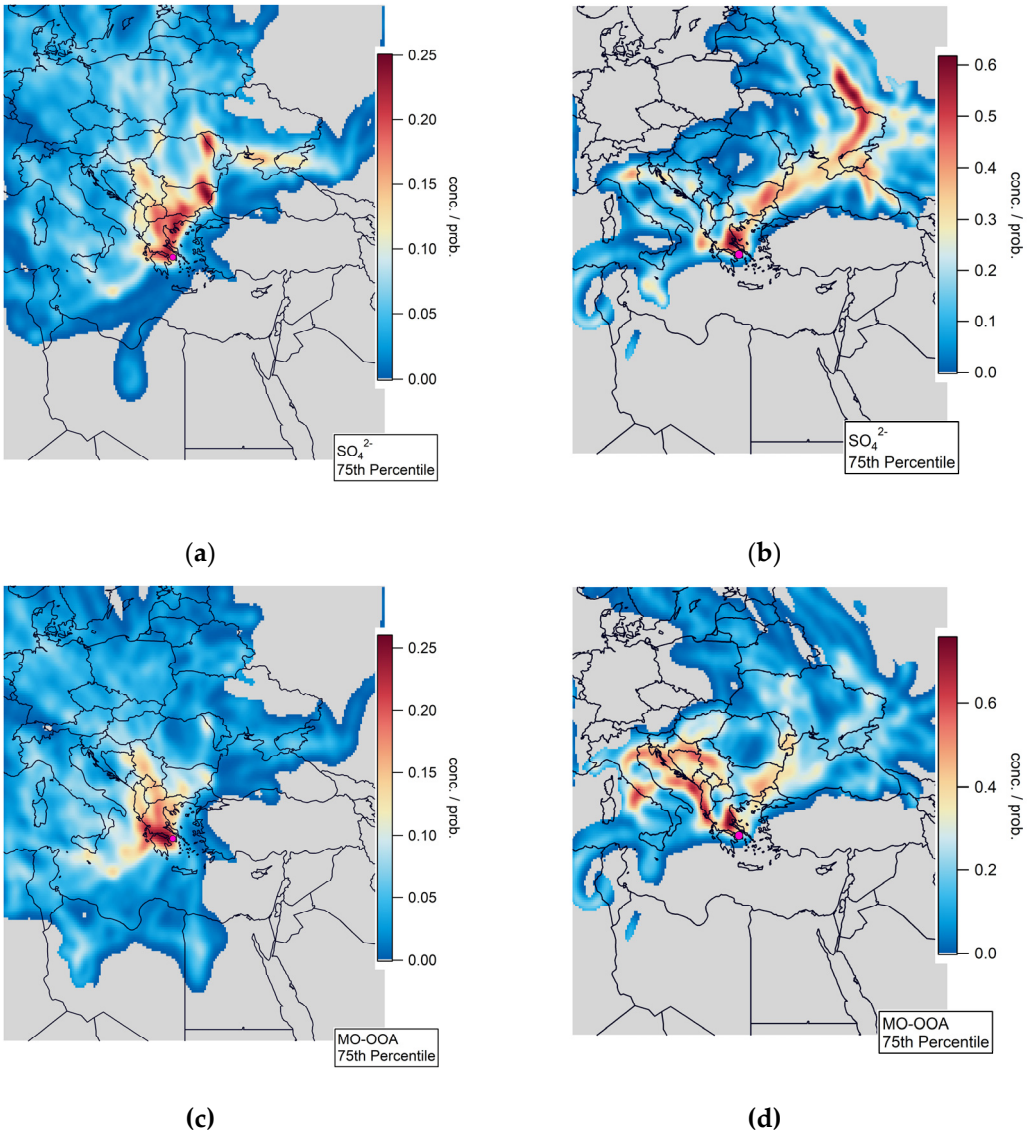


Figure S12. Potential Source Contribution Function calculations for sulfate during winter (**a**) and summer (**b**) as well as MO-OOA for winter (**c**) and summer (**d**), respectively.

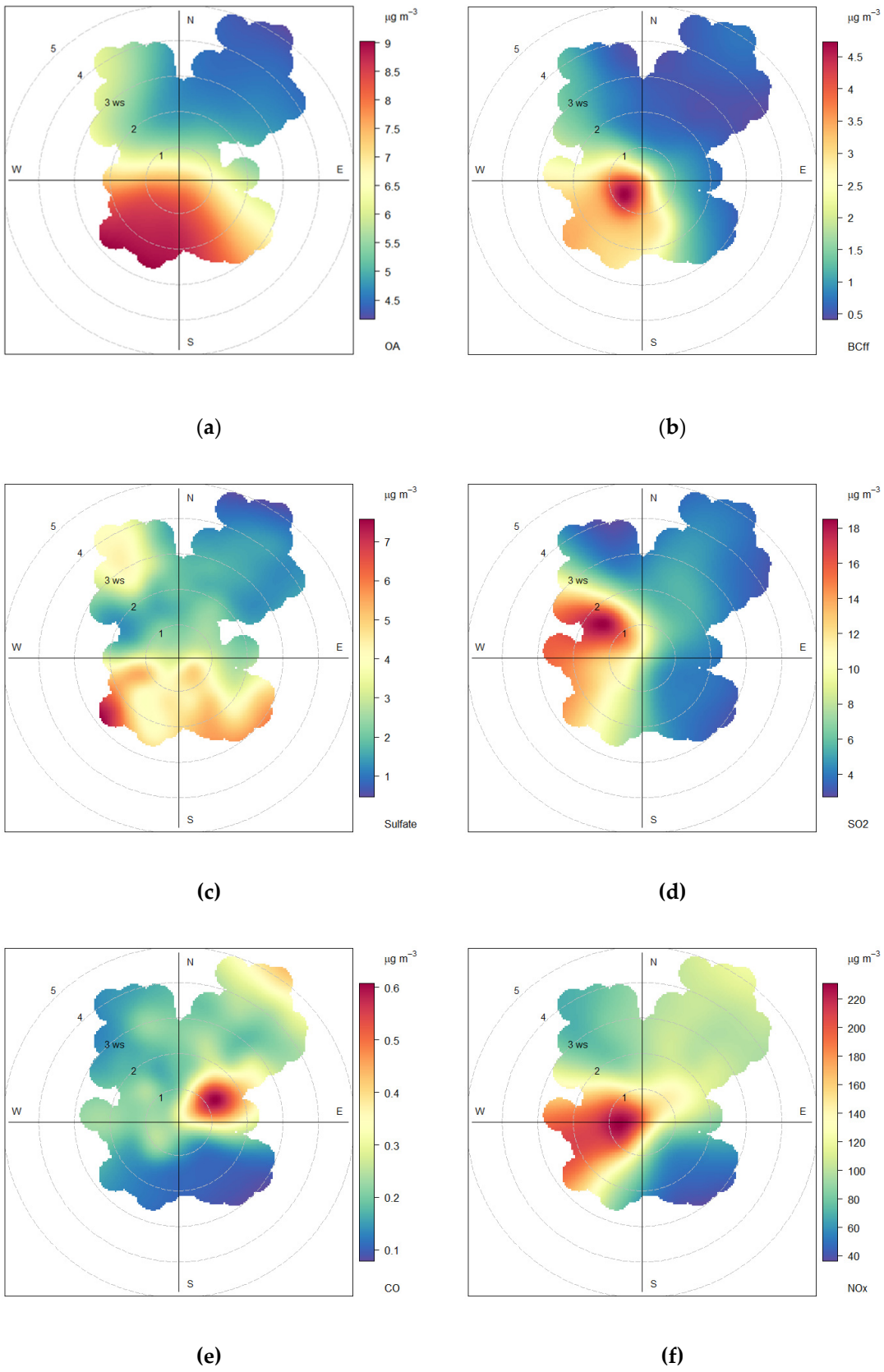


Figure S13. Bivariate polar wind plots for the effects of wind direction and wind speed (radial axis in m s^{-1}) to concentrations of OA (**a**), BC_{ff} (**b**), sulfate (**c**) as measured at P1 and SO_2 (**d**), CO (**e**), NO_x (**f**) as measured at PEI-1, during the summer period.

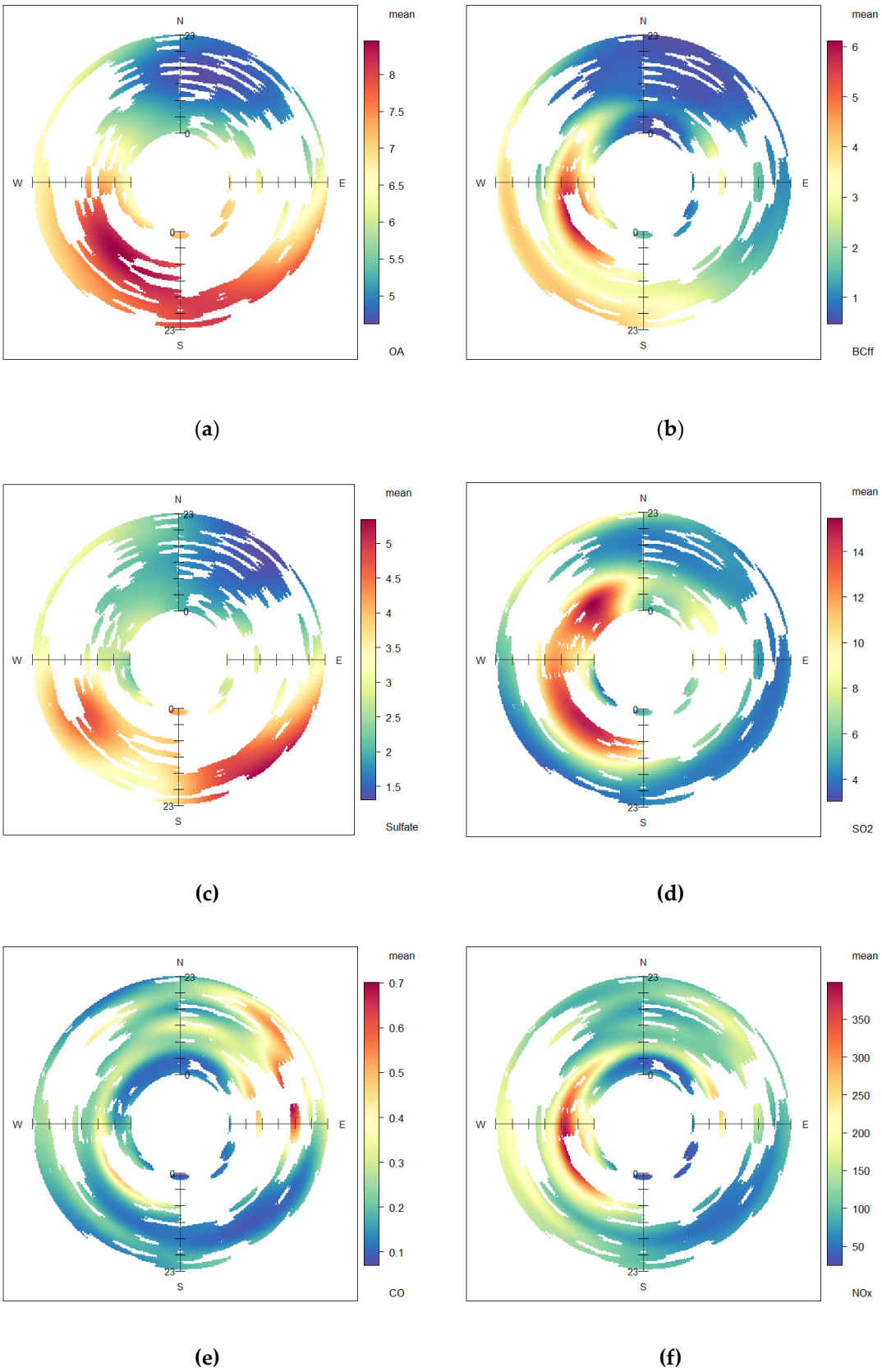
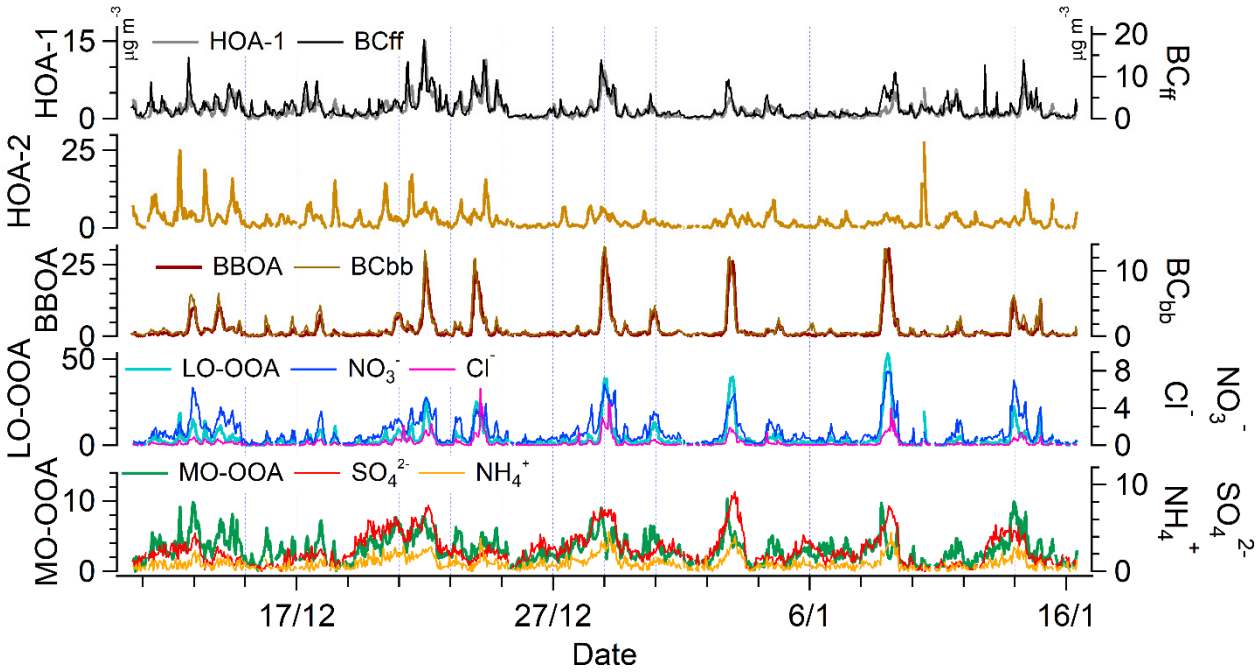
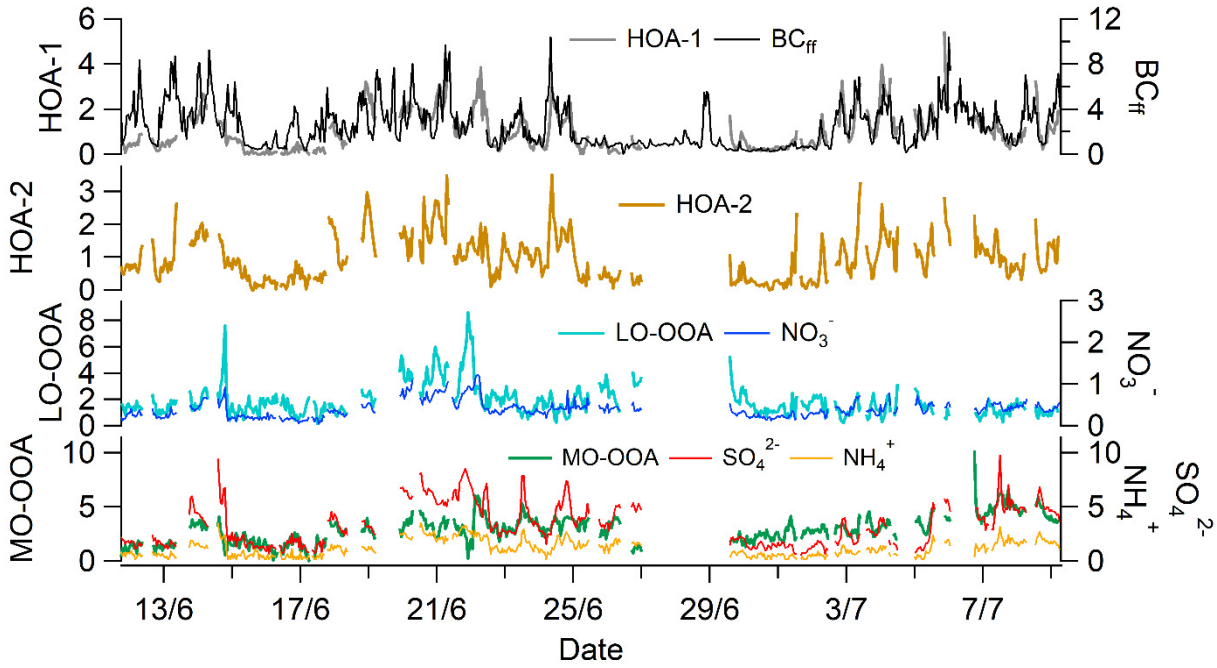


Figure S14. Annular wind plots for the effects of wind direction per hour of the day (radial axis, increments by 4 hr) to concentrations of OA (**a**), BC_{ff} (**b**), sulfate (**c**) as measured at P1 and SO_2 (**d**), CO (**e**), NO_x (**f**) as measured at PEI-1, during the summer period.



(**a**)



(b)

Figure S15. Time series of the PMF factors, along with key ACSM and aethalometer measurements for winter (**a**) and summer (**b**) campaigns.

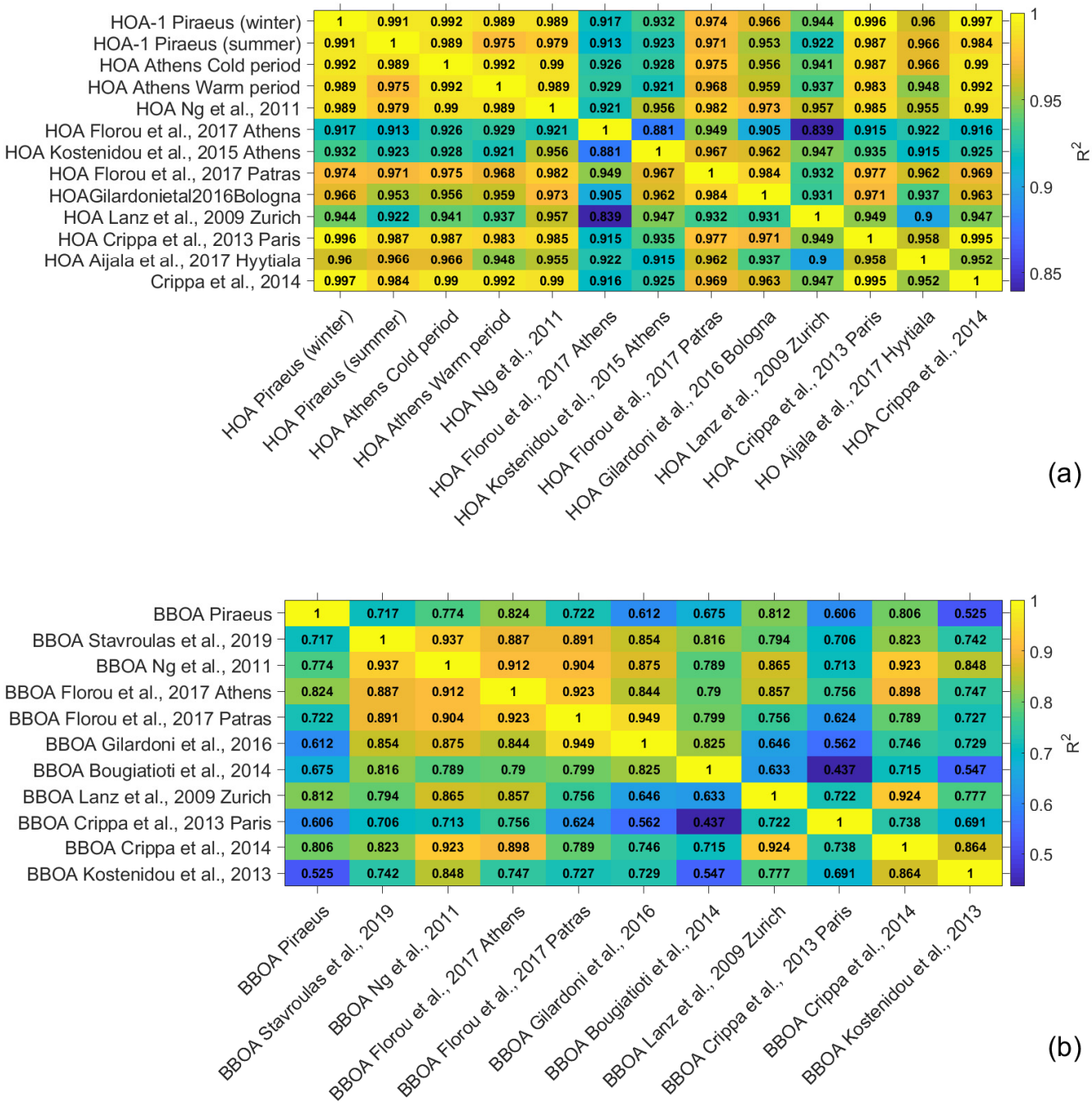


Figure S16. Correlation plots for the primary factor mass spectra HOA (**a**) and BBOA (**b**) obtained at Piraeus vs. a selection of published mass spectra in Greece and elsewhere [8–14].

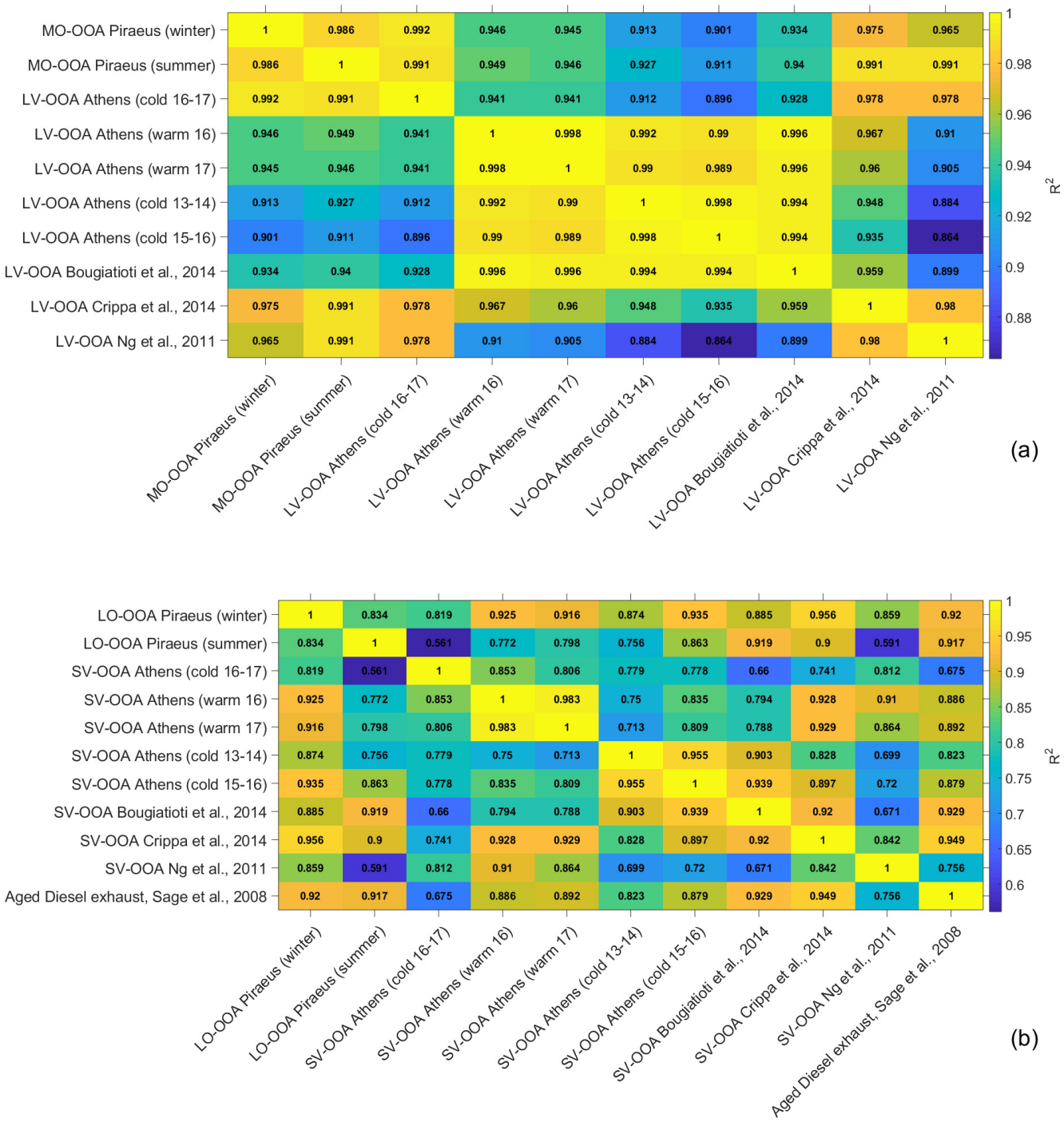


Figure S17. Correlation plots for the secondary factor mass spectra MO-OOA and LV-OOA (**a**), LO-OOA and SV-OOA (**b**) obtained at Piraeus vs. a selection of published mass spectra in Greece and elsewhere [8–15].

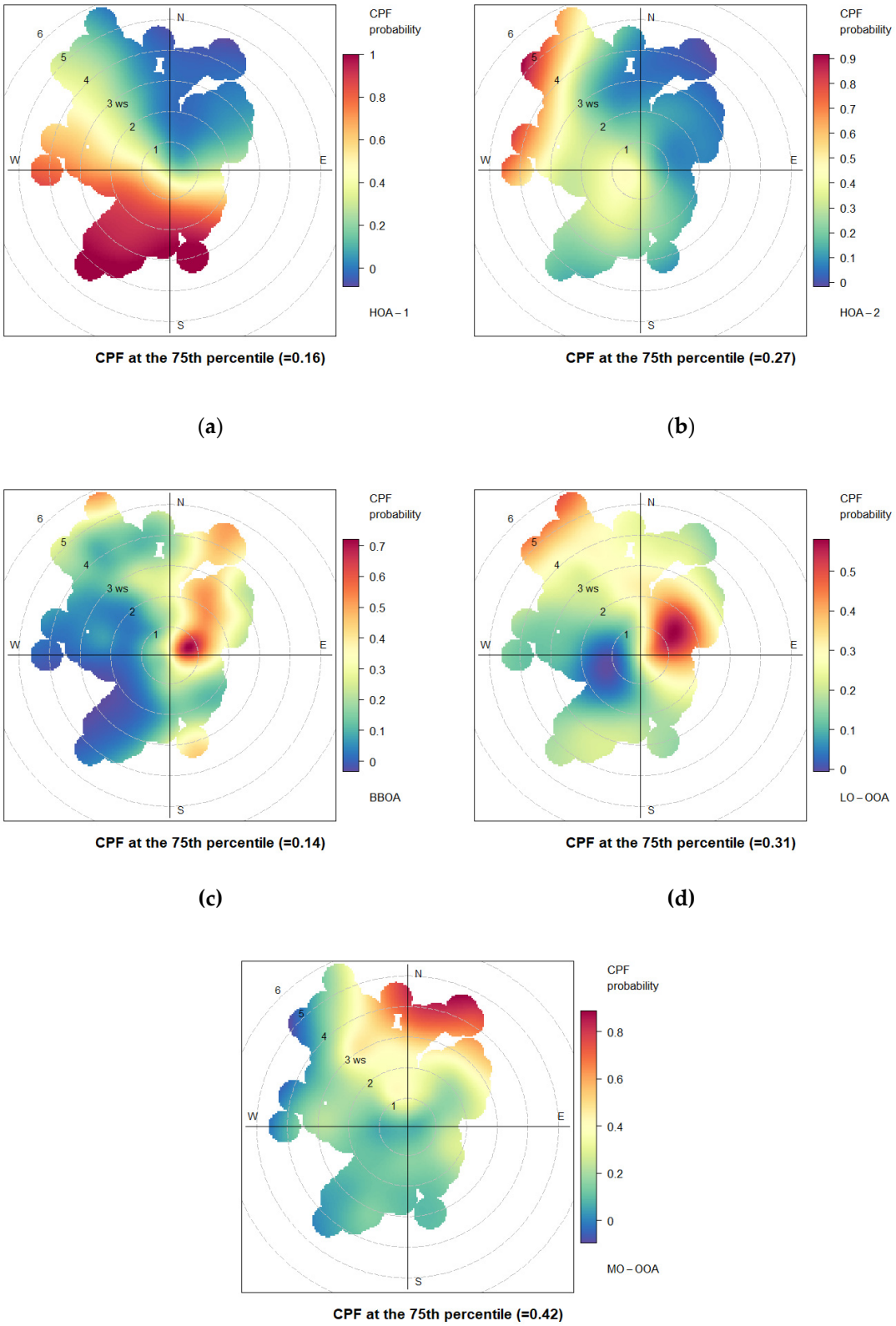


Figure S18. Bivariate polar plots for conditional probabilities of threshold exceedance (75th percentile), for fractional contributions of OA component in Piraeus, during the winter period. (**a**) HOA-1, (**b**) HOA-2, (**c**) BBOA, (**d**) LO-OOA, (**e**) MO-OOA.

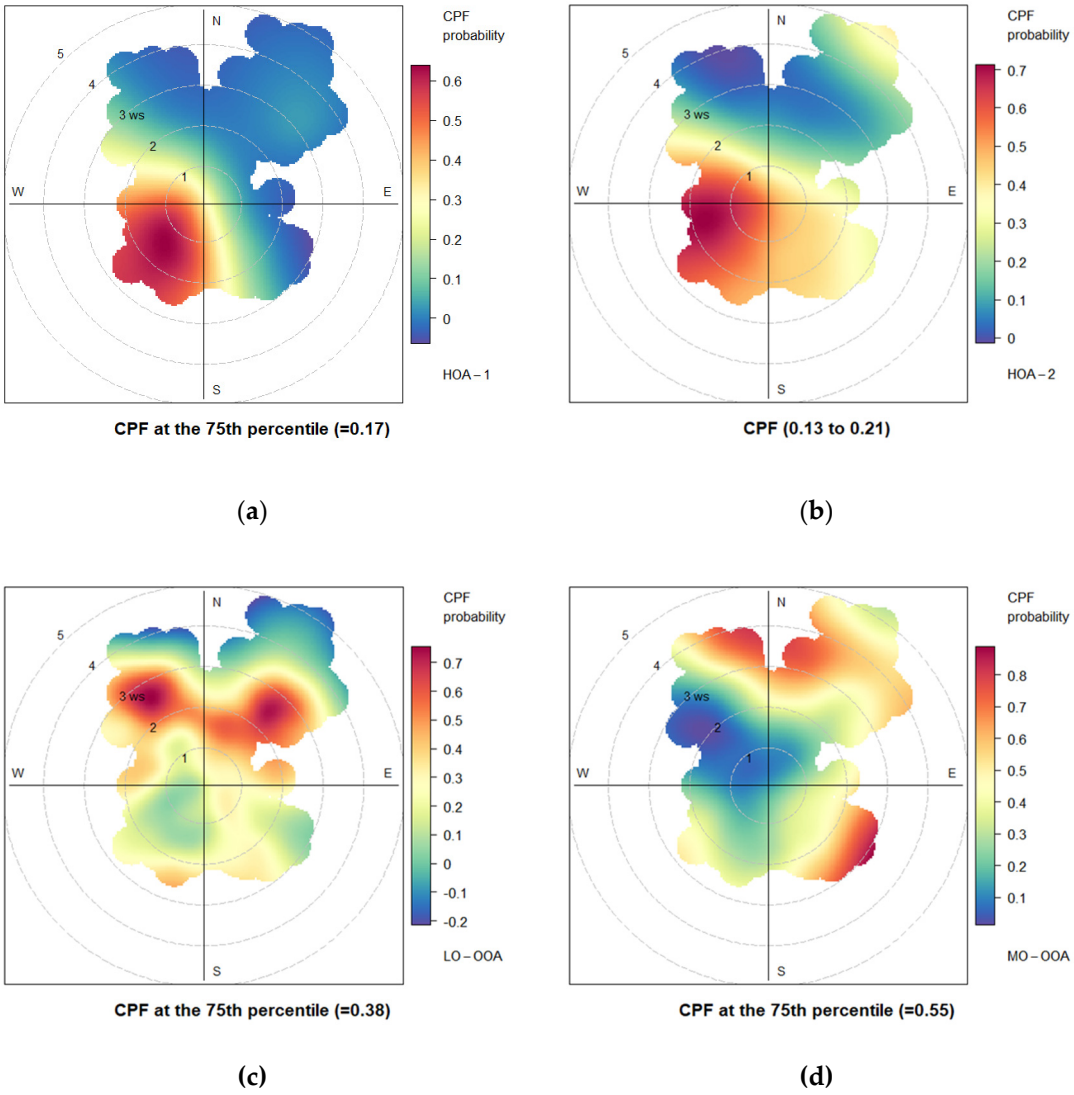


Figure S19. Bivariate polar plots for conditional probabilities of threshold exceedance (75th percentile), for fractional contributions of OA component in Piraeus, during the summer period. Displaying the 50th – 95th percentile range for (**a**) HOA-1, (**b**) HOA-2, (**c**) LO-OOA, (**d**) MO-OOA.

Table S2. Abbreviation Table.

Abbreviation	Meaning
AAE	Absorption Ångström Exponent
ACSM	Aerosol Chemical Speciation Monitor
BB	Biomass Burning
BBOA	Biomass Burning Organic Aerosol
BC	Black Carbon
BC_{bb}	Biomass Burning BC
BC_{ff}	Fossil Fuel Combustion BC
COA	Cooking Organic Aerosol
CPF	Conditional Probability Function
EC	Elemental Carbon
FP	Factor Profile
GAA	Greater Area of Athens
HDV	Heavy-Duty Vehicles

HOA	Hydrocarbon-like Organic Aerosol
LO-OOA	Less Oxidized Oxygenated Organic Aerosol
LV-OOA	Low-Volatility Oxygenated organic Aerosol
MAAP	Multi-Angle Absorption Photometer
MO-OOA	More Oxidized Oxygenated Organic Aerosol
OA	Organic Aerosol
OC	Organic Carbon
PMF	Positive Matrix Factorization
POA	Primary Organic Aerosol
PSCF	Potential Source Contribution Function
RWB	Residential Wood Burning
SOA	Secondary Organic Aerosol
SV-OOA	Semi-Volatile Oxygenated Organic Aerosol
WHO	World Health Organization
ΔBC_u	"Urban" Enhancement of BC ($BC_{660nm,P1} - BC_{637nm,P2}$)

References

1. Paatero, P.; Tapper, U. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **1994**, *5*, 111–126, doi:10.1002/ENV.3170050203.
2. Paatero, P.; Hopke, P.K.; Song, X.H.; Ramadan, Z. Understanding and controlling rotations in factor analytic models. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 253–264, doi:10.1016/S0169-7439(01)00200-3.
3. Canonaco, F.; Crippa, M.; Slowik, J.G.; Baltensperger, U.; Prévôt, A.S.H. SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data. *Atmos. Meas. Tech.* **2013**, *6*, 3649–3661, doi:10.5194/AMT-6-3649-2013.
4. Paatero, P.; Hopke, P.K. Rotational tools for factor analytic models. *J. Chemom.* **2009**, *23*, 91–100, doi:10.1002/CEM.1197.
5. Mohr, C.; Decarlo, P.F.; Heringa, M.F.; Chirico, R.; Slowik, J.G.; Richter, R.; Reche, C.; Alastuey, A.; Querol, X.; Seco, R.; et al. Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data. *Atmos. Chem. Phys* **2012**, *12*, 1649–1665, doi:10.5194/acp-12-1649-2012.
6. Crippa, M.; Canonaco, F.; Lanz, V.A.; Äijälä, M.; Allan, J.D.; Carbone, S.; Capes, G.; Ceburnis, D.; Dall'Osto, M.; Day, D.A.; et al. Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach. *Atmos. Chem. Phys.* **2014**, *14*, 6159–6176, doi:10.5194/ACP-14-6159-2014.
7. Crippa, M.; Haddad, I. El; Slowik, J.G.; DeCarlo, P.F.; Mohr, C.; Heringa, M.F.; Chirico, R.; Marchand, N.; Sciare, J.; Baltensperger, U.; et al. Identification of marine and continental aerosol sources in Paris using high resolution aerosol mass spectrometry. *J. Geophys. Res. Atmos.* **2013**, *118*, 1950–1963, doi:10.1002/JGRD.50151.
8. Ng, N.L.; Canagaratna, M.R.; Jimenez, J.L.; Zhang, Q.; Ulbrich, I.M.; Worsnop, D.R. Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data. *Environ. Sci. Technol.* **2011**, *45*, 910–916, doi:10.1021/es102951k.
9. Stavroulas, I.; Bougiatioti, A.; Grivas, G.; Paraskevopoulou, D.; Tsagkaraki, M.; Zarmpas, P.; Liakakou, E.; Gerasopoulos, E.; Mihalopoulos, N. Sources and processes that control the submicron organic aerosol composition in an urban Mediterranean environment (Athens): A high temporal-resolution chemical composition measurement study. *Atmos. Chem. Phys.* **2019**, *19*, 901–919, doi:10.5194/acp-19-901-2019.
10. Florou, K.; Papanastasiou, D.K.; Pikridas, M.; Kaltsonoudis, C.; Louvaris, E.; Gkatzelis, G.I.; Patoulias, D.; Mihalopoulos, N.; Pandis, S.N. The contribution of wood burning and other pollution sources to wintertime organic aerosol levels in two Greek cities. *Atmos. Chem. Phys.* **2017**, *17*, 3145–3163, doi:10.5194/ACP-17-3145-2017.
11. Gilardoni, S.; Massoli, P.; Paglione, M.; Giulianelli, L.; Carbone, C.; Rinaldi, M.; Decesari, S.; Sandrini, S.; Costabile, F.; Gobbi, G.P.; et al. Direct observation of aqueous secondary organic aerosol from biomass-burning emissions. *Proc. Natl. Acad. Sci.* **2016**, *113*, 10013–10018, doi:10.1073/PNAS.1602212113.
12. Bougiatioti, A.; Stavroulas, I.; Kostenidou, E.; Zarmpas, P.; Theodosi, C.; Kouvarakis, G.; Canonaco, F.; Prévôt, A.S.H.; Nenes, A.; Pandis, S.N.; et al. Processing of biomass-burning aerosol in the eastern Mediterranean during summertime. *Atmos. Chem. Phys.* **2014**, *14*, 4793–4807, doi:10.5194/ACP-14-4793-2014.
13. Äijälä, M.; Heikkinen, L.; Fröhlich, R.; Canonaco, F.; Prévôt, A.S.H.; Junninen, H.; Petäjä, T.; Kulmala, M.; Worsnop, D.; Ehn, M. Resolving anthropogenic aerosol pollution types - Deconvolution and exploratory classification of pollution events. *Atmos. Chem. Phys.* **2017**, *17*, 3165–3197, doi:10.5194/acp-17-3165-2017.
14. Kostenidou, E.; Kaltsonoudis, C.; Tsiflikiotou, M.; Louvaris, E.; Russell, L.M.; Pandis, S.N. Burning of olive tree branches: A major organic aerosol source in the Mediterranean. *Atmos. Chem. Phys.* **2013**, *13*, 8797–8811, doi:10.5194/acp-13-8797-2013.
15. Sage, A.M.; Weitkamp, E.A.; Robinson, A.L.; Donahue, N.M. Evolving mass spectra of the oxidized component of organic aerosol: results from aerosol mass spectrometer analyses of aged diesel emissions. *Atmos. Chem. Phys* **2008**, *8*, 1139–1152.