Decadal trends (2013–2023) in PM10 sources and oxidative potential at a European urban supersite (Alpine Valley, Grenoble, France)

Vy Ngoc Thuy Dinh[1], Jean-Luc Jaffrezo[1], Pamela A. Dominutti[1], Rhabira Elazzouzi[1], Sophie Darfeuil[1], Céline Voiron[1], Anouk Marsal[1], Stéphane Socquet[2], Gladys Mary[2], Julie Cozic[2], Catherine Coulaud[1], Marc Durif[3,4], Olivier Favez[3,4], Gaëlle Uzu[1]

[1] Université Grenoble Alpes, CNRS, IRD, INP-G, INRAE, IGE (UMR 5001), 38000 Grenoble, France
[2] Atmo Auvergne-Rhône-Alpes (Atmo AuRA), 69500 Bron, France
[3] INERIS, Parc Technologique Alata, BP 2, 60550 Verneuil-en-Halatte, France
[4] Laboratoire central de surveillance de la qualité de l'air (LCSQA), 60550 Verneuil-en-Halatte, France
Correspondence to: Gaëlle Uzu *gaëlle.uzu@univ-grenoble-alpes.fr*

S1. Multivariate imputation by chained equations (MICE)

Multivariate imputation by chained equations (MICE) is a flexible algorithm that could be combined with other statistical techniques for estimating missing values by using observed data (by performing the linear, multilinear regression or calculation of the mean/median to fill the missing) (Azur et al., 2011). The MICE algorithm employs multiple regression models, where each missing value is conditionally modeled based on the observed (non-missing) values. This imputation process provides a more complete database for PMF analysis, contributing to a reduction in recommendation errors in the results (Ocepek et al., 2015).

Here, the data preparation and imputation processes were implemented through 4 main steps. First, all missing values of each metal are replaced by the mean of its concentration, these replaced values are marked as "place-holders". Second step, the metal that has the fewest missing values was chosen, and the "place-holder" is put back to missing (called Px). In the third step, a multiple linear regression is applied, with the chosen metal set as a criterion and the other variables are predictors. Fourth step, the missing value Px is calculated by using the slope and intercepts of multiple linear regression in the third step. The processes between the second and the fourth steps are repeated until all "place-holders" are replaced by the regression prediction value; this repeated run is called the "cycle" of imputation. The imputation cycle is iterative implemented until the difference between the 2 last cycles is minimized. MICE was implemented in Python 3.9 using the package "*scikit-learn 1.2.0*" (Pedregosa et al., 2011).

30
31

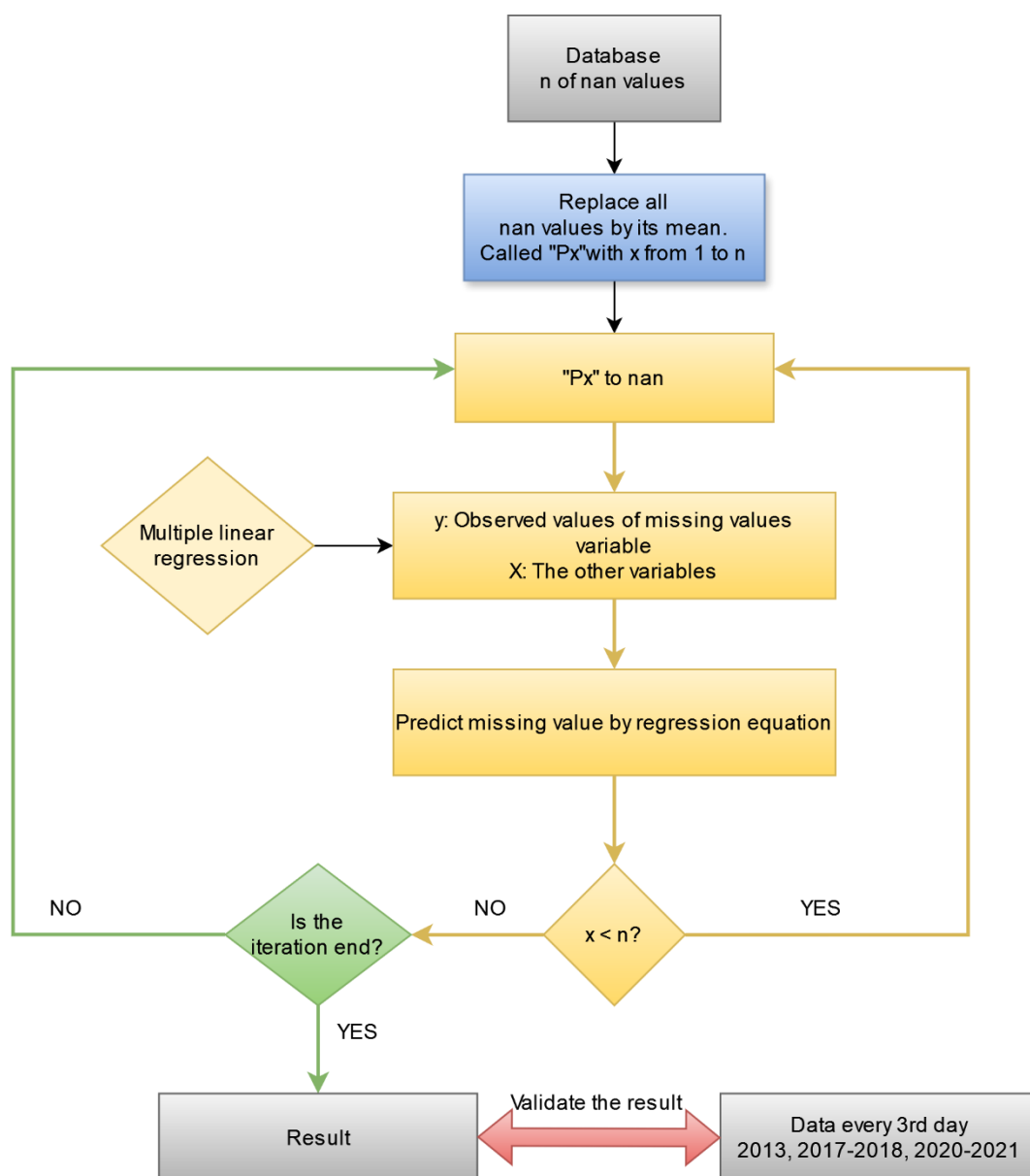
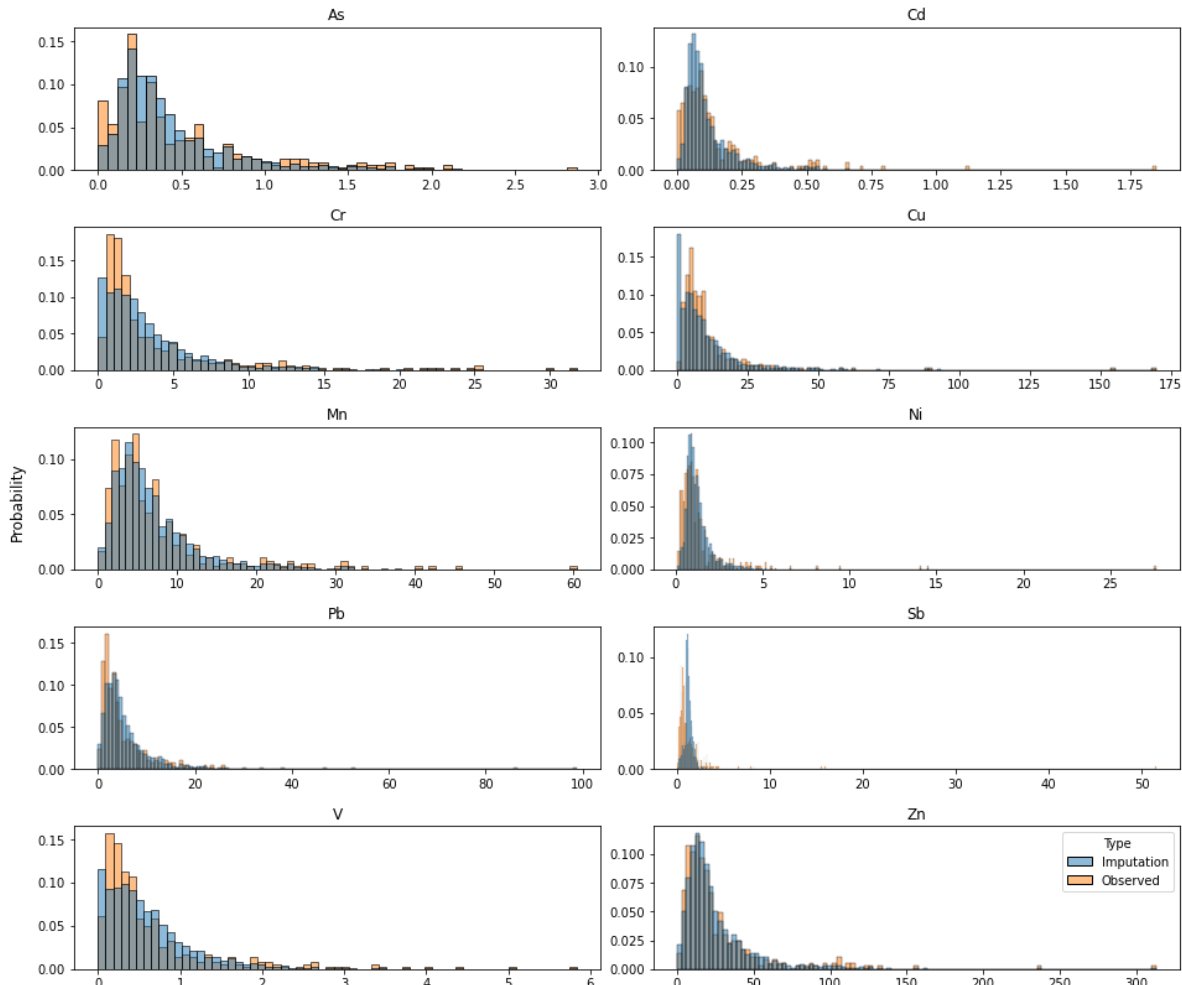



Figure S.1. Workflow of the MICE algorithm

33

34   **Table S 1. R^2 between imputed and measured metals**

	2013	**2017-2018**	**2020-2021**
As	0.43	0.6	0.2
Cd	0.41	0.55	0.1
Cr	0.32	0.34	0.38
Cu	0.63	0.4	0.42
Mn	0.33	0.35	0.5
Ni	0.14	0.33	0.14
Pb	0.33	0.31	0.15
Sb	0.37	0.50	0.2
V	0.41	0.27	0.52
Zn	0.26	0.35	0.20

35
36 **Figure S.2. Distribution of measured and imputed metals**

37

38 **S2. Some equations**

39 PMF equation:

$$X_{ij} = \sum_{k=1}^{p} G_{ik} * F_{kj} + E_{ij} \quad (Eq.(S1))$$

41
$$G_{ik} \geq 0, \ F_{ik} \geq 0 \ (Eq.(S2))$$

42 Where X is a (i × j) matrix of j chemical species in measured period i (daily) into p factors with a matrix (i x k)
43 representing the source contribution (G) and a matrix (k x j) representing the factor composition (F). E is the
44 residuals for each species. All the factor matrices G and F elements are constrained to be non-negative.

45

46 Variance Inflation Factor (VIF) represent the collinearity between the PM sources, which is calculated:

47
$$VIF_i = \frac{1}{1 - R_i^2}, i = 1, \dots, p - 1 \ (Eq.(S3))$$

48

49 Chemical Mass Closure:
50 PM10 ≈ EC + OM + [sea salt] + [nss-SO$_4^{2-}$] +[dust] + [non dust]

51 with:

52 [OM] = [OC]*1.8 (Favez et al., 2010)

53 [sea salt] = [Na$^+$]*1.47 + [Cl-] (Putaud et al., 2010)

54 [nss-SO42-] = [SO$_4^{2-}$] – [Na$^+$]*0.252 (Alexander et al., 2005)

55 [dust] = 5.6 * ([Ca^{2+}] – [Na$^+$]/26) (Putaud et al., 2004)

56 [non dust] = [Cu] + [Pb] + [V] + [Zn] (Salameh et al., 2015)

57 **S3. PMF input matrices**

58 The basic chemical species to be included in the input matrices are key components for the mass balance of the

59 PM: organic carbon (OC), element carbon (EC), major ions (nitrate, sulfate, ammonium), sea salt species (Na, Cl,

60 Mg), and mineral dust (Al, Fe, Ti). Then the input matrices should include a set of chemical tracers that allows the

61 discrimination of their sources. This list is very variable according to the analytical capabilities used in the research

62 program.

63

64 In other to avoid double counting of carbon mass in OC and in organic tracers added to the input data (ie

65 levoglucosan, mannosan, MSA, polyols in the classic PMF at IGE), we calculate a variable OC* by:

66 [OC*] = [OC] – [total mass C of organic tracers in PMF]

67

68 *Uncertainties of measured data*

69 A table of the uncertainties of each daily data for each chemical species should be constructed. Several ways for

70 their calculations exist. Here, we are using the formula described in the table below, resulting from all the work

71 developed in the SOURCE program and the post-doc work of D Salameh (published in Weber et al., 2019).

72 **Table S 2. Formula to calculate uncertainties**

Specie	Calculation by	Formula
OC*, EC, PM10	Fixed percentage	10%
Specie which has concentration < QL	Ratio of QL Polissar et al. (1998)	$\dfrac{5 \times QL}{6}$
Imputed metals	Gianini et al., 2012	$\sigma_{ij} = 2 * \sqrt{QL_j^2 + (CV_j \times x_{ij})^2 + \left(a \times x_{ij}\right)^2}$
The others	Gianini et al., 2012	$\sigma_{ij} = \sqrt{QL_j^2 + (CV_j \times x_{ij})^2 + \left(a \times x_{ij}\right)^2}$

73

74 With:

75 • QL : Quantification limit.

76 • CV : Coefficient of variation.

77 • a: Additional coefficient of variation

78 • x: Species concentration.

79 **S4. PMF criteria for validation and applied constraint**

80 According to the European guide on air pollution source apportionment with receptor models (Belis et al., 2014),

81 the validation criteria of a PMF solution include:

82   (1) Evolution of $Q_{true}/Q_{robust} < 1.5$: Indicate that the good result should not have more than 30% data outliers.

83   Generally, the final solution presents a much lower ratio.

84   (2) The chemical profile is clear: the concentration and percentage of trace species in the profile and the temporal

85   variability should be clear enough to identify a source.

86   (3) All factors should have a contribution > 1% to the total variable (PM10): to avoid a case where there is a source,

87   but it has almost no impact on the study area.

88   (4) The distribution of residuals: the distribution of residuals (differences between input data and reconstructed

89   data) is from -3 to 3, if there is any value is out of this range, that means that the reconstruction is not valid or that

90   there are outliers.

91   (5) Evaluate the species reconstruction: The correlation coefficient between measured and predicted concentrations

92   must be greater than 0.5.

93   (6) Bootstrap test: it indicates the stability of the solution: at least 70 runs per 100 runs for all factors where the

94   correlation between the base run and boot runs is greater than 0.6.

95   **Table S 3. The constraints applied for PMF**

Factor	Element	Type	Value
Industrial	Levoglucosan	Set to Zero	0
Industrial	Mannosan	Set to Zero	0
Industrial	PM10	Define Limits	0.1/0.4
MSA rich	MSA	Pull Up Maximally	NA
MSA rich	Polyols	Set to Zero	0
Biomass burning	Levoglucosan	Pull Up Maximally	NA
Biomass burning	Mannosan	Pull Up Maximally	NA
Primary biogenic	Polyols	Pull Up Maximally	NA
Primary traffic	Ba	Pull Up Maximally	NA
Primary traffic	Cu	Pull Up Maximally	NA
Primary traffic	OC*/EC	Ratio	0.44
Primary traffic	Cu/Sb	Ratio	12.6
Primary traffic	Cu/Mn	Ratio	5.7

96   **S5. PMF results**

97   **Table S 4. Q_{true}/Q_{robust} values**

	10 yeas	2013-2016	2017-2020	2021-2023
Base run	1.07	1.05	1.05	1.1
Constraint run	1.07	1.06	1.05	1.1

98

99

Table S 5. Bootstrap value before and after constraint

	10 yeas		2013-2016		2017-2020		2021-2023	
	Base	Constraint	Base	Constraint	Base	Constraint	Base	Constraint
Aged sea salt	100	100	100	100	100	100	100	100
Biomass burning	100	100	100	100	100	100	100	100
Industrial	100	100	100	100	100	100	100	100
MSA rich	100	100	97	100	100	100	93	100
Mineral dust	100	100	93	100	98	100	85	100
Nitrate rich	100	100	100	100	100	100	100	100
Primary biogenic	100	100	100	100	100	100	100	100
Primary traffic	80	98	82	96	93	100	70	88
Chloride rich	98	100	99	100	92	100	99	100
Sulfate rich	100	100	87	99	98	100	96	95

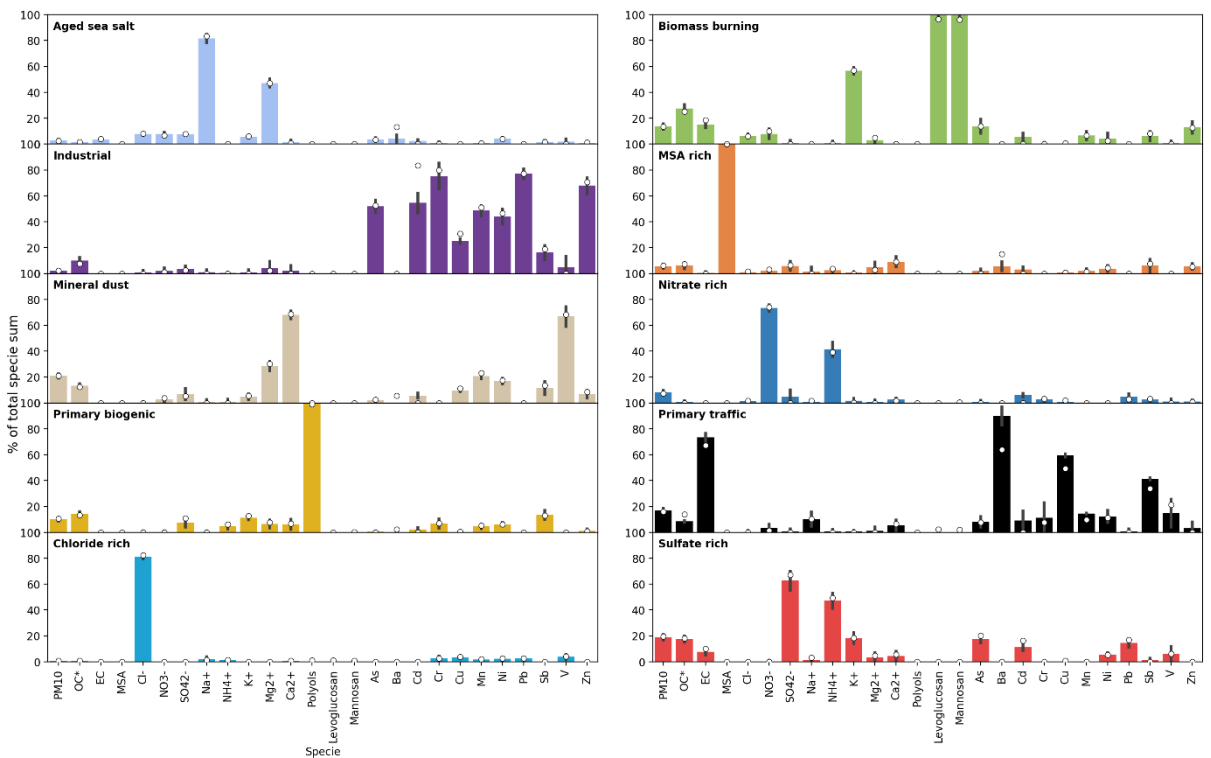
101
102

Table S 6. R2 between observed and predicted by PMF of PM concentration

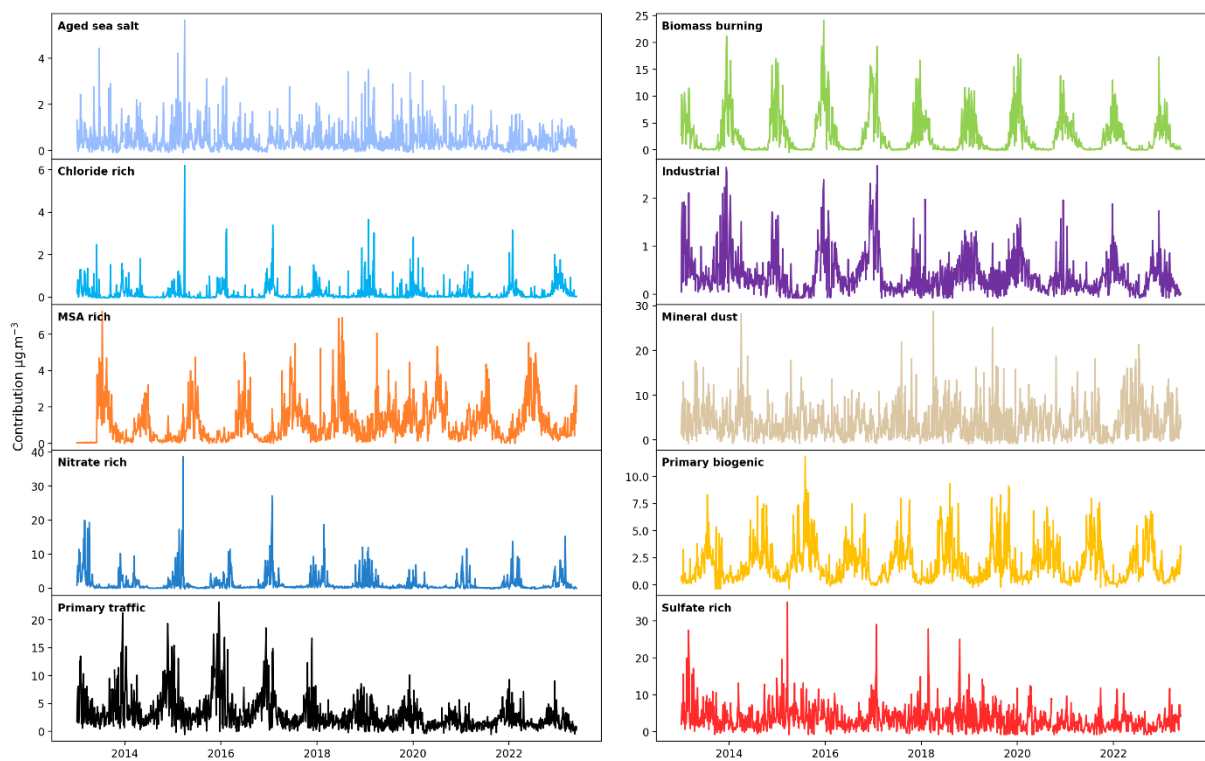
	10 yeas	**2013-2016**	**2017-2020**	**2021-2023**
Base run	0.97	0.98	0.97	0.98
Constraint run	0.97	0.98	0.97	0.98

103

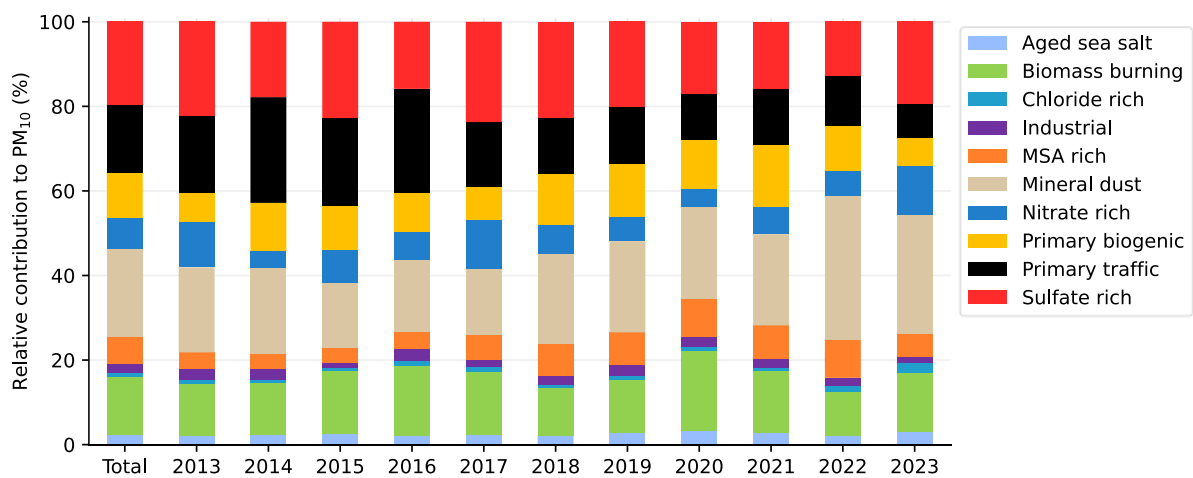
104 Displacement run: No warning for unstable or un-useable solution for PMF performed on 10-years dataset and
105 every 3 years datasets.

106

107
108 **Figure S.3. Chemical profiles of PM$_{10}$ sources, solution 10-year**

109

110

111 **Figure S.4. Temporal evolution of PM sources, solution 11 years**

112

113

114 **Figure S.5. The absolute average contribution of sources to PM$_{10}$ (%) for 11 years and every years.**

115

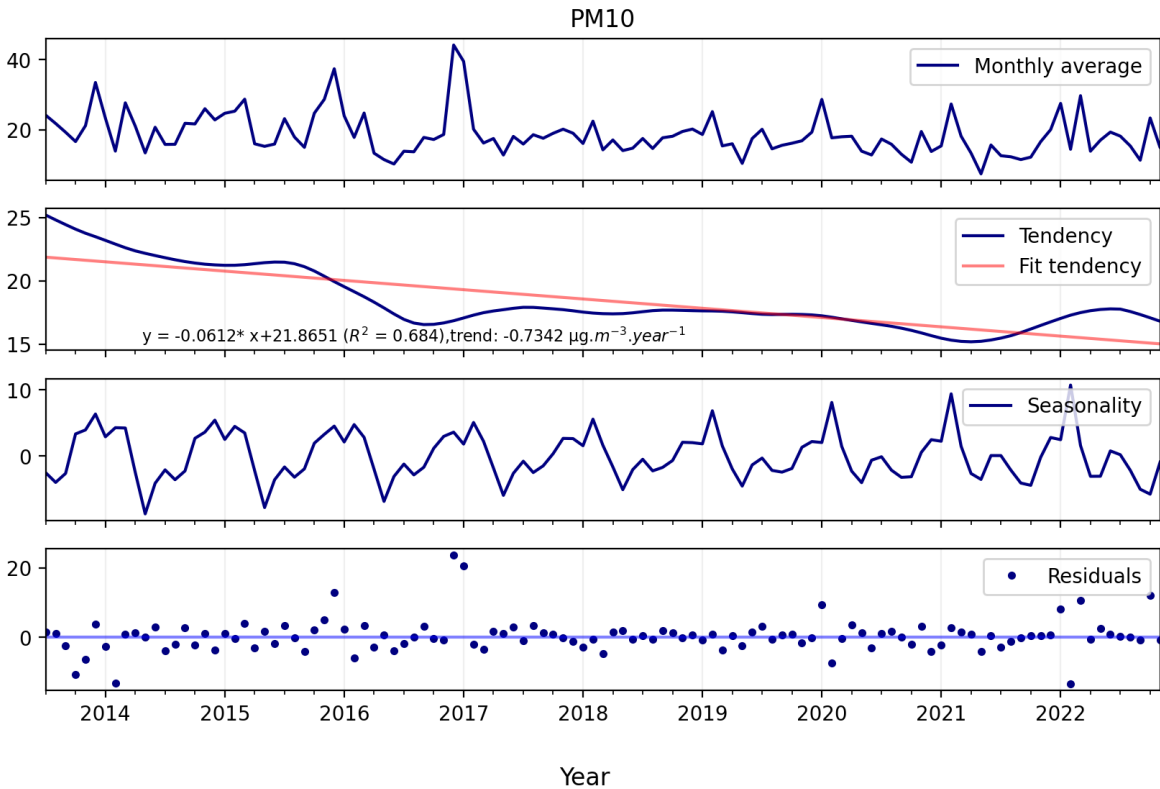
116

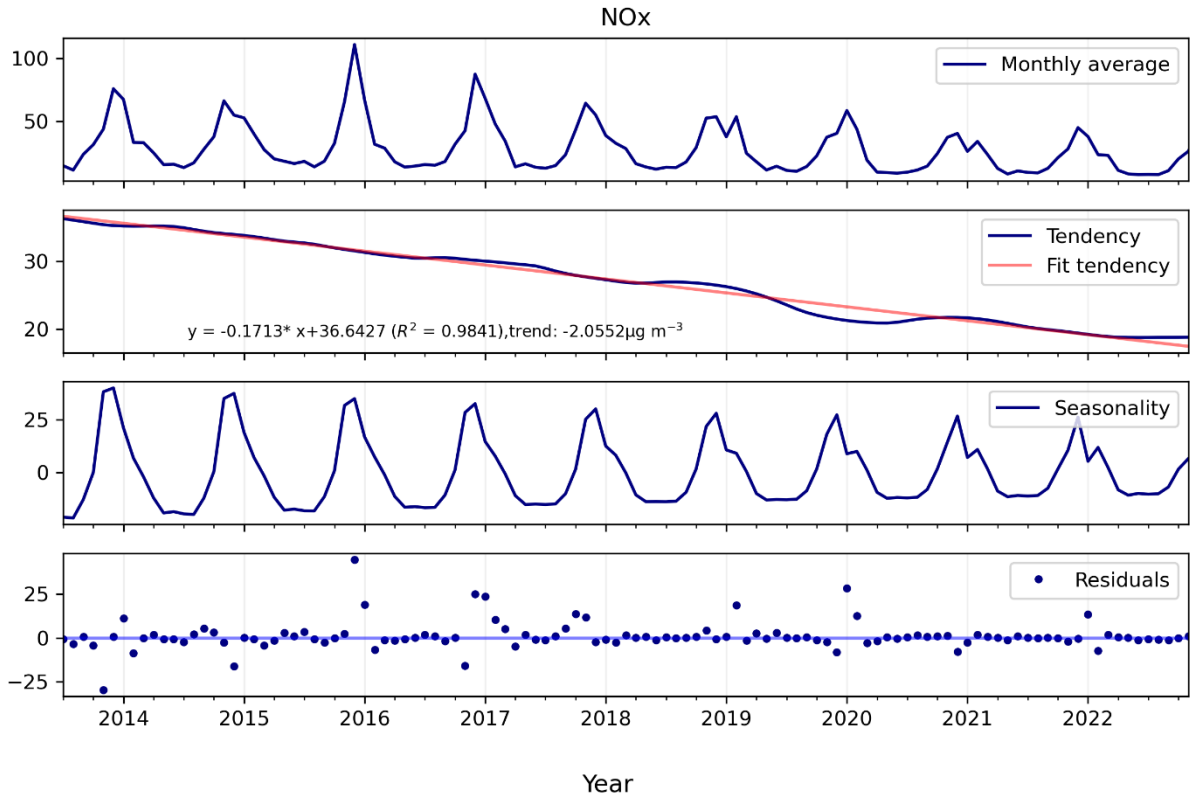
117

118

119

120

121 **S6. Tendency of PM10 and NOx**

122
123 **Figure S.6. STL decomposition of PM10**

124

125
126 **Figure S.7. STL decomposition of NO$_x$**

127

S7. Thermal inversion analysis

This analysis is conducted for the period Nov 2017 to May 2023, when the measurements on the mountain slopes are available. A good correlation is obtained between the PM_{10} and bulk temperature at the city level (r reaching 0.6, p<<0.001) for the winter months and it is even better when considering only the persistent inversion periods (r reaching 0.7, p<<0.001) for individual years (Table S7). Interestingly, the bivariate distribution between the daily PM_{10} concentration and daily average $\Delta T/\Delta z$ in winter months revealed that the majority of PM_{10} concentration peaks (exceedances above 40 µg m^{-3}) are from persistent inversion (Figure S8). The distribution of the day without and with persistent inversion (Figure S9) also shows that a few high PM_{10} concentration could be found in the days without persistent inversion, meanwhile a day with persistent inversion is not always associated to a high PM_{10} concentration. This result is not surprising, since the concentration of PM_{10} also depend on other meteorological conditions (precipitation, heat deficit) as well as variations in pollutant emissions (Carbone et al., 2010; Largeron and Staquet, 2016).

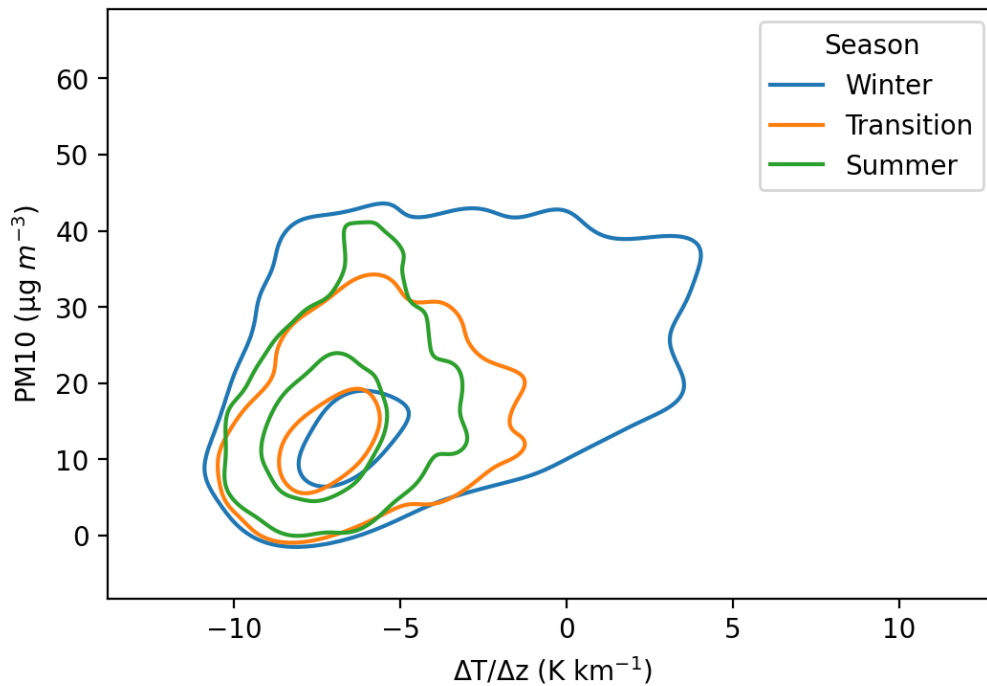



Figure S.8. Bivariate distribution between PM10 and bulk temperature in winter (blue contour, for the month of Dec, Jan, Feb), summer (green contour for the month of Jun, Jul, Aug) and transition season (orange contour for the remaining months).

Table S 7. Correlation between PM and bulk temperature in winter and in persistent inversion period. The star repressing the p-value. (* : < 0.001, **: < 0.01, * < 0.05, "ns": non-significant)**

	R^2 in winter	R^2 in persistent inversion period
2017	0.60**	0.50**
2018	0.19*	0.47**
2019	0.50***	0.33**
2020	0.57***	0.47***
2021	0.50***	0.67**
2022	0.30***	0.22*
2023	0.05 ns	0.05 ns

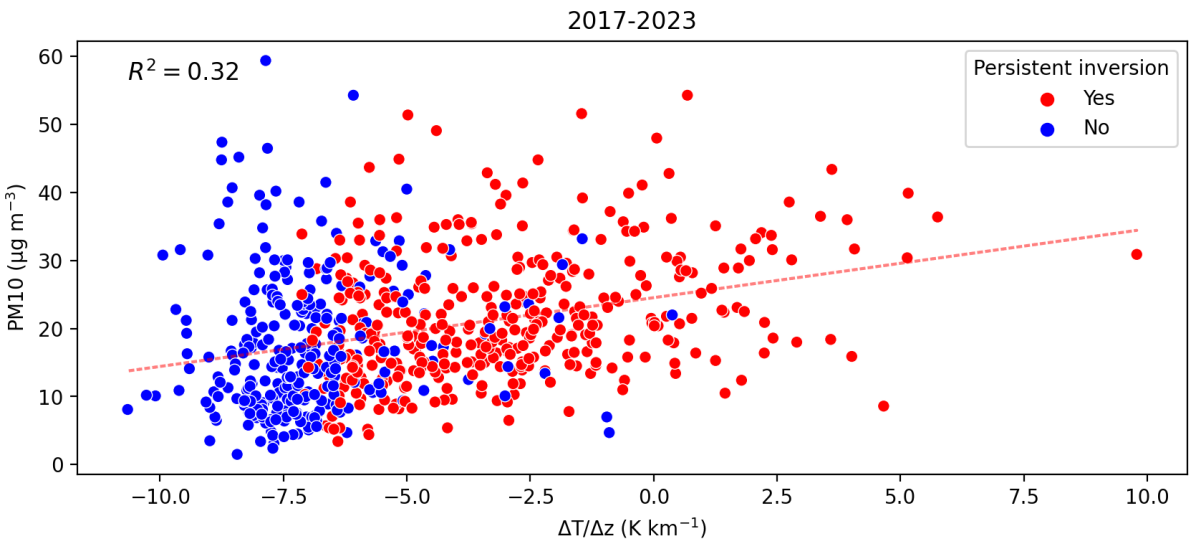



Figure S.9. Daily PM$_{10}$ concentration and daily average $\Delta T/\Delta z$ in winter months (from November to March) for the period of 2017 to 2023. The red points represent the winter days when the persistent inversion is detected, and the blue points represent the other winter days.

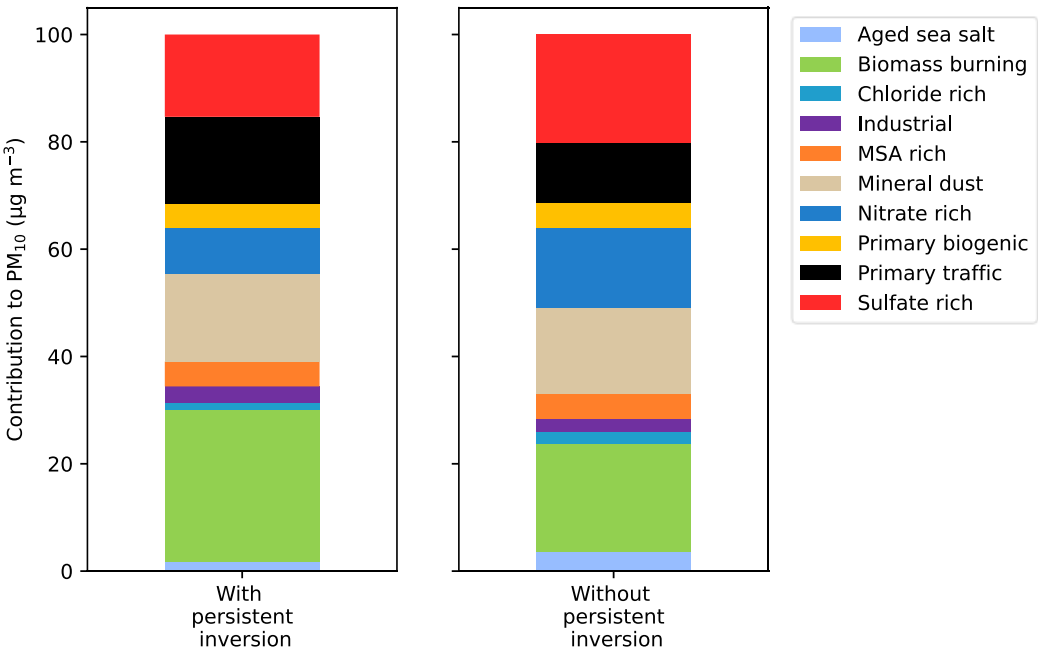



Figure S.10. Relative contribution of the different sources to PM$_{10}$ for days with persistent inversion vs non-inversion days of the winters 2017-2023.

156

157 Thermal inversion events were used to evaluate the relationship between biomass burning influence onPM_{10}

158 concentration and meteorological conditions. As shown in Figure S9, more than half of the high PM_{10}

159 concentration event is related to persistent inversion. Especially for biomass burning, which is emitted principally

160 in winter, could be highly enlarged under the influence of inversion. Figure 7 presents the contribution of biomass

161 burning and the bulk temperature in winter, reveals that 41% of the biomass burning contribution is explained by

162 the temperature gradient ($\Delta T/\Delta z$) between 2017 and 2023. better than considering PM_{10} ($R^2 = 0.3$). In particular,

163 the high contribution of biomass burning ($>10\,\mu g\,m^{-3}$) is always found in the episodes of persistent inversion. This

164 demonstrates that the occurrence of persistent inversion systematically decreasing the temperature in the city

165 enhances the use of heating as well as traps the aerosol and enlarges the contribution of biomass burning. Although

166 the relationship between persistent inversion events and biomass burning emissions is affirmed, the reduction trend

167 of biomass burning does not completely depend on the persistent inversion events. As shown in Figure 5, the

168 number of persistent inversions per year is nearly similar, while the contribution of biomass burning steadily

169 decreases over the years. The average annual biomass burning sources PMF-derived is compared to the local PM_{10}

170 emission inventory by residential heating (tonnes), provided by the Central air quality monitoring laboratory

171 (Atmo AuRA) to confirm the trend of biomass burning. The annual average of biomass burning is agreed with the

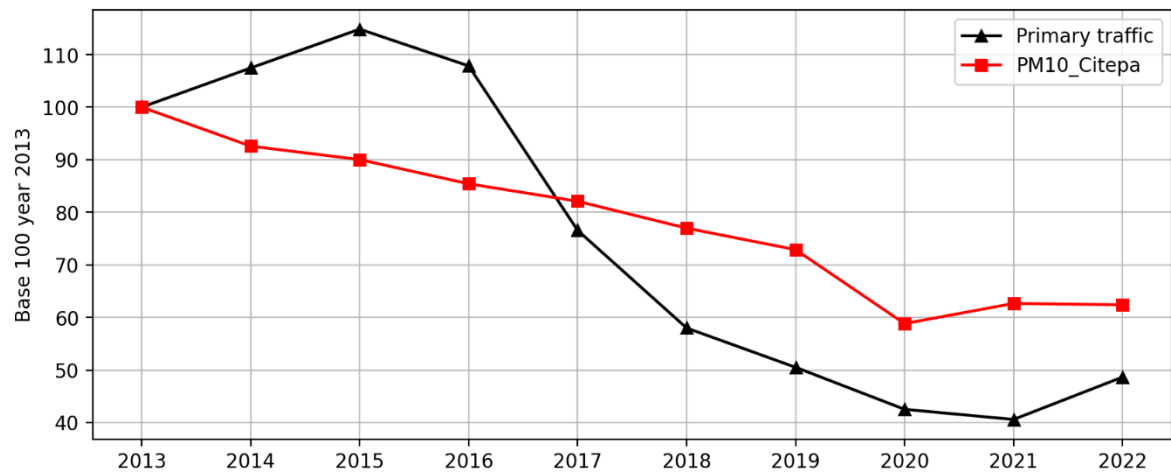
172 emission inventory, demonstrating the consistency between the sources observed by the PMF model and the local

173 inventory emission data. Since 2015, the Grenoble Metropolis has set up an air-wood bonus to encourage

174 households to renew their individual wood-burning appliance (fireplace or stove), and aim to replace all open

175 fireplaces with closed appliances in October 2024. The downward trend of biomass burning demonstrates the

176 effectiveness of the region's policy in changing residential heating equipment, to improve the city's air quality.

177

178
179 **Figure S.11. Comparison of the traffic contribution and PM10 emission by the transport sector in France.**

180
181

S8. OP PM10 SA result

Table S 8. The appropriate model of OP SA

	OP_{AA}	OP_{DTT}
Heteroscedasticity	Yes	Yes
Collinearity	No	No
The suitable model	wPLS, WLS	wPLS, WLS

Table S 9. The R^2, RMSE and MAE of the suitable models

	OP_{AA}		OP_{DTT}	
	wPLR	**WLS**	**wPLR**	**WLS**
R^2	0.69	0.70	0.61	0.61
RMSE	0.88	0.85	0.89	0.86
MAE	0.49	0.47	0.63	0.62

Tendency of OP_m and OP_v

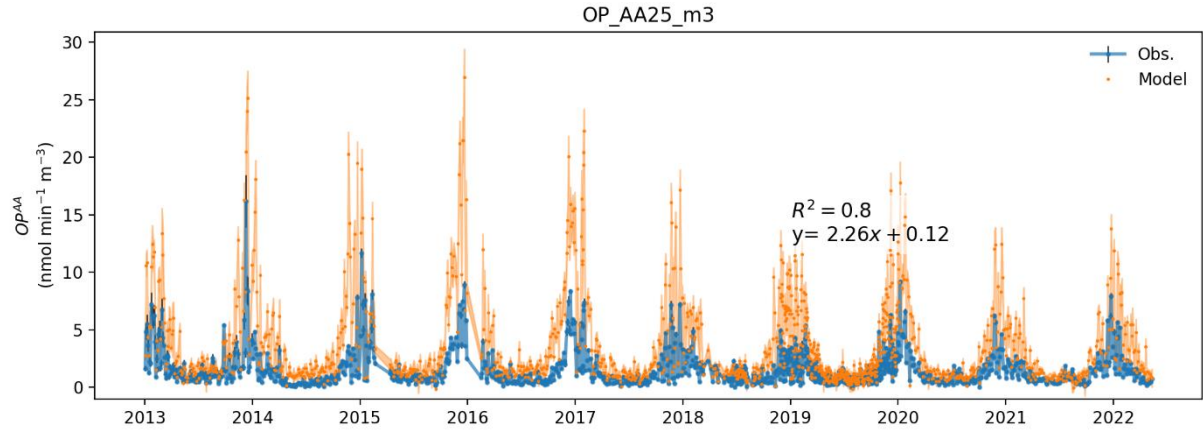



Figure S.12. The comparison between observed OP_{AA} and predicted OP_{AA}

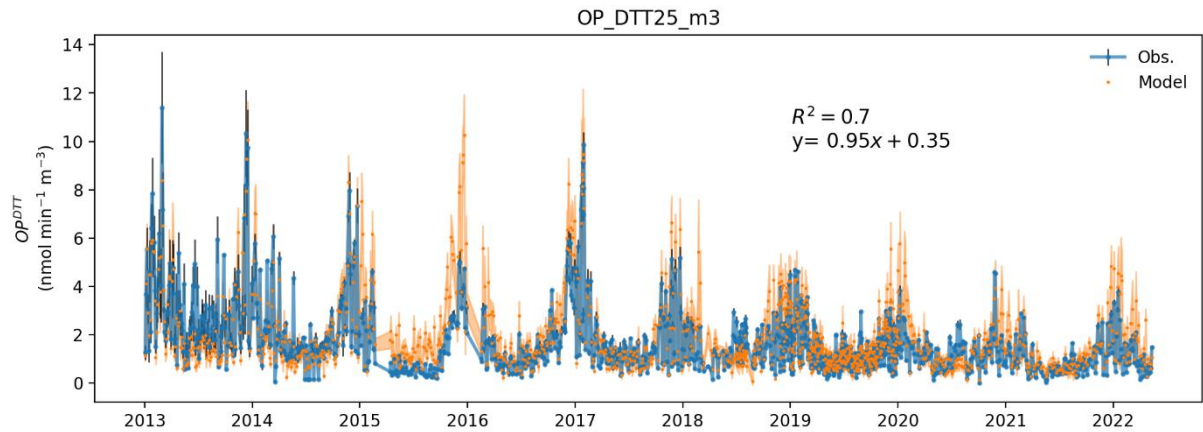



Figure S.13. The comparison between observed OP_{DTT} and predicted OP_{DTT}

194 **Table S 10. Accuracy metric of the testing and training dataset of Multiple layer perceptron (MLP) and Random**
195 **Forest (RF)**

	MLP		RF	
	Training	**Testing**	**Training**	**Testing**
R^2	0.75	0.75	0.90	0.72
RMSE	0.67	0.74	0.32	0.79
MAE	0.36	0.39	0.17	0.41

196

197 **Table S 11. Intrinsic OP_{AA} and intrinsic OP_{DTT} description**

OPtype	Source	count	mean	std	min	25%	50%	75%	max	IQR
AA	Aged sea salt	500	-0.02	0.07	-0.23	-0.06	-0.03	0.00	0.19	0.06
	Biomass burning	500	0.76	0.13	0.48	0.67	0.74	0.82	1.22	0.15
	Chloride rich	500	-0.07	0.09	-0.32	-0.09	-0.05	-0.02	0.18	0.07
	Industrial	500	0.48	0.14	-0.04	0.44	0.52	0.57	0.74	0.12
	MSA rich	500	0.20	0.04	0.03	0.18	0.20	0.23	0.30	0.05
	Mineral dust	500	-0.03	0.06	-0.17	-0.07	-0.04	0.01	0.19	0.07
	Nitrate rich	500	0.09	0.16	-0.43	0.00	0.13	0.20	0.51	0.20
	Primary biogenic	500	0.00	0.04	-0.11	-0.02	0.00	0.03	0.10	0.05
	Primary traffic	500	0.38	0.10	0.02	0.32	0.38	0.46	0.67	0.14
	Sulfate rich	500	-0.01	0.08	-0.22	-0.05	-0.02	0.04	0.23	0.09
DTT	Aged sea salt	500	0.03	0.02	0.00	0.02	0.03	0.05	0.13	0.02
	Biomass burning	500	0.14	0.09	0.00	0.08	0.13	0.17	0.45	0.09
	Chloride rich	500	0.01	0.02	0.00	0.00	0.00	0.00	0.12	0.00
	Industrial	500	0.52	0.08	0.31	0.47	0.51	0.55	0.86	0.08
	MSA rich	500	0.01	0.02	0.00	0.00	0.00	0.02	0.13	0.02
	Mineral dust	500	0.01	0.02	0.00	0.00	0.00	0.00	0.17	0.00
	Nitrate rich	500	0.11	0.12	0.00	0.02	0.07	0.12	0.50	0.10
	Primary biogenic	500	0.02	0.03	0.00	0.00	0.01	0.02	0.12	0.02
	Primary traffic	500	0.24	0.07	0.06	0.20	0.24	0.28	0.48	0.08
	Sulfate rich	500	0.09	0.04	0.00	0.06	0.08	0.11	0.22	0.05

198
199

200

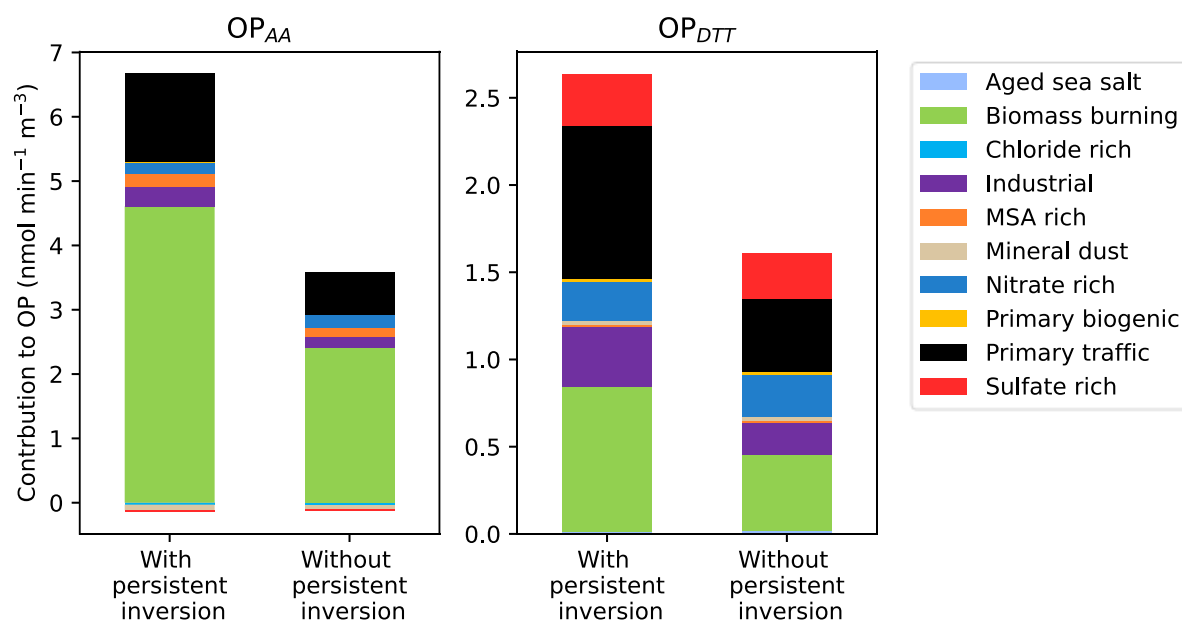



Figure S.14. Contribution of the different sources to OP$_v^{AA}$ (left) and OP$_v^{DTT}$ (right), for days with persistent inversion vs non-inversion days of the winters 2017-2023.

203

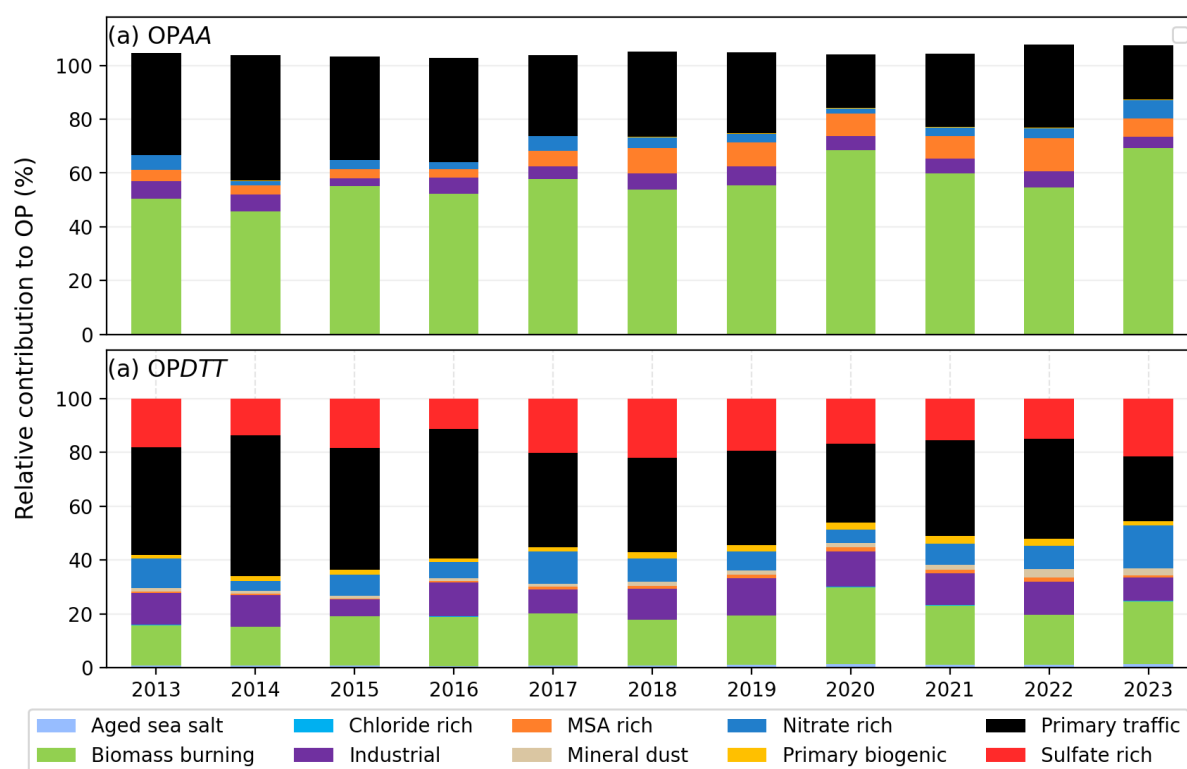
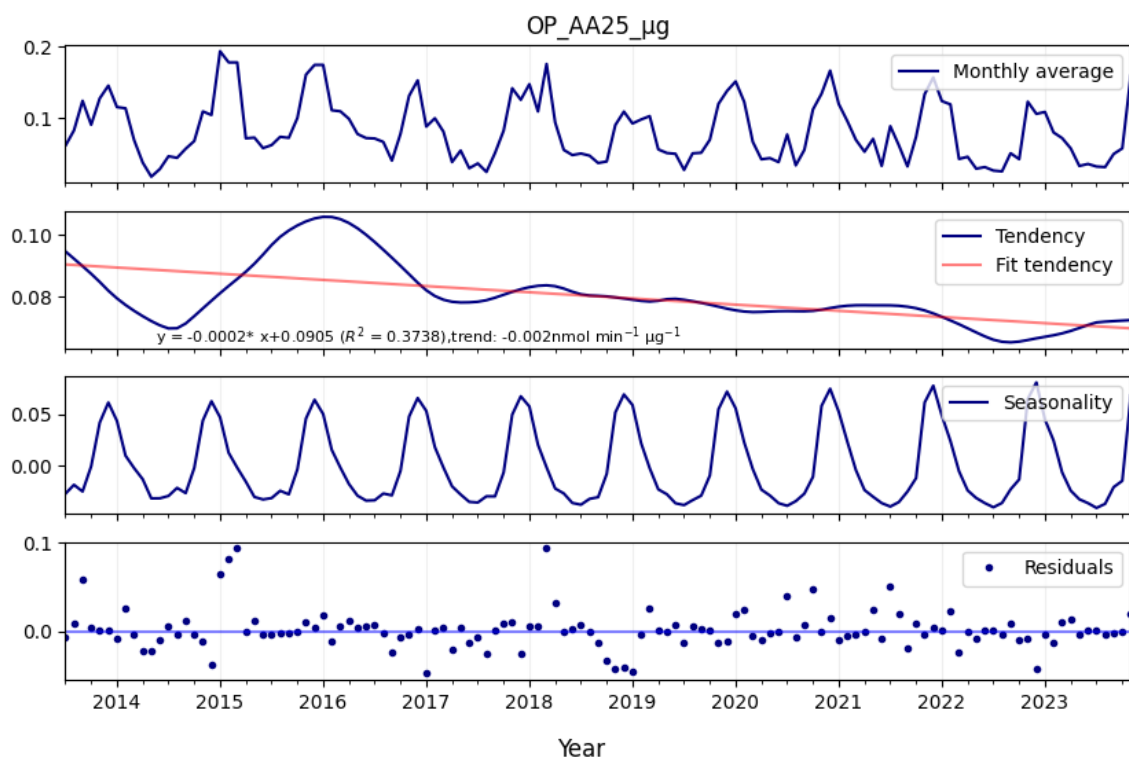
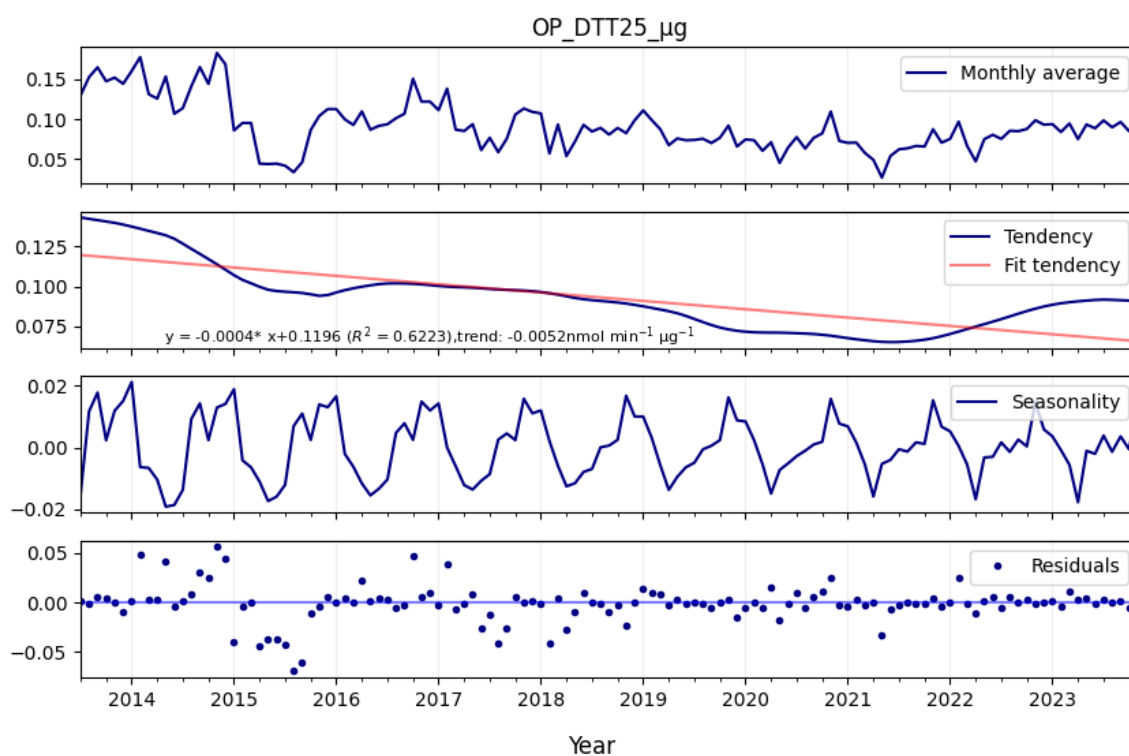



Figure S.15. Yearly average relative contribution of sources to (a) OP$_{AA}^v$ and (b) OP$_{DTT}^v$

205

206

207
208 **Figure S.16. STL decomposition of $OP_{AA}{}^m$**

209
210 **Figure S.17. STL decomposition of $OP_{DTT}{}^m$**

211

References

Alexander, B., Park, R. J., Jacob, D. J., Li, Q. B., Yantosca, R. M., Savarino, J., Lee, C. C. W., and Thiemens, M. H.: Sulfate formation in sea-salt aerosols: Constraints from oxygen isotopes, J. Geophys. Res. Atmospheres, 110, 2004JD005659, https://doi.org/10.1029/2004JD005659, 2005.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J.: Multiple imputation by chained equations: what is it and how does it work?, Int. J. Methods Psychiatr. Res., 20, 40–49, https://doi.org/10.1002/mpr.329, 2011.

Belis, C. A., Favez, O., Harrison, R. M., Larsen, B. R., Amato, F., El Haddad, I., Hopke, P. K., Nava, S., Paatero, P., Prévôt, A., Quass, U., Vecchi, R., Viana, M., and European Commission (Eds.): European guide on air pollution source apportionment with receptor models, Publications Office, Luxembourg, 1 pp., https://doi.org/10.2788/9332, 2014.

Carbone, C., Decesari, S., Mircea, M., Giulianelli, L., Finessi, E., Rinaldi, M., Fuzzi, S., Marinoni, A., Duchi, R., Perrino, C., Sargolini, T., Vardè, M., Sprovieri, F., Gobbi, G. P., Angelini, F., and Facchini, M. C.: Size-resolved aerosol chemical composition over the Italian Peninsula during typical summer and winter conditions, Atmos. Environ., 44, 5269–5278, https://doi.org/10.1016/j.atmosenv.2010.08.008, 2010.

Favez, O., El Haddad, I., Piot, C., Boréave, A., Abidi, E., Marchand, N., Jaffrezo, J. L., Besombes, J. L., Personnaz, M. B., Sciare, J., Wortham, H., George, C., and D'Anna, B.: Inter-comparison of source apportionment models for the estimation of wood burning aerosols during wintertime in an Alpine city (Grenoble, France), Atmospheric Chem. Phys., 10, 5295–5314, https://doi.org/10.5194/acp-10-5295-2010, 2010.

Largeron, Y. and Staquet, C.: Persistent inversion dynamics and wintertime PM10 air pollution in Alpine valleys, Atmos. Environ., 135, 92–108, https://doi.org/10.1016/j.atmosenv.2016.03.045, 2016.

Ocepek, U., Rugelj, J., and Bosnić, Z.: Improving matrix factorization recommendations for examples in cold start, Expert Syst. Appl., 42, 6784–6794, https://doi.org/10.1016/j.eswa.2015.04.071, 2015.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in {P}ython, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Putaud, J.-P., Raes, F., Van Dingenen, R., Brüggemann, E., Facchini, M.-C., Decesari, S., Fuzzi, S., Gehrig, R., Hüglin, C., Laj, P., Lorbeer, G., Maenhaut, W., Mihalopoulos, N., Müller, K., Querol, X., Rodriguez, S., Schneider, J., Spindler, G., Brink, H. T., Tørseth, K., and Wiedensohler, A.: A European aerosol phenomenology—2: chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe, Atmos. Environ., 38, 2579–2595, https://doi.org/10.1016/j.atmosenv.2004.01.041, 2004.

Putaud, J.-P., Van Dingenen, R., Alastuey, A., Bauer, H., Birmili, W., Cyrys, J., Flentje, H., Fuzzi, S., Gehrig, R., Hansson, H. C., Harrison, R. M., Herrmann, H., Hitzenberger, R., Hüglin, C., Jones, A. M., Kasper-Giebl, A., Kiss, G., Kousa, A., Kuhlbusch, T. A. J., Löschau, G., Maenhaut, W., Molnar, A., Moreno, T., Pekkanen, J., Perrino, C., Pitz, M., Puxbaum, H., Querol, X., Rodriguez, S., Salma, I., Schwarz, J., Smolik, J., Schneider, J., Spindler, G., Ten Brink, H., Tursic, J., Viana, M., Wiedensohler, A., and Raes, F.: A European aerosol phenomenology – 3: Physical and chemical characteristics of particulate matter from 60 rural, urban, and kerbside sites across Europe, Atmos. Environ., 44, 1308–1320, https://doi.org/10.1016/j.atmosenv.2009.12.011, 2010.

Salameh, D., Detournay, A., Pey, J., Pérez, N., Liguori, F., Saraga, D., Bove, M. C., Brotto, P., Cassola, F., Massabò, D., Latella, A., Pillon, S., Formenton, G., Patti, S., Armengaud, A., Piga, D., Jaffrezo, J. L., Bartzis, J., Tolis, E., Prati, P., Querol, X., Wortham, H., and Marchand, N.: PM2.5 chemical composition in five European Mediterranean cities: A 1-year study, Atmospheric Res., 155, 102–117, https://doi.org/10.1016/j.atmosres.2014.12.001, 2015.