SMH			Report
Author:	Client:		Report No.:
Olivier Tasse, Ulrica Sievert, Lars Gidhagen	SEPA (Naturvårdverket)		2021-61
Control date:	Controller:	Diary No.:	Version:
2021-11-17	Maria Grundström	2019/507/10.3	3.0

IMPAQ Project, component 3, Interim results from PMF 5.0 receptor modelling using the PM2.5 chemical analysis of the samples collected in 6 cities of BiH in winter 2020-2021

Olivier Tasse

Author	Contact								
SMHI	Olivier Tasse								
601 76 Norrköping	+4611-495 84 57								
Sweden	Olivier.tasse@smhi.se								
Client	Contact								
Swedich EDA	Andreas Lindström								
Swedish EPA	+4610-698 12 33								
	Andreas.lindström@naturvardsverket.se								
Distribution									
Swedish EPA, Andreas Lindström, Kristian Silver, Fl	HMZ, Enis Omercic, RHMZ, Ranka Radic								
Classification									
□ Allmän ⊠ Affärssekretess	□ Allmän ⊠ Affärssekretess								
Keywords									
Source apportionment, emission inventorying, dispersion modelling, Balkan, Bosnia									
Other									
Swedish EPA, Andreas Lindström, Kristian Silver, Fl Classification Allmän 🛛 Affärssekretess Keywords Source apportionment, emission inventorying, disper Other	HMZ, Enis Omercic, RHMZ, Ranka Radic								

Summary

Contributors to fine particulate matter (PM2.5) have been examined in six different cities in Bosnia and Herzegovina (Sarajevo, Tuzla, Zenica, Banja Luka, Bijeljina, and Brod), the investigation was carried out by a measurement campaign during the winter 2020-2021 followed by receptor-modelling from EPA's Positive Matrix Factorization (PMF) model. The results show high levels of PM2.5, on daily average $62 \,\mu g/m3$, and this study suggests that around 25% of it is emitted from wood and pellet burning, which are used for heating purposes. Furthermore, 20% of PM2.5 seems to come from fossil combustion including coal for heating purposes and vehicle engines. The portion of the background or long-range transport aerosol represent more than 25% of the total PM2.5. The results were then further investigated using local meteorological data, which in some cases increased the understanding of where and when different pollution sources originate from.

Table of content

SUMMARY0
TABLE OF CONTENT1
BACKGROUND
METHOD
The sampling campaigns3
The analytical analysis – gravimetric analysis5
Filter division before further analysis6
Elemental analysis6
Ion Chromatography analysis7
Analysis of the organic markers levoglucosan, manosan and galactosan 8
Analysis of the organic markers OC and EC9
Analytical results9
Analysis of total mass11
Positive Matrix factorisation (PMF) modelling12
Input data and settings in PMF analysis13
Optimisation of model solution17
RESULTS
Identification of different sources18
Sarajevo – constrained solution
Tuzla – constrained solution
Zenica – constrained solution
Banja Luka – constrained solution
Bijeljina – constrained solution ,
Brod – base solution
Secondary aerosols:
Biomass Burning:
Fossil Burning
Soil dust:
Other factors
CONCLUSION
REFERENCES
FIGURES
TABLES

APPENDIX	37
Sarajevo – factors contributions	38
Tuzla – factors contributions	39
Zenica – factors contributions	40
Banja Luka – factors contributions	41
Bijeljina – factors contributions	42
Brod – factors contributions	43

Background

Ambient air pollution is a global health problem and WHO estimates that it causes around 4,2 million peoples' death every year. In a recent report they also concluded that the global health risk from air pollution is as big as from unhealthy diet and tobacco smoking. WHO has set up guidelines about thresholds for clean air and according to these thresholds 99% of the world population lives in areas with poor air quality.

One among many other air pollutions is fine particulate matter (PM) which are problematic since they affect the lung capacity. PM is often divided into the categories PM2.5 and PM10 and these can be a complex combination of many pollution sources, for example combustion, sea salt and soil sources.

Bosnia and Herzegovina (BiH) among other countries struggle to mitigate the problem with high concentrations of particles in the air. It is believed that heating of homes explains a major part of PM2.5 as many households use combustion of wood or coal as heat sources. There is a challenge in describing exactly how much of PM2.5 that comes from these sources. One way of examining this problem is to chemically analyse samples of PM2.5, to determine its shares of different chemical species. Many of the species can in fact be connected to different emission sources, but the task is rather complex. Receptor modelling from EPA's Positive Matrix Factorization (PMF) can be used in order to discriminate the different emission sources.

In this study a daily collection of particles using filter for PM2.5 was done in six different cities in BiH during three winter months (November, December and January). A subsequent gravimetric and chemical analysis have been conducted on each filter for many specific elements, molecules and ions. This data could then be used as an input for the PMF model. From the model outputs attempts have been made in order to identify partitioned emission sources.

Method

The sampling campaigns

The samplings of PM2.5 was done at six different sampling sites in Bosnia and Herzegovina during three winter months from November 2020 to January 2021. The position of each site can be seen in Figure 1 and further information is found in Table 1. All of them are situated in cities. The sites in Sarajevo, Zenica, Tuzla, Bijeljina and Brod are classified as urban background while the site in Banja Luka is classified as urban traffic.



Figure 1: Sampling sites in Bosnia and Herzegovina during winter 2020-2021

Low volume samplers, Sven Leckel SEQ47/50-RV instruments were used for sampling, in accordance with SRPS EN12341:2015 standard reference method. Maintenance, installation and uninstallation of the samplers was provided by the official Sven Leckel distributors for Serbia.



Figure 2: Sven Leckel SEQ47/50-RV sampler and transport case

Whatman QM-A quartz filters, 47mm were used for the sampling campaign and 92 daily samples were collected during the measurement campaign.

Each sampler was equipped with two sets of filter magazines so that they could be filled with new set of unexposed filters in the controlled environment of the laboratory. At the sampling site the sampler was prepared and refilled with an interval of 14 days by an experimented team from the Institute of Public Health of Belgrade during a three days round trip. Each time the samplers were refilled one filter at the minimum was not to be sampled but remained in the magazine to serve as a field blank. Hence the reloading induced an interruption in the sampling series and daily sample is then missing each time.

The scheduled settings during the visits were:

- Exchange of the filter magazine of the sampler was done in laboratory. During transportation the magazines were covered and put into insulated containers to avoid external contamination and excessive heating.
- Change of the impaction plate for a cleaned and pre-greased from laboratory.
- Change of the nozzle for a clean one.
- Check of the sampler flow rate using a regularly calibrated ORIWLOW-reference flowmeter for samplers and leak check of the sampling system.

Back at the laboratory the sampled filters were stored at a suitable temperature ca 4°C so that loss of volatile and semi-volatile materials was minimized over the storage period.

	Sarajevo Bjelave meteorological site	Tuzla meteorological site	Zenica Brist meteorological site	Banja Luka Lazarevo meteorological site	Bijeljina meteorological site	Brod refinery meteorological site
Operator of station	Federalni hidrometeorološki zavod (FHMZ)	FHMZ	FHMZ	Republički hidrometeorološki zavod (RHMZ)	RHMZ	Optima Rafinerija nafte
Latitude, longitude	43.867743, 18.422950	44.542060, 18.685136	44.202056, 17.900428	44.793801, 17.205743	44.753659, 19.192754	45.135325, 17.982985
Туре	Urban background	Urban background	Urban background	Urban Traffic	Urban background	Urban background
View of the sampler at sites						
Sampling period	2020-10-30 to 2021-02-03	2020-10-29 to 2021-02-02	2020-10-30 to 2021-02-03	2020-10-31 to 2021-02-04	2020-10-29 to 2021-02-02	2020-10-31 to 2021-02-04
Missing days (samplers reloading days)	2020-11-17 2020-12-03 2020-12-22 2020-01-11 2020-01-28	2020-11-16 2020-12-02 2020-12-21 2020-01-10 2020-01-27	2020-11-17 2020-12-03 2020-12-22 2020-01-11 2020-01-28	2020-11-18 2020-12-04 2020-12-23 2020-01-12 2020-01-29	2020-11-16 2020-12-02 2020-12-21 2020-01-10 2020-01-27	2020-11-16 2020-12-04 2020-12-23 2020-01-10 2020-01-29

Table 1: characteristics of the sampling sites

It should be emphasized that all activities related to sampling were conducted in extraordinary circumstances during the pandemic of COVID-19 virus, with the epidemiological measures in force in Serbia and Bosnia and Hercegovina constantly changing, including tightening measures for crossing the state border (e.g. PCR testing when entering Serbia). Besides above mentioned, all activities related to sampling were conducted during the winter period accompanied with a significant amount of snowfall.

Gravimetric analysis

Gravimetric analysis of total mass concentration of particulate matter PM2.5 was performed by standard reference method SRPS EN 12341:2015, identical with EN 12341:2014 which guarantees that all of the requirements for the method performance and quality control are met.

Filter conditioning, sampling and weighing procedures included:

- Filter conditioning and weighing prior to sampling.
- Sampling procedure.
- Filter conditioning and weighing after sampling.
- Weighing room procedures.
- Filter blanks for quality control.

Uncertainty budget for the gravimetric analysis of total mass concentration was calculated considering all of the individual sources of uncertainty in accordance with SRPS EN 12341:2015

The result for expanded uncertainty for gravimetric analysis was: U=(0,5+0,07*x), with x being the calculated mass concentration of PM2.5.

The Method Detection Limit was $1 \mu g/m^3$.

Filter division before further analysis

After gravimetric analysis of total mass concentration was completed, further chemical and elemental analysis was performed. Each filter was split in half. From one half of the filter a punch for organic carbon (OC) and elemental carbon (EC) analysis was extracted and the rest of the filter was used for elemental analysis. From the second half of the filter a punch for ion analysis was extracted and the rest of the filter was used for the analysis.



Figure 3: filter cutting for analysis; 1/2 filter for elemental + punch for OC/EC (left), and 1/2 for anydrosugar + punch for ion chromatograhy (right)

Elemental analysis

Elemental analysis of As, Cd, Cr, Mn, Ni, Pb was based on the standard reference method SRPS EN 14902: 2008/AC:2013 Ambient air quality - Standard method for the measurement of the Pb,Cd, As and Ni in the PM10 fraction of suspended particulate matter for which the Institute of Public health of Belgrade is accredited.

The method of analysis included:

- Microwave digestion using Anton Paar equipment.
- ICP-MS analysis using Agilent ICP-MS, Series 7500, device for: As, Cd, Cr, Mn, Ni, Pb.
- ICP-OES analysis using Agilent ICP-OES model 5110SVDV for: Al, Co, Cu, Fe, V, Zn.
- Quality controls.

The uncertainties have been calculated considering random uncertainty, uncertainty of the sampling volume, non-random uncertainty of the analysis process. The results for the expanded uncertainty for elemental analysis is presented in the Table 2.

Element	U	Element	U	Element	U
As	0,00002+0,075* <i>x</i>	Cd	0,00002+0,071* <i>x</i>	Cr	0,00001+0,094* <i>x</i>
Mn	0,112* <i>x</i>	Ni	0,00003+0,117* <i>x</i>	Pb	0,00005+0,082* <i>x</i>
AI	0,257* <i>x</i>	Co	0,12* <i>x</i>	Cu	0,21* <i>x</i>
Fe	0,155* <i>x</i>	v	0,116* <i>x</i>	Zn	0,116* <i>x</i>

 Table 2: expanded uncertainties for elemental analysis, with x being the calculated mass concentration of each element

The method detection limits in $\mu g/m^3$ is presented in the Table 3:

Element	MDL (µg/m³)	Element	MDL (µg/m³)	Element	MDL(µg/m³)
As	0,001	Cd	0,0001	Cr	0,005
Mn	0,0024	Ni	0,003	Pb	0,005
AI	23,7	Со	23,7	Cu	23,7
Fe	23,7	v	11,8	Zn	2,4

Table 3: Method detection limit for the elemental analysis

Ion Chromatography analysis

Ion chromatography of: SO₄^{2–}, NO₃[–], NH₄⁺, Cl[–], Na⁺, Mg²⁺, K⁺, Ca²⁺ was performed by an inhouse method, based on the standard reference method SRPS EN 16913:2017 for which the Institute of Public health of Belgrade is accredited.

The method of analysis included:

- Ion chromatography using Methrom, model IC 930 Flex.
- Quality controls.

The uncertainties have been calculated considering uncertainty of the sampling volume, calculated recovery based on matrix spike sample, calibration of IC equipment, reference material on daily measurement. The results for the expanded uncertainty for ion chromatography is presented in the Table 4.

lon	SO4 ²⁻	NO ₃ -	NH4+	Cl-	Na+	Mg ²⁺	K+	Ca ²⁺
U	0,08* <i>x</i>	0,08* <i>x</i>	0,13* <i>x</i>	0,08* <i>x</i>	0,06* <i>x</i>	0,073* <i>x</i>	0,07* <i>x</i>	0,15* <i>x</i>

Table 4: expanded uncertainties for Ion Chromatography analysis, with x being the calculatedmass concentration of each ion

The method detection limits in $\mu g/m^3$ is presented in the Table 5:

lon	SO4 ²⁻	NO ₃ -	NH4 ⁺	Cl-	Na⁺	Mg ²⁺	K+	Ca ²⁺
MDL (µg/m³)	0,8	0,8	0,08	0,8	0,8	0,4	0,8	3,1

Table 5: Method detection limit for the Ion Chromatography

Analysis of the organic markers levoglucosan, manosan and galactosan

Analysis of the organic markers levoglucosan, manosan and galactosan was performed using a method that is based on the standard method VDI 2444, Ambient measurements of levoglucosan, Chromatographic method, march 2020.

The method of analysis included:

- Ultrasonic extraction.
- derivatization and GCMS quantification, using Agilent GCMS single quad 5975T.
- Quality controls.

Quality control was done according to the standard VDI 2444 and combined with quality control from the standard reference method SRPS EN 15549:2010, Air quality — Standard method for the measurement of the concentration of benzo[a]pyrene in ambient air.

It is relevant to highlight that the supervisor in charge for the Institute of Public health of Belgrade is a member of the CEN working group CEN/TC 264/WG 21 for the development of standard method: Ambient air – Determination of the concentration of Levoglucosan – Chromatographic method, upon the call from JRC Ispra, AQUILA group.

In order to improve quality of the data obtained for organic markers, IPH has participated in the Levoglucosan interlaboratory comparison study in the mentioned working group, which will be a final step before applying for accreditation of the method.

The uncertainties have been calculated considering uncertainty of the sampling volume, calculated recovery based on matrix spike sample, mass of sampled organic marker (sampling efficiency and stability, selectivity), mass of organic marker in blank sample.

The result for expanded uncertainty, with x being the calculated mass concentration of respective hydrocarbon, was:

- U=0,1448*x for Levoglucosan,
- U=0,162*x for Manosan,
- U=0,1448*x for Galactosan,

The related Method Detection Limit was:

- 0,001 µg/m³ for Levoglucosan
- 0,0009 μ g/m³ for Manosan
- $0,0009 \,\mu g/m^3$ for Galactosan

Analysis of the organic markers OC and EC

EC (Elemental Carbon) is a fraction of pure carbon usually emitted from combustion process.

OC (Organic Carbon) is a fraction of carbon blended with organic components and either emitted from combustion process, or as the result of atmospheric oxidation and/or condensation process.

Analysis of the organic markers OC and EC was performed by in-house method based on the standard reference method SRPS EN 16909:2017 using the EUSAAR protocol for which Institute of Public health of Belgrade is accredited.

The method of analysis included:

- Lab OC-EC Aerosol Analyzer, Sunset Laboratory Inc.
- Quality controls.

The uncertainties have been calculated considering uncertainty of the sampling volume, Peak area for the relevant carbon fraction (OC or EC) measured on the loaded filter sub-sample thermogram, Peak area for the calibration gas measured on the loaded filter sub-sample thermogram as well as on the external calibration standard thermogram, volume of external calibration standard solution analysed.

The result for expanded uncertainty, with x being the calculated mass concentration of respective carbon fraction, was:

- U=(0,5+0,1*x) for OC,
- U=(0,3+0,11*x) for EC.

The related Method Detection Limit was:

- $0,5 \,\mu g/m^3$ for OC
- 0,5 μ g/m³ for EC

Analytical results

Table 6 shows the average concentration in $\mu g/m^3$ and associated standard deviation (SD) for all the measured species.

Depending on the chosen analytical method, many species have not been able to be measured above the detection limit (ADL) often enough to be relevant for the PMF model and nor for the calculation of an average concentration. These species are identified in the Table 6 as:

- A minus sign (-) when all concentration data was below the Method Detection Limit (MDL),
- A plus sign (+) when the number of samples ADL was 29% or less.

Regarding the Elemental analysis, nearly all of the concentration levels were below MDL when ICP-OES method was used and therefore these results could not be included in this investigation.

In addition, Zn (the only element that was enough represented when ICP-OES method was used) showed unusual high levels. The average levels of Zn from the different sites were indeed 200 to 700 times the one recently measured in the region in similar urban areas. This high level together with other weak results lead to suspect a contamination. For these reasons none of the results issued from the ICP-OES method have thus been used here after.

Regarding the Ion Chromatography, the numerous results detected under the MDL highlighted the inadequacy of the analytical method for that type of low volume PM 2.5 sampling. Na+ and Cl- were among the missing ions and they are important tracers for sea salt, which usually is a small but significant part of PM2.5.

Regarding the anhydro-sugars on the other side, the analysis originally limited to the Levoglucosan has been successfully expanded in order to measure Manosan and Galactosan as well. All the results for anhydrosugar have been detected over the MDL.

	Sarajevo		Tuzla		Zenica		Banja L	uka	Bijeljina		Brod	
	Average (µg/m ³)	SD	Average (µg/m³)	SD	Average (µg/m ³)	SD						
Al	-		-		-		-		-		-	
As	0,00699	0,00563	0,00686	0,00387	0,02257	0,01284	0,00211	0,00077	0,01283	0,00749	0,00686	0,00387
Cd	0,00085	0,00072	0,00044	0,00030	0,00108	0,00074	0,00061	0,00025	0,00040	0,00022	0,00044	0,00030
Со	-		-		-		-		-		-	
Cr	-		+		+		+		+		+	
Cu	-		-		-		+		-		-	
Fe	-		-		-		-		-		-	
Mn	0,00453	0,00247	0,00375	0,00183	0,03610	0,03244	0,01133	0,01224	0,00291	0,00027	0,00375	0,00183
Ni	+		+		+		+		+		+	
Pb	0,01453	0,01026	0,00960	0,00478	0,02334	0,01180	0,00929	0,00340	0,00812	0,00330	0,00960	0,00478
V	-		-		-		+		-		-	
Zn	*11,35647	8,11133	*14,72071	7,96330	*9,80347	5,13138	*7,45335	3,77896	*13,68638	7,39706	*3,83592	1,10229
SO4 ²⁻	6,72088	5,66253	8,75496	6,37894	12,97391	7,09257	6,52068	3,35438	7,18864	3,77389	5,18687	4,08349
NO ₃ -	4,85086	4,07375	3,46188	2,10774	2,89716	1,57844	4,36488	2,25072	4,28666	2,32440	3,94316	2,51337
NH_{4}^{+}	3,12683	3,04969	3,42089	2,76581	4,32165	2,83087	2,24009	1,41273	3,36945	1,99995	2,52341	1,77741
Cl ⁻	+		+		+		+		+		+	
Na ⁺	+		+		+		+		+		+	
Mg ²⁺	-		-		-		-		-		-	
K ⁺	2,06985	1,45119	1,22536	0,42793	1,74805	0,79234	2,49399	1,62470	1,28594	0,48228	1,32356	0,57653
Ca ²⁺	-		-		-		-		+		-	
OC	16,56153	15,69400	16,98529	9,32303	20,45732	10,77313	18,84537	8,76522	15,98242	8,21045	10,50078	6,36503
EC	2,58778	1,27603	2,11548	1,06134	2,84121	0,99552	2,78565	1,00173	1,83905	0,68836	1,32261	0,81706
Levoglucosan	1,30186	0,95600	0,98490	0,53000	1,64989	0,87665	1,62815	0,80455	1,49462	0,83845	1,52928	1,05892
Mannosan	0,25671	0,21383	0,16144	0,09212	0,31444	0,15573	0,31444	0,15573	0,19356	0,10744	0,17801	0,14014
Galactosan	0,09554	0,08524	0,06257	0,03687	0,11429	0,06338	0,11429	0,06338	0,09395	0,05295	0,07711	0,06046
PM 2.5	60,56475	53,72316	60,76448	32,20991	75,99298	36,70262	73,01956	37,40793	61,72345	26,31241	41,74085	17,57572

Table 6: Averaged concentration, and standard deviation (SD), of PM 2.5 and chemical species

Analysis of total mass

Among the measured species, the OC is by far the most abundant and explains a large part of the whole PM2.5. The ratio between average OC and average EC (Table 7) is always high, from 6:1 for the two biggest cities of Sarajevo and Banja Luka to over 8:1 for the smaller cities. The traffic tends to even out this ratio and is probably most significant in the biggest cities. In the same way the coal and brown coal burnings that emit more EC than OC tends to even out this ratio as well. It is possible that the burning of such calorific solid fuel, which is easier to handle and store than wood, is more common in the biggest cities. In addition, burning oil for heating purpose is quite common depending of the cites and natural gas is used in Sarajevo as well.

The mass of OC is highly correlated with the mass of PM 2.5, except for Banja Luka. This last disconnection suggest that an important non-burning source of particles has influenced the results in Banja Luka. Furthermore, it is important to note that the percentage of average OC compared to the mass of PM 2.5 (Table 6) is quite constant between 25 to 27% for all the cities, including Banja Luka. This means that the non-correlation found in Banja Luka is more related to a time-series divergence between OC and PM 2.5 than the total amount of OC. In other words, there is a significant number of peaks of PM2.5 that are disconnected to the peaks of OC. A similar anomaly has been noticed at the same place in 2015 (Almeida, 2020).

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Population	642 000	80 000	75 000	250 000	50 000	71000 – incl. Slavonski-Brod
Ratio OC / EC	6,4	8,1	7,3	6,7	8,9	8,1
Correlation OC / PM 2.5	0,99	0,94	0,98	0,25	0,88	0,90
Correlation EC / PM 2.5	0,80	0,76	0,81	0,13	0,86	0,68
Σ of species [µg/m3]	37,3	37,0	46,8	38,8	35,5	26,2
% of mass explained	62%	61%	62%	53%	58%	63%

Table 7: parameters for the appreciation of total mass of PM 2.5

When comparing to a similar analysis from (Perrone, 2017), the results in BiH show 3:1 more OC among the PM 2.5 mass than in Zagreb in 2013. The same trend is noticed with the Levoglucosan and K^+ which are 4:1 to 8:1 over the levels observed in Zagreb. There is therefore a good reason to suspect the wide range of individual household stoves, with generally low temperature firebox and high PM 2.5 emission ratio, to be the most common source of particulate matters when burning wood, as well as coal and brown coal.

The sum of all the measured species reaches about 60% of the total mass of PM 2.5. But, as mentioned, some species that are used to identify soil sources are not among the detected ones. Other important species that have not been detected in this study are those that are included in sea salt, especially Na and Cl, which together make up almost 80% of the composition of sea salt. According to Peronne et al. 2017, the soil dust and sea salt reach a total of about $2 \mu g / m3$ in Zagreb 2013. This level can be assumed to be at least equivalent in Bosnia 2020-2021. In Peronne et al. 2017 66% of the total mass of PM2.5 was explained.

The lacking species are most importantly these that have a significant part of the of the overall PM2.5 weight, like calcium, silicon, iron, aluminium and that have not been measured. Measuring these crustal species is just difficult as they are common in the background, which interfere with the measure. In this study they have been considered uniformly distributed in all apportioned sources.

Positive Matrix factorisation (PMF) modelling

The source apportionment of the PM2.5 were estimated with using receptor modelling PMF (Paatero, 1994) with the USEPA PMF v3.0 software.

The underlying principle of receptor model is that mass conservation can be assumed and a mass balance analysis can be used to identify and apportion sources of airborne particulate matter in the atmosphere. This PMF is a multivariate factor analysis tool that uses two matrix of the measured concentrations and of the related uncertainties, provide families of solution that solve the mass balance equation.

 $X = G \times F + E$. where:

- X is the original matrix of measurements and uncertainty,
- F is a matrix whose vectors represent the profiles of *p* sources,
- G is a matrix whose columns represent the contributions of the *p* sources,
- E is the residual matrix.



The PMF factor analysis generally produce a batch of solutions with different G and F matrices. Each solution is unique and this is called the rotational ambiguity of the model. In order to be able to find the best-fit solution within the batch it is important to run about 100 random calculations.

Then the model assists to choose the best solution using the objective function Q that aims to minimize the difference between the real measurements and the modelled values. This difference is represented by the residual matrix. The residual matrix is typically influenced by the outliers which are extreme values that differ from the mean trend of all the data. These outliers can either be unwanted data-contaminations or be true outliers.

Choosing a solution with numerous factors lead to a quite sure solution, with a low Q. But the goal is also to connect the solution provided by the model to the reality of the environment were a few families of factor are expected, or can be explained (in this study a few families of air pollutant emitters).

A useful tool to decide for the best-fit solution among all the calculated solutions is then to compare the the Q expected (Q_{exp}) calculated using all the data, with the Q robust (Q_{robust}) calculated excluding the points beyond a decided uncertainty-scaled residual. The best-fit solution is when the difference is minimal, when there were a very few points that need to be excluded (Q_{robust} shall be less than 2:1 the Q_{exp}).

After deciding for the best-fit solution, the error estimation methods included in the PMF 5.0 software shall be used to confirm or reject the chosen solution.

The Bootstrap (BS) is the first one and help to detect and estimate possible random errors due to disproportionate effects of a small set of observations on the solution. It literally shows how strongly defined are the factors. An acceptable solution shall have 80% of the iterative calculations providing the same mapping of factors.

The Displacement (Disp) is the second one and defines the span of rotationally accessible space for the solution. The strong species have their value one by one "displaced" a little in the profile of each factor calculated. The effect on the other factors is then observed. The idea is to see how often factors change enough to exchange identities depending on the size of the displacement. An acceptable solution shall have no swap of identity for the minimal displacement. Within this study the result is then a number of factors (5 to 6) defined by the contribution of the factor to the weight of each species (and *vice versa* the concentration of the different measured species in the factor), and the related times-series defined by the contribution of each daily measurement to the overall weight of the factor during the period.



Figure 4: typical result with factor definition (left) and times-series of the factor (right)

Input data and settings in PMF analysis

The input data pre-processing and settings when following the PMF US-EPA guidelines are summarised in the Table 10.

The PMF analysis was run separately for each city. The amount of PM 2.5 samples was of 92 for each of the six sampling places. Few samples were excluded due to some very unusual and isolated events, in most cases only one isolated sample was excluded but in one occasion 5 consecutive samples were excluded. These last might have been related to a Sahara dust incursion episode that started during that period in early February 2021. The range of modelled samples was therefore between 93% to 100%, depending of the sampling location.

The values below method detection limit (MDL) were replaced by half of the detection limit (DL/2) accordingly to the guidelines.

The number of species actually used was between 12 and 13. The species were classified as "weak" when the number of samples ADL was < 55% and "bad" when the number of samples ADL was < 35%. However, only K⁺ have been set as "weak" until 17% due to its importance for the discrimination of the biomass burning.

The ratio signal to noise (S/N) was also used in order to classify the species. If S/N for one specie was lower 1, then it was classified as weak or bad if S/N<0.5.

The uncertainty has been provided by the IPH of Belgrade together with the measurement data and both concentrations and observation-based uncertainties was considered. The missing uncertainties related to the missing values were replaced by 5/6*DL as recommended by the guidelines.

In order to account for unknown sources of uncertainty, the analytical uncertainty provided was incremented by an extra-modelling uncertainty of 7% for all species. This is understood as making the final PMF solution stronger. For Sarajevo however, no extra uncertainty was added in order for the model to be able to converge to a solution.

A first estimation of the number of factors was accomplished by step-wise analysis of the Q value of multiple runs with increasing number of factors. The quality of the fit led to the decision of the best number of factors to be used (scaled residuals, fit of the observed versus predicted plots and histograms, $Q/Q_{expected}$ for the species). The interpretability of the results as well led to the decision of the number of factors (in terms of chemical profile and time series) as well as the decision to run a constrained model or not.

The best number of factors was either 5 (Bijeljina, Brod, Zenica, and Tuzla) or 6 (Banja-Luka and Sarajevo). As seen in the table, the PMF solution from Sarajevo did not quite reach the criteria from EPA's user guide. Though the results were improved by implementing constraints there is still one swap present from the DISP-analysis. Therefore, the results should be interpreted with this in mind.

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Period	2020-10-30 to 2021-02-03	2020-10-29 to 2021-02-02	2020-10-30 to 2021-02-03	2020-10-31 to 2021-02-04	2020-10-29 to 2021-02-02	2020-10-31 to 2021-02-04
Missing days (samplers reloading days)	2020-11-17 2020-12-03 2020-12-22 2020-01-11 2020-01-28	2020-11-16 2020-12-02 2020-12-21 2020-01-10 2020-01-27	2020-11-17 2020-12-03 2020-12-22 2020-01-11 2020-01-28	2020-11-18 2020-12-04 2020-12-23 2020-01-12 2020-01-29	2020-11-16 2020-12-02 2020-12-21 2020-01-10 2020-01-27	2020-11-16 2020-12-04 2020-12-23 2020-01-10 2020-01-29
Nb. of samples	92	92	92	92	92	92
% ADL S/N c	hosen category (Str	ong, Weak, Bad)	·		·
As	72% 5,4 Strong	89% 7,1 Strong	99% 9,7 Strong	37% 1,8 Weak	99% 9 Strong	89% 7,1 Strong
Cd	53% 4,5 Strong	50% 3,3 Weak	91% 8,5 Strong	59% 4,9 Strong	51% 3,5 Weak	50% 3,3 Weak
Mn	50% 4 Strong	24% 1,8 Bad	95% 7,5 Strong	85% 6,8 Strong	12% 0,9 Bad	24% 1,8 Bad
Pb	66% 4,7 Weak	74% 4,5 Strong	99% 8,2 Strong	72% 4,5 Weak	79% 4,6 Weak	74% 4,5 Weak
SO4 ²⁻	96% 9,5 Strong	99% 9,9 Strong	96% 9,7 Strong	96% 9,7 Strong	100% 10 Strong	98% 9,8 Strong
NO3 ⁻	79% 7,9 Strong	86% 8,6 Strong	87% 8,7 Strong	93% 9,5 Strong	99% 9,9 Strong	87% 8,7 Strong
$\mathbf{NH}_{4^{+}}$	78% 5,4 Strong	90% 6 Strong	78% 5,2 Strong	74% 5 Strong	95% 6,3 Strong	71% 4,7 Strong
K ⁺	49% 5,1 Strong	41% 4,1 Strong	66% 6,7 Strong	78% 7,9 Strong	65% 6,5 Strong	27% 2,6 Weak
OC	100% 5,7 Strong	100% 6,4 Strong	100% 6,7 Strong	100% 6,6 Strong	100% 6,3 Strong	100% 5,3 Strong
EC	100% 3,1 Strong	100% 2,7 Strong	100% 3,5 Strong	100% 3,5 Strong	100% 2,5 Strong	100% 1,8 Strong
Levoglucosan	100% 5,9 Strong					
Mannosan	100% 5,2 Strong					
Galactosan	100% 5,9 Strong					
PM 2.5	100% 9,7 Weak	100% 10 Weak	100% 10 Weak	100% 10 Weak	100% 10 Weak	100% 9,9 Weak

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Reason for unusual category if any	Pb as Weak since the specie was driving the high Q/Qexp trend which prevent the model to find a stable solution.	-	-	Pb as Weak since the specie was driving the high Q/Qexp trend which prevent the model to find a stable solution.	Pb as Weak since the specie was driving the high Q/Qexp trend which prevent the model to find a stable solution.	K+ as Weak instead of Bad even if only 27% ADL, but acceptable S/N and important role in biomass burning identification. Pb as Weak for the same reason as for the other sites.
Excluded from modelling	2020-11-24 2020-01-30 2020-01-31 2020-02-01 2020-02-02 2020-02-03	2020-11-23	2020-11-21	2021-01-22	-	2020-12-05
Reasons of exclusion	24/11suspicious Mn outlier + last 6 days disturbed by Sahara dust incursion	Isolated event that prevent the model to converge to an acceptable solution.	Inconsistent level of NO3 and a high Q/Qexp trend which prevent the model to converge to an acceptable solution.	High level of anhydrosugar inconsistent with very low OC and EC that prevent the model to converge to an acceptable solution.	-	Isolated event with very isolated source of EC that prevent the model to converge to an acceptable solution.
% of tot. samples modelled	93,5%	98,9%	98,9%	98,9%	100%	98,9%

Table 8: Input data and PMF 5.0 settings

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Base runs						
Nb. of runs	100	100	100	100	100	100
Seed number	22	22	22	22	22	22
Nb. of factors – tests solutions	4 to 7	4 to 6	4 to 7	4 to 7	4 to 6	4 to 6
Nb. of factors – final solution	6	5	6	6	5	5
Extra modelling Uncertainty	0%	7%	7%	7%	7%	7%
Choose of base run	86 instead of 80	13 instead of 22	27 as suggested	84 as suggested	81 as suggested	17 as suggested
Reason if not the suggested one	Better discrimination of the Mannosan.	Otherwise the SO4 influenced factor display a very high and isolated peak in November	-	-	-	-
Bootstrap (BS) a	analysis for base rur	1				
min. nb. of BS mapped / total	79 / 100	66 / 100	94 / 100	85 / 100	84 / 100	99 / 100
max nb. of swap on one other factor	6	21	6	11	6	1
Displacement (D	DISP) analysis for b	ase run				
Error code 2 nd valour first raw of swaps	0 -1,780 0 0 0 0 0 0 0	Not ran	0 -1,788 0 0 0 0 0 0 0	0 -0,003 0 0 0 0 0 0 0	0 -0,053 5 0 6 3 14	0 -0,014 0 0 0 0 0 0
Comment on error estimation for Base run	One BS under 80%	One BS was far under 80% with 21% of swap toward one other factor	Both BS and DISP are acceptable	Both BS and DISP are acceptable	A lot of swaps (14) toward a mixed burning source	Both BS and DISP are acceptable

Table 9: base runs diagnostics

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod		
Rotational tools – Fpeak								
Fpeak	0	0	0	0	0	0		
Rotational tools	– Constraints							
Applied constraints	Mannosan pulled down (soft pulling) for all factors other than Biomass burning	Mannosan pulled down (soft pulling) for all factors other than Biomass burning	Mannosan pulled down (soft pulling) for all factors other than Biomass burning	Mannosan pulled down (soft pulling) for all factors other than Biomass burning + a mixed burning	Mannosan pulled down (soft pulling) for all factors other than Biomass burning	-		
Reason for constraints	Mannosan as a reliable indicator of biomass burning should not be over- represented in other factors.	Same reason as on the left	Same reason as on the left	Same reason as on the left	Same reason as on the left	-		
% of dQ [should be close as possible than 1%]	1,07%	1,11%	0,70%	0,74%	0,98%	-		
Bootstrap (BS) a	analysis for constrai	ned run						
min. nb. of BS mapped / total	87 / 100	93 / 100	91 / 100	99 / 100	91 / 100	-		
max nb. of swap on one other factor	6	6	8	1	8	-		
Displacement (D	DISP) analysis for co	onstrained run						
Error code 2 nd valour first raw of swaps	0 -27,478 1 0 0 0 1 0	0 -0,905 0 0 0 0 0 0	0 -5,670 0 0 0 0 0 0 0	0 0,000 0 0 0 0 0 0 0	0 0,000 0 0 0 0 0 0	-		
Comment on error estimation for constraint run	Only one swap between a Cd related factor and the crustal factor.	Both BS and DISP are acceptable	Both BS and DISP are acceptable	Both BS and DISP are acceptable	Both BS and DISP are acceptable	-		

Table 10: Constrained runs settings and validations

Optimisation of model solution

In the Positive Matrix factorisation model, the best solution obtained is not unique. Because of the free rotation of matrices there is a family of solutions that are equally fit because of the so-called rotational ambiguity.

In this work most of the solutions have been constrained considering that the manosan is an accurate tracer of biomass burning even in region were brown coal (lignite) is a common fuel. In the chosen constrained solutions manosan has been pulled down maximally ("soft pulling", dQ = 0,50%) in factors where the contribution from biomass burning shouldn't appear.

Results

Display of the results

The result for each sampling places is a number of factors (5 to 6) defined by the contribution of the factor to the weight of each species (and *vice versa* the concentration of the different measured species in the factor), and the related times-series defined by the contribution of each daily measurement to the overall weight of the factor during the period.



Figure 5: example of result with factor definition (left) and times-series of the factor (right)

On the left side of Figure 5 the black dots show the contribution of the factor (here biomass burning) to the total weight of each species. In this example 60% of the OC is found in the biomass factor. This mean that the OC is meaningful to explain the sources that form the Biomass burning factor.

On the left side of Figure 5 the grey bars show the concentration of each species in the apportioned factor. In this example the concentration of levoglucosan is about $1 \mu g/m^3$ while the concentration of Mannosan is about $0.1 \mu g/m^3$.

On the right side of Figure 5 the dots show the part of the factor modelled that day compares to its cumulative weight during the whole period. In this so-called times-series example, while the daily portion of factor is 1% in average, a few days around the 24th of December show 3% of the whole biomass burning factor calculated for the 91 accumulated days. Since the scale of that graphic is adaptative, it means that when the scale lower (up to 4%) one can say that the modelled factor should be related to a rather continuous source, while if the scale is higher (about 8%) the modelled factor should be related to a discontinuous source.

The right side of Figure 5 also shows grey bars representing the weekends. A factor which emissions normally would be more substantial during business days is expected to show a decrease during these days. As an example, the emissions from traffic usually show a weekly decrease during at least one day of the week-end.

The Figure 6 shows the daily relative contribution figured in a coloured scale with similar boundaries as for the times-series itself. Each day is represented by a point which position depends on the wind direction origin (north is up) and the wind speed (from the centre). The points are then merged in a single pattern. These so-called polar-plots were produced while using the Openair R package developed for the purpose of analysing air quality data (Carslaw & Ropkins, 2012).

The range of colours is the result of a Nonparametric Wind Regression (NWR) gaussian smoothing that weight concentrations on a surface according to their proximity to defined wind speed and direction intervals. The NWR have been chosen since the number of measurements was limited. Because of the smoothing effect on extreme values, only the pattern of colour should be interpreted and not the dimension.



Figure 6: polar plot of the times-series

Identification of different sources

The identification and naming of the likely sources for each factor provided by the model solution has been made using the range of relative mass observed for these species in the chemical profiles of Specieeurope (Pernigotti, 2016) and various literature related to similar socio-geographic environments (see References, p35).

The discrimination between the factors was made by isolating known fingerprints from different sources. When it comes to regional sources, secondary aerosols are common and often in form of SO_4^- , NO_3^- and NH_4^+ . See more in the section about secondary aerosols.

Biomass burning was also one factor that used to be obvious in many cities. This factor is well identified through the high contribution of OC, usually higher than the EC and that typically go together with a clear contribution of K^+ . As illustrated in the literature the best marker for the biomass burning factor is the Manosan together with the Galactosan.

Another factor that is always present is soil dust. It is well identified by the large contribution of As. In the Balkan region, As is quite common in the mines and slopes cuttings, and top soils and therefore a good marker of crustal source.

Other factors show more variability in their finger prints in our study, probably because of a blend of sources that wasn't possible to discriminate further with the set of species that have been successfully measured.

In these latter factors are the fossil burning. The signal is however quite often a blend of different sources such as coal burning, possibly oil burning and probably traffic exhaust. The set of species makes it hard to discriminate the traffic from other fossil burning sources. But coal burning has a typical higher contribution of EC than OC together with a high contribution $SO_4^{2^2}$. It is however important to note that brown coal (lignite) emits less $SO_4^{2^2}$ than coal and that both types are burnt in the BiH.

Occasionally there are fingerprints with a high contribution from Mn and these factors have been attributed to the industry so far.

On one occasion (Banja-Luka) there was a fingerprint with a high K^+ contribution that couldn't be attributed to a known source so far.

Sarajevo - constrained solution



Tuzla - constrained solution



Zenica – constrained solution



Banja Luka - constrained solution



Bijeljina - constrained solution,







Secondary aerosols:

Two ions, sulphate (SO₄²⁻) and nitrate (NO₃⁻), were analysed from the filter samples and used as representatives for the secondary inorganic aerosols (SIA), principally existing together with ammonia ions. SIA is formed in the atmosphere through the transformation of gas phase precursors – principally NH₃, SO₂ and NOx - emitted both by anthropogenic and biogenic sources (Amato, 2016). The SIA gas-phase formation can take from a few hours to a few days, depending on the weather conditions (e.g.; temperature, humidity), solar radiation and the concentration of different oxidants. With high aerosol water content (AWC), e.g.; during haze conditions, there are also highly effective aqueous-phase transformation processes producing sulphate and nitrate (Chunrong Chen, 2021). One could thus expect variations in the SIA components that respond to different meteorological conditions (Ogulei, 2006) (Prakash, 2017) and stable atmospheric inversion episodes (Srivastava, 2018).

The FHMZ and RHMZ provided observations and oral communication about the weather situation that occurred during the monitoring campaign. November and December 2020 quite often showed a high relative humidity, fog and/or overcast weather conditions. January, on the other hand, showed more clear weather conditions with variable precipitation. Figure 7 shows that sulphate levels were significantly higher during November-December in all six cities. For nitrate the results were more diverse.



Figure 7: Concentrations (µg/m³) of sulphate and nitrate during November-December 2020 and January 2021

All combustion of sulphur-containing fuels will emit SO₂, but also a fraction of sulphate. The latter is normally considered small, in dispersion modelling the sulphur emissions from industrial ovens are typically set to 95% as SO₂ and 5% as SO₄²⁻. However, a study from China (Dai Q, 2019) from an area with extensive residential coal combustion indicated that the primary emitted sulphate could raise up to 40-50% during wintertime. Since we have a similar emission pattern in BiH, we could expect that a significant part of the sulphates is directly emitted and not secondarily formed in the atmosphere. This is likely to produce PMF results with more overlap between sulphate and other locally emitted residential heating markers such as OC and EC.

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Factor: Secondary sulphate aerosol	16%	23% (shared with coal burning sources)	21%	9%	18%	5%
Factor: Secondary nitrate aerosol	14% (shared with fossil burning sources)	10%	14%	24% (shared with fossil burning sources)	8%	24%

Table 11: secondary aerosols factors comparison

Brod:

The sampling place of Brod showed a factor strongly pointing to ammonium which explained about 80% of the whole measured ammonium. This factor could be named spontaneous primary ammonium because of this strong and isolated signal. The contribution of that factor to the SO_4^{2-} seems to be negligible. However, the relative concentration of SO_4^{2-} seems to be similar to the ones within the finger prints of secondary sulphate aerosols displayed for other places like Zenica. It is therefore possible that the strong emission of sulphate, specific to the Brod location, and related to the heavy oil has evened out the sulphate contribution in this first factor that would otherwise be much higher. That is why this first factor is called primary sulphate factor, even if the contribution for the SO_4^{2-} seems to be negligible.

Biomass Burning:

The biomass burning is often an important source of PM 2.5 particles in the Balkan region and frequently related to old fashion stoves, low temperature fireboxes, moist wood and hard wood. The classical fingerprints of the biomass burning when running PMF model is a significant contribution of OC and K⁺ together with anhydro-sugar like levoglucosan, manosan and galactosan. The biomass burning often shows an EC contribution as well but always with a lower contribution than of OC.

It is meaningful to underline that the most common burnt wood in BiH is the beech and that the combustion of that wood emits 10 times less levoglucosan per mass of burnt wood than other hard woods like the oak (Collet S, 2016). In addition, it has been shown in Poland that the combustion of brown coal (in particular lignite) can produce a quite high signal of Levoglucosan at the contrary for the Manosan and Galactosan (Rybicki, 2020). Since there are several brown coal mines in operation in the BiH and an active market of brown coal (Eurostat, 2021) these are a common source of energy even at individual household scale and particularly in towns. However, different policies and incitement programs conducted during the last decade accounted for a progressive swift to biomass burning and in particular pellets burning.

Because of these two specific reasons mentioned above, the Levoglucosan wasn't chosen as a primary indicator of biomass burning but instead the Manosan have been chosen since it is more closely related to the biomass burning.

In the table below the average contribution from biomass burning is listed and also weeks of peak events are listed. In some of the cases it can be seen that these peaks coincide with low temperatures, which should make sense since the need for heating increases. However, this is not always the case and there are other meteorological factors than temperature that can affect the factor contribution, such as events of atmospheric inversions, wind patterns and precipitation.

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Factor: Biomass burning	27%	34%	54% (shared with coal burning source)	19%	36%	44% (shared with coal burning source)
Peak period (weeks number)	48, 49, 51, 52	49, 51, 52	48, 52 , 3	48, 52, 3	52, 53, 1	52, 53, 1 , 2
Period with temperature near or below 0°C (weeks number)	48, 49, 51, 2, 3, 4	48 , 49, 51, 2, 3 , 4	48, 49, 51, 2, 3, 4	48, 49, 2	48, 49, 2	49 , 2, 3
				The importance of a non- burning source specific to the site in Banja Luka has probably even out the size of the biomass burning factor.		

Table 12:	Biomass	factor	comparison
-----------	----------------	--------	------------

Fossil Burning

The fossil burning factor is characterised by a higher contribution to the EC than to the OC and a generally low contribution of K^+ . The major sources of fossil burning particles in the BiH are the coal and brown-coal burning, in power-plant and district heating or individual household heating, together with diesel and gasoline engines. The range of measured species does not permit to clearly discriminate the coal sources from the liquid fuel sources unless their time series differ from each other. The coal burning sources are expected to have their peaks related to the coldest periods whereas the traffic exhaust sources are expected to decrease during the weekend. The fuel burning for heating purpose is however common in BiH which made even harder to isolate the traffic exhaust.

For these reasons the model often didn't succeed to separate the fossil burning sources from other sources. Therefore, the analysis has been divided and distributed in the below table in order to easily compare the underlying principles.

When the factor has a significant contribution to the SO_4^{2-} then the reasonably suspected fuel is coal or brown coal (Prcanovic, 2018).

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Fossil burning (coal/brown coal)		23% (shared with Secondary sulphate aerosol)	54% (shared with Biomass burning)	24% (shared with secondary nitrate aerosol)	36%	
		High contribution to NH4 ⁺ and SO4 ²⁻ that suggests a secondary sulphate aerosol source. But the time- series shows peaks correlated to the coldest weeks and the factor contribute to the EC as well. For these reasons this factor is suspected to be a blend of coal burning and sulphate aerosol.	Very high contribution to anhydrosugar and OC suggests Biomass burning source. But the very high contribution to EC shows that fossil-burning sources constitute a major part as well, probably about ½ of it. The time-series displays a quite constant source (there are no significant peaks, the max value was at 3%) Polar-plot shows a spreading from the direction of the university and up-hill settlements.	High contribution to NO ₃ ⁻ that suggest a secondary nitrate aerosol source. But the significant contribution to EC combined with the lack of other possible factors suggest that a fossil- burning source is blended in this factor.	Significant contribution to EC and to SO4 ²⁻ and As suggests a coal burning source. The time-series displays a quite constant source. Polar-plot shows a spreading mostly from the direction of the town.	

Table 13: fossil burning (coal / brown coal) factor comparison

Fossil burning & traffic	23%	20%		
	High contribution to EC and significant to metals & anhydrosugars suggests both exhaust and non-exhaust traffic sources. Time-series show a recurrent decreasing trend during week-ends.	High contribution to EC & significant to As suggests both exhaust and non- exhaust traffic sources. Time-series show a recurrent decreasing trend week- end.		
	Polar-plot shows a large spreading of the sources.	Polar-plot shows a spreading in the direction from the main settlements up-wind.		

Table 14: fossil burning & traffic factor comparison

Heavy oil burning primary sulphate in Brod:

The Heavy oil burning primary sulphate explains about 80% of the whole SO_4^{2-} and a small portion of EC. That finger-print, with very low K⁺ and low OC is consistent with the one named "petrochemical and heavy oil combustion suburban" in the finger-prints database of SpecieEurope (JRC, 2017). The origin of that factor is probably to a large degree from the refinery situated south of the sampler but might also be from some heating facilities in Brod and the neighbouring Croatian city of Slavonski Brod (Jeričević, 2019). It can also be seen from the polar plot that high concentrations of this factor often coincide with winds coming from the south, were the refinery is. It is important to underline however that the refinery had a very low activity during the sampling period and the last two years. This factor has a significant contribution of 21% of the total PM 2.5.

Soil dust:

The Soil dust factor is characterised by a recurrent significant contribution to the As concentration. In this study between 30% to 60% of this element is usually explained by the Soil dust factor. It is worth noting that the central part of the Balkan peninsula shows a complex geology and has some geological formations and ore deposits rich in As (Tarvainen T, 2013). In this context the mine tailings, infrastructure earthworks and road dust resuspension can be some direct sources of particulate matters with a fair amount of As. As well, the metallurgical and electric-power plants that process such ore can be some important sources of As which adsorb on the emitted particulate matters. Therefore, the As is used in this study as an indicator of Soil dust and Soil resuspension.

It seems that this factor is often driven by some local and intermittent sources that can vary a lot depending on the sampling localisation and place. These puffs of soil dust factors are depicted in the time-series by some sudden peaks, and in the polar plots by some preferred wind direction with contrasted contributions.

	Sarajevo	Tuzla	Zenica	Banja Luka	Bijeljina	Brod
Factor: Soil dust	8%	13%	6%	23%	2%	6%
	Probably driven by two unusual but non-identified events in the vicinity and otherwise by the traffic resuspension.	Probably driven by one significant non- burning source in the vicinity of the sampling site.	Probably driven by the traffic resuspension.	Probably driven by several significant non- burning sources in the vicinity of the sampling site.		Probably driven by the nearby road.

Table 15: Soil dust factor comparison

Other factors

Industry factor in Zenica and Banja Luka

The Industry factor explains about 80% of the Manganese. That metal used to be linked to the metal industry, at least when not together with a soil marker. The polar plot from Zenica shows a specific pattern with N-NW winds and the industrial steelwork area is situated 1,5 km upwind. For Banja Luka, the combination of the time series and the polar plot allows to conclude for general background enriched with a few concentrated events with two single days explaining each 12% of that factor. It is however not possible to point out a specific source. The portion of the Industry metal factor among the whole PM 2.5 is only about 5% both in Banja Luka and Zenica.

Cadmium-rich factor in Sarajevo:

One factor among the validated modelled solution explains about 90% of the Cadmium measured during the period. The portion of the Cadmium-rich factor among the whole PM 2.5 is rather high and reaches 12%. Both the time series and the polar plot indicate that the major sources should be linked to two events in late November and in mid-December, with one single day explaining 10% of that factor. Both isolated events seem closely related to the soil dust factor. A possible source may be some isolated fire of blended materials in the vicinity at that time or some more common sources such as fossil fuel combustion (Huremović, 2020) but the latter makes it difficult to explain the peaks.

Cadmium-rich factor in Zenica:

One factor among the validated modelled solution explains about 80% of the Cadmium measured during the period. The combination of the few peaks in the time series and the generally yellow polar plot allows to conclude for a general background enriched with a few concentrated events coming from two narrow places situated in the W-NW of the sampling place. It is not possible however to point out a specific source. The portion of the Cadmium factor among the whole PM 2.5 is only about 3% in Zenica.

Potassium-rich in Banja Luka:

The Potassium-rich factor explains about 60% of the Potassium whereas the Biomass burning explains only 20% of it. The combination of the time series and the polar plot allows to conclude for a quite continuous source in the background but that differs from the biomass burning factor. There is a significant correlation between the Potassium-rich factor and the Soil dust factor, probably related to the similar contribution time-series at least in November and December. It is worth considering to merge these two factors as a single Soil dust factor. But the Soil dust factor represent already 23% of the modelled PM 2.5 and sound therefore doubtful to even add the Potassium-rich factor that represent 20% of the modelled PM 2.5. It is not possible unless additional information or measurements to point out a specific sort of source for that amount of Potassium in Banja Luka.

Conclusion

This study demonstrates the interest of such result produced by the Positive Matrix Factorisation model when using the USEPA PMF 5.0 software, with the analytical results of the fine particulate matters PM2.5 daily collected on filters. The sampling campaign conducted for this study occurred during the winter 2020-2021 in the 6 cities of Bosnia and Hercegovina. Depending on the sampling location, between 5 to 6 different groups of sources of PM 2.5 have been isolated. The time-series has been analysed with consideration to the surrounding geography and human activities, together with the weather data. These gave some useful insights for the identification of major groups of sources of PM 2.5. The general apportionment of sources is similar to those available from other studies of the region (see References page35) and is shortly described here after.

The Secondary aerosols take a large part of the PM2.5 measured. Together the Secondary sulphate and Secondary nitrate aerosols represented between 20% to 35% of the measured PM2.5. These aerosols are the products of the natural atmospheric oxidation of the precursors that are emitted mostly by the anthropogenic activities (NH₄ by agriculture, NOx by all types of combustion, and SO₂ by coal combustion). Since this transformation takes between a few hours to a few days, Secondary aerosols might be considered to be originating from the areas and regions distant from the sampling place. However, the contribution of local anthropogenic emissions would be also possible during episodes of temperature inversions. These stable atmospheric conditions limit the air mass movement over the area and promote the accumulation of local emissions sources as well as their transformation *in-situ*.

The emission by the combustion of biomass like wood, pellet or shrub waste also represent a large part of the PM2.5. The portion of biomass burning group of sources seems to be larger in the small towns, where it represents about 35% of the PM2.5, than in the larger towns, where it represents between 19% to 25%. This might be related to a greater part of other local sources that reduces the portion of the Biomass burning. These can be heavier traffic load in the area of the sampler, more frequent use of coal combustion for residential heating in the larger cities than in the small cities, etc.

Due to a restricted number of measured species provided by the chemical analysis, the portion of PM2.5 emitted by the combustion of fossil fuel (coal, brown coal, diesel, gasoline) have been difficult to separate from some other groups of sources. In this study they are often blended in the results with other type of sources like Secondary aerosols and Biomass burning. It is possible however to say that the contribution of the combustion of coal and brown coal to the ambient PM2.5 is between 14% to about 25% and is sometime largely influenced by some large district heating installations or industrial installations in the neighbouring area.

The portion of PM2.5 issued from the vehicles have been isolated in two location only, Sarajevo and Tuzla, and includes both exhaust and non-exhaust emissions. The contribution of the traffic to the ambient PM2.5 is between 20% to 23% in these results.

The portion of PM2.5 related to Soil dust either naturally emitted, or resuspended by the human activity, has always been identified. The soil contribution depended on the sampling location and varied between 2% to 13%. One particulate sampling localisation in Banja-Luka has however a result largely over the range with a Soil dust contribution largely disturbed by a significant source of inorganic particles that haven't been possible to explain yet.

The total weight of all the measured species represent about 60% of the total mass of PM2.5 sampled. This ratio is common for the kind of analysis used. It means that some species that haven't been measured but that used to have a significant part of the of the overall PM2.5 - like calcium, silicon, iron, aluminium - have been uniformly distributed in all apportioned sources. But since these species are more related to the Soil dust factor than the others factors, it is therefore likely that the Soil dust source has a larger share of the PM2.5 than calculated by the PMF. It probably can even double and consequently reduce the portions of the other sources. The results should be interpreted with this in mind.

Finally, with the aims to conclude with a satisfactory analytical measurement, it has been agreed by the partners to conduct a second sampling campaign during the winter 2021-2022, with different technical arrangements than during the winter 2020-2021. These new arrangements will hopefully lead to a wider range of measured species and thus to a more accurate result of the PMF model. This campaign will be conducted in Sarajevo and Banja Luka, where the receptor model had the most difficulty to resolve a good quality source apportionment. Their very complex urban and topographical environments, especially in Sarajevo, require exceptionally accurate data for the PMF model to produce reliable results. The result will help then to get a better understanding of the source apportionment in Sarajevo and in Banja Luka but would also permit a better interpretation of the overall results.

All these source apportionment results will be compared in the future to the results of the dispersion models that will be conducted for the cities of Sarajevo and Banja Luka. These results should be available in the first half of 2022.

References

- Almeida, S. &.-I. (2020). Ambient particulate matter source apportionment using receptor modelling in European and Central Asia urban areas. *Environmental Pollution*, 266, 115199.
- Amato, F. A. (2016). AIRUSE-LIFE+: A harmonized PM speciation and source apportionment in five southern European cities. *Atmospheric Chemistry and Physics*, 16, 3289-3309.
- Carslaw, D., & Ropkins, K. (2012). *Function for plotting bivariate polar plots with smoothing*. Retrieved from openair: open source tools for air quality data analysis: https://davidcarslaw.github.io/openair/reference/polarPlot.html
- Chunrong Chen, H. Z. (2021). Aerosol water content enhancement leads to changes in themajor formation mechanisms of nitrate and secondary organicaerosols in winter over the North China Plain. *Environmental Pollution*, 287, 117625.
- Collet S, P. J. (2016). *Determination of pollutants emission factor of domestic wood stoves*. Angers, France: Ademe.
- Dai Q, X. B. (2019). Residential coal combustion as a source of primary sulfate in Xi'an, China,. *Atmospheric Environment*, 196, 66-76.
- Eurostat. (2021, august). *Production of lignite in the Western Balkans statistics*. Retrieved from Eurostat statistics explained: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Production_of_lignite_in_the_Western_Balkans_-_statistics
- Huremović, J. Ž. (2020). Analysis of PM10, Pb, Cd, and Ni atmospheric concentrations during domestic heating season in Sarajevo, Bosnia and Herzegovina, from 2010 to 2019. *Air Qual Atmos Health*, 13, 965–976.
- Jeričević, A. &. (2019). Identification of diverse air pollution sources in a complex urban area of Croatia. *Journal of Environmental Management*, 243, 66-67.
- JRC. (2017, July). *Specieurope*. Retrieved from Source Apportionment: https://sourceapportionment.jrc.ec.europa.eu/Specieurope/index.aspx#
- Ogulei, D. &. (2006). Source apportionment of Baltimore aerosol from combined size distribution and chemical composition data. *Atmospheric Environment*, 40, 396-410.
- Paatero, P. a. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5, 111-126.
- Pernigotti, D. &. (2016). SPECIEUROPE: The European data base for PM source profiles. *Atmospheric Pollution Research*, 7, 307-314.
- Perrone, M. &. (2017). Sources and geographic origin of particulate matter in urban areas of the Danube macro-region: The cases of Zagreb (Croatia), Budapest (Hungary) and Sofia (Bulgaria). Science of The Total Environment, 619 (4).
- Prakash, J. &. (2017). Chemical characterization of PM1.0 aerosol in Delhi and source apportionment using positive matrix factorization. *Environmental Science and Pollution Research*, 24, 445-462.
- Prcanovic, H. &. (2018). Seasonal variations of sulfur dioxide in the air in Zenica city during 11 years period 2006 2016. *International Journal of Advanced Research*, 6, 1133-1139.
- Rybicki, M. &. (2020). Composition of organic compounds from low-temperature burning of lignite and their application as tracers in ambient air. *Chemosphere*, 249, 126087.
- Srivastava, D. &.-L.-L. (2018). Speciation of organic fraction does matter for source apportionment. Part 1: A one-year campaign in Grenoble (France). Science of The Total Environment, 624, 1598-1611.
- Tarvainen T, S. A. (2013). Arsenic in agricultural and grazing land soils of Europe. *Applied Geochemistry*, 28, 2-10.

Figures

Figure 1: Sampling sites in Bosnia and Herzegovina during winter 2020-2021	3
Figure 2: Sven Leckel SEQ47/50-RV sampler and transport case	1
Figure 3: filter cutting for analysis; 1/2 filter for elemental + punch for OC/EC (left), and 1/2 fo anydrosugar + punch for ion chromatograhy (right)	r 5
Figure 4: typical result with factor definition (left) and times-series of the factor (right) 12	3
Figure 5: example of result with factor definition (left) and times-series of the factor (right) 1	3
Figure 6: polar plot of the times-series	3
Figure 7: Concentrations (µg/m ³) of sulphate and nitrate during November-December 2020 and January 2021	5
Figure 8: pie chart in appendixes	7

Tables

Table 1: characteristics of the sampling sites	5
Table 2: expanded uncertainties for elemental analysis, with x being the calculated mass concentration of each element	7
Table 3: Method detection limit for the elemental analysis	7
Table 4: expanded uncertainties for Ion Chromatography analysis, with x being the calculated mass concentration of each ion	7
Table 5: Method detection limit for the Ion Chromatography	8
Table 6: Averaged concentration, and standard deviation (SD), of PM 2.5 and chemical species 1	s 10
Table 7: parameters for the appreciation of total mass of PM 2.5 1	1
Table 8: Input data and PMF 5.0 settings 1	5
Table 9: base runs diagnostics 1	6
Table 10: Constrained runs settings and validations	17
Table 11: secondary aerosols factors comparison	26
Table 12: Biomass factor comparison	28
Table 13: fossil burning (coal / brown coal) factor comparison	29
Table 14: fossil burning & traffic factor comparison	30
Table 15: Soil dust factor comparison	31

Appendix

The following pie-charts, similar to the one in Figure 8, show two separated information for each sampling place.

- The first pie-chart shows the apportionment of the sources, or group of sources, that explain the measured weight of PM 2.5 during the sampling period. The colours of the named sources are the same in a page, but differ from one page to another.
- The other pie-charts show the distribution of each measured species weight within the different apportioned sources, or group of sources (see the colours of the first pie-chart for names). 0% is displayed when the contribution from a factor is under 1%.



Figure 8: pie chart in appendixes

This information is already shown by the black square-dots in the graphs displaying the composition of factors for each sampling places. But it is gathered by species here. Some species trend to be tightly related to only a few sources whereas some species are less specific.



Sarajevo - factors contributions





Zenica – factors contributions





Banja Luka – factors contributions



Bijeljina – factors contributions



