

## Article

# Fine Particulate Matter (PM<sub>2.5</sub>) Sources and Its Individual Contribution Estimation Using a Positive Matrix Factorization Model

Gahye Lee <sup>1</sup>, Minkyong KIM <sup>2</sup>, Duckshin Park <sup>2,\*</sup> and Changkyoo Yoo <sup>3,\*</sup><sup>1</sup> Seohaean Research Institute, Chung Nam Institute, Hongseong 32227, Republic of Korea<sup>2</sup> Korea Railroad Research Institute (KRRI), 176 Cheoldobakmulkwon-ro, Uiwang-si 16105, Republic of Korea<sup>3</sup> Department of Environmental Science and Engineering, Kyung Hee University, Yongin 17104, Republic of Korea

\* Correspondence: dspark@krii.re.kr (D.P.); ckyoo@khu.ac.kr (C.Y.)

**Abstract:** The effective management and regulation of fine particulate matter (PM<sub>2.5</sub>) is essential in the Republic of Korea, where PM<sub>2.5</sub> concentrations are very high. To do this, however, it is necessary to identify sources of PM<sub>2.5</sub> pollution and determine the contribution of each source using an acceptance model that includes variability in the chemical composition and physicochemical properties of PM<sub>2.5</sub>, which change according to its spatiotemporal characteristics. In this study, PM<sub>2.5</sub> was measured using PMS-104 instruments at two monitoring stations in Bucheon City, Gyeonggi Province, from 22 April to 3 July 2020; the PM<sub>2.5</sub> chemical composition was also analyzed. Sources of PM<sub>2.5</sub> pollution were then identified and the quantitative contribution of each source to the pollutant mix was estimated using a positive matrix factorization (PMF) model. From the PMF analysis, secondary aerosols, coal-fired boilers, metal-processing facilities, motor vehicle exhaust, oil combustion residues, and soil-derived pollutants had average contribution rates of 5.73 µg/m<sup>3</sup>, 3.11 µg/m<sup>3</sup>, 2.14 µg/m<sup>3</sup>, 1.94 µg/m<sup>3</sup>, 1.87 µg/m<sup>3</sup>, and 1.47 µg/m<sup>3</sup>, respectively. The coefficient of determination (R<sup>2</sup>) was 0.87, indicating the reliability of the PMF model. Conditional probability function plots showed that most of the air pollutants came from areas where PM<sub>2.5</sub>-emitting facilities are concentrated and highways are present. Pollution sources with high contribution rates should be actively regulated and their management prioritized. Additionally, because automobiles are the leading source of artificially-derived PM<sub>2.5</sub>, their effective control and management is necessary.

**Keywords:** conditional probability function; PM<sub>2.5</sub>; positive matrix factorization; receptor method

**Citation:** Lee, G.; KIM, M.; Park, D.; Yoo, C. Fine Particulate Matter (PM<sub>2.5</sub>) Sources and Its Individual Contribution Estimation Using a Positive Matrix Factorization Model. *Toxics* **2023**, *11*, 69. <https://doi.org/10.3390/toxics11010069>

Academic Editor: Md. Aynul Bari

Received: 1 December 2022

Revised: 1 January 2023

Accepted: 5 January 2023

Published: 11 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Among member countries of the Organization for Economic Cooperation and Development, the concentration of fine particulate matter—an atmospheric pollutant—in the Republic of Korea was recently reported to be the highest [1]. Although the Republic of Korea met its national standard for concentrations of fine particulate matter less than 2.5 µm in diameter (PM<sub>2.5</sub>; 15 µg/m<sup>3</sup> or less) in 2020, it failed to meet the recommended World Health Organization target of 10 µg/m<sup>3</sup> [2].

Human industrial and/or mechanical activities in rapidly-developing societies produce fine particulate matter [3], with such particles increasing the risks of disease and death as well as posing other threats to human health. Microparticles carrying heavy metals, sulfates, and nitrates are particularly influential in causing or exacerbating cardiovascular and respiratory diseases, cancer, and mental illness, resulting in serious adverse effects on the human body [4]. Moreover, ultrafine particulate matter directly penetrates cells, leading to severe damage [5]. Accordingly, long-term measures and policies to reduce PM<sub>2.5</sub> concentrations are needed in the Republic of Korea, as these particles are more harmful to the human body than coarser particles such as PM<sub>10</sub>.

The Republic of Korea has introduced many management policies to reduce PM<sub>2.5</sub> and uses distributed models and other methods to predict future air quality. Nevertheless, it is essential to identify major pollution sources and assess their contributions with an acceptance model to effectively predict and manage PM<sub>2.5</sub> concentrations, given that the chemical composition and physicochemical properties of PM<sub>2.5</sub> vary according to its spatiotemporal characteristics [6]. However, in the Republic of Korea, research is mainly limited to large areas such as Seoul and Jeju Island, and the majority of observational studies have focused on specific events [7].

“Air control zones” refer to areas with severe air pollution due to emissions from the capital region. The Republic of Korea has recently categorized the central, southeastern, and southern regions of the country as air control zones, with these zones including a total of 77 special and metropolitan cities and municipalities. For this study, Bucheon City in Gyeonggi Province, an air control zone, was selected as a PM<sub>2.5</sub> data collection area. The chemical composition of the particulate matter in this area was analyzed and its physical and chemical properties identified. The positive matrix factorization (PMF) model, the most commonly used receptor model worldwide, was employed to analyze the chemical contents. The PMF model does not rely on a correlation matrix of measured data but, rather, a weighted least squares fitting algorithm based on estimates of errors in the measured data and measurement information. It can accommodate data representing values below the detection limit (BDL) and estimate missing data through error evaluation. The model also minimizes the number of discarded measurements. The aims of the study were to identify the main sources of PM<sub>2.5</sub> and quantitatively estimate the contribution of each pollution source through PMF modeling. Conditional probability function (CPF) plots were generated to identify the potential locations of pollutants arriving from various wind directions using the pollutant contribution data calculated with the PMF model and meteorological data. The results of this study can help with more effective management of PM<sub>2.5</sub> generated not only in this study area, but also in other urban areas in the Republic of Korea, even without a pollutant classification table. Moreover, the data generated can be used as reference data for establishing measures to reduce fine particulate matter and improve air quality.

## 2. Materials and Methods

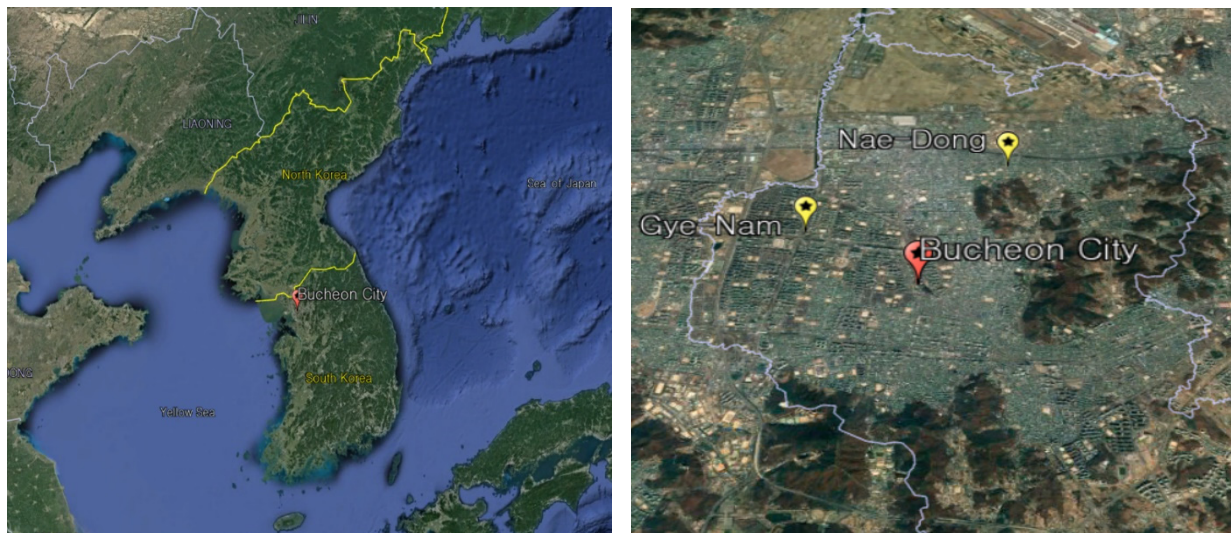
### 2.1. Research Period and Sampling Points

PM<sub>2.5</sub> samples were collected in Bucheon City, Gyeonggi Province, from 22 April to 3 July 2020. Bucheon City is a satellite city located adjacent to the southwestern part of Seoul Metropolitan City, east of Incheon Metropolitan City, and north of Siheung City. It is situated in the middle of the Korean Peninsula along the north–south axis and has a total area of 53.4 km<sup>2</sup>, accounting for 0.5% of Gyeonggi Province’s total area. According to Article 13 of the Enforcement Decree of the Republic of Korea Air Environment Conservation Act, business sites are classified into one to five types according to the number of pollutants generated from discharge facilities. Class 1 business refers to businesses with an annual total of more than 80 tons of air pollutants; Class 2 businesses are those with an annual amount of more than 20 tons to less than 80 tons; Class 3 businesses are those with an annual amount of more than 10 tons; and Class 4 businesses are those with an annual amount of more than 2 tons to less than 2 tons. Therefore, there are a total of 755 business sites emitting environmental air pollutants as of July 2020: 3 class 3 sites, 195 class 4 sites, and 558 class 5 sites in this research area. Table 1 shows the status of pollutant-emitting sites in each administrative district in Bucheon City.

**Table 1.** Air-pollutant-emitting facilities in Bucheon City [8].

Administrative District of Bucheon City	Area (km <sup>2</sup> )	Class 3	Class 4	Class 5	Total
Simgok-dong	2.6		1	4	5
Bucheon-dong	8.1		68	196	264
Sinjung-dong	4.5		3	6	9
Jung-dong	1.8		1	3	4
Sang-dong	3.6				0
Daesan-dong	4.1	1	1	8	10
Sosabon-dong	3.0		3	5	8
Beom-an-dong	5.7			3	3
Seong-gok-dong	7.6		2	3	5
Ojeong-dong	12.4	1	116	330	447
Total	53.4	2	195	558	755

Two sampling points were selected: the Songnae-daero monitoring station (Gyenam Park) in Jung-dong and the Nae-dong monitoring station in Samjeong-dong, located approximately 2.5 km apart (Figure 1).

**Figure 1.** Locations of the study area and sampling points.

## 2.2. Sampling and Analysis

Samples were collected using four sequential particulate matter samplers (Low-Volume Air Sampler, model PMS-104, APM Engineering Co., Ltd., Bucheon-si, Republic of Korea); at each sampling point, one sampler was placed within the monitoring station and one was placed outside. Each PMS-104 instrument sampled at a rate of 16.7 L/min over 24 h. A total of 74 samples—40 from Gyenam Park and 34 from Nae-dong—were collected. Each sampler was equipped with a Teflon filter (2.5- $\mu$ m PTFE membrane, 46.2 mm, Tisch Scientific, Cleves, OH, USA) pretreated for ion and heavy metal content analysis and a quartz filter (QM-A, 47 mm, Whatman PLC, Little Chalfont, UK) heat treated for carbon content analysis. Each Teflon filter was stored in an electronic desiccator (model 08-642-23C, Fisher Scientific, Waltham, MA, USA) under constant temperature ( $25 \pm 3$  °C) and humidity ( $35 \pm 5\%$ ) for 72 h to remove moisture before and after sampling. After drying to a constant weight with static electricity removed, the filters were weighed using an electronic balance (model XP6, No. 1123430327, Mettler Toledo, Seoul, Republic of Korea) with a sensitivity of 0.001 mg. Each quartz filter was heat treated in a furnace at 600 °C for about 6 h to remove organic matter that might have been present in trace amounts before sample collection. After the complete removal of organic matter, each filter was placed on a Petri dish, sealed

with parafilm, and stored in a desiccator before sampling. After sample collection, the filters were stored in a freezer ( $< -20\text{ }^{\circ}\text{C}$ ) to minimize the volatilization of particulate matter and effects from external factors. Analysis took place immediately once a certain number of samples were collected.

To analyze the water-soluble ionic content of  $\text{PM}_{2.5}$ , the filter on which dust was collected was first placed into a 50-mL beaker [9]. Then, 30 mL of tertiary distilled water was added to precipitate the particulate matter. Next, the sample was transferred to an ultrasonic extractor (model WUC-D03H, Daihan Scientific Co., Wonju-Shi, Republic of Korea) and eluted after mechanical shaking and ultrasonication extraction for 30 min. The extracted solution was filtered through a sterile membrane filter (Whatman PLC, UK) with a diameter of 47 mm (pore size  $0.45\text{ }\mu\text{m}$ ) to remove insoluble particles. Finally, to prepare for ion analysis, the solution was transferred to a vial and only the ionic component was eluted. Ion chromatography (model 861, Advanced Compact IC, Metrohm AG, Herisau, Switzerland) was used to analyze the water-soluble ion contents of the following 11 components in  $\text{PM}_{2.5}$ :  $\text{Br}^-$ ,  $\text{PO}_4^{3-}$ ,  $\text{F}^-$ ,  $\text{Cl}^-$ ,  $\text{NO}_3^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Na}^+$ ,  $\text{NH}_4^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ , and  $\text{Mg}^{2+}$ .

X-ray fluorescence (XRF) spectrometry (S2 Ranger X-ray spectrometer, Bruker, Germany) was used to analyze inorganic elemental contents in  $\text{PM}_{2.5}$ . Because XRF spectrometry is a non-destructive analytical method that does not damage filters, no pretreatment was necessary. The following 22 elements were analyzed: Al, Ti, Co, V, Se, Sn, As, SiO, Cl, Mg, Zn, K, S, Br, Ca, Ba, Pb, Cr, Cu, Mn, Fe, and Ni.

For carbon analysis, the thermal-optical transmittance method conforming to the National Institute of Occupational Safety and Health 5040 standard and following a temperature-specific protocol was employed. A thermos-optical OCEC analyzer (Sunset Laboratory Inc., Portland, OR, USA) was used to analyze organic carbon (OC) and elemental carbon (EC).

### 2.3. PMF Model

The PMF model is used in the field of applied statistical analysis. It considers standard deviation and measurement uncertainty through a mathematical algorithm that enhances negative factor loading during factor analysis. Hence, factor loads are calculated only as positive values [10], making it possible to quantitatively estimate pollutant emissions [11]. The basic PMF model is as follows:

$$X = GF + E \quad (1)$$

where  $X$  can be expressed as an  $m \times n$  matrix, with  $m$  representing the chemical species analyzed among  $n$  samples. Therefore, matrix  $X$  is structured as rows comprising the concentrations of chemical species across samples and columns that each represent a collected sample.  $G$  is an  $n \times p$  matrix representing the contribution of a pollutant source to each sample.  $F$  is a  $p \times m$  matrix representing the source profile for a specific pollutant. Here,  $p$  refers to the number of extracted elements. Each column in  $G$  then refers to the amounts of a specific pollutant emitted.  $E$  is a residual matrix and is expressed as shown in Equation (2) below. The most important step in PMF is to determine the appropriate number of elements, and a useful method is to minimize the  $Q$  value by repeatedly assigning a weight factor to  $G$  and  $F$  elements. The  $Q$  value can be obtained as follows:

$$E_{ij} = X_{ij} - \sum_{h=1}^p G_{ih}F_{hj} \quad (i = 1 \sim n, j = 1 \sim m, k = 1 \sim p) \quad (2)$$

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left( \frac{E_{ij}^2}{\sigma_{ij}^2} \right) = \sum_{i=1}^n \sum_{j=1}^m \frac{(x_{ij} - g_{ik}f_{kj})^2}{\sigma_{ij}^2} \quad (3)$$

EPA PMF version 5.0 (US Environmental Protection Agency) was used to run the PMF analysis. The raw data matrix ( $74 \times 35$ ) was reconfigured as part of data pretreatment. A sample with a difference of more than  $\pm 50\%$  between the  $PM_{2.5}$  mass concentration and the sum of all chemical contents was removed. Additionally, three species associated with low reliability—Sn, Cu, and  $F^-$ —and two chemical species—Ba and  $PO_4^{3-}$ —which had BDL values or missing data in 90% or more of all acquired samples, were not included in the analysis. For species that were analyzed using both ion chromatography and XRF spectrometry, the data that produced better results in the regression analysis were selected for PMF.

Data on  $PM_{2.5}$  mass concentrations, the analyzed contents in each sample, and the uncertainty associated with each analyzed concentration are required to perform PMF modeling. Missing concentration data were replaced with geometric mean values, and the corresponding error matrix values were replaced with values of four times the geometric mean. If an analyzed concentration was 0 or BLD, it was replaced by the method detection limit (MDL)/2, with the corresponding error matrix value replaced by  $(5/6) \times MDL$ . In addition, the error matrix corresponding to the calculated concentration values was derived using  $\sqrt{(species\ data \times 0.1)^2 + (\frac{1}{2} \times MDL)^2}$ , whereas the error matrix corresponding to  $PM_{2.5}$  mass concentrations was calculated by multiplying the mass concentration values by four. The original raw data matrix was reconfigured over several steps into a  $73 \times 25$  matrix. Table 2 shows the final input data used for PMF modeling.

Determination of the optimum number of factors is a critical step in the process of PMF modeling. Since the PMF model can be interpreted differently depending on the number of pollutants, repeated modeling should be performed (trial and error) to determine the optimal number of physically meaningful pollutants [12]. The most commonly used methods in determining the optimal number of pollutants in the PMF model include using the scaled residual matrix  $R$  and using the  $Q$  value. Since the purpose of the PMF model is to minimize the  $Q$  value, it is calculated under the condition that the  $Q$  value is minimized among the results modeled by various input variables. In the method of using the scaled residual matrix, the probability that the standardized residual value is in the interval between  $-3.0$  and  $+3.0$  must be at least 80%. The residual matrix  $R$  can be expressed as the following Equation (4).

$$R_{ij} = \frac{E_{ij}}{S_{ij}} \quad (4)$$

In addition, two variables, i.e., a maximum independent column mean ( $IM$ ) value and a maximum independent column standard deviation ( $IS$ ) value, can be calculated from the residual matrix  $R$  through Equations (5) and (6).

$$IM = \max_{j=1\dots m} \left( \frac{1}{n} \sum_{i=1}^n r_{ij} \right) \quad (5)$$

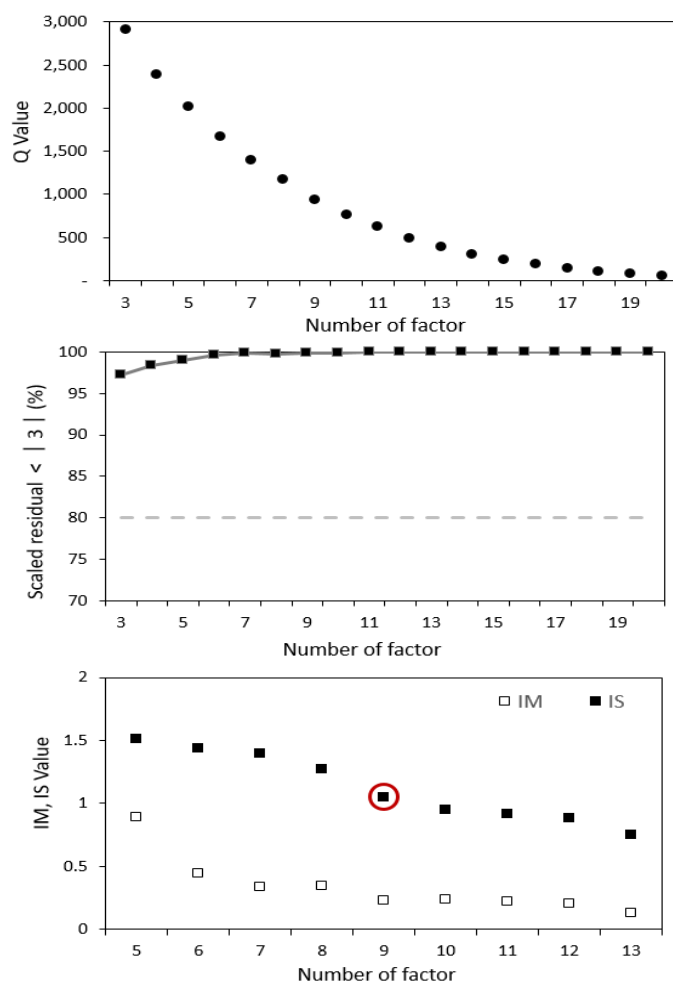
$$IS = \max_{j=1\dots m} \left( \sqrt{\frac{1}{n} \sum_{i=1}^n (r_{ij} - \bar{r}_i)^2} \right) \quad (6)$$

When the number of factors increases to a threshold value, the  $Q$  value decreases, and the  $IM$  value and the  $IS$  value also decrease rapidly (Figure 2). As a result, the optimal number of pollutants was determined to be 9.

**Table 2.** Summary of PM<sub>2.5</sub> concentrations and the concentrations of 25 chemical species used for PMF modeling.

Species	S/N Ratio DL <sup>(b)</sup>	Min	Concentration (µg/m <sup>3</sup> )			Max
			25th	50th	75th	
PM <sub>2.5</sub>	--	0.3066	14.5581	17.8298	22.3221	60.7302
Al	7.5 0.0060	0.0030	0.1199	0.2178	0.3749	2.2850
Ti	8.2 0.0004	0.0002	0.0094	0.0139	0.0205	0.0722
Co	7.8 0.3761	0.0000	0.0019	0.0044	0.0102	0.0481
V	7.0 0.0001	0.0001	0.0011	0.0029	0.0047	0.0092
Se	5.5 0.0004	0.0002	0.0009	0.0025	0.0043	0.0110
As	7.5 0.0001	0.0000	0.0017	0.0045	0.0080	0.0284
SiO	2.8 0.0687	0.2172	0.2172	0.2172	2.3837	7.8155
Mg	7.9 0.0075	0.0037	0.0822	0.1579	0.2422	1.0879
Zn	9.0 0.0002	0.0037	0.0440	0.0732	0.0910	0.1295
Br	8.9 0.0002	0.0020	0.0113	0.0186	0.0262	0.0519
Ca	6.7 0.0033	0.0017	0.0156	0.0419	0.0661	0.5559
Pb	8.6 0.0003	0.0001	0.0134	0.0207	0.0304	0.1302
Cr	8.4 0.0004	0.0021	0.0048	0.0061	0.0081	0.0339
Mn	8.1 0.0007	0.0003	0.0106	0.0142	0.0199	0.1024
Fe	8.9 0.0011	0.0102	0.2010	0.2625	0.3296	0.7181
Ni	6.9 0.0001	0.0000	0.0005	0.0012	0.0032	0.0211
Na <sup>+</sup>	5.3 0.0403	0.0448	0.1036	0.1480	0.2819	0.6206
NH <sub>4</sub> <sup>+</sup>	8.3 0.0311	0.0156	0.8914	1.7629	2.9646	8.6075
K <sup>+</sup>	3.6 0.0493	0.0246	0.0246	0.1654	0.2079	0.3672
Cl <sup>-</sup>	2.7 0.0983	0.0492	0.1044	0.1873	0.2932	0.7726
NO <sub>3</sub> <sup>-</sup>	6.4 0.1117	0.0559	0.4143	1.1186	2.5808	12.7241
SO <sub>4</sub> <sup>2-</sup>	3.3 0.1076	0.0538	0.0842	0.2841	0.4689	1.0071
OC	6.9 0.1164	2.1639	6.7140	7.9161	8.9695	13.4190
EC	7.3 0.0196	0.2119	1.1001	1.2565	1.4784	2.4926

<sup>(a)</sup> S/N ratio: Signal-to-noise ratio, <sup>(b)</sup> DL: Detection Limit.



**Figure 2.** The Q value, scaled residual, IM, and IS for determining the number of factors.

Thereafter, changes in the value of Q with each rotation of elements were checked every 0.1 steps from  $-1.0$  (instead of 0) to  $1.0$  to determine the optimal Fpeak value for converting the results of the derived final model into a simple form. The most physically meaningful result was derived when the Fpeak value was  $-0.1$ .

#### 2.4. CPF Model

The CPF model can be used to estimate the potential locations of emitted pollutants arriving from different wind directions by applying pollutant contribution data, calculated using the PMF model and meteorological data, to evaluate the local impacts of pollutants [13–16]. In CPF analysis, conditional probability values for different wind directions for a reference concentration can be calculated as follows:

$$CPF_{\Delta\theta} = \frac{m_{\Delta\theta}}{n_{\Delta\theta}} \tag{7}$$

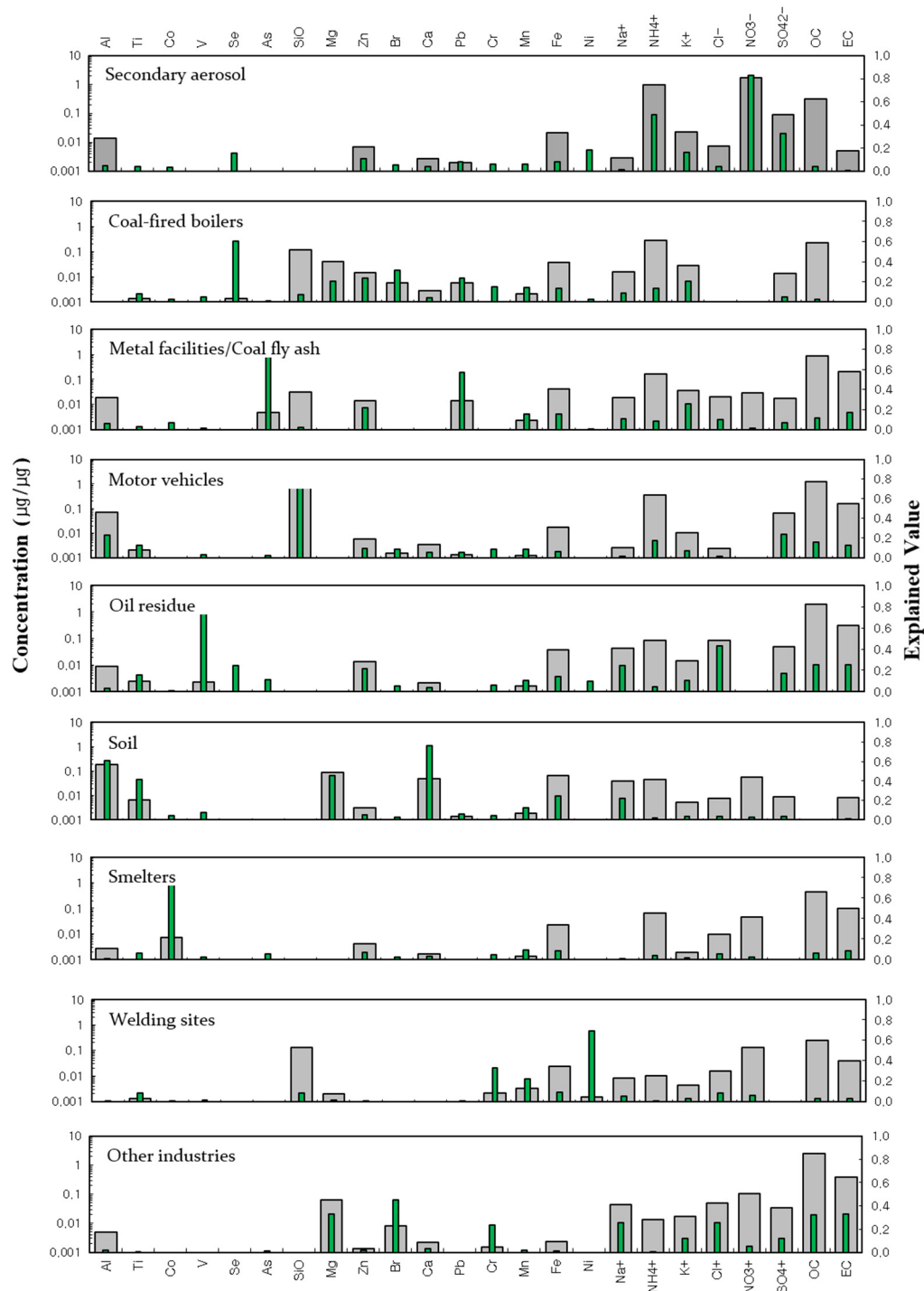
where  $\Delta\theta$  refers to a particular wind direction sector,  $n_{\Delta\theta}$  refers to the total frequency of wind blowing through  $\Delta\theta$ , and  $m_{\Delta\theta}$  refers to the number of times the pollutant concentration exceeds the appropriate standard when the wind is blowing through  $\Delta\theta$ .

The 80th percentile concentration value was used as the critical concentration standard and low wind speeds ( $\leq 1$  m/s) were excluded as, at such speeds, conditions are considered relatively “windless”, during which uncertainty about the wind direction is high. Because meteorological data for the Republic of Korea include 16 compass directions for wind direction, the number of wind direction sectors was determined to be 22.5.

### 3. Results and Discussion

#### 3.1. Source Classification and Identification through PMF Modeling

Pollution sources were identified by reviewing source profiles compiled from the analysis of the concentration and proportion (%) of each chemical species and the time series of elemental change patterns. Figure 3 shows the classification of pollutants based on modeling in this study.



**Figure 3.** Source profile resolved from PM<sub>2.5</sub> in the study area. Gray: Concentration of species. Green: Percentage of species (%).



Nine pollution sources were identified, of which the largest was that of secondary aerosols, comprised mostly of  $\text{SO}_4^{2-}$ ,  $\text{NO}_3^-$ , and  $\text{NH}_4^+$ . These species are considered secondary pollutants, with  $\text{SO}_4^{2-}$  produced from homogeneous and non-uniform reactions in air and  $\text{NO}_3^-$  produced from uniform reactions of gaseous  $\text{HNO}_3$ , which is generated from photochemical reactions and bonds strongly with  $\text{NH}_4^+$  [17,18]. These pollutants travel over long distances and can be derived from the conversion of aqueous solution states, such as in-cloud conversion [19].

The next largest pollution source was estimated to be that of coal-fired boilers, which release emissions containing Se, Br, Pb, Mg, Zn, and  $\text{K}^+$ . Although coal consumption has decreased gradually in recent years, coal is still used in many industrial boilers and furnaces [20]. The third-ranked pollution source was that of metal-processing facilities and coal fly ash related to the processing of iron and non-ferrous metals, which release Pb, As,  $\text{K}^+$ , and Zn as the main pollutants. Motor vehicles ranked fourth. The typical chemical components in automobile-derived pollutants are EC, OC,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ , and Si. Accordingly, these were found in the pollutant mix in this study. Around the study area, there is a high density of commercial districts, including a subway station and an intersection between Seoul's First Ring Expressway and Gyeongin Expressway; thus, traffic in the area is heavy and the floating population is large. This heavy traffic resulted in the classification of motor vehicles as the fourth-ranked pollution source.

The fifth-ranked pollution source was that of oil combustion residues from factories of all sizes in the study area. Pollutants were generated from the combustion of diesel or Bunker-C oil or the incomplete combustion of solid fuels, resulting in V, OC, EC, Na, and Cl being released into the air. Soil, which was found to release heavy metals such as Al, Ca, Mg, Ti, and Fe, was classified as the sixth-ranked pollution source. These elements are typically used to track soil-related pollution and were thus categorized as soil-derived pollutants in this study. Particles distributed in soil can be scattered onto roads, while those on sidewalks can be scattered into the atmosphere, including yellow dust. This study was conducted in spring and summer, when yellow dust is an issue [21].

Smelters—including rotary kilns and blast furnaces—in steelmaking and smelting facilities processing iron ore and non-ferrous metals near the study area were categorized as the seventh-ranked pollution source. Co was the largest contributor, whereas As, Zn, Mn, and Fe were secondary contributors. Together, these elements were the main indicators of pollutants produced by rotary kilns and blasting furnaces. The eighth-ranked pollution source was that of welding sites, with Ni, Mn, and Cr the primary metal components generated at steel-welding facilities. Finally, the ninth-ranked pollution source consisted of other industrial sites, which generated high concentrations of Br, Mg, total carbon,  $\text{Na}^+$ ,  $\text{Cl}^-$ , and Cr. Indeed, many air-pollutant-emitting industrial sites, especially manufacturing plants, are located near the study area.

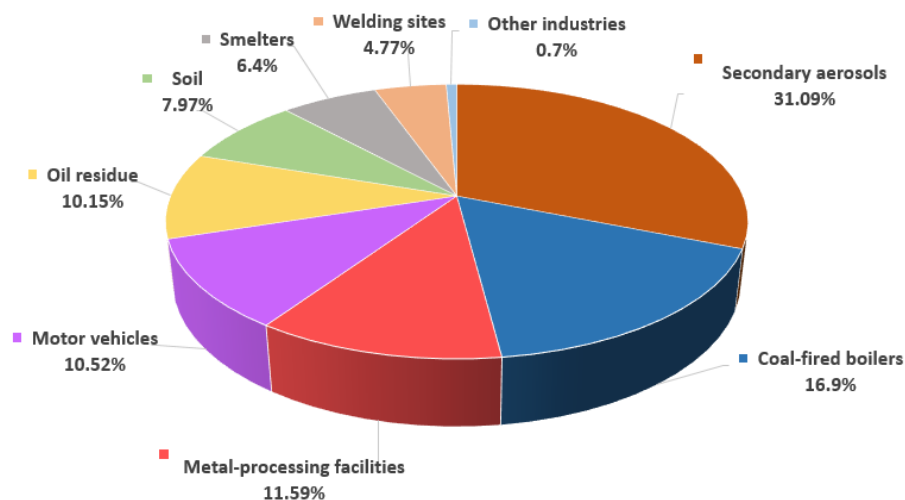
### 3.2. Quantitative Assessment of the Contribution of Each Pollution Source

The contribution of each pollution source was estimated using a scaled G matrix. Figure 4 shows the average mass contribution rate of each confirmed pollution source to  $\text{PM}_{2.5}$  collected in the study area from 22 April to 3 July 2020.

Secondary aerosols, coal-fired boilers, metal-processing facilities and coal fly ash, motor vehicle exhaust, oil combustion residues, contaminated soil, smelters, welding sites, and other industries had average contribution rates of 31.09% ( $5.73 \mu\text{g}/\text{m}^3$ ), 16.9% ( $3.11 \mu\text{g}/\text{m}^3$ ), 11.59% ( $2.14 \mu\text{g}/\text{m}^3$ ), 10.52% ( $1.94 \mu\text{g}/\text{m}^3$ ), 10.15% ( $1.87 \mu\text{g}/\text{m}^3$ ), 7.97% ( $1.47 \mu\text{g}/\text{m}^3$ ), 6.4% ( $1.18 \mu\text{g}/\text{m}^3$ ), 4.77% ( $0.88 \mu\text{g}/\text{m}^3$ ), and 0.7% ( $0.11 \mu\text{g}/\text{m}^3$ ), respectively.

As secondary aerosols contributed the most to  $\text{PM}_{2.5}$  pollution (31.09%, about 1/3 of the total), steps to manage this pollution source must be taken. Additionally, a large proportion of pollutants appeared to be emitted from workplaces. It is likely that many air-pollutant-emitting sites are located near the study area. Standards for air pollutant emissions are urgently needed, and other actions can include restrictions on coal use or the replacement of coal with cleaner fuels. Because motor vehicles are another leading

source of pollutants, specific measures such as restrictions on vehicle exhaust emissions, banning the use of old diesel vehicles, and the establishment of no-driving-day schemes for passenger vehicles (e.g., no driving every 2nd, 5th, or 10th day) should be introduced in the study area.



**Figure 4.** Contribution (%) of each identified source to PM<sub>2.5</sub> pollution.

The consideration of the seasonal variation in the analysis of pollutant contributions can ensure greater reliability in identifying pollutant sources. Accordingly, the contribution of each pollution source to the average PM<sub>2.5</sub> mass was estimated for each month and on weekdays versus weekends. Since the amount of data in July is smaller than in other months, there is a limit to providing accurate information in July. Therefore, it is difficult to generalize that the July contribution data are the average contribution in July, but the contribution was analyzed including July as a reference (Tables 3 and 4, Figures 5 and 6). In Tables 3 and 4, the first row of each pollutant refers to the concentration of the pollutant, and the second row refers to the percentage of each pollutant.

In July, May, April, and June, secondary aerosols accounted for 62.17% (24.35 µg/m<sup>3</sup>), 38.14% (6.03 µg/m<sup>3</sup>), 22.66% (3.61 µg/m<sup>3</sup>), and 20.81% (4.36 µg/m<sup>3</sup>) of total pollutants, respectively. The largest contribution was recorded in July, likely because of the active generation of secondary aerosols from photochemical reactions due to intense solar radiation in summer [22]. The percentage contribution was also higher on weekends than on weekdays (Figure 6).

For pollutants derived from coal-fired boilers, the contribution rates in April, May, June, and July were 30.04% (4.79 µg/m<sup>3</sup>), 16.89% (2.67 µg/m<sup>3</sup>), 11.40% (2.39 µg/m<sup>3</sup>), and 11.13% (4.36 µg/m<sup>3</sup>), respectively. The contribution rate was highest in April, when temperatures are lower than in summer. Despite the recent trend of gradually decreasing coal consumption, the high contribution rate in April can be attributed to greater consumption of coal for heating in winter/early spring. Similar to secondary aerosols, the contribution rate was higher on weekends than on weekdays.

For pollutants derived from metal-processing facilities and coal fly ash, the highest contribution rate was observed in June (14.23%, 2.98 µg/m<sup>3</sup>), followed by April (12.56%, 2.00 µg/m<sup>3</sup>), May (10.13%, 1.60 µg/m<sup>3</sup>), and July (4.74%, 1.86 µg/m<sup>3</sup>). There was no significant difference in the contribution rate between weekdays and weekends.

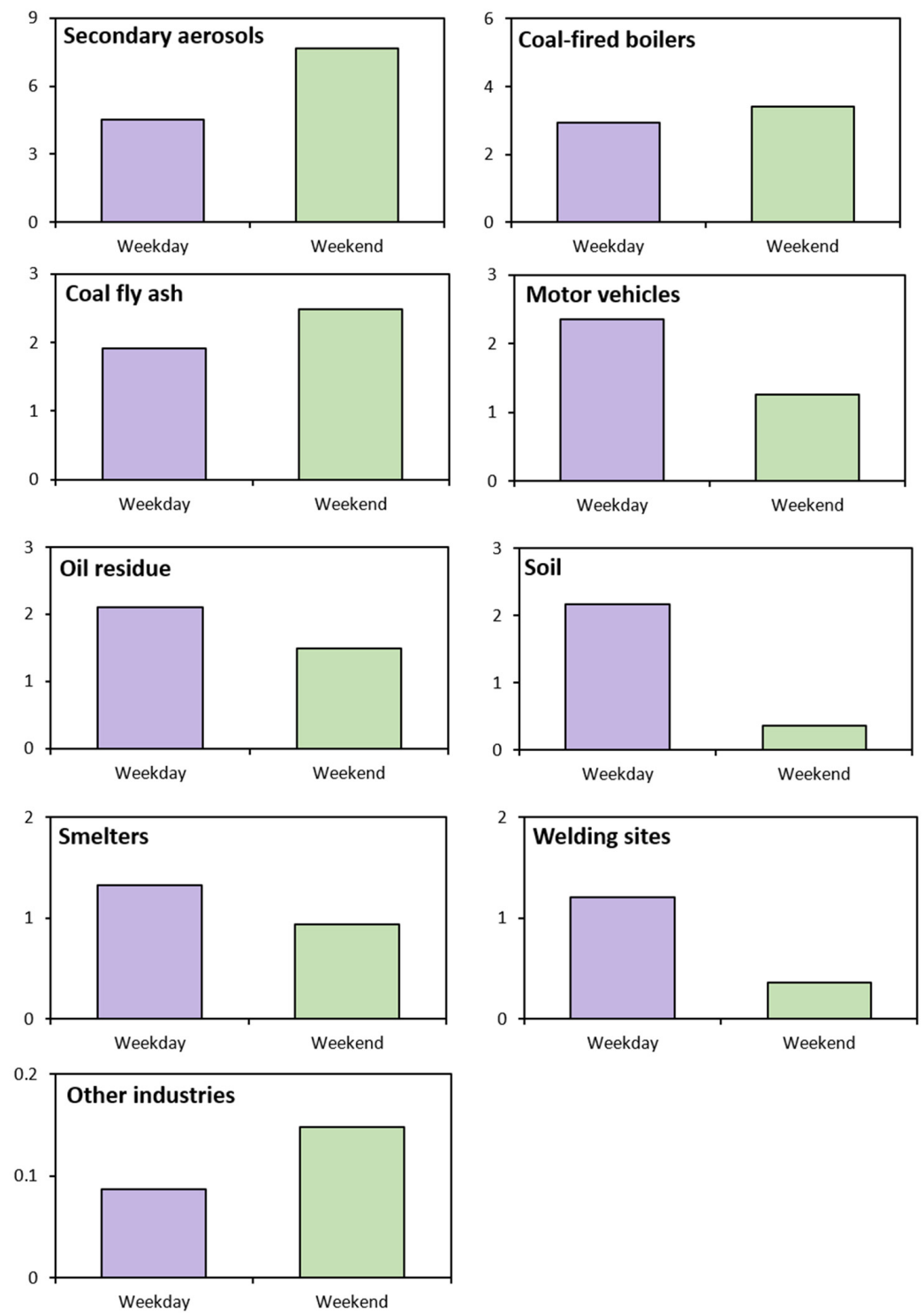
Motor vehicle pollutants contributed to 23.98% (5.02 µg/m<sup>3</sup>), 9.46% (3.70 µg/m<sup>3</sup>), 3.04% (0.48 µg/m<sup>3</sup>), and 0% (0 µg/m<sup>3</sup>) of total pollutants in June, July, May, and April, respectively. The contribution rate on weekdays was higher than on weekends.

**Table 3.** Contribution of each pollution source on weekdays and weekends (unit:  $\mu\text{g}/\text{m}^3$ , %).

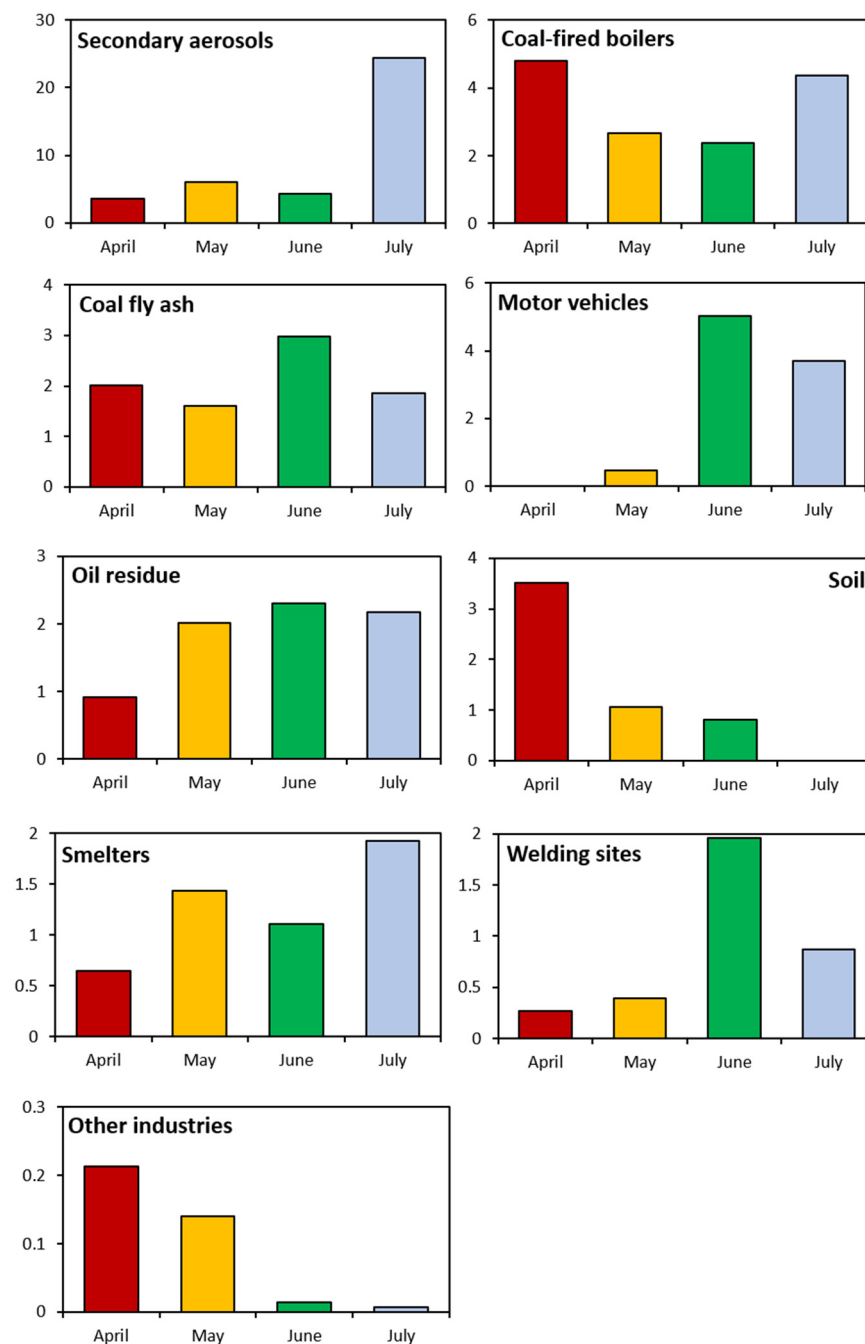
Source	Weekdays	Weekends
Secondary aerosols	4.53 24.34	7.65 42.29
Coal-fired boilers	2.93 15.74	3.41 18.86
Metal-processing facilities and coal fly ash	1.92 10.29	2.49 13.74
Motor vehicles	2.36 12.66	1.26 6.94
Oil combustion residues	2.11 11.31	1.49 8.22
Soil	2.16 11.61	0.36 1.97
Smelters	1.33 7.13	0.94 5.18
Welding sites	1.20 6.46	0.36 1.97
Other industries	0.09 0.47	0.15 0.82
Total	18.63100	18.09100

**Table 4.** Contribution of each pollution source by month (unit:  $\mu\text{g}/\text{m}^3$ , %).

Source	April	May	June	July
Secondary aerosols	3.61 22.66	6.03 38.14	4.36 20.81	24.35 62.17
Coal-fired boilers	4.79 30.04	2.67 16.89	2.39 11.40	4.36 11.13
Metal-processing facilities and coal fly ash	2.00 12.56	1.60 10.13	2.98 14.23	1.86 4.74
Motor vehicles	0.00 0.00	0.48 3.04	5.02 23.98	3.70 9.46
Oil combustion residues	0.91 5.72	2.01 12.71	2.30 11.00	2.18 5.56
Soil	3.52 22.05	1.05 6.64	0.81 3.89	0.00 0.00
Smelters	0.65 4.06	1.43 9.05	1.11 5.29	1.93 4.92
Welding sites	0.26 1.66	0.40 2.51	1.95 9.33	0.87 2.22
Other industries	0.21 1.33	0.14 0.88	0.01 0.07	0.01 0.02
Total	15.95 100	15.82 100	20.94 100	39.16 100



**Figure 5.** Difference in source contribution to PM<sub>2.5</sub> between weekdays and weekends.



**Figure 6.** Monthly variation in source contribution to PM<sub>2.5</sub>.

Pollutants from oil combustion residues accounted for 12.71% (2.01 µg/m<sup>3</sup>), 11% (2.3 µg/m<sup>3</sup>), 5.72% (0.91 µg/m<sup>3</sup>), and 5.56% (2.18 µg/m<sup>3</sup>) of total pollutants in May, June, April, and July, respectively. Although the contribution rate of these pollutants is generally high in winter due to oil use for heating [23], large manufacturing plants and small factories also use significant amounts of fuel. The contribution rate was higher on weekdays than on weekends, probably because the nearby factories operate primarily on weekdays.

The contribution rates of soil-derived pollutants were 22.05% (3.52 µg/m<sup>3</sup>), 6.64% (1.05 µg/m<sup>3</sup>), 3.89% (0.81 µg/m<sup>3</sup>), and 0% (0.0 µg/m<sup>3</sup>) in April, May, June, and July, respectively. That the contribution rate was highest in April was likely due to yellow dust being present in the study area in April and May. Yellow dust has a significant effect on soil-derived pollution in spring. The contribution rate was higher on weekdays than on weekends, indicating the presence of many industrial facilities near the study area. The

higher rate may also be due to particulate matter on sidewalks dispersing into the air via the activities of the large floating population on weekends.

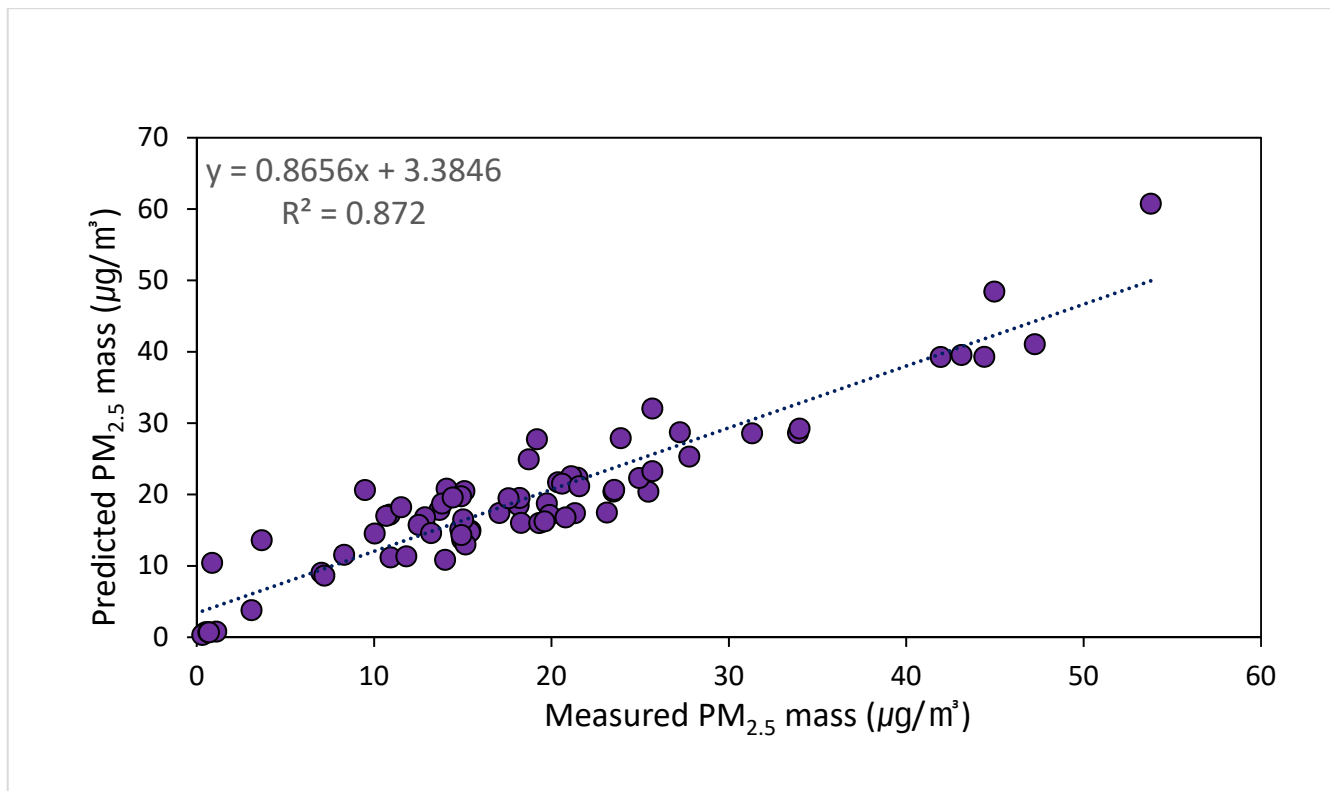
The highest contribution rate of 6.64% ( $1.05 \mu\text{g}/\text{m}^3$ ) for pollutants from smelters was recorded in May. No significant differences in the rate were observed between April, June, and July. The contribution rate was higher on weekdays than on weekends due to kilns, furnaces, and other industrial facilities operating on weekdays rather than on weekends.

The contribution rate of welding-derived pollutants was highest in June (9.33%,  $1.95 \mu\text{g}/\text{m}^3$ ), whereas no significant differences were observed between April, May, and July. The contribution rate was higher on weekdays, when workplaces operate, than on weekends, when they do not operate.

The contribution rates of other industrial pollutants in April, May, June, and July were 1.33% ( $0.21 \mu\text{g}/\text{m}^3$ ), 0.88% ( $0.14 \mu\text{g}/\text{m}^3$ ), 0.07% ( $0.01 \mu\text{g}/\text{m}^3$ ), and 0.02% ( $0.01 \mu\text{g}/\text{m}^3$ ), respectively. There was no significant difference in the monthly rate between weekdays and weekends, indicating that these pollutants continuously affect the study area. Although many air pollutants contribute to the overall  $\text{PM}_{2.5}$  concentration, oil combustion residues and motor vehicles were found to be the major pollution sources, implying that intensive management is needed to reduce emissions from these sources.

### 3.3. Assessing PMF Model Reliability

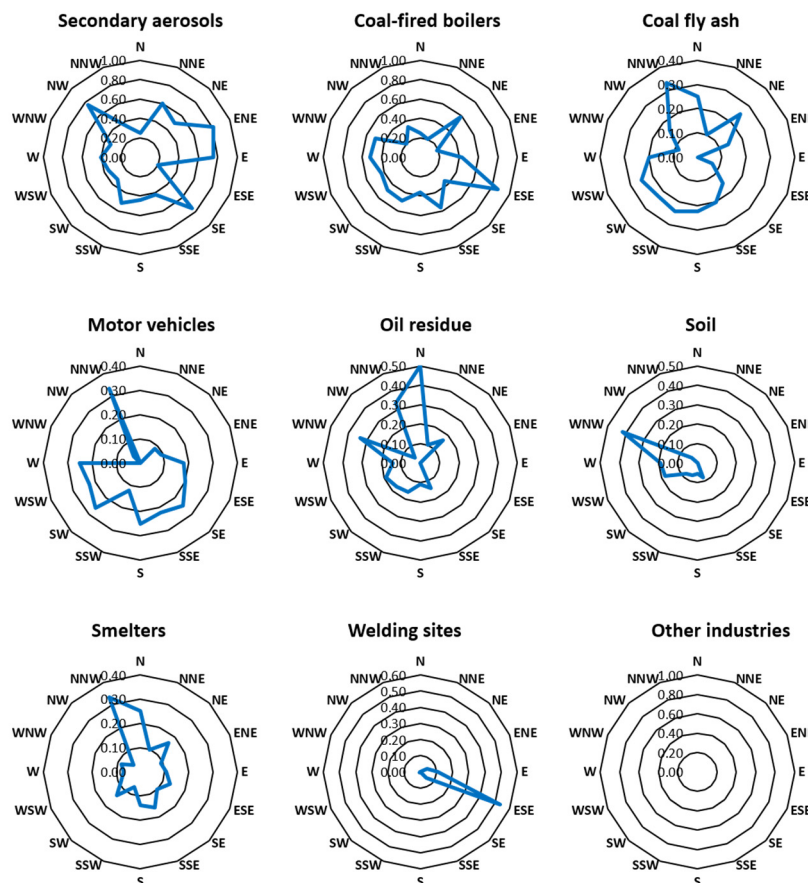
It is important to verify the reliability of results of PMF modeling. This can be completed via correlation analysis of the estimated and actual mass concentrations of secondary aerosols [12]. Figure 7 displays a scatter plot of measured  $\text{PM}_{2.5}$  concentrations and concentrations calculated using PMF modeling during the study period. The coefficient of determination ( $R^2$ ) was 0.872, indicating that the estimated  $\text{PM}_{2.5}$  concentrations can explain 87.2% of the variability in actual concentrations.



**Figure 7.** Mass concentrations of  $\text{PM}_{2.5}$  predicted from PMF modeling versus measured mass concentrations.

### 3.4. Using CPF Modeling to Verify the Source Locations of Pollutants

CPF modeling was used to evaluate the directions from which incoming identified came by combining the contribution rates calculated from the PMF model with detailed weather observation data (AWS; Automatic Weather Station, which automatically transmits or records observations obtained from the measuring instrument.) by region. Figure 8 shows the results of CPF modeling.



**Figure 8.** CPF plot for each source of PM<sub>2.5</sub> identified in the PMF analysis.

Secondary aerosol pollutants largely came from the east and northwest, likely due to vehicles traveling along highways in the northwest and emissions of SO<sub>x</sub> from facilities of various sizes in the east. Additionally, because secondary aerosol pollutants are generated through chemical reactions, they may originate from far away. Therefore, it is possible that some of the aerosols were generated in China.

Pollutants from coal-fired boilers were mostly transported from the southeast. Indeed, a number of industrial worksites, including the Seoul Onsu General Industrial Complex, are located to the southeast of the sampling sites. According to their CPF plot, pollutants from metal-processing facilities and coal fly ash mainly originated in the northwest, likely from emissions from several industrial sites concentrated northwest of the sampling sites. Motor-vehicle-derived pollutants came from most directions, except from the east, in particular the northwest and southwest. The Gyeongin Expressway runs in an east–west direction to the north of the study area, while the First Loop Expressway in the metropolitan region runs in a north–south direction to the west. The two highways also intersect to the northwest.

Pollutants from oil combustion residues arrived mainly from the north. Ojeong General Industrial Complex and many other pollutant-emitting facilities are concentrated to the north of the sampling sites and are considered the source of these pollutants. Soil-derived pollutants are likely to have originated in the northwest. Much construction was underway during the study period, including interior construction, spatial rearrangements,

and the construction and demolition of apartment complexes in the northwest. Yellow dust from Incheon and China are assumed to have added to the amount of soil-derived pollutants, as those two locations are northwest of the study area.

Smelter-derived pollutants were observed to originate mainly from the northwest, where industrial sites with metal-processing facilities and those emitting coal fly ash can be found. Pollutants from welding sites appeared to be transported from the southeast, possibly from several worksites that emit air pollutants. The source locations of pollutants from other industries could not be determined as the threshold value (80th percentile) was not met on any of the measurement dates.

From CPF modeling, it was determined that most pollutants came from the northwest, where traffic congestion on the expressways frequently occurs. This releases a high concentration of pollutants that is added to by air pollutants emitted from industrial sites in the northwest and southeast.

#### 4. Conclusions

As mentioned in the introduction, air quality studies on PMF models are being conducted mainly in large cities such as Seoul and Jeju Island in the Republic of Korea. Even in small cities, such as the research areas in this study, the PMF model should be applied to identify the source of pollution.

In this study, PMS-104 samplers were used to collect 74 PM<sub>2.5</sub> samples at two study sites between 22 April and 3 July 2020, in an effort to identify sources of particulate matter and suggest ways to manage particulate matter concentrations more effectively. A total of 34 chemical components in the collected PM<sub>2.5</sub> samples were analyzed. PMF modeling showed that secondary aerosols contributed the most to PM<sub>2.5</sub> pollution (31.09%, 5.73 µg/m<sup>3</sup>), followed by coal-fired boilers (16.9%, 3.11 µg/m<sup>3</sup>), metal-processing facilities and coal fly ash (11.59%, 2.14 µg/m<sup>3</sup>), motor vehicles (10.52%, 1.94 µg/m<sup>3</sup>), oil combustion residues (10.15%, 1.87 µg/m<sup>3</sup>), soil (7.97%, 1.47 µg/m<sup>3</sup>), smelters (6.4%, 1.18 µg/m<sup>3</sup>), welding sites (4.77%, 0.88 µg/m<sup>3</sup>), and other industries (0.7%, 0.11 µg/m<sup>3</sup>).

The PMF model showed an R<sup>2</sup> value of 87.2. Other papers judged that results are reliable when the R<sup>2</sup> value is greater than 80, and this study provides results similar to those of other studies. Therefore, the results of this study were deemed reliable. CPF modeling was performed to identify the locations of pollution sources and showed that most pollutants came from areas northwest of the sampling sites. This finding is consistent with the presence of major expressways in the northwest. Furthermore, air-pollutant-emitting worksites are concentrated in the southeast. It was confirmed that most of the coal boiler pollutants were transported from the southeast. This was also consistent with the concentration of air pollutant emission businesses in the southeast.

The emission of pollutants with high contribution rates should be actively regulated and their management prioritized. In particular, the contribution rate was high for “motor vehicle pollutants”, a representative artificial pollutant. This research area is implementing a number of policies to reduce diesel vehicle emissions for automobile pollutants. For example: (1) restriction on the operation of emission grade 5 vehicles from 6 a.m. to 9 p.m. on weekdays; (2) removal of old diesel vehicles from road use; (3) notification of the operation restriction system to owners of vehicles that have not taken measures for low pollution; (4) owners of a vehicle who have not taken measures against low pollution should quickly attach a reduction device. In addition specific mitigation measures are needed, such as no-driving-day schemes, the promotion of electric vehicles, and the installation of smog-reduction devices.

Given the brevity of the study period, the identification of pollutants and the determination of pollutant contributions using PMF modeling were limited to one season. Further research is needed to identify additional sources of pollution and pollutant contribution rates in each season. Moreover, the relatively low number of samples was likely to have affected the reliability of the results; thus, future studies should include more samples and identify more chemical components to enable investigations of seasonal pollution and



improve model reliability. The use of potential source contribution function, distributed, and geographic information system models to analyze long-distance pollution sources would also facilitate more effective air quality management and better pollution-related predictions.

The Republic of Korea's air quality standards, which are laxer than those of other countries, should first be strengthened to manage PM<sub>2.5</sub> emission sources more effectively. A legal foundation for applying strict particulate matter standards is necessary to ensure public health [24]. The effect of pollutants on air quality can vary according to pollutant concentration and species; thus, measures that specifically address particular risks are needed, such as the Special Act on Fine Dust Reduction and Management and a scheme to control the total pollutant quantity. Currently, the Republic of Korea has few regional measures for addressing air pollution. However, overall air pollution can only be effectively reduced if measures specific for each region are established, with authority granted to the relevant regional management units.

**Author Contributions:** D.P. and C.Y. designed the study; G.L. was principally responsible for writing the manuscripts; project administration, D.P. and C.Y.; G.L. and M.K. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was conducted with support from the AI-Based Development of Technology to Reduce Exposure of Subway Users to Particulate Matter program RS-2019-KA152306 of the Subway Fine Dust Reduction Technology Development Project of the Ministry of Land.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This research was conducted with support from the AI-Based Development of Technology to Reduce Exposure of Subway Users to Particulate Matter program RS-2019-KA152306 of the Subway Fine Dust Reduction Technology Development Project of the Ministry of Land, Infrastructure and Transportation and APM Engineering Co., Ltd. The authors wish to express appreciation for this support.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest.

## References

1. AirVisual. 2019 *World Air Quality Report*; AirVisual: Goldach, Switzerland, 2019.
2. IQAir. 2020. Available online: <https://www.iqair.com/ko/world-most-polluted-countries> (accessed on 20 June 2021).
3. Kim, E.J. *Legislative Study for the Integrated Management of Air Pollutants*; Korea Legislation Research Institute: Seoul, Republic of Korea, 2019; pp. 1–446. Available online: <http://klri.re.kr:9090/handle/2017.oak/9157?mode=full> (accessed on 29 January 2019).
4. Russell, A.G.; Brunekreef, B. A Focus on Particulate Matter and Health. *Environ. Sci. Technol.* **2009**, *43*, 4620–4625. [[CrossRef](#)] [[PubMed](#)]
5. Jo, J.H.; Kim, E.J. *Analysis of Technology Market Trends Related to Indoor Fine Dust Removal*; Korea Institute of Science and Technology Information: Daejeon, Republic of Korea, 2014; pp. 1–109. Available online: <https://repository.kisti.re.kr/handle/10580/6571> (accessed on 31 December 2014).
6. Zhao, W.; Hopke, P.K. Source apportionment for ambient particles in the San Gorgonio wilderness. *Atmos. Environ.* **2004**, *38*, 5901–5910. [[CrossRef](#)]
7. Heo, J.; Kim, C.; Min, Y.; Kim, H.; Sung, Y.; Kim, J.; Lee, K.; Heo, J. Source Apportionment of PM<sub>10</sub> at Pyeongtaek Area Using Positive Matrix Factorization (PMF) Model. *J. Korean Soc. Atmos. Environ.* **2018**, *34*, 849–864. [[CrossRef](#)]
8. Bucheon. 2020. Available online: <https://www.bucheon.go.kr/site/program/board/basicboard/list?boardtypeid=26716&menuid=148004003004> (accessed on 8 February 2021).
9. Lee, K.-B.; Kim, S.-D.; Kim, D.-S. Ion Compositional Existence Forms of PM<sub>10</sub> in Seoul Area. *J. Korean Soc. Environ. Eng.* **2015**, *37*, 197–203. [[CrossRef](#)]
10. Lee, T.-J.; Huh, J.-B.; Yi, S.-M.; Kim, S.-D.; Kim, D.-S. Estimation of PM<sub>10</sub> source contributions on three cities in the metropolitan area using the PMF model. *J. Korean Soc. Atmos. Environ.* **2009**, *25*, 275–288. [[CrossRef](#)]
11. Nam, B.H. Pattern Classification of PM<sub>10</sub> Sources and Their Mass Contributions to the Indoor Environment. Master's Thesis, Kyunghee University, Yongin, Republic of Korea, 2001; pp. 1–77.

12. Song, X.-H.; Polissar, A.V.; Hopke, P.K. Source of fine particle composition in the north eastern US. *Atmos. Environ.* **2001**, *35*, 5277–5286. [[CrossRef](#)]
13. Begum, B.A.; Kim, E.; Biswas, S.K.; Hopke, P.K. Investigation of sources of atmospheric aerosol at urban and semi-urban areas in Bangladesh. *Atmos. Environ.* **2004**, *38*, 3025–3038. [[CrossRef](#)]
14. Kim, E.; Hopke, P.K. Comparison between conditional probability function and nonparametric regression for fine particle source directions. *Atmos. Environ.* **2004**, *38*, 4667–4673. [[CrossRef](#)]
15. Zhou, L.; Hopke, P.K.; Paatero, P.; Ondov, J.M.; Pancras, J.P.; Pekney, N.J.; Davidson, C.I. Advanced factor analysis for multiple time resolution aerosol composition data. *Atmos. Environ.* **2004**, *38*, 4909–4920. [[CrossRef](#)]
16. Kim, E.; Hopke, P.K.; Edgerton, E.S. Source identification of Atlanta aerosol by positive matrix factorization. *J. Air Waste Manag. Assoc.* **2003**, *53*, 731–739. [[CrossRef](#)] [[PubMed](#)]
17. Chow, J.C.; Liu, C.S.; Cassmassi, J.; Watson, J.G.; Lu, Z.; Pritchett, L.C. A neighborhood-scale study of PM10 source contributions in Rubidox, California. *Atmos. Environ. Part A Gen. Top.* **1992**, *26*, 693–706. [[CrossRef](#)]
18. Chang, Y.S.; Carmichael, G.R.; Kurita, H.; Ueda, H. An investigation of the formation of ambient NH<sub>4</sub>NO<sub>3</sub> aerosol. *Atmos. Environ.* **1986**, *20*, 1969–1977. [[CrossRef](#)]
19. Shim, S.Y. Identification of Ambient PM2.5 Sources with PMF Receptor Model and Estimation of their Source Contribution in the Gwangju Area. Master's Thesis, Chonnam National University, Gwangju, Republic of Korea, 2013; pp. 1–58.
20. Hwang, I.J. Estimation of Quantitative Source Contribution of Ambient PM10 Using the PMF Model. Ph.D. Thesis, Kyunghee University, Yongin, Republic of Korea, 2003; pp. 1–228.
21. Jeong, Y.J. Source Apportionment of PM2.5 in Gyeongsan Using Positive Matrix Factorization. Master's Thesis, Daegu University, Gyeongsan, Republic of Korea, 2014; pp. 1–74.
22. Park, M.B. PM2.5 Characterization and Source Apportionment based on Real-Time Seoul Monitoring Data by PMF Model. Master's Thesis, Kyunghee University, Yongin, Republic of Korea, 2017; pp. 1–103.
23. Oh, M.S. Quantitative Source Apportionment of Airborne Particulate Matter Using an 8-Stage Cascade Impactor in Suwon. Master's Thesis, Kyunghee University, Yongin, Republic of Korea, 2008; pp. 1–132.
24. Lee, G.H. Identification of pollution sources and estimated contribution of fine particulate matter (PM2.5) in the atmosphere using multivariate statistical analysis and PMF Model—Climate change adaption policy. Master's Thesis, Kyunghee University, Yongin, Republic of Korea, 2021; pp. 1–130.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.