

Polychlorinated Biphenyl Congeners in River Sediments: Distribution and Source Identification Using Multivariate Factor Analysis

Ronghu Jin · Seok-Un Park · Jong-Eun Park ·
Jong-Guk Kim

Received: 5 April 2011 / Accepted: 17 October 2011 / Published online: 4 November 2011
© Springer Science+Business Media, LLC 2011

Abstract To assess the level of polychlorinated biphenyl (PCB) contamination and identify their sources, surface sediments were collected from selected locations along Nakdong River, Korea, and analyzed for 209 PCB congeners using high-resolution gas chromatography/high-resolution mass spectroscopy. PCB levels ranged from 0.124 to 79.2 ng/g dry weight (coplanar PCBs 0.295 to 5720 pg/g dry), which were similar to those of three other major rivers (Han, Geum, and Youngsan rivers) in Korea but slightly lower than those in neighboring countries. Regarding homologue composition, tetra-CBs were most abundant in most samples, but some samples with much higher PCBs concentrations had relatively lower proportions of tetra-CBs and higher proportions of penta- to hepta-CBs. To identify the sources of PCBs in sediment

samples, principal component analysis/absolute principal component scores (PCA/APCS), positive matrix factorization (PMF), and multiple linear regression (MLR) were used with the congener composition of aroclors (1242, 1248, 1254, and 1260) and the flue gas of waste incinerators (data obtained from a previous article) as source profiles. Results showed that the three models showed similar source apportionments. Most sediment samples with lower PCB concentrations had higher proportions of incineration-derived materials, and some sediment samples with much higher PCB concentrations had higher proportions of aroclor 1260. This occurred because many industrial facilities, such as landfill leachate-treatment facilities, were gathered around sampling points with high PCB concentrations, and high-chlorinated PCBs are more stable in the elution process of landfill leachate than the incineration process. PCB concentrations estimated by APCS, PMF, and MLR were similar to the measured values with coefficients of determination ranging from 0.77 to 0.99.

Electronic supplementary material The online version of this article (doi:10.1007/s00244-011-9722-7) contains supplementary material, which is available to authorized users.

R. Jin
Department of Chemical Engineering, Yanbian University,
977 Gongyuan Road, Yangi 133002, Jilin, China

S.-U. Park
Quality and Innovation Team, KEPCO Research Institute,
105 Munji-Ro Yuseong-Gu, Daejeon 305-760,
Republic of Korea

J.-E. Park
Department of Environmental Engineering,
Chonbuk National University, 664-14 Dukjin-dong Duckjin-gu,
Jeonju, Jeonbuk 561-756, Republic of Korea

J.-G. Kim (✉)
Center for Chemical Safety Management,
Chonbuk National University, 664-14 Dukjin-dong Duckjin-gu,
Jeonju, Jeonbuk 561-756, Republic of Korea
e-mail: kjongguk@jbnu.ac.kr

Polychlorinated biphenyls (PCBs) have been used extensively in various industries, such as dielectric fluids in capacitors and transformers, resins, wax extenders, flame retardants, dedusting agent, adhesives, inks, and pesticides (Kim et al. 2004; Lee et al. 2001; Loganathan and Lam 2011; UNEP/CHEMICALS 1999). Although the use and production of PCBs has been prohibited due to their significant adverse effects on the environment and human health, pollution is still of great concern in Korea. In particular, liquid wastes, including insulating oils in transformers and capacitors, have been regulated at relatively lower levels (2 mg PCBs/L liquid waste) than in other countries. The National Institute of Environmental Research (NIER) of Korea has conducted annual surveys of

endocrine-disrupting compounds (including PCBs) in environmental media, such as air, soil, and water, etc. However, the surveys were performed less intensively (focused on other environmental matrices, such as ambient air, soil, etc.) on sediments, although sediment is an important environmental matrix as the final destination of environmental pollutants and food supplier to organisms in the food chain (NIER 2007).

Nakdong River is the second largest river in Korea, and its basin covers approximately 74% of the southeastern part of Korea. Its total length is 521.5 km, and the basin area is 23,817 km². Approximately 1.32 billion tons of water are used annually by industries, with approximately 381 million tons used per day for residential purposes and approximately 2.86 billion used annually for agricultural purposes. In addition, a wide variety of flora and fauna inhabit Nakdong River; therefore, its role in the environment around the southeastern area of Korea is crucial. The pollution of Nakdong River has been a long-term concern because various sources of pollution, such as industrial or agricultural wastewater-treatment facilities, landfills, and livestock night soil-treatment facilities, etc., exist along the river. In particular, the river-restoration project is now proceeding in four major rivers (Han, Nakdong, Geum, and Youngsan rivers) in Korea. Therefore, the hydrospheric environment of Nakdong River has been disturbed and will be under the potential influence of accumulated pollutants. As with many pollutants, PCBs emitted from sources also tend to be adsorbed and accumulated by soils and sediments, depending on the surface area and properties of the particulates, because they are hydrophobic and denser than water (Donna and Ralph 1996; Bonifazi et al. 1997; Kodavanti et al. 2008). In addition, rain, floods, and river discharges can result in the accumulation of PCBs in sediments (Smith et al. 1988).

Therefore, assessment for PCB contamination levels and their source(s) are essential to prevent further contamination and to protect the valuable living resources of Nakdong River. The objective of this study was to determine congener concentrations in sediments from upstream, midstream, downstream, and marsh areas and to identify PCB sources using multivariate factor analysis.

Materials and Methods

Sample Collection

Sampling points were selected by considering tributaries, vicinity of pollution sources, and availability of sample collection. Sample collection was performed intensively from May to June 2006. A total of 21 sampling points were selected and distinguished by their positions along the river

as well as the density and location of pollution sources: upstream (6 points), midstream (7 points), downstream (5 points), and marsh (3 points). Three duplicate samples were collected at each point. Surficial sediment was collected using a grab-type sampler at a sampling depth of approximately 2 cm from the surface. Collected samples were preserved in wide-mouth glass bottles (1,000-mL volume; precleaned using acetone and *n*-hexane), protected from light, and stored at -4°C , without disturbing. The sampling points are shown in Fig. 1.

Sample Pretreatment and Instrumental Analysis

Collected samples were air-dried indoors, and their water content was measured periodically (approximately every 2 to 3 days). Dried samples were pulverized and then sieved using a 2-mm sieve. Extraction of samples was performed for >16 h using a Soxhlet extractor and toluene as the solvent. Before extraction, a surrogate standard (EC-4977; Cambridge Isotope Laboratories) was injected. The solvent of the extracts was completely changed from toluene to *n*-hexane, and then sulfuric acid treatment was performed until the extract was colorless. Multilayer column cleanup was then performed. Adsorbents were filled into a glass apparatus from the bottom as follows: 1 g of neutral silica gel (activated at 130°C for >19 h), 2 g of 30% potassium hydroxide-coated silica gel, 1 g of neutral silica gel (activated at 130°C for >19 h), 4 g of 44% sulfuric acid-coated silica gel, 1 g of neutral silica gel (activated at 130°C for >19 h), 1 g of 10% silver nitrate-coated silica gel, and 2 g of anhydrous sodium sulfate. The adsorbent-filled column was precleaned using 50 mL of *n*-hexane before loading the extracts, and 120 mL of *n*-hexane was used for elution. After multilayer column, copper column was performed to remove sulfur, which is easily contained under anaerobic conditions. Then alumina column [activated at 190°C for >19 h (14 g)] was finally performed. Before elution, the alumina column was precleaned using 50 mL of *n*-hexane. The prefraction (90 mL of *n*-hexane) was discarded, and the postfraction (30 mL of 50% dichloromethane/*n*-hexane) was used for instrumental analysis. After injection of an internal standard (EC-4979; Cambridge Isotope Laboratories), high-resolution gas chromatography/high-resolution mass spectroscopy (HRGC/HRMS) analysis was performed. HRGC/HRMS analysis was performed in the electron impact/selected ion monitoring mode, and the resolution was >10,000. The column was a DB-5MS capillary column (60 m \times 0.25 mm \times 0.25 μm ; J&W Scientific). All reagents and organic solvents used in this study were of PCB or pesticide analytical grade. The detailed procedures for pretreatment of the samples and instrumental analyses were based on the Analytical Methods of Endocrine Disrupting Chemicals (NIER 2005) and United

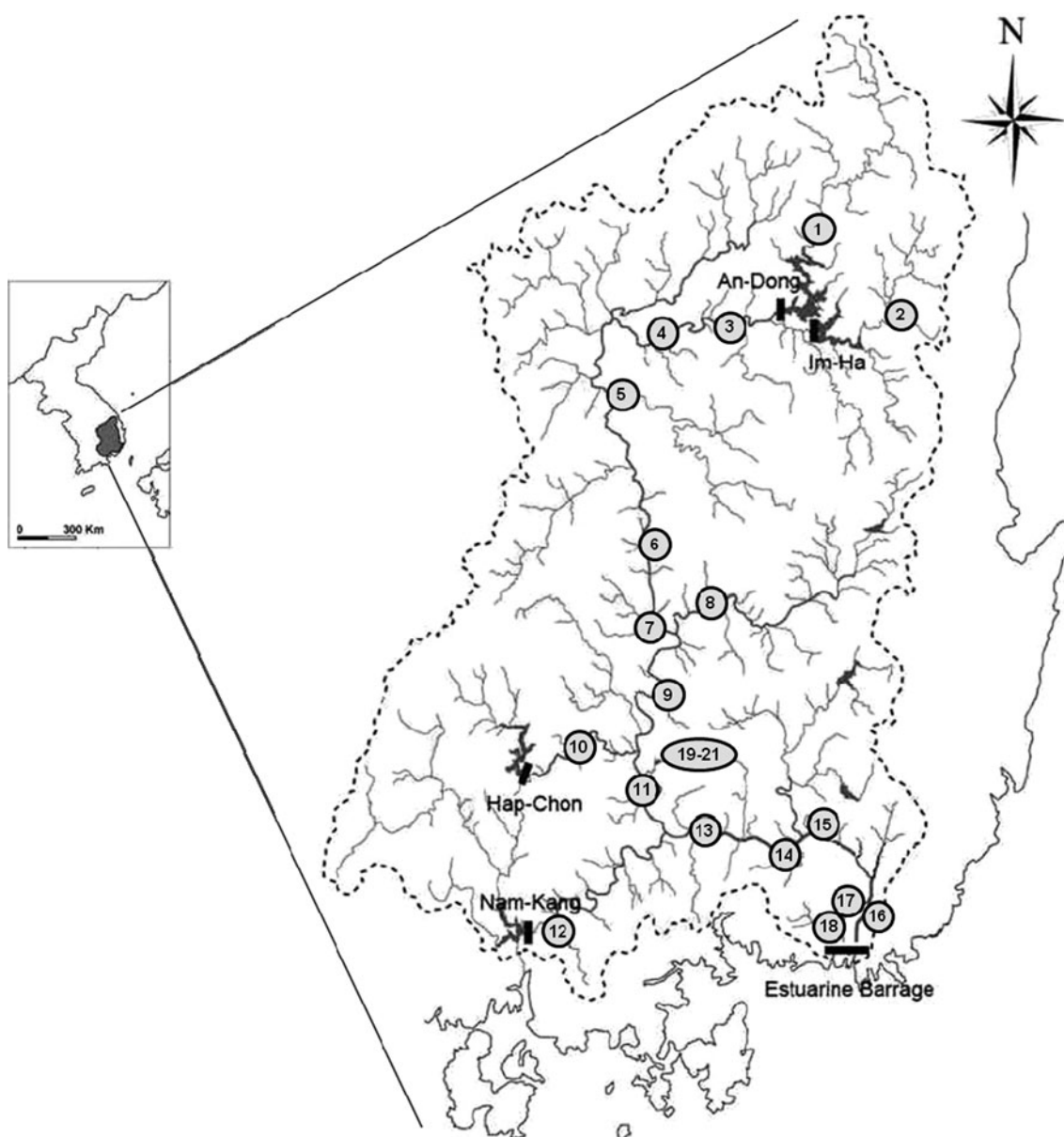


Fig. 1 Sampling points along the Nakdong River. Samples 1–6 = upstream; samples 7–13 = midstream; samples 14–18 = downstream; and samples 19–21 = marsh

States Environmental Protection Agency (USEPA) method 1668A. The toxic equivalent (TEQ) concentration was calculated using toxic equivalency factors of the World Health Organization (1998).

Multivariate Factor Analysis

The basic assumption for receptor models is that the concentration of a pollutant at a receptor for a given sample is the linear sum of the products of the emission profile and the contribution of sources; samples are well mixed with chemicals from different sources, and chemicals are relatively stable during transport from the emission source to

the receptor site. PCA, the most widely used tool in environmental science, decreases the number of variables while retaining as much of the original information as possible. In general, each factor extracted from PCA is associated with a source and characterized by its most representative chemical(s) (or PCB congener[s] in this study). Principal component analysis/absolute principal component scores (PCA/APCS), the details of which have been described elsewhere (Guo et al. 2004b; Thurston and Spengler 1985), was used in this study.

Positive matrix factorization (PMF) is multivariate factor analysis, with nonnegative constraints, that decomposes a matrix of sample data into two matrices (factor

contributions and factor profiles) using oblique solutions; in other words, the results are constrained so that no sample can have a negative source contribution. This method defines the concentration matrix of chemical species measured at receptor sites as the product of source composition and contribution factor matrices with a residual matrix. PMF allows each data point to be weighed individually. This feature allows the analyst to adjust the influence of each data point depending on confidence in the measurement. For details on PMF, please refer to Paatero (1997) and Paatero and Tapper (1994).

In contrast, unlike PCA/APCS or PMF, previous information (congener composition of source) is necessary for multiple linear regression (MLR). However, if source profiles have already been obtained, MLR can be a useful method for source identification of an individual sample, even with only few data. The detailed procedure of MLR used in this study has previously been described in Kim (2004) and Park et al. (2009).

Data Preparation for Multivariate Factor Analysis

Using all 209 congeners could be useful for source identification, but there were too many for both the PCA/APCS and PMF models to compare with the number of sediment samples in this study. Thus, some congeners were selected from all possible congeners for both PCA/APCs and PMF, with the others being excluded from these receptor models. In this study, two criteria were applied for congener selection: detection magnitude in the source profile (i.e. predominance in congener composition of source profiles) and distinction from other sources.

According to these two criteria, 37 congeners (25 congener groups by coelution) (#10&4, #9&7, #14, #11, #12&13, #15, #17, #72&71, #84, #97&86, #85&120, #151, #135&144, #141, #128, #179, #182&187, #183, #174&181, #177, #180&193, #170&190, #199, #203&196, and #194) were selected. Sixteen samples (e.g., S3-2, S3-3, S5-1, S5-2, S5-3, S12-1, S12-2, S19-1, S19-2, S19-3, S20-1, S20-2, S20-3, S21-1, S21-2, and S21-3) that exhibited too many missing values (refer to the Supplementary information for details on missing values) were excluded from the raw data set. Therefore, the final data set comprised 25 rows (congener groups) and 47 columns (sediment samples).

Missing values or data below the detection limits may cause uncertainty in source identification. However, these can be changed to another statistical value, e.g., data below detection limits can be changed to half the detection limits, and missing values can be replaced with the average of the other samples' data. The purpose of this was not only to retain the originality of the raw data set but to also prevent invalid statistical interpretation. In this study, the ratio of

missing values or data below the detection limits in the selected data set was <10%, and these were substituted for other statistical values using the method of Polissar et al. (1998). Regarding the final data set for multivariate factor analysis and the detection limits for individual PCB congeners, please refer to Tables S1 and S2 of the Supplementary information.

Results and Discussion

Concentration Distribution of PCBs in Sediment Samples

The levels of total PCBs ranged from 0.124 to 79.2 ng/g dry weight [average 6.88 (median 0.335)]. In terms of coplanar PCBs (12 congeners: #81, #77, #123, #118, #114, #105, #126, #167, #156, #157, #169, and #189), the concentrations ranged from 0.295 to 5,720 pg/g dry weight [average 518 (median 13.9)], and TEQ values ranged from 0.000251 to 3.18 pg WHO-TEQ/g dry weight [average 0.404 (median: 0.0196)]. Total recovery rates were approximately 38 to 119%, but mono- and di-CBs had relatively lower ranges (approximately 38 to 56%), and the others (tri- to deca-CBs) had higher recovery rates (approximately 76 to 119%).

According to the annual survey on EDCs in environmental media by the NIER in Korea (NIER 2007), the levels of coplanar PCBs in sediments from four major rivers ranged from not detectable (ND) to 5,060 pg/g (average 938), which was similar to the results obtained in this study. The concentration distribution of PCBs in sediments from rivers in Korea has been reported previously. Koh et al. (2004) and Kim et al. (2009) analyzed PCB concentrations in sediment samples from the Hyeongsan and Han rivers, and the results ranged from 1.0 to 170 ng/g dry weight (14 congeners) and from 0.415 to 4.53 ng/g dry weight (12 coplanar PCBs), respectively. Jeong et al. (2001) investigated PCB levels in Nakdong River sediments, and levels ranged from 1.1 to 141 ng/g dry weight, which was higher than found in this study, even though only 43 congeners were analyzed. However, note that the sampling date was approximately 7 years earlier (May 1999) than this study. In addition, Ren et al. (2009) reported levels of coplanar PCBs in East River sediment in China that ranged from 48 to 270 pg/g dry weight; therefore, they were lower than those found in this study. In addition, Liu et al. (2007), Zhang et al. (2010), and Hung et al. (2006) reported higher levels of PCBs than those reported in this study: 0.775 to 154 ng/g dry weight (18 congeners in Haihe River, China), 44.3 to 154 ng/g dry (18 congeners in Dagu Drainage River, China), 1.85 to 1,076 ng/g dry (18 congeners in Liaohe River, China), and

ND to 83.9 ng/g dry (85 congeners in Danshui River, Taiwan). Generally, PCB concentration distributions in Korean river sediments are regarded as being slightly lower than those in neighboring countries. Table 1 lists the PCB concentrations found in this study and in neighboring countries.

In terms of homologue composition, tri- to hexa-CBs predominated, with tetra-CBs having the highest concentrations. In particular, sample S-8, which showed the highest PCB concentration (48.5 to 79.2 ng/g dry weight), had higher proportions of hexa- and hepta-CBs than the other samples. However, the isomer composition of each homologue was similar for almost all samples, with high-chlorinated biphenyls (hexa- to octa-CBs) being more prominent. Similar results have been reported in previous articles on atmospheric and soil samples (Kim 2004; Park et al. 2009).

PCA/APCS

Before performing PCA, information matching each extracted factor to a specific source was needed. Therefore, the congener composition of each source (e.g. aroclors 1242, 1248, 1254, 1260, and incineration) obtained from the USEPA Web site (http://www.epa.gov/toxteam/pcb/download/aroclor_frame.xls) was used. There is no proof that aroclor is the only PCB product used in Korea, but aroclors (1242, 1248, 1254, and 1260) are being used as qualification and quantification standards in the Korean Waste Official Test Method. The compositions of aroclors and kanechlors (PCB products made in Japan) are similar. Under these circumstances, inputting aroclors as a PCBs source in this study was regarded as reasonable. In contrast, some previous articles have reported that PCBs could also be emitted by way of thermal processes involving chlorine-containing materials or combustion byproducts, such as fuel combustion, waste incineration, and iron oxidation, etc. (Ikonomou et al. 2002; Ishikawa et al. 2002). Shin

et al. (2006) reported the concentration and congener patterns of all 209 PCBs in flue gases of industrial and municipal incinerators in Korea. The composition of incineration data was obtained from this reference. Attempts were made in this study to distinguish the 209 PCBs into 2 or 3 groups by way of hierarchical cluster analysis (HCA) to avoid unreasonable exclusion of incineration from the independent variables using statistical calculations of the receptor models. As a result of HCA, 2 or 3 source profiles for incineration were obtained, but the correlation of these 2 or 3 groups was high, with almost no difference from the apportioned results: Mean values came from a set of 15 incineration data. Thus, the average of 15 incineration data was used as an independent variable. The fraction composition of each source is shown in Fig. 2.

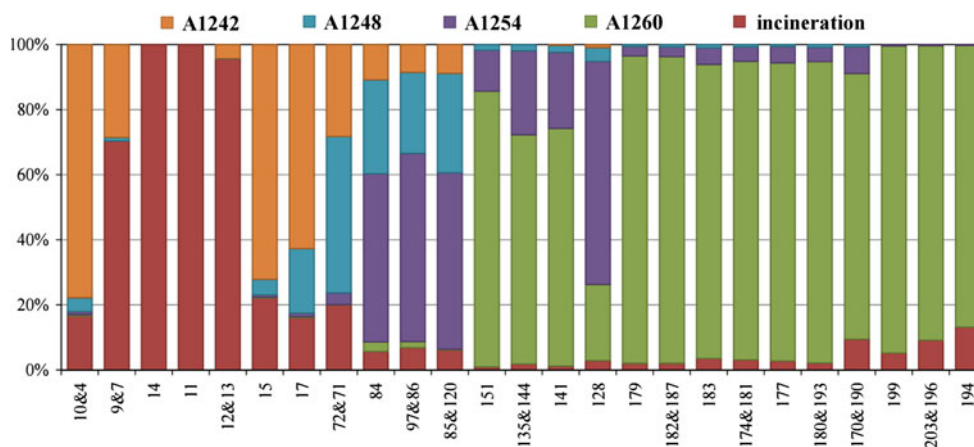
PCBs #9&7, #14, #11, and #12&13 had higher proportions in incineration material, and #10&4, #15, and #17 were predominant in aroclor 1242. PCB #72&71 had higher proportion in aroclor 1248, and PCBs #84, #97&86, and #85&120 were predominant in aroclor 1254. In addition, most high-chlorinated PCBs (such as #151, #135&144, #141, #179, #182&187, #183, #174&181, #177, #180&193, #170&190, #199, #203&196, and #194) had higher ratios in aroclor 1260.

When conducting PCA using data sets with 25 rows (congener groups) and 47 columns (sediment samples), a relatively obscure result for the source identification was obtained: The regression coefficients of some factors were negative; there were seldom-extracted factors that matched source profiles of aroclors and incineration; and these factors showed a merged pattern of plural aroclors, if any. Consequently, the coefficient of determination (including the changed coefficient of determination) of the regression analysis was low, and the significance level of the regression coefficients for each extracted factor was high ($p > 0.05$). Some investigators, such as Guo et al. (2004a), Ito et al. (2004), and Park et al. (2010), classified a raw data set into two or three groups using sampling location or sampling

Table 1 PCBs concentrations found in this study and in other river sediments of neighboring countries

Region	Sampling date	PCBs concentration (ng/g dry)	Coplanar-PCBs (ng/g dry)	No. of PCB congeners	References
Nakdong River, Korea	May–June 2006	0.124–79.2	0.295–5720	209	This study
Nakdong River, Korea	May 1999	1.1–141	–	43	Jeong et al. (2001)
Masan Bay, Korea	–	1.24–41.4	–	22	Hong et al. (2003)
Hyeongsan River, Korea	February 2001	1.0–170	–	14	Koh et al. (2004)
Han River, Korea	November 2005	–	41.5–4530	–	Kim et al. (2009)
East River, China	April 2007	–	48–270	–	Ren et al. (2009)
Haihe River, China	July 2003	0.775–154	–	18	Liu et al. (2007)
Dagu Drainage River, China	July 2003	44.3–154	–	18	
Liaohe River, China	July 2004	1.85–1076	–	18	Zhang et al. (2010)
Danshui River, Taiwan	November 2004	ND–83.9	–	85	Hung et al. (2006)

Fig. 2 Fraction composition of each source with the 25 congener groups



period. Grouping a raw data set in this way can be useful, i.e., in the case where obvious results cannot be obtained from a raw data set as mentioned previously. In this study, however, applying those criteria (sampling location or sampling period) was regarded as unreasonable to some extent because the sampling points were scattered around the river, and the sampling period was relatively short.

In contrast, separated homologue patterns were observed with some samples as mentioned previously. Most samples had the highest proportions of tetra-CBs in their homologue compositions, but some had relatively low proportions of tetra-PCBs but higher proportions of penta- to hepta-CBs. When considering the similarity between the isomer compositions for each homologue, regardless of the total PCB concentration and the hypothesis for a receptor model, i.e., “the behavior of chemicals are relatively stable during the transport from the emission source to the receptor site,” the difference in the homologue composition may imply different source-related conditions (for example, the existence of an unknown source) (Kim 2004). In addition, some samples had extremely high total concentrations compared with others. Because they could be outliers by way of statistical interpretation with a raw data set, the result could be affected by these outliers if not grouped separately or excluded. Thus, our raw data set was classified into two groups by considering both the homologue composition and total PCB concentrations of 25 congener groups; thereafter, one-way multivariate analysis of variance was conducted, and a statistical difference was found between the two groups: Wilks’ lambda (λ) was 0.175, and the significance level (p) was <0.05 . It was thought that the result of the source identification may become clearer by grouping the raw data set; therefore, further receptor modeling was separately conducted using each group. Samples no. S1, S2, S3, S4-1, S6, S7, S10, S11, S12, S13, S14, S15, and S16 [$n = 33$ (including duplicates)] were classified as group 1, and sample S4-2, S4-3, S8, S9, S17, and S18 [$n = 14$ (including duplicates)] were classified as group 2, in which the samples contained

relatively lower proportions of tetra-CBs and higher proportions of penta- to hepta-CBs.

SPSS 12.0 (SPSS Inc.) was used for PCA, in which factors with Eigenvalue >1.0 were extracted. The axis of the extracted factors was rotated using VARIMAX method, which is a commonly used orthogonal rotation method. Table 2 lists the rotated factor loadings of each factor from PCA, which were compared with the fraction composition of each source. Four factors were extracted, accounting for 84.97 and 95.82% of the variances in groups 1 and 2, respectively. Factor loadings >0.55 are highlighted in boldface; these were compared with the fraction profile in Fig. 2 to identify each extracted source.

Group 1

Factor 1 (56.09% of the total variance) showed high correlations for high-chlorinated biphenyls (#84, #85&120, #151, #135&144, #141, #128, #179, #182&187, #183, #174&181, #177, #180&193, #170&190, #199, 203&196, and #194) and indicated a mixed profile for aroclors 1254 and 1260. Factor 2 (15.01% of the total variance) was highly loaded in PCBs #11, #12&13, #15, and #17 and also showed a favorable correlation with PCB #14. Therefore, this factor was identified as incineration material. Factor 3 (7.32% of the total variance) showed good loadings for PCBs #10&4, #9&7, and #11, indicative of aroclor 1242. Factor 4 showed good correlation with PCBs #10&4, #72&71, and #97&86; in particular, the loading value of PCB #72&71 was highest among the four factors. Thus, factor 4 was matched to aroclor 1248.

Group 2

When conducting PCA with group 2 sample data, a relatively unfavorable result was obtained, even though group 2 was grouped separately from the raw data set. The proportion of the factor matched to a mixed form of aroclors 1254 and 1260 was negative for most samples in group 2,

Table 2 Rotated factor loadings for each extracted factor using PCA

Congener group	Group 1				Group 2			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
10&4	−0.14	0.26	0.55	0.72		0.32	0.38	0.68
9&7	0.01	0.29	0.78	0.08	0.36	0.06	0.91	−0.02
14	−0.44	0.47	−0.27	0.27	0.23	0.12	0.90	0.30
11	0.23	0.65	0.57	−0.01	−0.29	0.14	0.92	0.06
12&13	0.18	0.71	0.48	−0.18	0.19	−0.05	0.96	−0.08
15	0.21	0.89	0.14	0.04	0.84	−0.19	0.28	−0.41
17	0.07	0.93	0.13	−0.04	0.25	0.48	0.74	−0.15
72&71	−0.10	0.05	−0.07	0.57	0.36	0.49	−0.20	0.73
84	0.84	0.36	−0.18	−0.01	0.26	0.91	0.11	0.27
97&86	−0.07	−0.18	0.05	0.66	0.19	0.93	0.08	0.30
85&120	0.74	0.63	0.04	−0.01	−0.05	0.79	0.32	−0.34
151	0.96	0.16	−0.07	−0.05	0.76	0.64	0.14	0.05
135&144	0.94	0.21	−0.04	−0.03	0.55	0.81	0.13	0.14
141	0.96	0.17	−0.04	−0.02	0.60	0.77	0.11	0.15
128	0.86	0.34	0.16	−0.02	0.26	0.88	0.01	0.39
179	0.98	−0.03	−0.02	−0.13	0.97	0.16	0.14	0.04
182&187	0.94	−0.01	0.11	−0.17	0.91	0.26	0.13	0.27
183	0.98	0.03	−0.06	−0.02	0.91	0.26	0.02	0.24
174&181	0.99	0.07	0.06	−0.08	0.91	0.28	0.08	0.27
177	0.99	0.06	0.06	−0.08	0.90	0.32	0.01	0.27
180&193	0.99	0.05	0.05	−0.09	0.98	0.15	0.05	0.11
170&190	0.97	0.12	0.09	−0.03	0.92	0.30	0.01	0.18
199	0.83	−0.03	0.23	−0.27	0.90	0.25	0.30	0.03
203&196	0.90	0.02	0.18	−0.24	0.95	0.17	0.21	−0.07
194	0.89	0.11	0.24	−0.23	0.98	0.15	0.09	0.01
Eigenvalue	14.02	3.75	1.83	1.64	11.37	5.95	4.57	2.07
% Variance	56.09	15.01	7.32	6.55	45.46	23.78	18.29	8.29
Cumulative variance	56.09	71.10	78.42	84.97	45.46	69.24	87.53	95.82

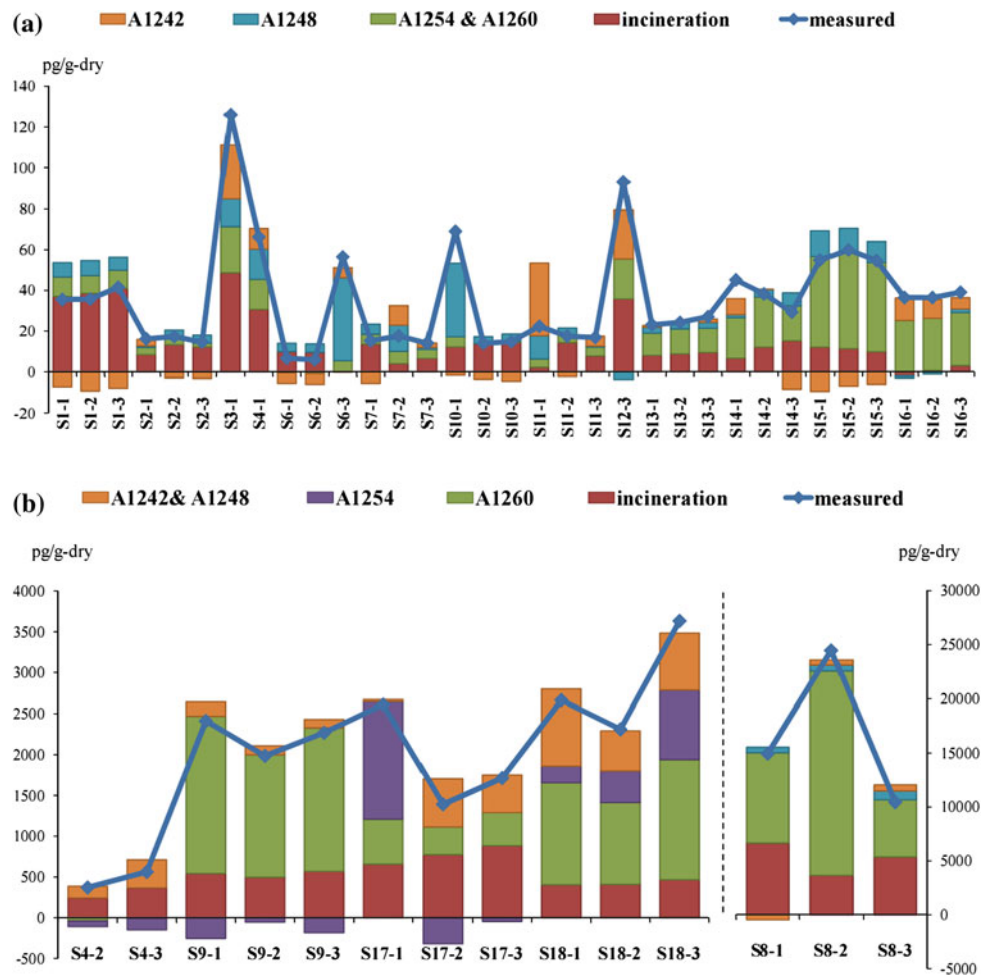
Note: VARIMAX rotation method was used. Factor loadings >0.55 are highlighted in bold face. In the case of group 2, factor loadings of PCA without the outliers (samples S8-1, S8-2 and S8-3) are presented

without the exceptions of samples S8-1, S8-2, and S8-3. The contribution of incineration was too high compared with other models' results (described later in the text), which may have been due to outliers (samples S8-1, S8-2, and S8-3), for which the total concentration was much higher than for the other samples in group 2. In the PCA, the variables were equally weighted for the extraction of factors, which resulted in a relatively huge negative contribution when rotated orthogonally because this maximizes the most correlated factor (Park et al. 2010; Sofowote et al. 2008). Thus, PCA was performed without these outliers, and the results are listed in Table 1. Factor 1 was highly loaded in congeners of aroclor 1260, and factor 2 (high correlation in PCBs #84, #97&86, and #85&120) was matched to aroclor 1254. Factor 3 showed a pattern of incineration, and factor 4 showed a mixed form of aroclors 1242 and 1248.

To apportion the extracted sources to the total PCB concentration in each sediment sample, regression analysis was performed using APCS. The APCS value was calculated by subtracting the component score of the true zero value from the component score of each sediment sample (Guo et al. 2004a). The apportionment result is shown in Fig. 3.

Generally, the proportion of incineration was higher than of other sources in group 1. However, in terms of samples located downstream, the incineration proportion became less, and those of aroclors 1254 & 1260 became relatively higher. Some samples (samples S6-3, S10-1, and S11-1) showed higher contributions of aroclors 1242 and 1248, but their overall proportions were insignificant. Even though PCA for group 2 was performed without samples S8-1, S8-2, and S8-3, their apportionment results are also depicted in Fig. 3 for the sake of comparison with the

Fig. 3 Source apportionment of PCBs in each sediment sample from Nakdong River using PCA/APCS: **a** group 1, **b** group 2



results from the other models. Apportionment results for samples S8-1, S8-2, and S8-3 were determined from the result of PCA with group 2, which included samples S8-1, S8-2, and S8-3.

PMF

One of the many advantages of PMF is that each datum can be weighted individually by its PMF uncertainty (Sofowote et al. 2008). Thus, unlike with PCA/APCS or MLR, PMF needs an uncertainty data set in addition to the concentration data for each sample. There are various methods for calculating PMF uncertainty (Nicolas et al. 2008; Polissar et al. 1998; Sofowote et al. 2008; Shelly et al. 2002), but of the many formulas, the easiest and simplest method is to use the analytical or method uncertainty. However, due to the absence of analytical uncertainty data, the method of Park et al. (2010) was used in this study. Park et al. (2010) down-weighted some congeners as weak when the scaled

residuals (automatically obtained parameter when carrying out PMF) fell outside -2 and $+2$; therefore, their uncertainties were multiplied by 3 (Norris et al. 2008). The input data of the PMF uncertainty is shown in Table S3 of the Supplementary information.

With PMF, the most important step is determining the number of factors. A trial-and-error approach, using a different number of factors, is usually used to determine the most optimal number of factors. The Q -value, or maxima individual column mean (IM) and standard deviation (IS), can be used (Lee et al. 1999; Paatero 1997; Paatero and Tapper 1994). As a result of examining the appropriate number of factors from the Q -value, IM, and IS, the optimal number was 4 to 5 for both groups 1 and 2. After setting up the number of factors to either 4 or 5, the PMF was rerun with a different multistage (seed number 1 to 15). The standard deviation of each factor's proportion from the 15 multistages was negligible compared with the average value. Thus, only the results of the apportionment with a random seed are presented herein.

Before apportioning each factor, F_{peak} rotation was performed by exploring F_{peak} values of -1.0 to $+1.0$ (by 0.1). A positive F_{peak} value sharpens the factor profile and smears the factor contribution; a negative value sharpens the factor contribution and smears the factor profile (Norris et al. 2008). In this study, the most appropriate result was obtained for an F_{peak} value of 0.1 . Note that F_{peak} rotation is not always needed for PMF because a unique solution can be obtained, albeit rarely. The factor profiles extracted using PMF were compared with each source's fraction composition shown in Fig. 2. EPA PMF 3.0 (U.S. EPA) was used in this study, and the source apportionment results are shown in Fig. 4.

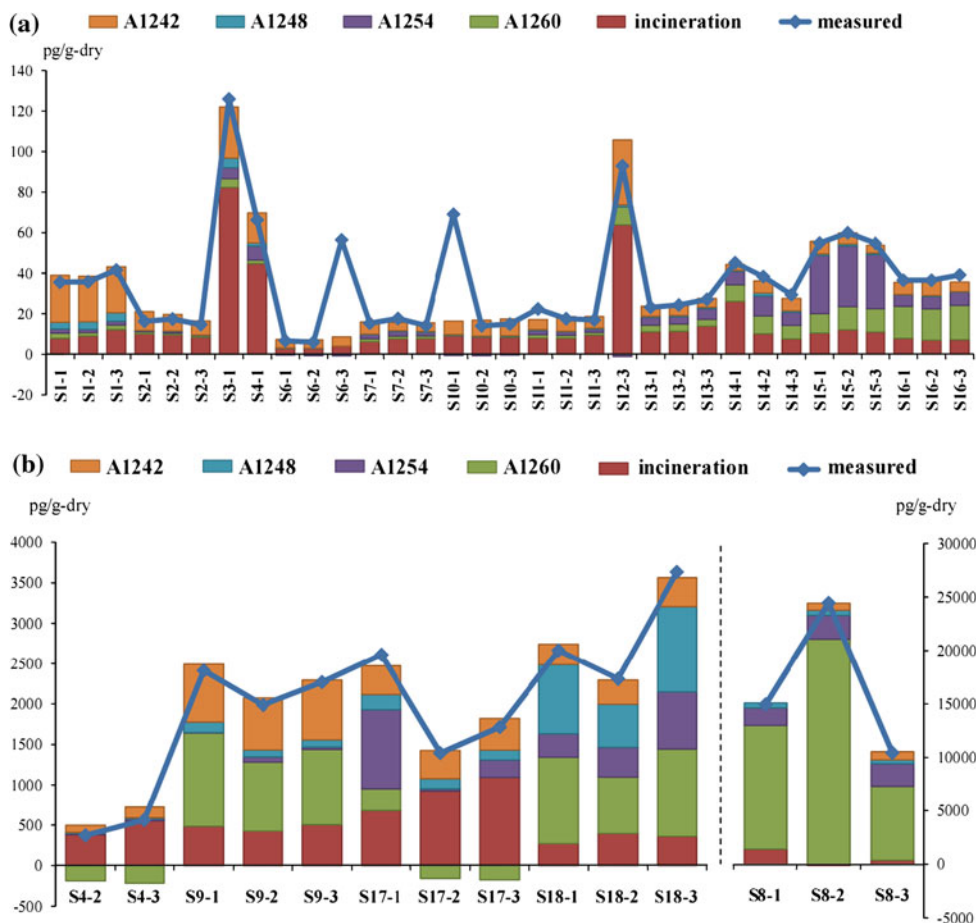
Figure 4 shows the results with only five factors. When setting the number of factors to four, some aroclors could not be distinguished, as with PCA/APCS. In group 1, aroclors 1242 and 1248 were not separated, and a factor with a mixed pattern of aroclors 1248 and 1254 was obtained in group 2. However, they were separated favorably by way of PMF with five factors. The proportion of each source was almost the same as that with the PMF with four factors. Samples S6-3 and S10-1 had relatively larger

residuals between the estimated and measured values, which showed a much higher proportion of aroclor 1248 than the other samples in PCA/APCS. In the case of group 2, PMF showed relatively more stable results than PCA/APCS. It appears that PMF tends to be less affected by outliers (samples S8-1, S8-2, and S8-3) and thus gave favorable apportionment results without excluding the outliers from group 2. Overall, the apportionment results using PMF showed no large differences from those using PCA/APCS. Incineration was predominant in group 1, and its proportion became less in the downstream samples.

MLR

The similarities in the isomer composition in each homologue, regardless of the concentration, should be acutely considered because this may mean that the behavior of PCBs in an environmental matrix depends on to which homologue the congener belongs. Initially, MLR was performed separately for each homologue. Mono- and di-CBs were excluded due to analytical uncertainty from the relatively low recovery rates. In the case of tri-CBs, the

Fig. 4 Source apportionment of PCB in each sediment sample from Nakdong River using PMF: **a** group 1, **b** group 2

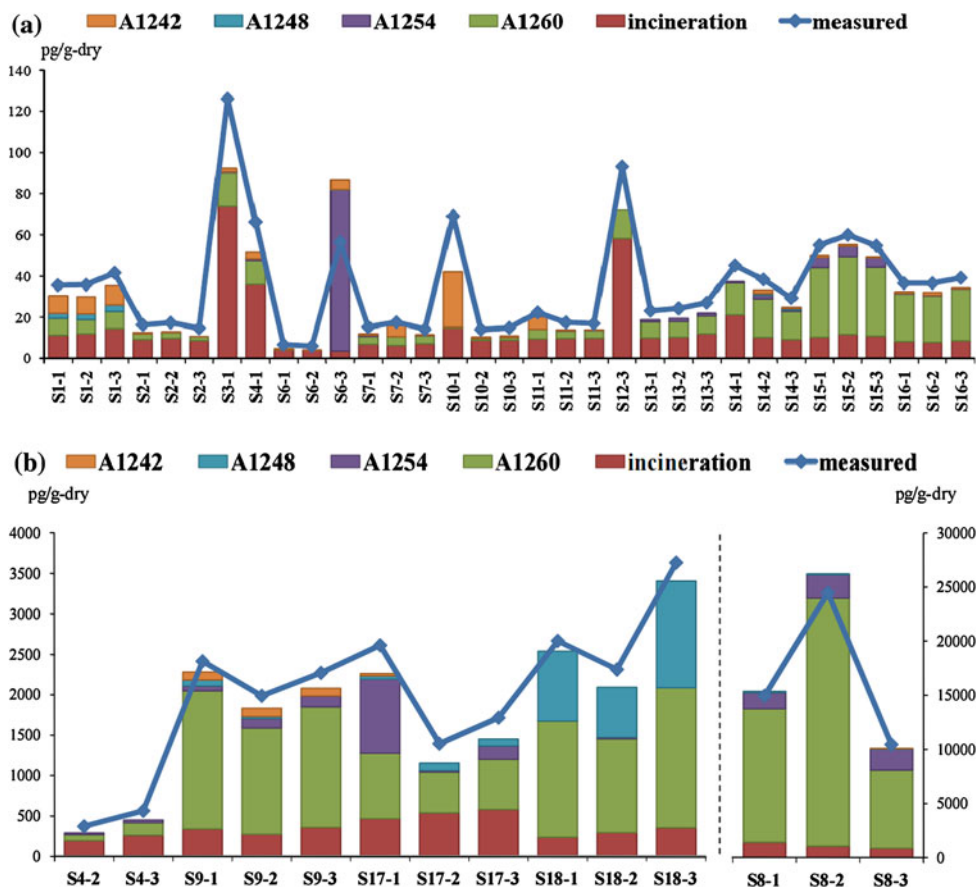


isomer composition of each source was too well correlated with one another; therefore, the MLR results (regression coefficient value of each source) had no statistical significance due to multicollinearity. Therefore, MLR was performed with the data for tetra- to deca-CBs. The isomer compositions from the USEPA and from Shin et al. (2006) were employed in this study. In MLR, the input of all independent variables can cause a negative regression coefficient for some sources (Park et al. 2009). Thus, the stepwise method was also used to input independent variables, with the most appropriate regression coefficient selected by comparison with the results of the other methods. In addition, MLR was also conducted by selecting 25 congener groups. As a result, these two kinds of MLR showed similar features; therefore, only the MLR results with the 25 congener groups is presented in Fig. 5. Aroclor 1260 predominated in the samples from group 2, but the proportion of incineration was relatively higher than other sources in group 1. The proportion of aroclor 1254 was much higher in sample S6-3 as was aroclor 1242 in sample S10-1. Generally, source apportionment by way of MLR was similar to that of PCA/APCS and PMF.

Comparison of Receptor Models' Results

Some samples had negative contributions in PCA/APCS, but their magnitude was negligible. Overall, the source contribution estimates of the three receptor models were similar. Incineration predominated in group 1; the mean contribution was 42.1% (38.6% by PCA/APCS, 42.2% by PMF, and 45.5% by MLR). In contrast, aroclor 1260 was shown to be the main source in group 2; the mean contribution was 66.3% (61.3% by PCA/APCS, 62.8% by PMF, and 74.8% by MLR). The mean contribution of incineration in group 2 was 16.9% (29.7% by PCA/APCS, 10.9% by PMF, and 10.2% by MLR). There was no evidence that a specific aroclor was used in any specific industry in Korea. PCBs have been detected as both singular and mixed forms in transformer oil, which is regarded as the main use of PCB products. Thus, it would not be appropriate to discuss the proportion of each aroclor separately. The source contribution to total PCB concentration in sediment samples is shown in Fig. 6. The results of the two groups are depicted separately. The source contribution of all 47 sediment samples was almost the same as that of

Fig. 5 Source apportionment of PCBs in each sediment sample from Nakdong River using MLR: **a** group 1, **b** group 2. Samples S8-1, S8-2, and S8-3 are depicted separately for comparison with the results of PCA/APCS and PMF



group 2. Note that the proportion of each source in group 2 was calculated by including samples S8-1, S8-2, and S8-3.

Correlation between the measured PCB concentrations and the estimated values by each receptor model was high (values of 0.77 to 0.99; see Fig. S1 in the Supplementary information). In PMF for group 1, the coefficient of determination (0.77) was lower than that for PCA/APCS and MLR due to the inclusion of samples S6-3 and S10-1. These samples exhibited large differences between their measured and estimated values in PMF. When these samples were excluded from group 1, the coefficient of determination was 0.99. In the case of group 2, the coefficients of determination ranged between 0.99 and 1.00, which appeared to happen because the concentration of outliers (samples S8-1, S8-2, and S8-3) was so much higher than for the other samples; the results, with the exclusion of the outliers, are shown in Fig. S1 in the Supplementary information.

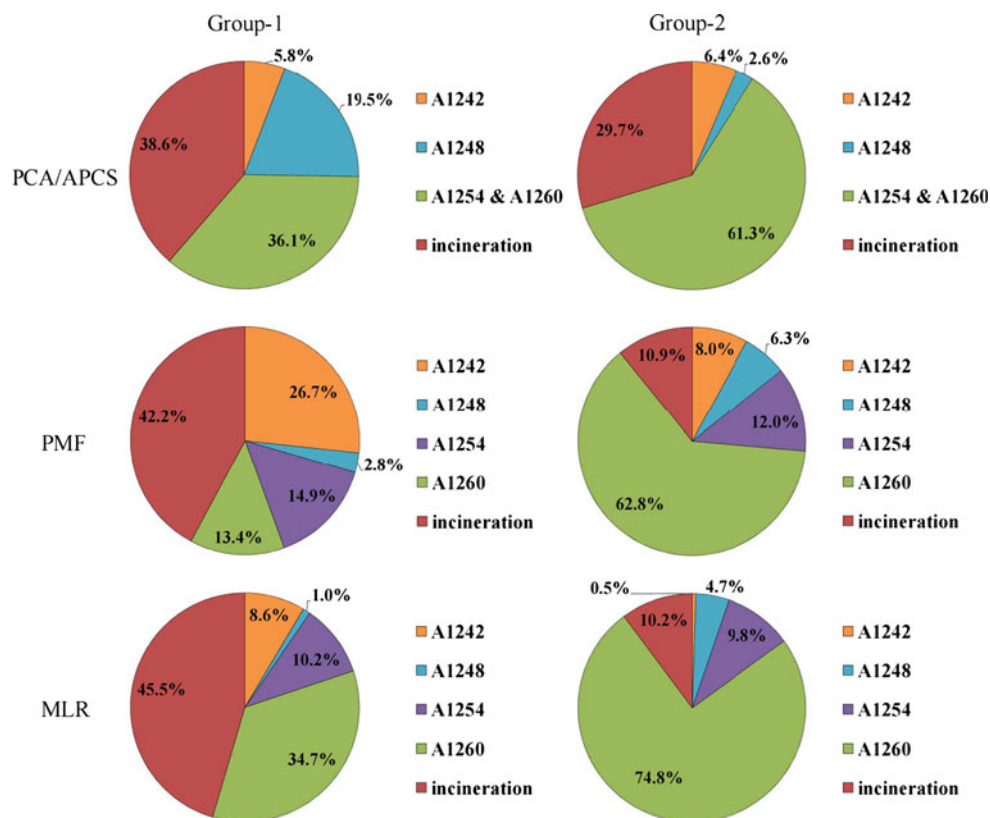
It appeared that there was not a large variation in total PCB concentration in sediments in the entire Nakdong River region, with the exceptions of some samples (samples S4, S8, S9, S17, and S18). The samples in group 2 were collected in a region having greater numbers of industrial facilities than that comprising group 1 samples. In particular, approximately 20% of the total landfill

leachate-treatment facilities were located around the area from which sample S8 was taken. Overall, PMF gave a more stable result in the source apportionment. Only a few negative regression coefficients and extracted factor profiles were similar to the references compared with PCA/APCS, which can be attributed to the characteristics of PCA, in which variables were equally weighted for extracting factors. However, in this study, no such large difference in source contributions was observed with PCA/APCS, PMF, and MLR.

Conclusion

Many studies and social efforts on POPs have been made in Korea. Of these, PCBs have been of great concern due to their widespread use in numerous industries, and PCB wastes have been managed at lower regulatory levels than in other countries. In particular, NIER of Korea has conducted an annual survey on EDCs in environmental media, but this was performed less intensively on river sediments. Therefore, surficial sediments of the Nakdong River were collected, concentration distributions and source identifications determined, and the results examined using multivariate factor analysis.

Fig. 6 Source contributions to total PCB concentrations in the sediment samples as estimated by way of PCA/APCS, PMF, and MLR



The level of total PCBs (209 congeners) ranged from 0.124 to 79.2 ng/g dry weight, which was regarded to be similar or slightly lower than that in other neighboring countries. In terms of coplanar PCBs, their levels were similar to those of four other major rivers in Korea. Most samples had a high proportion of tetra-CBs, but some had relatively much higher PCB concentrations for different homologue compositions, in which the proportion of tetra-CBs was less but those of hexa- and hepta-CBs higher. However, the isomer compositions in each homologue were similar.

To identify the sources of PCBs in the sediments of the Nakdong River, three receptor models (PCA/APCS, PMF, and MLR) were used, and the results of the three models were similar. In the case of samples with lower PCB concentrations, incineration sources predominated in most, with lower proportions found in downstream samples. However, it was shown that some samples with much higher PCB concentrations had the highest proportion of aroclor 1260. Around the sampling points having high PCB concentrations, many industrial facilities, e.g., landfill leachate-treatment facilities, existed. Therefore, the samples are more likely than others to contain high-chlorinated PCBs because various kinds of wastes could be buried and because high-chlorinated PCB products, such as aroclor 1260, are more stable in the elution process of landfill leachate than in the incineration process.

The correlation between estimated and measured PCBs concentrations in sediments was high, with coefficients of determination ranging from 0.77 to 0.99. As in other previous research, PMF showed relatively more stable results than PCA/APCS. Tools for multivariate factor analysis are complementary to each other, and it is important to apply several receptor models and compare their results for robust source identification.

Acknowledgment This research was supported by research funds from Chonbuk National University in 2010.

References

- Bonifazi P, Pierini E, Bruner F (1997) Solid-phase extraction of PCBs from water containing humic substances. *Chromatographia* 44:595–600
- Donna LB, Ralph JM (1996) Characterization of the polychlorinated biphenyls in the sediments of woods pond: evidence for microbial dechlorination of aroclor 1260 in situ. *Environ Sci Technol* 30:234–245
- Guo H, Wang T, Louie PKK (2004a) Source apportionment of ambient non-methane hydrocarbons in Hong Kong: application of a principal component analysis/absolute principal component scores (PCA/APCS) receptor model. *Environ Pollut* 129: 489–498
- Guo H, Wang T, Simpson IJ, Blake DR, Yu XM, Kwok YH et al (2004b) Source contributions to ambient VOCs and CO at a rural in eastern China. *Atmos Environ* 38:4551–4560
- Hong SH, Yim UH, Shim WJ, Oh JR, Lee IS (2003) Horizontal and vertical distribution of PCBs and chlorinated pesticides in sediments from Masan Bay, Korea. *Mar Pollut Bull* 46:244–253
- Hung CC, Gong GC, Jiann KT, Yeager KM, Santschi PH, Wade TL et al (2006) Relationship between carbonaceous materials and polychlorinated biphenyls (PCBs) in the sediments of the Danshui River and adjacent coastal area, Taiwan. *Chemosphere* 65:1452–1461
- Ikonomou MG, Sather P, Oh JE, Choi WY, Chang YS (2002) PCB levels and congener patterns from Korean municipal waste incinerator stack emissions. *Chemosphere* 49:205–216
- Ishikawa Y, Noma Y, Yamamoto T, Mori Y, Sakai S (2002) PCB decomposition and formation in thermal treatment plant equipment. *Chemosphere* 67:1383–1393
- Ito K, Xue N, Thurston GD (2004) Spatial variation of PM_{2.5} chemical species and source-apportioned mass concentrations in New York City. *Atmos Environ* 38:5269–5282
- Jeong GH, Kim HJ, Joo YJ, Kim YB, So HY (2001) Distribution characteristics of PCBs in the sediments of the lower Nakdong River, Korea. *Chemosphere* 44:1403–1411
- Kim KS (2004) Study on the behavior and mass balance of PCBs in ambient air. Doctoral thesis, University of Yokohama, Yokohama, Japan (in Japanese)
- Kim KS, Hirai Y, Kato M, Urano K, Masunaga S (2004) Detailed PCB congener patterns in incinerator flue gas and commercial PCB formulations (kanechlor). *Chemosphere* 55:539–553
- Kim KS, Lee SC, Kim KH, Shim WJ, Hong SH, Choi KH et al (2009) Survey on organochlorine pesticides, PCDD/Fs, dioxin-like PCBs and HCB in sediments from the Han River, Korea. *Chemosphere* 75:580–587
- Kodavanti PR, Senthilkumar K, Loganathan BG (2008) Organohalogen pollutants and human health. In: Heggenhougen HK, Quah S (eds) *Encyclopedia of public health*, vol 4. Academic Press, San Diego, pp 686–693
- Koh CH, Kim JS, Kannan K, Villeneuve DL, Senthilkumar K, Giesy JP (2004) Polychlorinated dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDF), biphenyls (PCBs), and polycyclic aromatic hydrocarbons (PAHs) and 2,3,7,8-TCDD equivalents (TEQs) in sediment from the Hyeongsan River, Korea. *Environ Pollut* 132:489–501
- Lee E, Chan CK, Paatero P (1999) Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmos Environ* 33:3201–3212
- Lee KT, Tanabe S, Koh CH (2001) Contamination of polychlorinated biphenyls (PCBs) in sediments from Kyeonggi Bay and nearby areas, Korea. *Mar Pollut Bull* 42:273–279
- Liu H, Zhang Q, Wang Y, Cai Z, Jiang G (2007) Occurrence of polychlorinated dibenzo-p-dioxins, dibenzofurans and biphenyls pollution in sediments from the Haihe River and Dagu Drainage River in Tianjin City, China. *Chemosphere* 68:1772–1778
- Loganathan BG, Lam PKS (eds) (2011) *Global contamination trends of persistent organic chemicals*. CRC Press, Boca Raton
- National Institute of Environmental Research (2005) *Analytical methods of endocrine disrupting chemicals*. NIER, Korea
- National Institute of Environmental Research (2007) *The 8th monitoring results on endocrine disrupting chemicals in Korea*. NIER, Korea
- Nicolas J, Chiari M, Crespo J, Orellana IG, Lucarelli F, Nava S et al (2008) Quantification of Saharan and local dust impact in an arid Mediterranean area by the positive matrix factorization (PMF) technique. *Atmos Environ* 42:8872–8882
- Norris G, Vedantham R, Wade K, Brown S, Prouty J, Foley C (2008) EPA positive matrix factorization (PMF): 30 fundamentals &

- user's guide. United States Environmental Protection Agency, Office of Research and Development, Washington, DC
- Paatero P (1997) Least squares formulation of robust non-negative factor analysis. *Chemometr Intell Lab* 37:23–35
- Paatero P, Tapper U (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5:111–126
- Park SU, Kim JG, Masunage S, Kim KS (2009) Source identification and concentration distribution of polychlorinated biphenyls in environmental media around industrial complexes. *Bull Environ Contam Toxicol* 83:859–864
- Park SU, Kim JG, Jeong MJ, Song BJ (2010) Source identification of atmospheric polycyclic aromatic hydrocarbons in industrial complex using diagnostic ratios and multivariate factor analysis. *Arch Environ Contam Toxicol*. doi:10.1007/s00244-010-9567-5
- Polissar AV, Hopke PK, Paatero P, Malm WC, Sisler JF (1998) Atmospheric aerosol over Alaska 2 elemental composition and sources. *J Geophys Res* 103:19045–19057
- Ren M, Peng P, Chen D, Chen P, Li X (2009) Patterns and sources of PCDD/Fs and dioxin-like PCBs in surface sediments from the East River, China. *J Hazard Mater* 170:473–478
- Shelly LM, Melissa JA, Eileen PD, Jana BM (2002) Source apportionment of exposures to volatile organic compounds. I. Evaluation of receptor models using simulated exposure data. *Atmos Environ* 36:3629–3641
- Shin SK, Kim KS, You JC, Song BJ, Kim JG (2006) Concentration and congener patterns of polychlorinated biphenyls in industrial and municipal waste incinerator flue gas, Korea. *J Hazard Mater* A133:53–59
- Smith JA, Witkowski PJ, Chiou CT (1988) Partition of nonionic organic compounds in aquatic system. *Rev Environ Contam Toxicol* 103:125–151
- Sofowote UM, Mccarry BE, Marvin CH (2008) Source apportionment of PAH in Hamilton Harbour suspended sediments: comparison of two factor analysis methods. *Environ Sci Technol* 42:6007–6014
- Thurston GD, Spengler JD (1985) A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmos Environ* 19:9–25
- UNEP (United Nations Environment Programme) CHEMICALS (1999) Guidelines for the identification of PCBs and materials containing PCBs, first issue. <http://www.chem.unep.ch/Publications/pdf/GuidIdPCB.pdf>
- Zhang H, Zhao X, Ni Y, Lu X, Chen J, Su F et al (2010) PCDD/Fs and PCBs in sediments of the Liaohu River, China: levels, distribution, and possible sources. *Chemosphere* 79:754–762