NOVEL SOURCE APPORTIONMENT METHOD BASED ON GIS RECEPTOR MODEL FOR AQUATIC PERFLUORINATED COMPOUND (PFC) POLLUTION -A CASE STUDY IN THE BASIN OF TOKYO BAY, JAPAN-

Zushi Y and Masunaga S
Graduate School of Environment and Information Sciences, Yokohama National University (YNU), 79-7 Tokiwadai, Hodogaya, Yokohama 240-8501, Japan

Introduction

Perfluorinated compounds (PFCs) have been detected from various environmental matrices even in Arctic[1], and known to be ubiquitous pollutants. Fluorotelomer alcohols (FTOHs) and fluorooctane sulfonamidoethanol (FOSE), which can be degraded to perfluorooctanoate (PFOA) and perfluorooctane sulfonate (PFOS), respectively, spread to all over the globe through atmosphere and are considered to be the contributing factor to the ubiquitous pollution by PFCs[1, 2]. In addition to this, nonpoint source type pollution by PFCs must have contributed to the widespread pollution, especially in urban region[3-6]. The nonpoint source pollution should have been caused by the use of distributed consumer products containing PFCs[7], and the absence of the declining trend of PFC pollution level after various regulations on their production and manufacturing[8] might be due to the contribution of the PFC loading from nonpoint sources. However, nonpoint sources of PFCs have not been well characterized and their contributions to the environment compared with point sources have not been sufficiently evaluated. Thus in this study, we developed a GIS-based technique for source identification of the nonpoint sources of PFC, and evaluation of the source contribution to the PFC pollution. The results based on the data obtained from intensive survey in Tokyo Bay Basin showed the usefulness of the GIS-based receptor modeling for source identification and apportionment.

Methods

Study area and data for the analysis

Tokyo Bay Basin in Japan was selected for the GIS-based analysis. In our previous study[9], a survey of river water PFC pollution was conducted in the basin. River water samples (n=50), which were collected from the down-stream end of a river in each watershed[9] should represent the characteristics of the watershed, such as population density, level of industrial development and so on[7]. The PFC concentrations in the samples quantified in our previous study[9] were used for the GIS-based analysis.

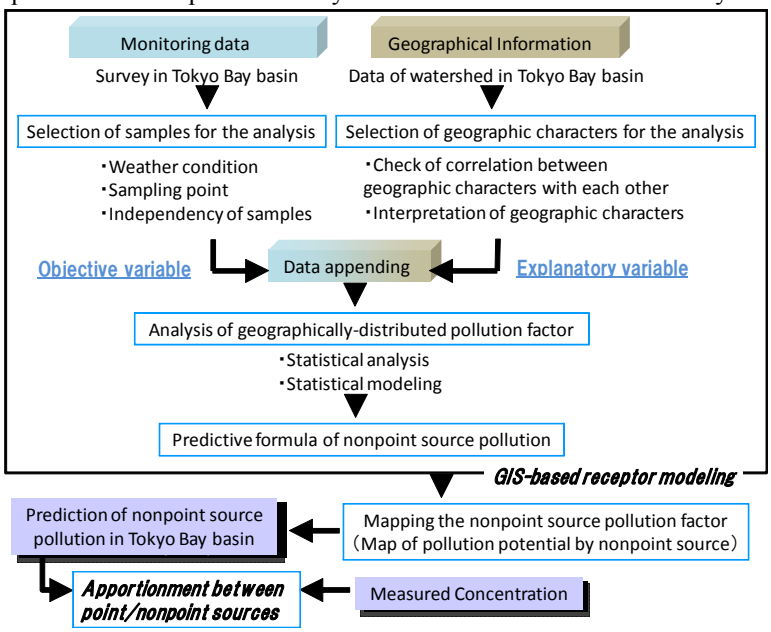

Fig. 1 Scheme of source identification and apportionment of PFCs using GIS-based receptor model

Construction of GIS database

For the development of source apportionment of PFCs using GIS-based receptor model, GIS database in the study area was constructed. The database contains watershed boundary, land-use, population density, sewage-treatment area, locations of sewage treatment plants (STPs), waste disposal sites, train stations and airports, elevation, number of establishments/companies in each basin, etc. on the base map. The 11 candidates of explanatory variables (in other words, candidates of pollution factor) for source identification of PFCs were selected. These candidates were percentages of agricultural area, the area except forest and waste land, arterial traffic area, river and lake area, golf field, other land use area, sewage-treatment area, and catchment area of sewage water in

each treatment plant, buffer area (50 m in radius) of train station, number of waste disposal sites, and population in each watershed and they were extracted from GIS database.

GIS-based receptor modeling for nonpoint source identification

From the results of analysis of correlation between PFC concentration and geographic characters, geographically-distributed sources of PFCs were identified. The scheme of the source identification is shown in Fig. 1. Dataset of PFC concentrations was prepared from the survey in Tokyo Bay basin for a multiple regression analysis as objective variables. Geographic characters were prepared as explanatory variables in the multiple regression analysis. Those variables were selected after checking the correlation between all obtained explanatory variables to avoid multicollinearity. The variables whose meanings were inexplicable were omitted from the analysis. The dataset of PFC concentrations and geographic characters in respective watersheds were appended using the watershed ID as indicator, then, the data matrix for the multiple regression analysis was obtained. The multiple regression analysis was conducted to obtain the pollution factors in PFCs pollution. Because the PFC concentrations in the environment have been considered to have lognormal distribution, they were converted to logarithmic values. Also, the relationships between log-transformed PFC concentrations and geographic characters were log-linear in this study, thus, the log-transformed values of geographic character were used. Finally, linear multiple regression model with log-transformed values of objective and explanatory variables were conducted. Stepwise procedure (p-in and p-out: 0.05) was used for variable selection in the analysis, then, the obtained multiple regression model was reverted to antilogarithm of PFC concentration. The final regression equation for each PFC homologue is shown as equation (I).

$$PFC_{conc} = \prod_{j=0}^{n} f_j^{\beta_j} \quad \ldots(\mathrm{I})$$

Where, f_j is score of pollution factor j, β_j is estimated parameter derived from the multiple regression analysis. In the case of j=0, f_j=e is accepted. PFC_{conc} means the formed PFC concentration by the pollution factors existed in each watershed.

Median + 3σ, which are calculated from the distribution of predicted PFC_{conc} in the process of the analysis, were used as threshold value in determining the outliers among the monitored PFC concentrations. We considered that the outliers were generated as a result of the loading from point sources, such as PFC-using facilities, and those outliers were confirmed by checking the basic map, individually. Then, the outliers, which were caused by the point sources, were omitted from the multiple regression analysis, and reanalyzed. Thus, the obtained model represents the pollution by nonpoint sources. Map of pollution potential was drawn, exhibiting the PFC_{conc} in each watershed.

Source apportionment between point and nonpoint source

PFC_{conc} was calculated for exhibiting spatial distribution of pollution factors, and represented only PFC concentration in individual watershed without dilution and mixing process of PFC concentration along the flowing down of river to downstream (or from upstream). Therefore, virtual PFC concentrations in the river caused by the nonpoint sources were calculated based on equation (I), considering the dilution and mixing along the flowing down of the river. The processes of loss/generation by degradation, sedimentation of PFCs during the flowing down were not considered in the calculation. Measured PFC concentration in the river must be formed by the contributions of both point and nonpoint sources, thus, the point source contribution can be estimated by subtracting the calculated PFC concentration caused by the nonpoint sources from the measured PFC concentration. Then the source apportionment between point and nonpoint sources in the river of Tokyo Bay basin was accomplished.

Results & Discussion

The pollution factors and its parameters were obtained by the multiple regression analysis. Among the 35 PFC homologues, only 18 homologues (PFHxA, PFHpA, PFOA, PFNA, PFDA, PFUnDA, PFOAisomer, PFNAisomer, PFDAisomer, PFUnDAisomer, PFPeS, PFHxS, PFOS, PFOSisomer1, PFOSisomer2 and FOSA, NMeFOSAA, NEtFOSAA; See reference[9] for those abbreviations) were analyzed, because others were below LOQ for more than 55% of their samples. Among the 18 homologues, R^2 (adjusted for the degrees of freedom) for PFOAisomer, PFDAisomer, PFUnDAisomer, PFPeS and NMeFOSAA, was low (< 0.5). PFPeS and NMeFOSAA were the lowest two among the 18 homologues in regard to the number of >LOQ samples (52 and 48% of all samples, respectively). This shows that the variance of concentrations could not be well explained

Table 1 The multiple regression equations for the major PFCs

		Selected pollution factor j	Partial regression coefficient β	Adjusted R^2
PFOS	j=1	Ratio of arterial traffic area	0.83	0.84
	j=2	Population density	0.42	
	j=0	Constant	1.6	
PFOA	j=1	Ratio of area except forest and waste land	1.3	0.84
	j=0	Constant	2.2	
PFNA	j=1	Ratio of area except forest and waste land	1.4	0.68
	j=2	Ratio of catchment area of sewage wastewater	0.15	
	j=0	Constant	4.4	

by the geographic characters for a homologue with the low number of quantified samples. For most of branched isomers of perfluorocalboxylate (PFCA) could not be well explained by the geographic characters, it might be due their special behavior compared with other PFCs[9]. The results of the regression analysis for PFOS, PFOA and PFNA, which were important homologues, considering the PFC pollution[9], are summarized in Table 1. The percentage of arterial traffic area was chosen for PFOS, its isomers, NEtFOSAA and FOSA (PFOS-related compounds) as the most dominant variable among the explanatory variables. Thus, the traffic-related activities/constructions probably contributed to the PFOS-related pollutions. The percentage of area except for forest and waste land was chosen for PFCAs. This variable can be regarded as the indicator of uniform contribution of artificial land use to the pollution. This variable was selected only in the case of PFOA and PFHpA. It means that the nonpoint sources of these homologues were probably more widely distributed than those of other homologues. In the case of PFNA, the percentage of catchment area of sewage water in each treatment plant was also selected. If STP existed in the watershed, the percentage of its catchment area of its STP in the watershed was calculated as the variable and it should represent the impact of STP effluents. Thus, PFNA pollution was shown to be contributed from STP effluents. This variable was selected only for PFNA, and this result was consistent with the result of survey of the PFC pollution in Tokyo Bay basin[9].

The result of the analysis for PFOS was shown with measured value in Fig. 2. The pollution by PFOS was well explained by the ratio of arterial traffic area and 84% of variance of measured PFOS concentration was explained by 2 geographic variables shown in Table 1. Three outliers were determined in the analysis. Those samples were obtained in the down-stream end of rivers whose basin include square of self-defense force, many facilities of electric and electronic equipment, and improper landfill site. The map representing spatial distribution of the PFC_{conc} (PFOS) formed by pollution factors (in other words, pollution potential) of PFOS nonpoint source is shown in Fig. 3. We could see that the PFOS pollution potential was high in the center of the capital. On the other hand, the pollution potential was spatially-uniform for PFHpA

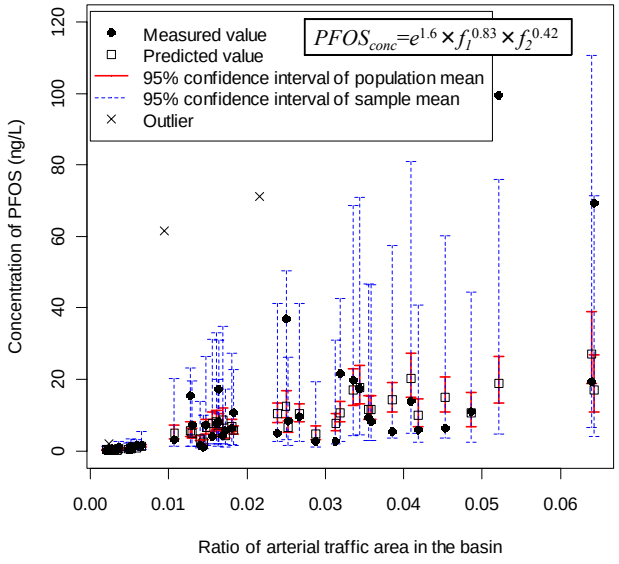

Fig. 2 Predicted concentration by the model and measured concentration of PFOS

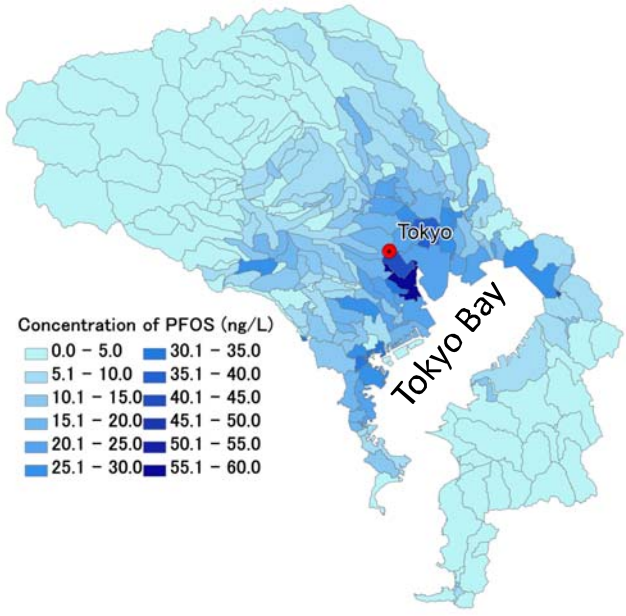

Fig. 3 Map of pollution potential by nonpoint sources in Tokyo Bay basin

and PFOA (Fig. 4). PFCAs were likely to be spatially-uniform compared with PFOS and its related compounds. High concentration spots, which were due to the existence of STPs, were observed for PFNA.

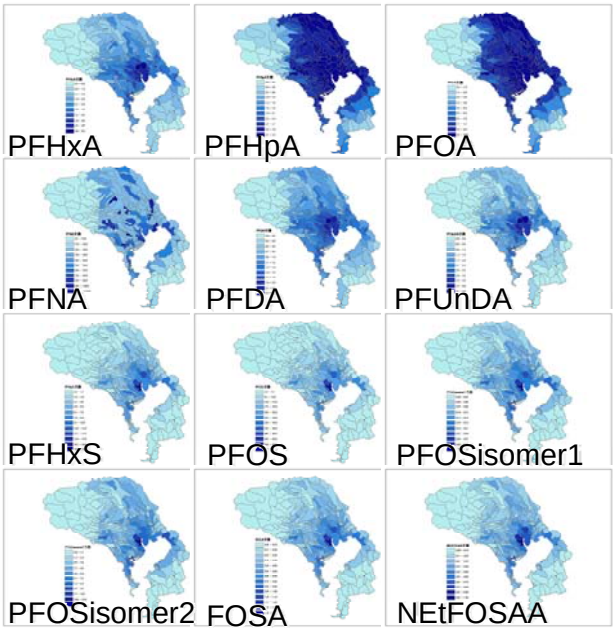

Fig. 4 Maps of pollution potential for various PFCs

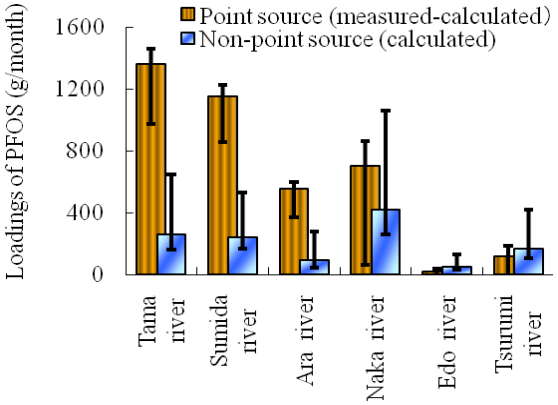

Fig. 5 Source apportionment between point and nonpoint source for PFOS in the major rivers in Tokyo Bay basin

Error bar shows the range of standard deviation derived from the prediction by the model. The deviation in measured value was not included because only one time measurement was conducted.

Finally, we conducted the source apportionment between point and nonpoint source using the result calculated from the receptor model. The results for PFOS are shown in Fig. 5. Six rivers, which are the first-class rivers in Tokyo Bay basin, have the total of their flow covering most of the input into the Bay. In some river basins, the contribution of PFOS loadings from nonpoint source was comparable to that from point sources. When the input of PFOS loadings by the rain runoff[10] is taken into account, the contribution of nonpoint source would be larger.

Source identification and apportionment using GIS-based receptor model was shown to be effective in this case study, especially for the ubiquitous type pollution like nonpoint source pollution.

Acknowledgement

The data used in this study were obtained in a survey conducted in collaboration with the Center for Environmental Science in Saitama and the Tokyo Metropolitan Institute of Public Health. This study was supported by JSPS Research Fellowships for Young Scientists (ID: 213467).

Reference

1. Shoeib M., Harner T. and Vlahos P. *Environ. Sci. Technol.* 2006; 40:7577.
2. Stock N. L., Furdui V. I., Muir D. C. G. and Mabury S. A. *Environ. Sci. Technol.* 2007; 41:3529.
3. Zushi Y., Takeda T. and Masunaga S. *Chemosphere* 2008; 71:1566.
4. Kim S. K. and Kannan K. *Environ. Sci. Technol.* 2007; 41:8328.
5. Murakami M., Shinohara H. and Takada H. *Chemosphere* 2009; 74:487.
6. Adams J., Houde M., Muir D., Speakman T., Bossart G. and Fair P. *Marine Environmental Research* 2008; 66:430.
7. Zushi Y. and Masunaga S. *Environ. Toxicol. Chem.* 2009; 28:691.
8. Zushi Y., Tamada M., Kanai Y. and Masunaga S. *Environ. Pollut.* 2010; 158:756.
9. Zushi Y., Ye F., Masunaga S., Motegi M., Nojiri K., Hosono S., Suzuki T., Kosugi Y. and Yaguchi K. *Organohalogen Compd.* 2010; 72:In Press.
10. Zushi Y. and Masunaga S. *Chemosphere* 2009; 76:833.