Nontarget screening using passive air and water sampling with a level II fugacity model to identify unregulated environmental contaminants

In-Young Chung1, Yu-Mi Park1, Hyun-Jeoung Lee2, Hyuk Kim1, Dong-Hoon Kim1, Il-Gyu Kim1, Sang-Min Kim1, Young-Sun Do1, Kwang-Seol Seok1, Jung-Hwan Kwon2,*

1. Chemical Research Division, Environmental Health Research Department, National Institute of Environmental Research, Incheon 22689, Republic of Korea
2. Division of Environmental Science and Ecological Engineering, Korea University, Seongbuk-gu, Seoul 02841, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 31 March 2017
Revised 26 June 2017
Accepted 29 June 2017
Available online 11 July 2017

Keywords:
Nontarget screening
GC × GC-TOFMS
Passive sampling
Pollutant release and transfer register (PRTR)
Fugacity model

A B S T R A C T

It is thought that there are many unregulated anthropogenic chemicals in the environment. For risk assessment of chemicals, it is essential to estimate the predicted environmental concentrations. As an effort of identifying residual organic contaminants in air and water in Korea, nontarget screening using two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS) was conducted at 10 sites using polyurethane foam passive air sampler and at 6 sites using polydimethyl siloxane (PDMS) passive water sampler in three different seasons in 2014. More than 600 chemical peaks were identified satisfying the identification criteria in air and water samples, respectively, providing a list for further investigation. Chemical substances with reported national emission rates in 2014 \((n = 149)\) were also screened for potential existence in the environment using a level II fugacity model. Most of chemical substances classified as not detectable were not identified with detection frequency greater than 20% by nontarget screening, indicating that a simple equilibrium model has a strong potential to be used to exclude chemicals that are not likely to remain in the environment after emissions from targeted monitoring.

© 2017 The Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences. Published by Elsevier B.V.

Introduction

For the evaluation of environmental risks of chemical substances, it is required to know the predicted environmental concentration (PEC). Two typical methods used to obtain the PEC values are targeted chemical analysis after environmental sampling and the estimation using fate models based on the anticipated emission rate. The former is used for chemical substances under the existing regulations such as persistent organic pollutants under the Stockholm Convention (UNEP, 2008) and pharmaceuticals of the watch-list under the Water Framework Directive (European Commission Directive 2013/39/EU, 2013). It is also acknowledged that the capacity for evaluating the environmental levels of known chemical substances falls short of the number of chemical substances being used in our modern society. The latter is more commonly adopted for screening of environmental risks of industrial chemicals for registration purposes (RIVM, 2004). Advantage of modeling approach is low experimental cost and required data but the predicted values are in general of higher uncertainties than those obtained by chemical analyses.
With increasing concerns about chemical substances in the environment, there are attempts to add more chemicals under environmental regulations to reduce risks on human and ecosystem health. For the identification of emerging organic pollutants in the environment, analytical methods using two-dimensional gas chromatography coupled to mass spectrometry (GC × GC-MS) have been developed (Allan et al., 2013; Hernández et al., 2011; Hug et al., 2014; Zushi et al., 2016). For the identification of semi-volatile and hydrophobic analytes in the environment, nontarget screening methods can be coupled with equilibrium passive sampling (Allan et al., 2013; Van Stee et al., 2002). In our previous nontarget screening study using passive air and water samplers coupled to two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS), up to 90,000 different chemical peaks were detected (Chung et al., 2015; Appendix A Table S1). Although chemicals that can be identified by nontarget analysis coupled with passive samplers are mainly neutral, semi-volatile and hydrophobic organic chemicals whereas many industrial chemicals are very polar and ionizable as was in European Union’s Registration, Evaluation, and Authorization of Chemicals (EU REACH) inventory (Franco et al., 2010), this screening tool is still useful to identify unregulated environmental contaminants within the hydrophobic and semi-volatile chemical space.

On the other side, environmental models have been developed for the estimation of environmental concentration of chemicals at various levels of complexity. For example, Mackay (2001) classified multimedia environmental models into four levels. Lower tier models, level I and level II, are equilibrium models between environmental phases with minimal inputs of chemical properties. Higher tier models are steady-state (level III) or unsteady-state models (level IV) with various spatial and temporal resolutions. The predictability of models strongly depends on the estimation of emission. In Korea, pollutant release and transfer register (PRTR) system is operated to collect and disseminate information on environmental releases and transfers of priority chemical substances from industries and other facilities (NIER, 2017). Under the Korean PRTR system, emission and transfer of 226 chemical substances (149 organic chemicals) were voluntarily reported as of 2014 (NIER, 2017). Although the number of chemical substances with reported emission rates is much less than the number of chemicals in use, it would be a good starting point to estimate PEC of chemicals that might remain in the environment using fate models and the emission data in Korea.

Environmental monitoring using nontarget screening and fate models are complementary. If the occurrence of chemicals was identified by nontarget screening, models can be used for the estimation of emission strength. Model results can also guide the sampling strategy for target and nontarget environmental analyses. Consequently, we attempted to compare gaps between nontarget chemical screening using passive air and water samplers with GC × GC-TOFMS and the PECs obtained using a level II multimedia environmental model for South Korea. Nontarget screening analyses using GC × GC-TOFMS were conducted at 10 sites using polyurethane foam (PUF) passive air sampler and at 6 sites using polydimethyl siloxane (PDMS) passive water sampler in three different seasons in 2014. Values of PECs from the level II fugacity model were multiplied by estimated partition coefficients between sampler material and air or water to obtain estimated chemical concentration at equilibrium in passive samplers. Using a threshold concentration in passive samplers above which chemicals are likely to be identified by the GC × GC-TOFMS screening method, PRTR chemicals are divided into two groups and they are compared with nontarget screening results in 2014 at 8 sampling locations. Gaps between two approaches to identify potentially persistent chemicals are identified and discussed.

1. Materials and methods

1.1. Materials

A polyurethane foam disk (Ziemer chromatographie, Germany; 14 cm diameter, 1.35 cm thickness) was used as a receiving phase in a diffusion type passive sampler. It consists of two stainless steel bowls connected with an external metal bolt and PUF is placed in the bottom bowl with the help of a stainless steel support. Passive water sampling was performed using a PDMS film (7.5 cm width × 30 cm length × 1 mm thickness) as a receiving phase. PUF disks were cleaned using high-performance liquid chromatography grade n-hexane and acetone from J. T. Baker (Phillipsburg, NJ, USA) and all other solvents (hexane, acetone, dichloromethane, ethyl acetate, methanol) were used from Wako (Tokyo, Japan). Silica gel for clean-up was purchased from Merck (Darmstadt, Germany).

1.2. Selection of PRTR data

Reported nation-wide emission data under the Korean PRTR system were collected for 208 chemicals in 2014. Of those, mixtures, inorganic chemicals, and organic chemicals with zero emission were excluded. For the modeling, annual emission data for 149 organic chemicals were used. All values are listed in the MS Excel file, Supporting Material.

1.3. Level II fugacity model

SimpleBox Korea, a recently developed multimedia fate model for the evaluation of PEC in local and regional scales (NIER, 2014), was simplified to a level II fugacity model to obtain PEC for the evaluation of PEC in local and regional scales (NIER, 2014). The volumes and advection rates were taken from default values in SimpleBox Korea. Detailed values of environmental parameters are listed in Table 1.

<table>
<thead>
<tr>
<th>Table 1 – Environmental parameters for level II fugacity model.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Air</strong></td>
</tr>
<tr>
<td>Advection rate (m/h) *</td>
</tr>
<tr>
<td>Volume (m³) *</td>
</tr>
<tr>
<td>Density (kg/m³) *</td>
</tr>
<tr>
<td>Fraction organic carbon *</td>
</tr>
<tr>
<td>Concentration inflow (mol/m³) *</td>
</tr>
</tbody>
</table>

\* SimpleBox Korea v. 2.0 (NIER, 2014).
Half-lives in air and in water were estimated using AOPWIN and BIOWIN4 programs in EPISuite™ (US EPA, 2012). AOPWIN gives atmospheric half-lives based on the estimated rate constant with hydroxyl radicals and BIOWIN4 is an expert judgment model giving a range of environmental half-lives. The range estimated by BIOWIN4 was then converted to a single value according to Aronson et al. (2006). Soil half-lives were assumed to be double the water half-lives, as was done in previous studies (Fenner et al., 2005; Lee and Kwon, 2016).

Input chemical properties are Henry's law constant, octanol–water partition coefficient ($K_{OW}$), and half-lives in environmental compartments. Experimental values of Henry's law constants and $K_{OW}$ in EPISuite database were given priority over predicted values. For chemicals without experimental data in the EPISuite database, predicted values from KOWWIN (US EPA, 2012) and the bond method in HENRYWIN (US EPA, 2012) were used. All chemical specific input values are listed in the MS Excel file, Supporting Material. With these input data, equilibrium concentrations in air and water were obtained. The environmental concentrations were then multiplied by octanol–air ($K_{OA}$) or octanol–water partition coefficient ($K_{OW}$) as surrogates for partition coefficients between sampling phase and air or water. Concentrations of a chemical in passive air or water sampler at equilibrium were estimated by multiplying $PEC$ with $K_{OA}$ or $K_{OW}$. Assuming an order-of-magnitude detection limits of the GC × GC-TOFMS screening method of $10^{-5}$ g/m³ in both passive air and water samplers, chemicals under the PRTR list were divided into two groups – one that is likely to be identified by the nontarget screening with the concentration in passive samplers greater than $10^{-5}$ g/m³ and the other that is unlikely to be identified.

1.4. Air and water sampling campaign

The sampling sites of ambient air were eight background reference sites and two potential pollution hotspots (Fig. 1). They are Ulleung (U (N37°1′20.39″, E130°47′48.02″)), Jeju (J1 (33°21′37.7″, E126°27′40.1″), J2 (33°17′37.7″, E126°09′42.3″), J3 (33°20′49.13″, E126°23′19.2″))), Baengnyeong Islands (B (37°57′51.7″, E124°37′50.3″)) and upstream of Han River (H1 (N 38°07′2.0″, E127°46′36.4″), H2 (N37°56′52.8″, E127°49′5.0″), H3 (N37°50′13.0″, E127°40′34.4″)) as reference sites and Siheung (S1 (N 37°19′56.9″, E 126°42′20.3″) and S2 (N 37°19′3.9″, E 126°44′59.5″)) as sites near pollution sources. Samplers were deployed during 34–70 days. Detailed sampling time for each campaign is shown in Table S2.

Water sampling was conducted at three sites near potential pollution hotspots (S1, S2, and S3 (N 37°17′51.7″, E 126°36′8.5″)) and at three reference sites (H1, H2, and H3) (Fig. 1). Detailed sampling time for each campaign is shown in Appendix A Table S2.

1.5. Passive sampling and extraction of chemicals

Passive air sampler was customized according to the US Geological Survey suggestions (Alvarez, 2010). Before sampling, PUF was cleaned by Soxhlet extraction using n-hexane and acetone for 16 hr, respectively. Cleaned PUF was dried in a vacuum drying oven (WOV-70, Daihan scientific Co. Ltd., Wonju, Korea) at an ambient temperature for 6 hr and stored in a desiccator until use. One cleaned PUF disk was placed in a custom-made sampler holder and one sampler holder was deployed at each sampling point (detailed description of the deployment of the PUF sampler, Appendix A Fig. S1). After deployment, one disk of PUF was extracted for 24 hr in 400 mL dichloromethane using a Soxhlet extractor. Sample extract was concentrated to 5 mL using a rotary evaporator. Half milliliter of concentrated sample extract was loaded in an activated silica gel column. Fifty milliliter of hexane was passed through the silica column and the eluate was discarded. Then, the column was extracted using each 100 mL of dichloromethane/hexane (25:75), dichloromethane/hexane (50:50), and dichloromethane. Three extracts were combined and the volume was concentrated to 100 μL using a rotary evaporator and a nitrogen gas concentrator for GC × GC-TOFMS analysis.

PDMS was cleaned by Soxhlet extraction using ethyl acetate for 60 hr and stored in a desiccator until use. One sheet of PDMS was placed in a sampler holder and the sampler was submerged in water (Appendix A Fig. S1). After sampling, PDMS were extracted in a Soxhlet extractor for 24 hr in 400 mL methanol/acetone (1:1) solution. Sample extracts were concentrated to 5 mL using a rotary evaporator. Clean-up and elution procedures using an activated silica gel column were the same as passive air samples.

1.6. Nontarget analysis

Two-dimensional gas chromatography (7890A, Agilent Technologies, USA), time-of-flight mass spectrometry (Pegasus 4D, LECO Corporation, St, Joseph, MI, USA) was equipped with an auto sampler (COMBI PAL, LEAP Technologies, USA) and fitted with a two-dimensional column set consisting of a DB-5MS (30 m x 0.25 mm x 0.25 μm film thickness) as the primary column and Rtx-1701 (1 m x 0.10 mm x 0.10 μm film thickness) as the secondary column. One microliter of extracted sample was injected in an injector at 250°C on a splitless mode. The main oven containing column 1 was held at 35°C for 10 min, and ramped at 5°C/min from 35 to 80°C, and ramped at 5°C/min from 80 to 100°C and held at 100°C for 3 min and ramped at 10°C/min from 100 to 270°C and held at 270°C for 20 min. The secondary column oven containing column 2 was started at 55°C, and held for 10 min, and ramped at 5°C/min from 55 to 100°C, and ramped at 5°C/min from 100 to 120°C and held at 120°C for 3 min and ramped at 10°C/min from 120 to 290°C and held at 290°C for 20 min.

Time-of-flight mass spectrometer was scanned in the mass range between 35 and 550 m/z with the acquisition rate of 100 spectra per second. Source temperature was 220°C and ionization was conducted using electron impact mode at 70 eV. Data acquisition and analysis were performed using ChromaTOF software, version 4.44 (LECO Corporation, St. Joseph, MI, USA).

For peak identification, a signal-to-noise (S/N) threshold was set at 1000:1 and resulting peaks were identified by a forward search of the NIST 2011 library. A forward match score of at least 700 was required for putative compound identification (Rees et al., 2016; Prebihalo et al., 2015). For detected peaks from each sample, peaks originated from field blank, column bleeding were removed. In order to avoid duplicated detection, peaks within 25 sec of the primary retention time with mass spectrum matching greater than
70% were regarded as the same chemical peaks. Aligned chemicals were nominated as first ranked hit name by the NIST library. There was a limitation of the identification of aligned chemicals, because the aligned chemicals were matched by integer m/z values.

2. Results and discussion

2.1. Occurrence of chemicals in air and water samples

Number of chemical peaks identified per one passive air sample, satisfying the identification criteria (S/N ratio > 1000 and similarity > 700), ranged from 84 to 408. Numbers of chemical peaks were in general greater in S sites near Shihwa–Banwol industrial complex than other sites although the differences were not very notable. Thus, three sampling campaigns at 10 different sampling sites were pooled to evaluate chemicals identified in passive air samples as a whole. Twelve chemicals were identified in all 30 passive air samples (ten chemical structures in Fig. 2). They are mostly low molecular weight and volatile organic chemicals, mainly used as solvents and intermediates for chemical synthesis.

Numbers of identified peaks in water samples, on the other hand, showed noticeable differences between H and S sites. Whereas they ranged 48–155 in H sites, more chemical peaks, ranging 69–357, were identified in S sites. This would be likely due to heavy load of organic chemicals from Shihwa–Banwol.

Fig. 1 – Air and water sampling sites using passive samplers during the monitoring campaigns (2014.03–2014.11).
industrial complex to the sites. Chemicals identified in all sites regardless of seasons are presented in Fig. 2. Diethylene glycol monoethyl ether, identified in all air samples, was also identified in all water samples. Fatty acids, methyl palmitate and methyl stearate, were identified in all water samples. Because fatty acids are synthesized in organisms, the frequent occurrence of fatty acids might not be due to industrial uses. The major uses of all other chemicals identified in all water samples are plastic plasticizers, fragrances, intermediates, solvents, food flavors, and non-ionic surfactants. Because S sites and H sites are not in the same watershed (Fig. 1), it is likely that they have many wide-spread sources. One interesting chemical not shown in Fig. 2 is tris(1-chloro-2-propyl) phosphate (TCPP). It was identified from all S site samples regardless of sampling time, whereas it was not identified in any H site samples. TCPP is an organophosphorous flame retardant, replacing the use of brominated flame retardants. The frequent occurrence of TCPP in water samples affected by industrial sources suggests that targeted monitoring of TCPP may be required because TCPP is one of the chemical substances of high concerns and a suspected carcinogen (Li et al., 2014).

Fig. 2 – Chemical identified in all passive air and water samples with their representative uses. *Diethylene glycol monoethyl ether was identified in both air and water samples.

2.2. Comparison with model predictions

Among 149 organic substances with emission data under the PRTR system, only 12 and 7 chemicals were identified with detection frequency ≥20% from passive air (≥6/30) and water samples (≥4/18) (Table 2). Sixty two and sixty three chemicals were classified by the level II fugacity model to be detectable in air and water, respectively, satisfying that the predicted concentration in passive sampling material is greater than 10<sup>-5</sup> g/m<sup>3</sup>. All model simulations are found in Appendix A. Although deterministic criterion was used in this study, uncertainties associated with the model prediction could be assessed using a Monte-Carlo simulation. The variation of model input parameters were assumed as uniform around the...
selected values. The range of log $K_{OW}$ and log $H$ was one log unit with the selected values as geometric means and that of degradation half-lives was between single values right above and below the selected values. An example histogram is shown in Appendix A Fig. S2 for naphthalene. Without significant differences in the range of input parameters, the upper 95% limit of the Monte-Carlo simulation is approximately one order-of magnitude higher than the deterministic model results. Thus, the following comparison was simply based on the deterministic simulation results.

The usefulness of the level II fugacity model was demonstrated by comparing the type I and type II errors of model predictions. The false negative prediction of the model was defined as a type I error and the false positive prediction was a type II error. Because the level II fugacity model is a screening tool for the existence of anthropogenic chemicals substances in the environment, a type II error is not of significant concerns. The probabilities of having a type I error were 4.6% in air samples and 2.3% in water samples. If the model is used for a rejection tool for targeted analysis, the obtained type I errors indicate the probabilities of the model failure. Because it is common to use the 95% confidence level for many environmental decisions, the level of type I errors of the model may be regarded acceptable.

Although the performance of the level II fugacity model is acceptable when rejecting chemical substances for targeted analysis, it is useful to look through the cases that the model failed. Four substances of the false negative prediction for passive air samples are aniline, naphthalene, 4-methyl-$m$-phenylenediamine, and tert-butyl methyl ether and the false negatively predicted chemicals for passive water samples are tributylamine and S-benzyl diisopropyl phosphorothioate (Table 2).

Aniline and tributylamine were only found in passive air and water samples near Shiheung area (S sites). Although they were predicted to be lower concentrations in the environment due to relatively short atmospheric and water half-lives (MS Excel file, Appendix A), the frequent occurrence in samples from S sites might be affected by strong local sources from Shihwa–Banwol industrial complex. Naphthalene, 4-methyl-$m$-phenylenediamine, and tert-butyl methyl ether were frequently detected in passive air samples from both S and H sites and the underestimation of PEC by the level II fugacity model might be due to the incorrect emission estimation. The PRTR system only counts emissions from industrial point sources. Thus, the emission of chemicals with various emission sources such as naphthalene may be much greater than reported value under the PRTR system. S-benzyl diisopropyl phosphorothioate (iprobenfos) is a fungicide widely used for controlling rice blast in Korea (Kim et al., 2008) and the emission rate under the PRTR did not count small uses by professional farmers.

### Table 2 - The predictability of the level II fugacity model demonstrated by type I and type II errors of prediction for 149 organic chemicals under the PRTR system.

<table>
<thead>
<tr>
<th></th>
<th>Detection frequency ≥ 20% using nontarget screening*</th>
<th>Detection frequency &lt; 20% using nontarget screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive air samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted detectable by the model</td>
<td>12.9% (8/63) Propan-2-ol Dibutyl phthalate Ethylbenzene Styrene Toluene Bis(2-ethylhexyl) phthalate 2,6-Di-tert-butyl-$p$-cresol 4-(1,1,3,3-tetramethylbutyl)phenol</td>
<td>87.1% (54/62)</td>
</tr>
<tr>
<td>Predicted not detectable by the model</td>
<td>4.6% (4/87) Aniline Naphthalene 4-Methyl-$m$-phenylenediamine tert-Butyl methyl ether</td>
<td>95.4% (83/87)</td>
</tr>
<tr>
<td>Passive water samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted detectable by the model</td>
<td>6.3% (4/63) $N,N$-dimethylformamide Dibutyl phthalate 2,6-Di-tert-butyl-$p$-cresol Nonylphenol</td>
<td>93.7% (59/63)</td>
</tr>
<tr>
<td>Predicted not detectable by the model</td>
<td>2.3% (2/86) Tributylamine S-Benzyl diisopropyl phosphorothioate</td>
<td>97.7% (84/86)</td>
</tr>
</tbody>
</table>

PRTR: pollutant release and transfer register.

* Chemicals under the PRTR system identified using nontarget screening with identification frequency ≥20% are listed.

2.3. Implications of the study toward a new chemicals regulation in Korea

Although only 149 organic substances were evaluated using the level II fugacity model due to the limited availability of emission data, the model holds a potential to be used to guide environmental monitoring. In 2015, Korean government
enforced a new-age chemicals regulation, Act on the Registration and Evaluation, etc. of Chemical Substances (ARECS) (MOE, 2013), benchmarking EU REACH (European Union, 2006). According to this new regulation, manufacturers and importers of chemicals are obliged to submit data for registration. With the successful settling of ARECS, much more emission data or use information that can be used to estimate emission would be available in the near future. Unlike the emission data in the PRTR system, emission estimation under ARECS uses emission factors by default and considers all possible stages of the life-cycle of chemicals. When emission factors are used, environmental release is not likely to be underestimated because of the underlying worst-case assumptions for environmental protection. Thus, the likelihood of a model failure that chemical concentration in the environment is much higher than the PEC estimated by a model might be even lower than those in this study.

The level II model used in this study used the same environmental variables in SimpleBox Korea, a steady-state (level III) multimedia model and a default model for ARECS, although it used different estimation methods for environmental degradability from SimpleBox Korea. When SimpleBox Korea is used for the regulatory estimation of the environmental concentration, many PEC values for industrial chemicals would be available. As tested in this study, the use of estimated PEC values would be helpful for guiding environmental monitoring strategy and the addition of chemicals in environmental guidelines based on PEC.

3. Conclusions

Two screening methods, nontarget screening using GC × GC-TOFMS and a level II fugacity model with reported emission data, were performed for identifying unregulated chemicals that are likely to be found in the environment. Although further monitoring is required to understand the occurrence of anthropogenic chemicals in the environment and quantify chemical concentrations in environmental samples, chemicals identified in this study provide a valuable list to be investigated. In addition, the usefulness of a simple screening fate model was also proven when a model is used to reject chemical substances that are not likely remained above detectable levels in the environment.

Acknowledgments

This research was supported by the National Institute of Environmental Research (No. NIER-RP-2014-335).

Appendix A. Supplementary data

Photos of passive air and water samplers are shown in Fig. S1. An example Monte-Carlo simulation for naphthalene is presented in Fig. S2. List of frequently identified chemicals in our earlier study (Chung et al., 2015) is presented in Table S1. Detailed sampling time during each campaign is shown in Table S2. Chemicals identified with high detection frequency are listed in Tables S3–S6. All model simulation results are available in a separate Microsoft Excel file. They are available online. Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.jes.2017.06.036.

REFERENCES


