In silico screening for unmonitored, potentially problematic high production volume (HPV) chemicals prone to sequestration in biosolids†

Randhir P. Deo and Rolf U. Halden*

Received 25th January 2010, Accepted 27th July 2010
DOI: 10.1039/c001559h

Thousands of high production volume (HPV) chemicals are used in the US at rates exceeding 450 000 kg (1 million pounds) per year, yet little is known about their fates during wastewater treatment and upon release into the environment. We utilized a recently introduced empirical model to predict the fraction of the mass loading (in raw sewage) that is expected to persist in digested sludge following conventional municipal treatment of chemical-laden sewage. The model requires only two readily available input parameters, a compound’s log $K_{OW}$ value and a dimensionless curve fitting parameter ($p_H$). Following refinement of the fitting parameter and cross-validation of the model using the Jackknife method, we predicted the mass fractions of 207 hydrophobic HPV chemicals (log $K_{OW}$ of $\geq 4.0$) that are expected to accumulate in digested municipal sludge during conventional wastewater treatment. Using this screening approach in conjunction with information from toxicity databases, we identified 11 HPV chemicals that are of potential concern due to (i) their propensity to accumulate and persist in sludge (>50% of mass loading), (ii) unfavorable ecotoxicity threshold values, and (iii) structural characteristics suggestive of environmental persistence following release of these HPV chemicals on land during biosolids recycling. The in silico screening approach taken in this study highlights existing environmental monitoring needs and may guide risk management strategies for biosolids disposal.

Introduction

Several thousands of organic chemicals are produced or used in the US in quantities exceeding 450 000 kg or one million pounds per year. These so-called high production volume (HPV) chemicals are used in residential, institutional, and commercial/industrial settings and are discharged in whole or in part into municipal sewage. Their fate in wastewater treatment plants (WWTPs) is mostly unknown.1,2

Of particular concern are the hydrophobic organic chemicals (HOCs) with high sorption potential, i.e., those having a logarithmic $n$-octanol–water partitioning coefficient (log $K_{OW}$) of $>4.0$.3 Hydrophobic chemicals are known to preferentially sorb to particulate matter and become part of the primary and secondary sludge produced during conventional wastewater treatment. Sequestration of chemicals in sewage sludge limits their availability to aerobic biodegradation in the activated sludge treatment step and during anaerobic digestion.4-6 Hydrophobic compounds in raw sewage entering WWTPs have the potential to become enriched in digested sewage sludge (biosolids) to concentrations of several orders of magnitude higher than those found in raw sewage.7 For HPV chemicals, this enrichment process can result in the occurrence of pollutants in digested sludge at parts per million (ppm) concentrations.8 Disposal of digested municipal sludge on land as soil conditioner or fertilizer can provide a mechanism by which problematic wastewater constituents are reentering the environment, thereby providing a potential human and ecological health hazard.

Similar to the REACH regulations in Europe,9 the US Environmental Protection Agency (EPA) is responsible for protecting the human health and the environment through earlier identification of potentially toxic chemical substances. However, despite ongoing monitoring programs for biosolids, currently there are

Environmental impact

Over 4000 chemicals are used in the US at rates exceeding 450 000 kg per year. The majority of these have never been tested for ecotoxicity and human health risks in greater detail. Testing and monitoring all mass-produced chemicals are cost-prohibitive and impractical. In the present work, an empirical model was applied to predict which chemicals contained in raw sewage may persist during wastewater treatment and become sequestered in digested sewage sludge destined for application on land. The presented model identified a small subset of compounds that exhibit unwanted qualities such as pronounced persistence to biodegradation and significant toxicity to environmental receptor organisms. Chemicals identified in this modeling exercise deserve further attention by environmental scientists concerning risks posed to humans and the environment.
no requirements by the US Environmental Protection Agency (EPA) to examine the presence of organic chemicals in sludge prior to application on land. Previous studies attempting to assess risks associated with land application of sludge highlighted a lack of knowledge regarding the identity, occurrence, concentration, fate and toxicity of sludge-borne contaminants. While the risk assessment of these chemicals is essential for avoiding adverse impacts on human health and ecosystems, measuring these chemicals in sludge is not routinely performed because it is expensive, technically challenging, and in general it is unclear which compounds should be focused on. An informative and economically attractive method for screening of chemicals prone to accumulate in sludge is the qualitative modeling using in silico approaches.

A previous study had predicted the concentration of 34 randomly selected HPV chemicals in sewage sludge using their physicochemical properties such as volatility, organic carbon partition coefficient ($K_{OC}$), biodegradation and hydrolysis. Other models have been developed for the EPA and find limited applications as online tools. Deterministic models can provide nuanced estimates but their applicability frequently is limited by a lack of the input parameters required. A recent study examined the behavior of pharmaceuticals and personal care products (PPCPs) during sewage treatment to suggest potential causes for discrepancies between predicted fates and results from actual measurements.

Recently, an empirical model was introduced that can predict the fraction of the total mass loading of a given hydrophobic organic chemical that is expected to persist in biosolids after treatment. The model was developed in fitting exercises that sought to produce the best approximation of actual measurements made at full-scale sewage treatment plants. The best approximation was obtained with a mathematical equation that emphasizes sorption as a dominant process that apparently is dictating the fate and persistence of organic compounds in biosolids. The model’s only input requirements are the log $K_{OW}$ value of the contaminant of interest and an empirically determined fitting parameter that reflects the combined effect of all potential removal processes including biodegradation.

In the present study, we sought to validate the aforementioned model with statistical analyses, and applied it to forecast the fraction of selected HPV chemicals that can be expected to accumulate in biosolids upon discharge into raw sewage and processing by conventional municipal wastewater treatment plants. A total of 207 HPV chemicals were examined in this study. Estimations of chemical sequestration in sludge were interpreted in conjunction with halogenated chemicals and toxicity data (LD$_{50}$ values) to identify compounds of potential human health and ecological concern.

**Materials and methods**

**Empirical model**

We adopted the empirical model proposed by Heidler and Halden, based on the following parameters:

$$f_{sludge} = \frac{p_{in} K_{OW}}{1 + p_{in} K_{OW}}$$  \hspace{1cm} (1)

where $f_{sludge}$ is the mass fraction of chemical(s) expected to accumulate in sludge ($\%$), $p_{in}$ is a dimensionless fitting parameter, and $K_{OW}$ is the 1-octanol–water partition coefficient.

The best value for $p_{in}$ ($6.51 \times 10^{-4}$) was determined using the “Solver” function of Microsoft Excel by minimizing the error between the model output values and experimental data obtained empirically in 11 studies (Table S1†).

Validation of the empirical model was done by first performing a paired $t$-test to compare the predicted versus measured percent mass fractions of chemicals that accumulate in sludge. Second, the Jackknife estimation method was performed, where data for one compound were deleted from the full dataset and the $p_{in}$ value was estimated with the reduced dataset using the “Solver” function of Microsoft Excel. The newly estimated $(n - 1)p_{in}$ value obtained in this leave-one-out approach was then used to predict the deleted compound’s mass fraction in sludge. This approach, chosen to minimize spurious self-correlation, was repeated for each of the 11 chemicals considered to obtain a set of predicted values that could then be compared to the set of actual measurements made using a paired $t$-test.

A standard error of the $p_{in}$ value was calculated by dividing the standard deviation of the $p_{in}$ value by the square root of the total number of measurements made.

**Source of HPV chemicals**

HPV chemicals were obtained from the High Production Volume Information System (HPVIS) database, which is part of the HPV Challenge program of the EPA that encourages companies to compile and release to the agency toxicity and environmental fate data for chemicals produced or imported into the United States in quantities exceeding 1 million pounds (450 000 kg) per year. Two hundred and seven HPV chemicals having a log $K_{OW}$ of >4 were included.

**Results and discussion**

**Validation of the empirical model**

We validated a recently introduced empirical model that predicts the fraction of the total mass loading of HOCs arriving at WWTPs in influent that persists in biosolids. Shown in Fig. 1 are fractions of 11 organic wastewater compounds that have been investigated using this mass balance approach (Fig. 1 and Table S1†). Each data point for a particular chemical is annotated with a unique symbol and color that depict, respectively, the identity of the compound and the chemical group it belongs to.

Monitoring data reported in the literature for full-scale WWTPs with the best model fit are shown in Fig. 1. The three estrogens, estrone, 17z-ethinylestradiol and 17β-estradiol that feature log $K_{OW}$ values of 3.13, 3.67 and 4.01, respectively, were sequestered into sludge to a similar degree (4–11% of the initial WWTP loading). Fragrances represented by galaxolide (log $K_{OW} = 5.90$) and tonalide (log $K_{OW} = 5.70$) were sequestered into sludge with a yield ranging from 48 to 80%. Antimicrobials for which monitoring data were available featured log $K_{OW}$ value range of 0.90 to 4.90. According to the wide spectrum of hydrophobicity covered, their sequestration into digested sewage sludge was observed to be in the range of 0.2 to 83%.
Fig. 1 Analysis of research findings reported for 11 organic wastewater compounds in 13 peer-reviewed mass balance studies. The fraction of the mass loading of chemicals persisting in digested sludge was found to be primarily a function of the compounds’ sorption potential, expressed as the logarithm ($\log_{10}$) of the 1-octanol–water partition coefficient ($K_{OW}$). The empirical data were fit to an S-shaped model that yielded the best fit. Each data point is annotated with a unique symbol and color that depict, respectively, the compound’s identity and principal use. Also shown in the plot are open circle data points depicting the results of a cross-validation of the empirical model using the Jackknife estimation method.

The empirical data were fit to an S-shaped model that yielded the best fit. Each data point is annotated with a unique symbol and color that depict, respectively, the compound’s identity and principal use. Also shown in the plot are open circle data points depicting the results of a cross-validation of the empirical model using the Jackknife estimation method.

The distribution of empirical data points in Fig. 1 also allows for an examination of the variability of chemical accumulation in sludge as a function of plant-specific treatment efficiency. For example, multiple measurements available for nonylphenol and the two fragrances showed a high propensity of these compounds to persist in sludge, although some plants apparently performed better than others in removing these substances. Fitting of these empirical data to the empirical model produced an S-shaped curve that nicely tracked the monitoring information available for organic chemicals during full-scale treatment (Fig. 1). On average, predictions matched experimental observations within a factor of 1.1-fold (standard deviation of ±1.4). Results from a paired t-test that compared predicted to calculated values showed the two datasets to be statistically indistinguishable at the 90% confidence level, with coefficient of determination equal to 0.74.

Also shown in Fig. 1 are black open circles that depict the results of sequestration predictions obtained with the Jackknife method in the cross-validation of the empirical model. A visual inspection of Fig. 1 shows a good correlation of mass fractions in sludge between those calculated using multiple $p_{hit}$ values (Jackknife method, $n-1$ dataset, open circles) and single $p_{hit}$ value (empirical model, full dataset, fitted curve). Whereas the predicted values from two methods were within a factor of 1.0-fold (standard deviation of ±0.1), paired t-test between the two predicted datasets was statistically indistinguishable at the 90% confidence level. Results from this cross-validation show that the mass fractions of chemicals in sludge can be calculated with reasonable accuracy even for compounds for which no monitoring information is available. Further evaluations showed the model to be robust with respect to fitting of $p_{hit}$ value and that the quality of predictions obtained typically is good for compounds featuring a log $K_{OW}$ value of four or greater (see ESI, Fig. S1†). This opened the door to predicting the mass fraction in sludge of moderately to highly hydrophobic HPV chemicals ($log K_{OW} \geq 4$) for which monitoring information is not currently available.

Selection of HPV chemicals examined

From the HPVIS database, we initially selected 316 hydrophobic HPV chemicals that featured log $K_{OW}$ values of >4.0 and whose mineralization to CO₂ over a period of 28 days was predicted to be less than 60% when utilizing the ready-biodegradability test of the Organization for Economic Cooperation and Development (OECD). This list was shortened by excluding surfactants and those compounds previously monitored in sludge, which left a total of 207 HPV chemicals (ESI, Table S2†).

The range and distribution of compounds examined are depicted in Fig. 2. Log $K_{OW}$ values in the range of 4 to <12 were grouped in 0.5-log intervals, and those with log $K_{OW}$ values of ≥12 were lumped together as a single group. Visual inspection of the graph shows that most of the HPV chemicals are in the group featuring log $K_{OW}$ values of 4 to <4.5 (total of 43) and 4.5 to <5 (total of 44). Together, these constituted ~40% of the total number of HPV chemicals included in this study. This implies that approximately two-thirds of the total number of HPV chemicals included in this study feature log $K_{OW}$ of ≥5, which is approximately the threshold log $K_{OW}$ value above which >50% of the mass loading is predicted to persist in sludge (Fig. 1).

Qualitative prediction of HPV chemicals in sludge

The distribution of mass fractions of HPV chemicals predicted to persist in digested sewage sludge upon entry into a WWTP is shown in Fig. 3. Predictions cover the log $K_{OW}$ range of 4 to 25. Data points are annotated with colored symbols to define the uses of the chemicals considered. Due to the considerable range of pH values observed in municipal and industrial wastewater, no attempts were made to adjust the $K_{OW}$ values for pH effects. When the pH regime of particular treatment plants is known, the use of pH-adjusted $K_{OW}$ (i.e., $D_{OW}$) values would be preferred and is recommended, which also changes the $p_{hit}$ value. For example, at pH 7.5 (the average pH of most influent wastewater in the USA†), the new $p_{hit}$ value is 1.76 $\times 10^{-6}$. The pie chart, inset in Fig. 3, shows that ~77% of the HPV chemicals selected in this screening study are used as industrial chemicals. Their log $K_{OW}$ values and mass fractions predicted to become sequestered in digested sludge ranged from 4 to 25 and 6 to 100%, respectively. Five of these industrial chemicals have log $K_{OW}$ values between 14 and 25. Seventeen percent of the compounds examined fell into the category of flavors and fragrances. About 6% of the compounds are used as fuels and oils. Whereas industrial chemicals are distributed throughout the displayed log $K_{OW}$ range, flavors and fragrances (log $K_{OW} < 7$), and fuels and oils (log $K_{OW} < 5$) showed a more limited range of hydrophobicity.

Chemicals within the category of ‘industrial chemicals’ dominate the total number of HPV chemicals included in this study and showed a wide range of predicted persistence in sludge. These compounds are employed as antioxidatnts, metal chelators, intermediates, by-products, catalysts, flame retardants,
phenylating agents, plasticizers, heat storage and transfer agents, lubricants, solvents, anticorrosive agents, etc. Mass produced chemicals used as flavors and fragrances also are predicted to accumulate in biosolids, with five of these expected to persist in sludge to 50% or greater of their mass loading in raw sewage.

Toxicity assessment of the HPV chemicals predicted to accumulate in sludge

We used literature information to evaluate the toxicity of HPV chemicals identified in the initial screening approach (Fig. S2†). Information on toxic threshold values was limited. Most of the information available referred to the lethal dose producing a 50% kill rate in animal studies (LD50 value). Among the HPV chemicals considered, suitable LD50 values were available for 156 chemicals. Reported values ranged from 196 to 62 080 mg kg−1/C0 with a mean of 7089 mg kg−1/C0. The toxicity values showed no correlation with chemical persistence during wastewater solids treatment (Fig. S2†). We examined the dataset for potentially problematic compounds featuring a low LD50 value (high toxicity) and a high (>50%) accumulation in sludge.

According to accepted guidelines,16 we grouped the LD50 values (mg kg−1) into four categories, which yielded no compounds of high toxicity (LD50 < 50), six compounds of moderate toxicity (50 ≤ LD50 < 500), 55 compounds of low toxicity (500 ≤ LD50 < 5000), and 95 compounds of very low toxicity (LD50 ≥ 5000).

Four HPV chemicals judged to be of moderate toxicity were predicted to accumulate to >50% in sludge. These include 2,4-di-tert-pentylphenol, triphenylborane, N,N,N',N'-tetrabutylhexane-1,6-diamine and a reaction product of 4-methylphenol with dicyclopentadiene and isobutylene (Fig. 4B). All of these chemicals are used primarily as starting materials (precursors) for the production of other industrial chemicals or products. Thus, if an environmental path to wastewater exists, it likely is dominated by effluent of wastewaters from the chemical manufacturing industry. The occurrence of 2,4-di-tert-pentylphenol, for example, has been reported for sediments obtained from throughout Lake Erie, including in dated sediment cores reflective of chemicals deposited in the mid-1930s to the late 1980s.20 Their great hydrophobicity (log Kow range of 5.52 to 7.67) suggests, however, that their bioavailability in sediments may be diminished which would significantly reduce their potential to cause harm to aquatic organisms or human populations.

Assessment of organohalogen HPV chemicals

We also analyzed the HPV chemicals that are organohalogen since most of them have a high tendency to persist in the environment.22 Altogether seven organohalogen compounds were predicted to accumulate in biosolids to more than 50% of their initial mass loading. Of these, five contained bromine, and one each contained chlorine and fluorine substituents (Fig. 4A). All of them were predicted to become sequestered in sludge at yields of >70%. Their environmental half-lives, using the EPA’s PBT Profiler prediction software, were estimated to range from 120 to 360 days in soil (Fig. 4, panel B). The pronounced persistence of these compounds in both engineered and natural environmental systems could be of concern.

The potentially problematic halogenated organic compounds identified in this study (Fig. 4A) find wide industrial applications as components in plastics, electrical and electronic equipment, and as flame retardants. The latter use may be particularly problematic because of the very large volumes used and the relatively strong persistence of these compounds. In 2004, organobromines accounted for 25% of the flame retardants used globally.23 With their widespread use and persistent behavior, brominated flame retardants are destined to occur in abiotic and
biotic samples. For example, 1,2-bis(2,4,6,tribromophenoxy)ethane (Fig. 4A, CAS# 37853-59-1) has been reported in bird tissues, in egg pools of herring gulls, and in dust samples. The ubiquity of brominated flame retardants demands additional investigations into their toxicological and ecological risks. Monitoring of biosolids for these and other compounds identified in this screening study may aid in identifying potential problems early and is essential for understanding the entry and mass flow of these compounds in the environment.

Conclusions

This study served to validate a previously proposed empirical model that performed favorable despite requiring only two simple input values, a chemical’s log $K_{\text{OW}}$ value and a universal, dimensionless, non-linear curve fitting parameter, $P_{\text{fit}} (6.51 \pm 0.585 \times 10^{-4})$. An assessment of 207 HPV chemicals indicated that two thirds of these compounds are projected to accumulate in digested sludge to greater than 50% by mass relative to their...
initial loading in raw sewage received by conventional wastewater treatment plants. Further evaluation revealed that four of these chemicals are moderately toxic, while seven were found to represent halogenated compounds with half-lives in soil estimated to range from 120 to 360 days. Compounds of concern identified in this study include the nonhalogenated, moderately toxic compounds 2,4-di-tert-pentylphenol (CAS# 120-95-6), triphenylboron (CAS# 960-71-4), N,N',N'-tetradecylhexane-1,6-diamine (CAS# 27090-63-7), and reaction products of 4-methylphenol with dicyclopentadiene and isobutylene (CAS# 68610-51-5). Organohalogen compounds of concern identified include 1,2-bis(2,4,6-tribromophenoxy)ethane (CAS# 37853-59-1), Dechlorane 605 (CAS# 13560-89-9), hexabromocyclododecane (CAS# 3194-55-6), bromophthal (CAS# 632-79-1), 59-1), Dechlorane 605 (CAS# 13560-89-9), hexabromocy-

This exemplary application illustrates the model’s value as an in silico screening tool for identifying potentially persistent compounds that are prone to accumulate in municipal biosolids destined to be spread on land. The principal benefit of the model is to narrow down the expansive list of mass-produced chemicals of potential concern that may require additional laboratory investigations to ensure their environmental safety.

However, since chemical predictions are not always accurate, particularly when stemming from simplistic models such as the one employed here, it is essential to follow up on any type of in silico screening with actual environmental monitoring studies. These often will require the development and application of new analytical methods that are suitable for affirming or disproving the results of predictions obtained in silico.

Acknowledgements
This study was supported in part by grant 1R01ES015445 by the National Institute of Environmental Health Sciences (NIEHS) and by the Johns Hopkins University Center for a Livable Future. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIEHS or the National Institutes of Health.

References