



# On the replacement of empirical parameters in multimedia mass balance models with QSPR data

Tomasz Puzyn\*

University of Gdansk, Faculty of Chemistry, Laboratory of Environmental Chemometrics, ul. Sobieskiego 18/19, 80-952 Gdansk, Poland

## ARTICLE INFO

### Article history:

Received 26 October 2010

Received in revised form 18 April 2011

Accepted 24 May 2011

Available online 17 June 2011

### Keywords:

Quantitative structure–property relationships

QSPR

Multimedia models

Long range transport potential

Overall persistence

## ABSTRACT

Official OECD recommendations give the highest priority to application of purely empirical phys/chem data (partition coefficients, environmental half-live times etc.) in multimedia mass balance modeling of environmental overall persistence and long-range transport for potentially hazardous chemicals. We have demonstrated that the replacement of the empirical data with those predicted by employing Quantitative Structure–Property Relationships (QSPR) technique did not significantly decrease the performance of *The Tool 2.0* – the OECD multimedia mass balance model. To prove this, we compared each other the output results (overall persistence –  $P_{OV}$ ; characteristic traveling distance –  $CTD$  and transport efficiency  $TE$ ) obtained from 6 of multimedia models. The models utilized combinations of experimentally determined and QSPR-predicted values of the partition coefficients and half-live times. For predicting phys/chem data, we utilized 2 QSPRs developed in our laboratory and the EPI Suite package (US EPA). We did not observe any statistically significant ( $p < 0.05$ ) differences between the models. This conclusion is important, because it leads to reducing time and costs of laboratory studies required during the risk assessment procedure. Moreover, regarding the obtained results, we proposed to replace the single-threshold approaches established by majority of international regulations to screen substances for persistence, bioaccumulation and long-range transport potential with the approaches taking into account uncertainty of the results and/or probability of passing a given threshold.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Multimedia mass balance (MM) models play an important role in the exposure assessment for variety of environmental pollutants. Among others, the Eulerian, so-called ‘box’ models, in which the environment is segmented into a number of spatially homogeneous boxes, representing particular environmental components (e.g., air, water, soil etc.), are of special meaning. They can be employed to estimate the long-range transport potential (LRTP) and the overall persistence ( $P_{OV}$ ) of a chemical in relation to its intrinsic properties [1,2].

Multimedia models have been widely used for characterizing mainly hydrophobic Persistent Organic Pollutants (POPs). Members of the POPs family (i.e., polychlorinated biphenyls, polychlorinated naphthalenes, polychlorinated dibenzo-*p*-dioxins, pesticides etc.) characterize by high environmental persistence, lipophilicity and, in effect, high bioaccumulation potential, which – in combination with confirmed toxicity – makes them hazardous for human and the

environment [3]. POPs became also ‘model compounds’ for creating a PBT category of chemicals (P: persistent, B: bioaccumulative, T: toxic) established in many international regulations, including the European REACH system [4,5].

For a given chemical the MM models require a set of parameters that characterize its chemical nature. Typically, for hydrophobic POPs, they are: *n*-octanol/water partition coefficient ( $K_{OW}$ ), *n*-octanol/air partition coefficient ( $K_{OA}$ ), air/water partition coefficient ( $K_{AW}$ ) and measures of the persistence in particular media, such as half-lives ( $t_{1/2}$ ) or degradation rate constants ( $k$ ) in air, water, soil and/or sediments [6]. These input data are usually determined empirically and many of them are available in various databases, for instance in the excellent handbook published by Mackay et al. [7].

However, the number of novel chemicals (potential environmental pollutants) is rapidly increasing and records of the databases are becoming limited. Thus, modelers are forced to calculate (interpolate) the parameters and use them instead of the empirical ones. Among a vast range of theoretical methods for predicting the parameters, one has found the largest number of applications: Quantitative Structure–Property Relationships (QSPR) technique. The QSPR approach is based on mathematical dependencies between the variance in molecular structures, encoded by so-called molecular descriptors (e.g., number of atoms,

\* Corresponding author. Tel.: +48 58 523 5451; fax: +48 58 523 5472.  
E-mail address: [t.puzyn@qsar.eu.org](mailto:t.puzyn@qsar.eu.org)

dipole moment etc.), and the variance in a given physicochemical property (e.g.,  $\log K_{OA}$ ) within a set of similar compounds. Thus, whenever data are available only for a part of this group, one is able to interpolate the lacking data from the molecular descriptors and a suitable mathematical model [8]. The examples of QSPRs capable to predict the properties of interest are fragment-based models implemented in the EPI Suite package designed by the US Environmental Protection Agency [9] and many models published by individual authors, for example [10–13].

There is an important question, how the origin of the physicochemical properties influences the results of multimedia mass balance modeling? According to the official recommendations by OECD [14,15], the empirical data should be utilized always when possible. But, on the other hand, if the replacement of the empirical data with the QSPR-calculated ones does not significantly decrease the performance of multimedia models, why not to reduce time and cost of laboratory studies? This contribution was aimed to give answers on these two practical questions.

## 2. Materials and methods

### 2.1. Multimedia model

In this investigation, we employed The OECD  $P_{OV}$  and LRTP Screening Tool, ver. 2.2 (short: *The Tool*) – the software officially recommended by OECD and publicly available on the website: [http://www.oecd.org/document/17/0,3343,en\\_2649\\_34373\\_40754961\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/17/0,3343,en_2649_34373_40754961_1_1_1_1,00.html). *The Tool* software includes a fugacity-based steady-state multimedia mass balance model, in which the environment is represented by three bulk compartments: soil surface layer, seawater surface layer, and troposphere.

The model requires partition coefficients and persistence data (in terms of half-lives) as inputs. The outputs from the model are: overall persistence and two measures of LRTP: characteristic travel distance and transfer efficiency. Overall persistence ( $P_{OV}$ , in days) is a measure of the chemical's degradation time in the whole environment. It is calculated according to the formula (1):

$$P_{OV} = \frac{M_{TOT}}{F_{DEG,A} + F_{DEG,W} + F_{DEG,S}} \quad (1)$$

where  $M_{TOT}$  is the total mass at steady-state (kg) and  $F_{DEG,A}$ ,  $F_{DEG,W}$  and  $F_{DEG,S}$  are degradation mass fluxes (in kg/h) in air (A), water (W) and soil (S), respectively. Characteristic travel distance (CTD, in km) is a transport-oriented LRTP indicator that quantifies the distance from the point of release to the point at which the concentration decreases to  $1/e$  (about 37%) of its initial value. It is calculated as (2):

$$CTD = \frac{M_{TOT}}{F_E} \times \frac{M_i}{M_{TOT}} \times v \quad (2)$$

where  $F_E$  is the emission mass flux,  $M_i/M_{TOT}$  is the (dimensionless) mass fraction in the mobile medium (the same as the medium that receives the emissions), and  $v$  (km/h) is the transport velocity. Transfer efficiency (TE, in percents) is a target-oriented LRTP measure that describes the percentage of the emitted amount of the substance capable to reach a give target (other part of the environment, when we divide the whole model environment into two symmetrical parts and assume that the substance was released in the first part). The value of TE is calculated according to Eq. (3):

$$TE = \frac{F'_D}{F_E} \times 100\% = \frac{F_D \times F_A}{F_E^2} \times 100\% \quad (3)$$

where  $F_A$  is the advective long-range transport flux and  $F_D$  and  $F'_D$  are deposition fluxes in the first part (the part of emission) and the second part (the target) of the model environment. More detailed description of the model can be found either in the manual or in

**Table 1**

Description of the multimedia models compared within the study.

Model	Origin of the partition coefficients	Origin of the persistence data
I	Experimental	Experimental
II	Predicted with L-QSPR models	Experimental
III	Predicted with EPI-Suite	Experimental
IV	Experimental	Predicted with PCA-kNN QSPR model
V	Predicted with L-QSPR models	Predicted with PCA-kNN QSPR model
VI	Predicted with EPI-Suite	Predicted with PCA-kNN QSPR model

the contribution currently published by the authors of the model [16].

### 2.2. Comparison scheme

The main idea of the present study was to compare the output results ( $P_{OV}$ , CTD and TE) obtained from a series of multimedia models (*The Tool* software) that based of combinations of experimentally determined and QSPR-predicted values of the partition coefficients and the half-live times (Table 1). We tried to verify, if there were any significant differences between the model utilizing experimental data as inputs and the models that utilized the computationally predicted properties instated. Since the model I (Table 1) was based only on the experimental input data, we treated that as a reference model, and systematically compared the output from I with the outputs calculated for the rest of models (II–VI). The pairwise comparisons were performed with standard statistical procedures, including pairwise Student's  $t$ -test, and the analysis of rank correlations based on  $\rho$ -Spearman's and  $\tau$ -Kendall's correlation coefficients. The comparisons performed within the present study were based on data for the family of POPs.

### 2.3. Experimental data

Experimental values of the partition coefficients utilized to develop multimedia models nos. I and IV were collected from the Mackay's handbook [7] and also from five peer-reviewed papers [17–21]. For details, please refer to the electronic supplementary data file (SD). Experimental data on persistence in air, water and soil (used in models nos. I–III) were taken from the handbook [7]. Note that the real half-live value for a compound depends not only on its intrinsic properties, but also on the environmental conditions. Since the conditions (sunlight intensity, microbial population etc.) are different in different regions of the earth, it is impossible and misleading to document a concrete, reliable half-live value, for instance, in water. Instead of measuring a single  $t_{1/2}$  value, Mackay et al. [7] proposed to divide the compounds into semi-quantitative classes by giving only a range of the particular half-live ( $t_{1/2}$  in air, water, soil and sediment) for each class. For the purpose of multimedia modeling, the mean value calculated from the minimum and maximum  $t_{1/2}$  in a particular class can be used, because this is the most probable value when assuming the normal distribution of the data.

### 2.4. L-QSPR models

Two linear QSPR (L-QSPR) models were adapted from our previous contribution [13]. The adapted models (Eqs. (4) and (5)) predicted the values of  $\log K_{OW}$  and  $\log K_{OA}$  for 1436 chloro- and bromo-analogues of various POPs, including: dibenzo-*p*-dioxins, dibenzofurans, biphenyls, naphthalenes, diphenyl ethers and benzenes. They utilized four molecular descriptors ( $\mu$ : dipole moment, SAS: solvent accessible surface area,  $\epsilon_{HOMO}$ : energy of the highest

occupied molecular orbital, and  $\alpha$ : average polarizability) calculated at the semi-empirical level of the quantum-mechanical theory with very accurate PM6 method [22]:

$$\log K_{OW} = -0.3587 - 0.1220\mu + 0.0247SAS \quad (4)$$

$$\log K_{OA} = 7.3108 + 0.7408\varepsilon_{HOMO} + 0.2862\alpha \quad (5)$$

The third partition coefficient ( $\log K_{AW}$ ) was calculated as a difference between the logarithmic values of  $K_{OW}$  and  $K_{OA}$ . In the same previous work [13], by employing hierarchical clustering method, we identified five classes of similar partitioning in the environment. In the present study, the predicted values of  $\log K_{OW}$  and  $\log K_{AW}$  were used as inputs for the multimedia models nos. **II** and **V**, and the 5 partitioning classes – to select representative compounds for the multimedia models' comparison.

### 2.5. PCA-kNN model

Multimedia models **IV–VI** utilized persistence data originated from the other preceding work [11]. QSPR models developed in [11] with  $k$ -Nearest Neighbors classifier combined with Principal Component Analysis (PCA-kNN) delivered data on the half-live times ( $t_{1/2}$ ) of brominated and chlorinated POPs in air, water, soil and sediments. PCA-kNN classifications were based on 24 molecular descriptors of two types: constitutional and quantum-mechanical ones (from semi-empirical PM6 method). The persistence data were expressed in a semi-quantitative mode, similar to these originally developed by Mackay et al. (see Section 2.3). Note that in the present study, we took the average value calculated from the minimal and maximal  $t_{1/2}$  in each class to have a single  $t_{1/2}$  value required for multimedia modeling with *The Tool* software.

### 2.6. EPI Suite software package

Multimedia models **III** and **VI** based on the partition coefficients calculated with the EPI Suite package. The package designed by the US EPA is publicly available free of charge on the agency's website and its detailed description can be found elsewhere [9].

### 2.7. Selection of the representative test compounds for comparing the performance of the studied multimedia models

For making the comparisons of multimedia models within this study more reliable, we choose a set of test compounds having various partitioning characteristics ( $\log K_{OW}$  and  $\log K_{OA}$ ). The selection procedure was as follows. At first, we extracted only those compounds, for which the experimental values of both partition coefficients have been available. Then, by applying the fast version of the Kennard–Stone algorithm [23], we tried to select three representative compounds from each of the partitioning classes, predefined in [13]. The Kennard–Stone algorithm is designed to select the specified number of the most dissimilar (based on the values of the partition coefficients in this case) compounds. In effect, the selected compounds were evenly distributed in the space defined by the partition coefficients.

## 3. Results

### 3.1. Representative test compounds

Initially, we selected 42 compounds, for which the experimentally measured values of both  $\log K_{OW}$  and  $\log K_{OA}$  have been available (for details please refer to SD). Since we have not found such data for the class containing the least mobile substances (according to the classification from [13]), we decided to omit that class in our current research. Then, by applying the Kennard–Stone

**Table 2**

Comparison of the experimentally determined (Exp) with calculated (L-QSPR and EPI Suite) values of the partition coefficients by pairwise Student's  $t$ -test ( $\alpha = 0.050$ ).

Exp. vs. L-QSPR	Exp. vs. EPI Suite
$\log K_{OW}$	
$\bar{d} = 0.24$	$\bar{d} = -0.30$
$t = 1.939$	$t = 1.690$
$p = 0.078$	$p = 0.119$
$\log K_{OA}$	
$\bar{d} = -0.06$	$\bar{d} = 0.15$
$t = 0.676$	$t = 0.807$
$p = 0.513$	$p = 0.437$

$\bar{d}$  = average residual value.

algorithm, we selected three representative compounds per each of the four remaining partitioning classes (predefined in [13]). In effect, the finally selected set of test compounds was evenly distributed in the space of the partition coefficients (Fig. 1). We utilized these test compounds for comparing the performance of the studied multimedia models.

### 3.2. Partitioning data

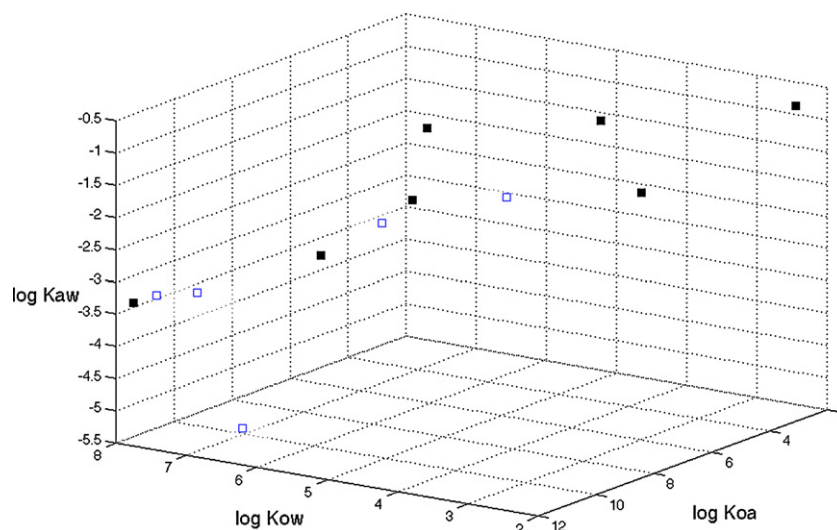
At first, we compared the values of partition coefficients determined empirically with those predicted by employing the L-QSPR and EPI Suite models (Table 2). The L-QSPR model usually underestimated the values of  $\log K_{OW}$  (the average residual  $\bar{d} > 0$ ), whereas it overestimated the values of  $\log K_{OA}$  (the average residual  $\bar{d} < 0$ ). Interestingly, EPI Suite predictions, in most cases, were characterized by the opposite trend: overestimated  $\log K_{OW}$  and underestimated  $\log K_{OA}$ . Nevertheless, the observed differences were not statistically significant ( $p > 0.050$ ).

Logarithmic values of the third partition coefficient ( $\log K_{AW}$ ) were calculated as follows:

$$\log K_{AW} = \log K_{OW} - \log K_{OA} \quad (6)$$

One can ask, if there is any important difference in the values of  $\log K_{AW}$  calculated this way and the values measured/modelled directly. Theoretically, there should be, regarding that uncertainties of  $\log K_{OW}$  and  $\log K_{OA}$  measuring/modeling are propagated on the calculated  $\log K_{AW}$ . However, when compared the calculated  $\log K_{AW}$  values to the values measured empirically (Table 3), and when performed paired Student's  $t$ -test, we found those differences insignificant ( $t = 0.842$ ,  $p = 0.42$ ). Note that, regarding the lack of directly measured values of  $\log K_{AW}$  in the literature, the comparison was performed based on nine compounds only. Also interestingly, dissimilarity between the measured and calculated value for the last compound (octachloronaphthalene) was very high – equal to 3.71. This can be explained by the existence of two extremely different, but both empirically measured, values of  $\log K_{OA}$ . The first one ( $\log K_{OW} = 6.42$ ) is recommended by Opperhuizen et al. [17]. This value has been utilized in our study, since it correlates better with the values for other, less chlorinated naphthalene congeners. The second one ( $\log K_{OW} = 8.50$ ), provided by Lei et al. [24], fits better to the empirically determined value of  $\log K_{AW}$  given in Table 3.

The lowest and the highest values of  $\log K_{AW}$  calculated from L-QSPR-predicted  $\log K_{OW}$  and  $\log K_{OA}$  (Table 3) were underestimated, whereas the  $\log K_{AW}$  values for the rest of compounds were slightly overestimated. However, both the average residual value and the variance were relatively low. On the contrary, the residuals variance of  $\log K_{AW}$  obtained from  $\log K_{OW}$  and  $\log K_{OA}$  previously predicted with EPI Suite was ten times higher; the residuals varied from  $-3.72$  to  $0.69$ . By applying ANOVA we confirmed that the residuals characterizing both methods of  $\log K_{AW}$  calculation



**Fig. 1.** Distribution of the representative compounds in the space of the partition coefficients. Empty squares (□) indicate representative compounds, for which only experimental partitioning data are available (experimental persistence data are unavailable).

(based on QSPR and based on EPI Suite predictions) differed significantly ( $F = 5.387$ ,  $p = 0.030$ ).

The observations above suggest that L-QSPR and EPI Suite provide partitioning data of different internal consistency. The problem of lacking consistency in partitioning data, i.e., (7):

$$\log K_{AW} - \log K_{OW} + \log K_{OA} \neq 0 \quad (7)$$

has been discussed in details by many authors [25,26]. Recently, Schenker et al. [26] proposed a least squares adjustment procedure for harmonizing physicochemical properties of organic chemicals. We employed this procedure to verify internal consistency of L-QSPR and EPI-Suite predicted data. For this exercise, we took partition coefficients ( $\log K_{OW}$  and  $\log K_{OA}$ ) previously predicted with L-QSPR and EPI Suite and  $\log K_{AW}$  calculated from empirically measured  $\log K_{OW}$  and  $\log K_{OA}$ . Then, from calculating the final adjusted values (FAVs) according to the least square procedure and comparing them with the original ones, we were able to make conclusions about the consistency of both data sets. Majority of data from L-QSPR required the adjustment of the moderate level (Fig. 2a) – similar to that reported in the original work by Schenker et al. [26] for PCBs. Only in case of octachloronaphthalene, the adjustment was higher than 90%, and this was related to the problem discussed above (very different values of exper-

imentally measured  $\log K_{OW}$  recommended by two authors). In contrast, data predicted with EPI Suite (Fig. 2b) were very consistent for majority of compounds (the adjustment of less than 20%). However, the adjustments required for three chemicals (including octachloronaphthalene) were extremely high (between 132% and 1223%). Thus, the predictions from statistically based L-QSPR models led to slightly higher, but more systematic, inconsistency of the data set, whereas the predictions from the fragment-based method implemented in EPI Suite, for some compounds gave almost perfect consistency of the data set, but data for others were highly inconsistent. This means, from practical viewpoint, that the EPI Suite data for a series of compounds, when not adjusted, would influence multimedia modeling results much stronger than not adjusted data from L-QSPR.

### 3.3. Persistence data

Similarly to the previous step, we compared experimentally determined data on persistence (expressed in terms of  $t_{1/2}$ ) in air, water, soil and sediments with data estimated by employing PCA-kNN QSPR model. Note, for 5 compounds (indicated with empty squares on Fig. 1), there were no experimental data on persistence ( $t_{1/2}$ ) available. In consequence, these cases had to be excluded

**Table 3**  
Comparison of the experimental and calculated values of  $\log K_{AW}$ .

Compound	Experimental		L-QSPR		EPI-Suite	
	Meas.	Calc.	Pred.	Res.	Pred.	Res.
Benzene	-0.64	-0.67	-0.47	-0.20	-0.78	0.11
1,2,4-Trichlorobenzene	-1.24	-0.91	-1.16	0.25	-1.60	0.69
1,2,3,4,5,6-Hexachlorobenzene	-1.16	-1.04	-1.74	0.70	-1.03	-0.01
2,2',5-Trichlorobiphenyl	-1.99	-2.00	-2.62	0.62	-1.85	-0.15
2,3,3',4',6-Pentachlorobiphenyl	n/a	-2.76	-3.21	0.45	-1.66	-1.10
1,2,3,4,7-Pentachlorodibenzo- <i>p</i> -dioxin	-3.98	-3.27	-3.82	0.55	-3.82	0.55
1,2,3,4,7,8-Hexachlorodibenzo- <i>p</i> -dioxin	-3.79	-3.31	-3.94	0.63	-3.38	0.07
1,2,3,4,6,7,8-Heptachlorodibenzo- <i>p</i> -dioxin	-2.15	-3.42	-4.25	0.83	-1.50	-1.92
Naphthalene	-1.75	-1.86	-1.89	0.03	-1.88	0.02
1,5-Dichloronaphthalene	n/a	-1.95	-2.17	0.22	1.77	-3.72
1,4,6,7-Tetrachloronaphthalene	n/a	-2.32	-2.75	0.43	-2.25	-0.07
Octachloronaphthalene	-1.35	-5.06	-4.19	-0.87	-1.70	-3.36
$\bar{d}$				0.30		-0.74
$s^2$				0.20		2.02

Meas.: measured directly, Calc.: calculated according to Eq. (6) based on the empirically measured values of  $\log K_{OW}$  and  $\log K_{OA}$ , Pred.: calculated according to Eq. (6) based on the values of  $\log K_{OW}$  and  $\log K_{OA}$  predicted with an appropriate model, Res.: residual values, and n/a: data not available.



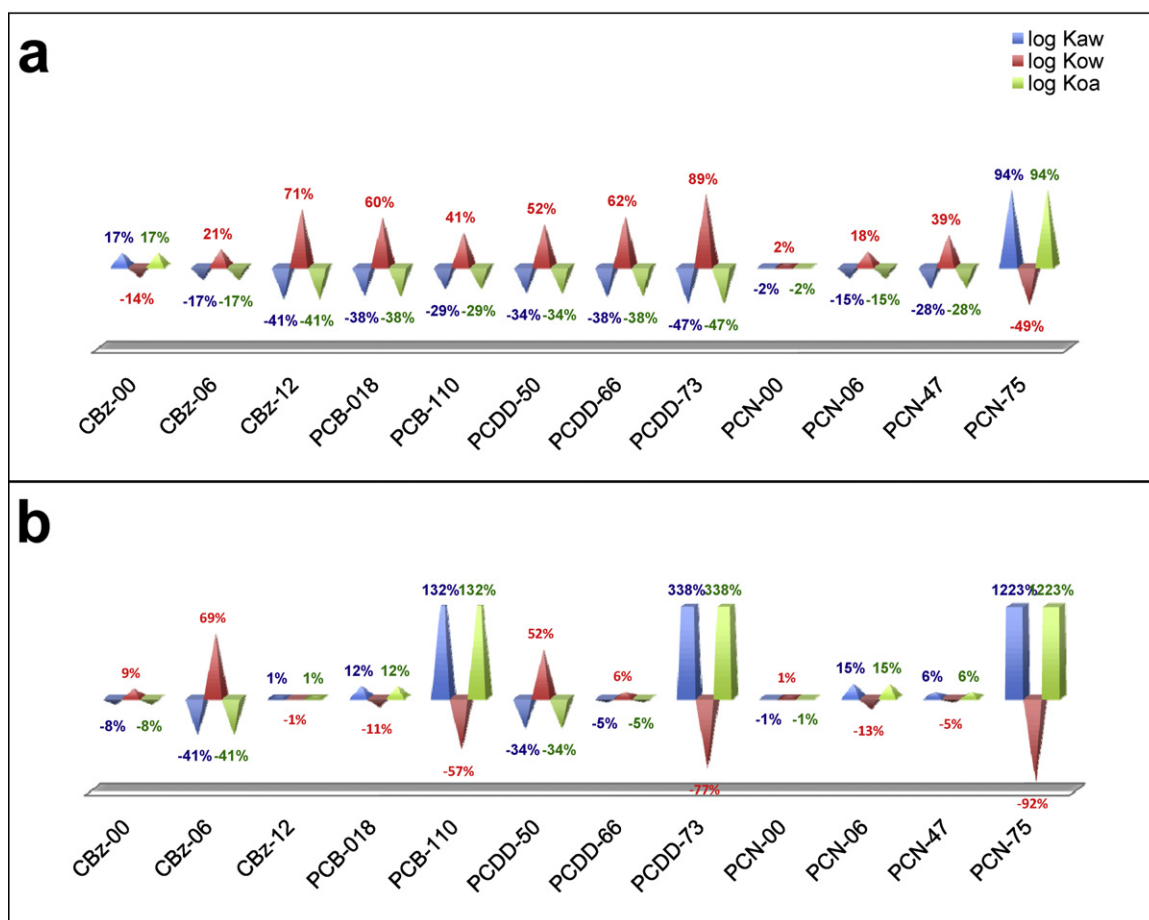


Fig. 2. Least square adjustment of L-QSPR (a) and EPI Suite (b) data.

from the comparison. When applied pairwise Student's  $t$ -test for the remaining compounds, we did not notice any significant ( $p < 0.050$ ) differences between both origins of data. Appropriate statistics were as follows:  $\bar{d} = 389$ ,  $t = 0.953$  and  $p = 0.377$  for  $t_{1/2}$  in air,  $\bar{d} = 703$ ,  $t = 0.520$  and  $p = 0.622$  for  $t_{1/2}$  in water,  $\bar{d} = 5426$ ,  $t = 0.999$  and  $p = 0.356$  for  $t_{1/2}$  in soil, and  $\bar{d} = 5429$ ,  $t = 0.997$  and  $p = 0.357$  for  $t_{1/2}$  in sediments.

### 3.4. Multimedia mass-balance models

Finally, we investigated how the replacement of empirically determined properties with the properties estimated by means of QSPR techniques could influence the results of multimedia mass balance modeling. Since, as mentioned, we had no experimental data on the persistence of 5 compounds, it was impossible to calculate the output values ( $P_{OV}$ , CTD and TE) from the multimedia models I–III for these cases.

We noticed that rank correlations between the outputs from particular multimedia models based on the predicted data (models II–VI) and the outputs from the model utilizing only the experimental properties (model I) were significant in all cases (Fig. 3). However, evidently lower correlation coefficients for  $P_{OV}$  in pairs: I–IV, I–V, and I–VI suggest strong influence of the origin of the persistence data used in modeling (experimental or predicted with PCA-kNN QSPR model) on the overall persistence determined by the multimedia model. Indeed, when *The Tool* utilized the experimentally determined persistence data (models II and III), the rank correlation with model I was almost perfect. On the contrary, the origin of the partition coefficients did not influence the output from *The Tool*. Neither  $P_{OV}$  nor both LRTP measures (CTD and TE) differed

between the models based on the empirical and QSPR-predicted partitioning data. Although one can conclude that the use of L-QSPR-based partition coefficients led to better predictions of CTD and, simultaneously, the data from EPI Suite reproduced better TE, one should remember that the observed slight differences in the values of  $\rho$  and  $\tau$  resulted from only single compounds differing in the ranks (for details please refer to SD). Thus, making such strong recommendations without performing a more comprehensive study might be unjustified.

	$P_{OV}$	CTD	TE
I-II	$\rho = 0.96$ $\tau = 0.90$	$\rho = 1.00$ $\tau = 1.00$	$\rho = 0.96$ $\tau = 0.90$
I-III	$\rho = 1.00$ $\tau = 1.00$	$\rho = 0.96$ $\tau = 0.90$	$\rho = 1.00$ $\tau = 1.00$
I-IV	$\rho = 0.75$ $\tau = 0.62$	$\rho = 0.96$ $\tau = 0.90$	$\rho = 1.00$ $\tau = 1.00$
I-V	$\rho = 0.75$ $\tau = 0.62$	$\rho = 0.96$ $\tau = 0.90$	$\rho = 0.96$ $\tau = 0.90$
I-VI	$\rho = 0.79$ $\tau = 0.71$	$\rho = 0.96$ $\tau = 0.90$	$\rho = 0.99$ $\tau = 0.98$

Fig. 3. Rank correlations between the outputs from multimedia models based on the predicted phys/chem. data (II–VI) and the outputs from the model utilizing only the experimental properties (I).

**Table 4**Comparison of  $P_{OV}$ , CTD and TE in pairs of the models by pairwise Student's  $t$ -test ( $\alpha=0.050$ ).

I-II	I-III	I-IV	I-V	I-VI
$P_{OV}$				
$\bar{d} = -23$	$\bar{d} = 16$	$\bar{d} = -314$	$\bar{d} = -348$	$\bar{d} = -297$
$t = 0.4198$	$t = 0.3842$	$t = 0.9826$	$t = 1.1028$	$t = 0.9189$
$p = 0.6893$	$p = 0.7140$	$p = 0.3638$	$p = 0.3124$	$p = 0.3936$
CTD				
$\bar{d} = 13,770$	$\bar{d} = -1538$	$\bar{d} = 1936$	$\bar{d} = 12,211$	$\bar{d} = 404$
$t = 1.0861$	$t = 1.1614$	$t = 0.8391$	$t = 1.0710$	$t = 0.1332$
$p = 0.3191$	$p = 0.2896$	$p = 0.4336$	$p = 0.3254$	$p = 0.8984$
TE				
$\bar{d} = -49.14$	$\bar{d} = 16.57$	$\bar{d} = 12.48$	$\bar{d} = -69.22$	$\bar{d} = 28.95$
$t = 0.9990$	$t = 1.0828$	$t = 1.0063$	$t = 0.9900$	$t = 1.5184$
$p = 0.3564$	$p = 0.3205$	$p = 0.3531$	$p = 0.3604$	$p = 0.1797$

 $\bar{d}$  = average residual value.

In addition to the rank analysis, we have compared the particular values of  $P_{OV}$ , CTD and TE in pairs of the models (I-II, I-III, I-IV etc.). Interestingly, we did not notice statistically significant ( $p < 0.05$ ) differences based on pairwise Student's  $t$ -test (Table 4). The observation leads to the conclusion that differences between the multimedia modeling results based on the experimentally derived and QSPR-predicted phys/chem data, although statistically insignificant, could influence the order of compounds in the ranking. Thus, in general, ranking-based methods in such exercises are much more sensitive measures than the conventional, parametric statistics.

#### 4. Discussion

A year ago Wittekindt and Goss [27] initiated an interesting discussion on quality of partitioning data derived from various available software packages and their impact on the results of high throughput screening for persistent (P), bioaccumulative (B) and toxic (T) chemicals. The authors compared each other the values of  $\log K_{OW}$  and  $\log K_{AW}$  predicted for 8560 compounds with both EPI Suite and COSMOtherm computer programs. COSMOtherm is quantum-chemistry-based software, successfully employed for calculating equilibrium partitioning and solving various environmental problems [28,29]. They noticed a root mean square error (RMSE) of 0.7 log units in the predicted  $\log K_{OW}$  and RMSE of 1.8 log units in the predicted  $\log K_{AW}$ . Wittekindt and Goss explained such a high variation of  $\log K_{AW}$  by the fact that EPI Suite predictions, from definition, were based on the two-dimensional representation of the molecular structure, whereas COSMOtherm accounts also information of the three-dimensional structure and stereochemistry. More recently, Zhang et al. [30] performed an extensive comparison of the partitioning properties of 529 chemicals, derived with four prediction methods: EPI Suite, COSMOtherm, SPARC, and ABSOLV. SPARC is a freely available online phys/chem calculator (<http://ibmlc2.chem.uga.edu/sparc/>) hosted by the University of Georgia and, it seems, currently competing with EPI Suite [31,32]. The ABSOLV module is a part of ADMETBoxes package designed by Pharma Algorithms, Inc. that calculates solute descriptors (so-called Abraham descriptors). The descriptors can be then utilized to develop poly-parameter free energy relationships (pp-LFER) for predicting  $\log K_{OW}$ ,  $\log K_{AW}$  and  $\log K_{OA}$  [33]. The authors used the four sets of predictions to screen the chemicals against various LRTP and B criteria. As they concluded, screening results were the same for only about 70% of 529 compounds.

However, no one should expect exactly the same results provided by different computational methods, developed from different assumptions and schemes. Even experimental data could be seriously biased because of many reasons [34] and one should remember that the computational models are usually calibrated on

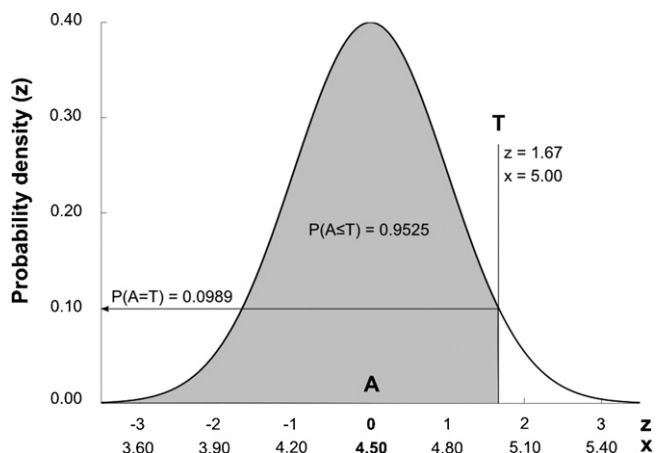
these biased data. Moreover, the predicted data, from definition, will always contain additional noise originated from mathematical methods of interpolation (e.g., the least squares method) employed for development of any QSPR.

On one hand, as we demonstrated, the differences either between the phys/chem data predicted with QSPRs and the experimental values itself or between the outputs from multimedia models ( $P_{OV}$ , CTD, and TE) utilizing those QSPR data and the outputs from the models utilizing only experimentally determined phys/chem data were not statistically significant. On the other hand, those 'insignificant' differences influenced rankings of the compounds according to:  $P_{OV}$ , CTD and TE. Hence, should we trust in the results when the empirical data are replaced with computationally predicted ones or not?

Both, the exercises performed by the cited authors [27,30] and our study demonstrated significant weakness of the assumptions currently used in the risk assessment for POPs and, more generally, for all PBT substances. First, whenever the substances are ranked along with a given property (i.e., P, B, T or LRTP), and the substances are of similar structure (i.e., congeners), it is highly probable than the randomly distributed error that is related to measuring/calculating of this property would have a strong impact on the ranking results. In such cases, we would rather recommend to employ classification or grouping techniques instead of simple rankings. Second, several screening procedures are based on single thresholds criteria. For instance, according to regulations of the Stockholm Convention on POPs [35], chemicals with  $\log K_{OW} > 5$  are regarded as bioaccumulative. As shown, the measured/predicted values of the partition coefficient can vary dependently on the applied method, and the standard variation can be equal to 0.7 log unit or even more. Thus, evidently, classifying compounds with values of  $\log K_{OW}$  very close to 5 as bioaccumulative ones or not is somewhat of a lottery. To solve this problem, as Zhang et al. [30] pointed out, screening approaches should be based on quantifying numerical hazard estimates that take the uncertainty of the predicted properties into account. In addition, we would suggest deriving novel criteria with use of the fuzzy set theory [36,37], in which a compound is classified as – for instance – bioaccumulative with a given probability (e.g., bioaccumulative with probability of 80%, not bioaccumulative with probability of 20%).

The concept of fuzzy classification can be illustrated by a simple example. As mentioned, all chemicals passing the threshold of  $\log K_{OW} = 5$  are labeled as bioaccumulative [35]. According to this criterion, a compound having  $\log K_{OW} = 8$  must, without doubts, be bioaccumulated easily. However, when the  $\log K_{OW}$  value of a chemical is very close to the threshold (for instance when  $\log K_{OW} = 4.5$ ), the classification becomes more uncertain. How to assess probability of passing the threshold in this case?

Let us assume that the residuals in the validation set of the QSPR model for  $\log K_{OW}$  applied in this study [13] are normally



**Fig. 4.** Probability density plot that illustrates the way of calculating probability that the value  $A$  of  $\log K_{OW}$  predicted with QSPR model passes the threshold  $T$ . The values of  $\log K_{OW}$  are provided with use of the original scale ( $x$ ) and after standardization ( $z$ ), according to Eq. (8).

distributed around the zero value. Such type of dispersal proves the random distribution of errors, and that condition should be fulfilled for any regression model. Moreover, it is better to analyze the errors within the external validation set to avoid overoptimistic conclusions, in case they base only on the compounds utilized for calibrating the model (training set). Then, of course, we can calculate the standard deviation of residuals ( $s_E = 0.30$ ) that characterizes the distribution.

In the next step, we can simulate probability of obtaining the values of  $\log K_{OW}$  that differ from that of 4.50 by 1, 2 and 3 standard deviations  $s_E$  (Fig. 4). Finally, based on the tabularized values of the probability density function and of the cumulative distribution function for the standardized normal distribution, we are able to calculate probability of passing the threshold. The calculations are as follows:

- 1) Both the predicted value of  $\log K_{OW} = 4.50$  (here denoted as  $A$ ) and the value of the  $\log K_{OW}$  threshold = 5.00 (here denoted as  $T$ ) should be standardized first according to (8):

$$z = \frac{x - A}{s_E} = \frac{x - 4.50}{0.30} \quad (8)$$

The standardized value of  $A = 0$  (expressed in  $z$ -scale). The standardized value of the threshold  $T = 1.67$ , which means that  $T$  differs from  $A$  by 1.67 standard deviations ( $s_E$ ).

- 2) Directly from the probability density function (Fig. 4), we can derive probability that  $A = T$  (9):

$$P(A = T) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{T^2}{2}\right) = 0.0989 \quad (9)$$

- 3) From definition, probability that  $A \leq T$  is equal to the value of the cumulative distribution function, which is the integral from minus infinity to  $T$  (10):

$$P(A \leq T) = \int_{-\infty}^T \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = 0.9525 \quad (10)$$

- 4) Therefore, probability that  $A \geq T$  can be calculated as (11):

$$\begin{aligned} P(A \geq T) &= P(A > T) + P(A = T) = 1 - P(A \leq T) + P(A = T) \\ &= 1 - 0.9525 + 0.0989 = 0.1464 \end{aligned} \quad (11)$$

Finally, we can provide the result in a way of the fuzzy set theory, concluding that the compound, for which predicted

$\log K_{OW} = 4.50$  is bioaccumulative with probability of 85% and non-bioaccumulative with probability of 15%.

## 5. Conclusions

In this contribution, we tried to answer the question: How the origin of the physicochemical properties influences the results of multimedia mass balance modeling? We demonstrated that the replacement of the empirical data with the QSPR-predicted ones did not significantly decrease the performance of *The Tool* multimedia model. We are conscious our study was performed on only one multimedia model (*The Tool*) with a limited number of compounds. However, since it was possible to successfully substitute the empirical data with the QSPR-predicted ones in one case, it would be also possible to do the same in case of other models. Thus, we strongly recommend revising the OECD recommendations that give the highest priority to the application of purely empirical data in environmental mass balance modeling. Moreover, regarding that data of both types (empirical and predicted) are always determined with uncertainty, we suggest to modify the screening approaches implemented in majority of international regulations. The approaches based on single thresholds might be changed for those taking into account uncertainty of a given property and/or enabling to express probability that the substance passed the threshold or not. The last condition could be fulfilled by employing the fuzzy sets theory. Our observations and proposals in that matter are consistent with previously published works.

## Acknowledgements

Tomasz Puzyn thanks Mrs. Aleksandra Morstrag-Szlichtyng for her valuable comments on the text. The author is grateful to the Foundation for Polish Science for granting him with a fellowship and a research grant in frame of the HOMING Program supported by Norwegian Financial Mechanism and EEA Financial Mechanism in Poland. This work was supported by the Polish Ministry of Science and Higher Education, grant no. DS/8430-4-0171-1.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhazmat.2011.05.078.

## References

- [1] J. Klasmeier, M. Matthies, M. Macleod, K. Fenner, M. Scheringer, M. Stroebe, A.C. Le Gall, T. McKone, D. Van De Meent, F. Wania, Application of multimedia models for screening assessment of long-range transport potential and overall persistence, *Environ. Sci. Technol.* 40 (2006) 53–60.
- [2] A. Hollander, M. Scheringer, V. Shatalov, E. Mantseva, A. Sweetman, M. Roemer, A. Baart, N. Suzuki, F. Wegmann, D. van de Meent, Estimating overall persistence and long-range transport potential of persistent organic pollutants: a comparison of seven multimedia mass balance models and atmospheric transport models, *J. Environ. Monit.* 10 (2008) 1139–1147.
- [3] K.C. Jones, P. de Voogt, Persistent organic pollutants (POPs): state of the science, *Environ. Pollut.* 100 (1999) 209–221.
- [4] H. Sanderson, P. Fauser, M. Thomsen, P.B. Sorensen, PBT screening profile of chemical warfare agents (CWAs), *J. Hazard. Mater.* 148 (2007) 210–215.
- [5] Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, and amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, in: *Official Journal of the European Union*, L 396/1, The European Parliament and the Council of the European Union, Brussels, 2006.
- [6] K. Fenner, M. Scheringer, M. Macleod, M. Matthies, T. McKone, M. Stroebe, A. Beyer, M. Bonnell, A.C. Le Gall, J. Klasmeier, D. Mackay, D. Van De Meent, D. Pennington, B. Scharenberg, N. Suzuki, F. Wania, Comparing estimates of persistence and long-range transport potential among multimedia models, *Environ. Sci. Technol.* 39 (2005) 1932–1942.

- [7] D. Mackay, W.Y. Shiu, K.-C. Ma, S.C. Lee, *Handbook of Physical–Chemical Properties and Environmental Fate for Organic Chemicals*, second ed., Taylor & Francis, Boca Raton/London/New York, 2007.
- [8] M.T.D. Cronin, Quantitative structure–activity relationships, (QSARs) – applications and methodology, in: T. Puzyn, J. Leszczynski, M.T.D. Cronin (Eds.), *Recent Advances in QSAR Studies – METHODS and Applications*, Springer, Dordrecht/Heidelberg/London/New York, 2010.
- [9] EPISuite, Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.00, United States Environmental Protection Agency, Washington, DC, USA, 2010.
- [10] E. Papa, P. Gramatica, Screening of persistent organic pollutants by QSPR classification models: a comparative study, *J. Mol. Graphics Modell.* 27 (2008) 59–65.
- [11] T. Puzyn, M. Haranczyk, N. Suzuki, T. Sakurai, Estimating persistence of brominated and chlorinated organic pollutants in air, water, soil, and sediments with the QSPR-based classification scheme, *Mol. Divers.* 15 (2011) 173–188.
- [12] A. Mostrag, T. Puzyn, M. Haranczyk, Modeling the overall persistence and environmental mobility of sulfur-containing polychlorinated organic compounds, *Environ. Sci. Pollut. Res.* 17 (2010) 470–477.
- [13] T. Puzyn, N. Suzuki, M. Haranczyk, How do the partitioning properties of polyhalogenated POPs change when chlorine is replaced with bromine? *Environ. Sci. Technol.* 42 (2008) 5189–5195.
- [14] Report of the OECD/UNEP Workshop on the Use of Multimedia Models for Estimation Overall Environmental Persistence and Long-Range Transport in the context of PBTs/POPs Assessment, Series on Risk Management 36, OECD Environment, Health and Safety Publications, Paris, France, 2002.
- [15] Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-Range Transport, Series on Testing and Assessment 45, OECD Environment, Health and Safety Publications, Paris, France, 2004.
- [16] F. Wegmann, L. Cavin, M. MacLeod, M. Scheringer, K. Hungerbuhler, The OECD software tool for screening chemicals for persistence and long-range atmospheric transport potential, *Environ. Modell. Soft.* 24 (2009) 228–237.
- [17] A. Opperhuizen, E.W. van der Velde, F.A.P.C. Gobas, D.A.K. Liem, J.M.D. van der Steen, Relationships between bioconcentration in fish and steric factors of hydrophobic chemicals, *Chemosphere* 14 (1985) 1871–1896.
- [18] P. Kömp, M.S. McLachlan, Octanol/air partitioning of polychlorinated biphenyls, *Environ. Toxicol. Chem.* 16 (1997) 2433–2437.
- [19] Y.S. Su, Y.D. Lei, G.L. Daly, F. Wania, Determination of octanol–air partition coefficient (K-OA) values for chlorobenzenes and polychlorinated naphthalenes from gas chromatographic retention times, *J. Chem. Eng. Data* 47 (2002) 449–455.
- [20] T. Harner, T.F. Bidleman, Measurement of octanol–air partition coefficients for polycyclic aromatic hydrocarbons and polychlorinated naphthalenes, *J. Chem. Eng. Data* 43 (1998) 40–46.
- [21] T. Harner, N.J.L. Green, K.C. Jones, Measurements of octanol–air partition coefficients for PCDD/Fs: a tool in assessing air–soil equilibrium status, *Environ. Sci. Technol.* 34 (2000) 3109–3114.
- [22] T. Puzyn, N. Suzuki, M. Haranczyk, J. Rak, Calculation of quantum-mechanical descriptors for QSPR at the DFT level: is it necessary? *J. Chem. Inf. Model* 48 (2008) 1174–1180.
- [23] M. Daszykowski, B. Walczak, D.L. Massart, Representative subset selection, *Anal. Chim. Acta* 468 (2002) 91–103.
- [24] Y.D. Lei, F. Wania, W.Y. Shiu, D.G.B. Boocock, HPLC-based method for estimating the temperature dependence of n-octanol–water partition coefficients, *J. Chem. Eng. Data* 45 (2000) 738–742.
- [25] A. Beyer, F. Wania, T. Gouin, D. Mackay, M. Matthies, Selecting internally consistent physicochemical properties of organic compounds, *Environ. Toxicol. Chem.* 21 (2002) 941–953.
- [26] U. Schenker, M. MacLeod, M. Scheringer, K. Hungerbuhler, Improving data quality for environmental fate models: a least-squares adjustment procedure for harmonizing physicochemical properties of organic compounds, *Environ. Sci. Technol.* 39 (2005) 8434–8441.
- [27] C. Wittekindt, K.U. Goss, Screening the partition behavior of a large number of chemicals with a quantum-chemical software, *Chemosphere* 76 (2009) 460–464.
- [28] F. Eckert, A. Klamt, COSMOtherm, Version C2.1, Release 01.05, COSMOlogic GmbH & Co. KG, Leverkusen, Germany, 2005.
- [29] T. Puzyn, A. Mostrag, J. Falandysz, Y. Kholod, J. Leszczynski, Predicting water solubility of congeners: chloronaphthalenes—a case study, *J. Hazard. Mater.* 170 (2009) 1014–1022.
- [30] X. Zhang, T.N. Brown, F. Wania, E.S. Heimstad, K.U. Goss, Assessment of chemical screening outcomes based on different partitioning property estimation methods, *Environ. Int.* 36 (2010) 514–520.
- [31] H.P. Arp, S.T. Droge, S. Endo, W. Giger, K.U. Goss, S.B. Hawthorne, S.A. Mabury, P. Mayer, M.S. McLachlan, J.F. Pankow, R.P. Schwarzenbach, F. Wania, B. Xing, More of EPA's SPARC online calculator—the need for high-quality predictions of chemical properties, *Environ. Sci. Technol.* 44 (2010) 4400–4401.
- [32] R. Boethling, C. Davies, C. Fehrenbacher, E. Lavoie, L. Libelo, E. Sommer, T. Henry, P. Howard, W. Meylan, L. Morlacci, J. Tunkel, T. Webb, K. Mayo, W. Waugh, M. Zeeman, Comment on “More of EPA's SPARC online calculator—the need for high-quality predictions of chemical properties”, *Environ. Sci. Technol.* 44 (2010) 7745 (pp. 7746–7747 author reply).
- [33] T.N. Brown, F. Wania, Development and exploration of an organic contaminant fate model using poly-parameter linear free energy relationships, *Environ. Sci. Technol.* 43 (2009) 6676–6683.
- [34] C. Shunthirasingham, Y.D. Lei, F. Wania, Evidence of bias in air–water Henry's law constants for semivolatile organic compounds measured by inert gas stripping, *Environ. Sci. Technol.* 41 (2007) 3807–3814.
- [35] Stockholm Convention on Persistent Organic Pollutants, United Nations Environment Programme, Geneva, Switzerland, 2001.
- [36] N. Musee, C. Aldrich, L. Lorenzen, New methodology for hazardous waste classification using fuzzy set theory Part II. Intelligent decision support system, *J. Hazard. Mater.* 157 (2008) 94–105.
- [37] N. Musee, L. Lorenzen, C. Aldrich, New methodology for hazardous waste classification using fuzzy set theory Part I. Knowledge acquisition, *J. Hazard. Mater.* 154 (2008) 1040–1051.