Strategies to Reduce Uncertainties from the Best Available Physicochemical Parameters Used for Modeling Novel Organophosphate Esters Across Multimedia Environments

Changyue Xing[1, 2], Jianxin Ge[1, 2], Rongcan Chen[1, 2], Shuaiqi Li[1, 2], Chen Wang[3], Xianming Zhang[4], Yong Geng[1], Kevin C. Jones[5], and Ying Zhu[1, 2*]

[1]School of Environmental Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

[2]The Key Laboratory of Environmental Health Impact Assessment for Emerging Pollutants, Ministry of Ecology and Environment of the People's Republic of China, Shanghai 200240, China

[3]Shenzhen Key Laboratory of Precision Measurement and Early Warning Technology for Urban Environmental Health Risks, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China

[4]Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke Street West, Montreal, QC H4B 1R6, Canada

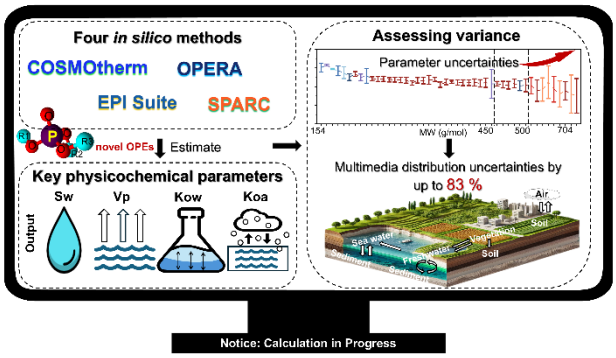
[5]Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, United Kingdom

*Corresponding author

Email address: yzhu16@sjtu.edu.cn (Ying Zhu)

Abstract

Organophosphate esters (OPEs) raise growing environmental and human health concern globally. However, numerous novel OPEs lack data on physicochemical properties, which are essential for assessing environmental fate, exposure and risks. This study predicted water solubility (S_w), vapor pressure (V_p), octanol-water partition coefficient (K_{ow}) and octanol-air partition coefficient (K_{oa}) at 25 ℃ for 46 novel OPEs by identifying optimal *in silico* tools and establishing prediction strategies based on molecular weights (MWs). Prediction discrepancies between *in silico* tools increased with MWs and structural complexity. Method evaluations for compounds with MWs >450 g/mol suggest that COSMOtherm is advantageous in predicting S_w and V_p for alkyl-OPEs, while SPARC is better for predicting V_p for aryl- and halogenated-OPEs. For compounds with MWs >500 g/mol, COSMOtherm and SPARC are recommended for K_{ow} and K_{oa} prediction, respectively. For smaller OPEs, average values from the top three of COSMOtherm, SPARC, EPI Suite and OPERA, ranked by validation on traditional flame retardants, are recommended. Using improper software could cause deviations in multimedia distribution and overall persistence in environment by up to 83% and 350%, respectively. The present data and prediction strategy are useful to enhance reliability of environmental fate, exposure and risk assessments of various OPEs and emerging contaminants.



42

43 **Synopsis:** Key physicochemical parameters are predicted for 46 novel
44 organophosphate esters lacking experimental data, and a prediction strategy is
45 established.

46 **Keywords:** novel organophosphate esters, physicochemical properties, water
47 solubility (S_w), octanol-water partition coefficient (K_{ow}), vapor pressure (V_p), octanol-
48 air partition coefficient (K_{oa}), environmental multimedia distribution, overall
49 persistence.

Introduction

Organophosphate esters (OPEs) are synthetic organic chemicals widely used as flame retardants (FRs) and plasticizers etc. in a wide variety of sectors such as construction, textiles, and electronics manufacturing.[1, 2] Global production had grown to one million tonnes by 2018[3], as they were initially marketed as "environmentally friendly" replacements of polybrominated diphenyl ethers (PBDEs) flame retardants. PBDEs have been listed under the Stockholm Convention in 2009 and 2017.[3-6] However, increasing evidence indicates that OPEs also have persistent, bioaccumulative and toxic (PBT) properties, raising concerns about their status as a "regrettable replacement".[7] Furthermore, recent studies highlight additional risks from OPE environmental transformation products, which may exhibit higher toxicity than the parent compounds, as well as novel OPEs (NOPEs) derived from organophosphite antioxidants (OPAs).[8-10] Despite this, research on the physicochemical properties of these NOPEs remains scarce. OPEs are typical semi-volatile organic compounds (SVOCs) with diverse physicochemical properties.[11] Compound physicochemical properties, such as water solubility (S_w), vapor pressure (V_p) and octanol-water partition coefficient (K_{ow}), are essential for understanding OPE multimedia distribution and their primary exposure routes to humans and environmental organisms.

In silico simulation tools have become essential for predicting physicochemical properties of substances for which experimental data and laboratory measurements are lacking. COSMOtherm,[12] EPI Suite,[13, 14] SPARC,[14] and OPERA[15] etc. are examples of software tools that have been extensively used for compounds including

polychlorinated biphenyls (PCBs), PBDEs, per- and polyfluoroalkyl substances (PFAS), and certain OPEs well-studied as FRs (often referred to as "traditional" chemicals).[16-22] Each tool exhibits varying performance across different chemicals and properties. For instance, COSMOtherm has demonstrated superior accuracy in predicting S_w, V_p for PBDEs and PFAS, and K_{ow}, octanol-air partition coefficient (K_{oa}), air-water partition coefficient (K_{aw}) for PCBs and PFAS.[16-21] Meanwhile, SPARC has performed well for K_{ow} and K_{aw} predictions for brominated flame retardants (BFRs), while OPERA has exceled at S_w, K_{ow}, and K_{oa} predictions for traditional OPEs.[21, 23] EPI Suite could provide parameter estimates that allow for more accurate calculation of K_{aw} for traditional OPEs.[17] In a screening study conducted by Zhang et al. (2010) using four software applications, consistent predictions for bioaccumulation and long-range transport potential across the four tools were obtained for only 70% of 529 substances analyzed, highlighting significant variability among the software applications.[24]

This study aims to identify the optimal *in silico* tools for predicting V_p, S_w, K_{ow}, K_{oa} for 46 NOPEs. Uncertainties of chemical multimedia distribution and overall persistence induced by variability of physicochemical parameters are illustrated by a well-developed and validated multimedia environmental fate model developed for China. This work highlights the impact of selecting inappropriate software on chemical fate predictions, which has rarely been discussed and demonstrated quantitatively before. The study provides valuable data and insights for researchers conducting either experimental or modelling research on environmental fate, exposure, and health risks of OPEs and other chemicals.

Methods

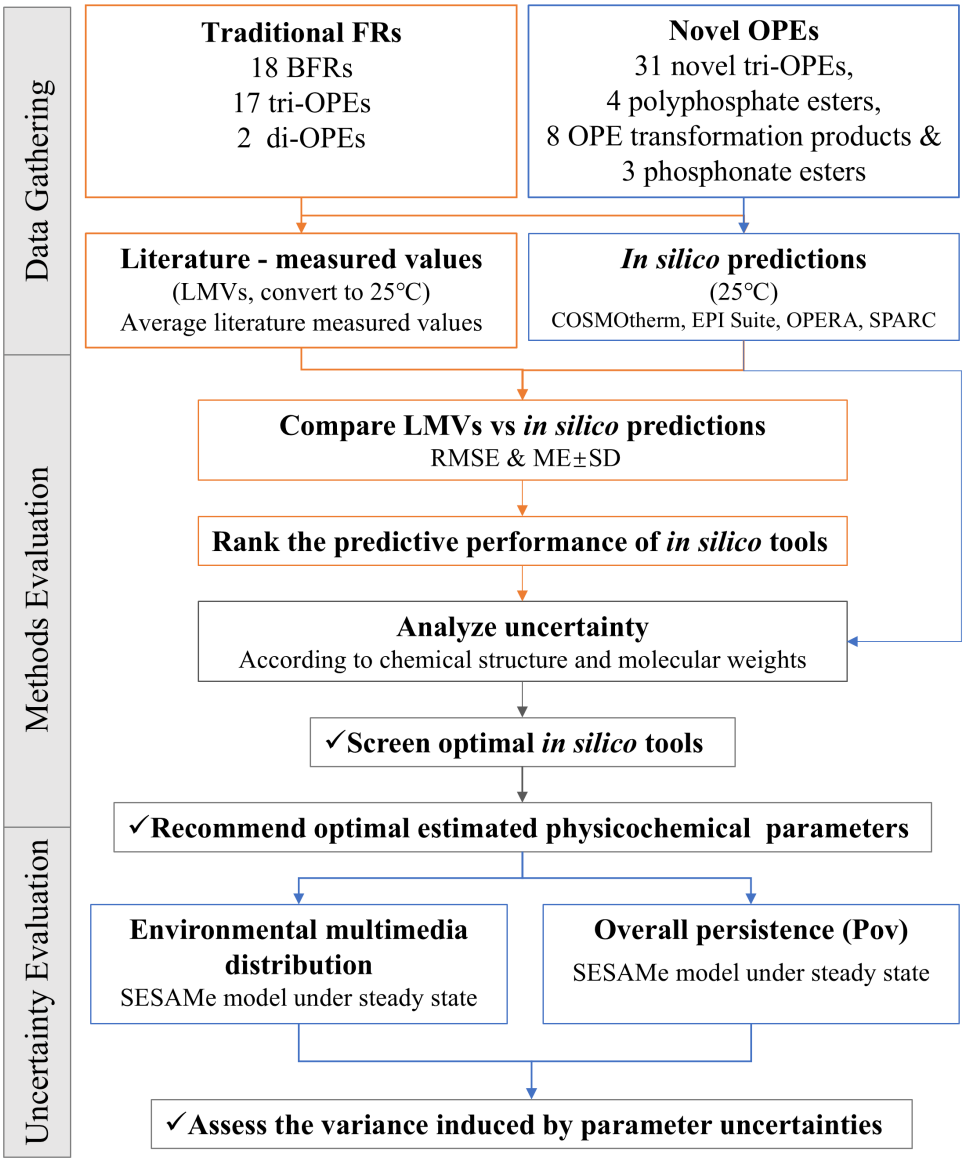


Figure 1. Flowchart of the workflow.

Figure 1 demonstrates the workflow and method of this study, with details stated as follows.

Chemical selection and data gathering.

Two groups of chemicals were included in this study: Group (1) consists of 37 traditional FRs for model evaluation, including 18 BFRs and 19 traditional OPEs (5 aryl-organophosphate triesters (tri-OPEs), 7 alkyl-tri-OPEs, 5 halogenated-tri-OPEs

and 2 alkyl-organophosphate diesters (di-OPEs)); and Group (2) comprises 46 NOPEs without reported experimental data of above physicochemical parameters, including 31 novel tri-OPEs, 4 polyphosphate esters, 8 environmental transformation products of OPEs and 3 phosphonate esters. The targeted OPE transformation products, except MDPP and 2,4DtBP (phenol), are di-OPEs. The pairing of OPEs and their transformation products is shown in Figure S1 in the Supporting Information (SI).

Information of chemicals in Group (2) is given in SI Text S1 and Tables S1-S2. Most NOPEs are primarily used as flame retardants and plasticizers (Table S2). Some of them are commercially used as alternatives to traditional organophosphate FRs. For example, o-CDPP, m-CDPP, P-CDPP (the isomers of CDPP), RDP, IDDPP, BPDPP and BPA-BDPP are replacements of TPHP while V6 and RDP are the alternatives to TEP, TCEP and TCIPP.[1, 25] Additionally, four NOPEs, i.e. AO168=O, TNPP, TiDeP, AO626=O_2, are oxidation products of OPAs. While AO168=O is used as a processing stabilizer for polymers in limited quantities, there is no direct industrial application for the other NOPEs (Table S2).[25, 26] Many of the target OPE transformation products also have direct production for industrial application (Table S2). The three phosphonate esters (DEEP, mono-PMMMP and di-PMMMP) have been newly identified in environmental matrices; and meanwhile, mono-PMMMP and di-PMMMP are transformation products of widely-used commercial organophosphate esters containing phosphorus-oxygen or phosphorus-sulfur bonds, probably exhibiting high environmental abundance.[27, 28] Some NOPEs have been detected in various environmental media and wildlife, with concentrations higher or comparable to those

of traditional OPEs.[25, 29-32] A few of them have demonstrated high acute or chronic toxicity to aquatic organisms, such as BPA-BDPP and CDPP.[33, 34]

The chemical selection in Group (1) has covered the functional groups contained in the compounds in Group (2) (Table S3) for better representativeness and reasonable guidance in identification of the optimal model for novel compounds. The molecular weights (MWs) range from 140 to 959 g/mol, reflecting a broad variation of chemical structural complexity. The literature-derived data of physicochemical parameters for the 37 traditional compounds in Group (1) were compiled and further harmonized to only include more reliable experimental data for better validation of the *in silico* tools, given that laboratory measurements may also have uncertainties.[35] The selection of literature-derived experimental data referred to the suggested rules by previous studies, such as the standardization of experimental methods and reasonable parameter ranges obtained by specific experimental methods etc (Table S4).[35, 36] Subsequently, to ensure thermodynamic consistency across the dataset, the literature values not originally measured at 25°C were converted to this temperature using equation S1, thereby unifying all values to reflect measurements at 25°C.[21, 37] If more than one value was found for individual parameters of each chemical, the average values were taken for the validation. The detailed method and selected experimental data are given in the SI Text S2 and Tables S5-S8.

***In silico* prediction of physicochemical properties and method evaluation.**

For all traditional and novel compounds, S_w, V_P, K_{ow}, and K_{oa} at 25 °C were predicted using COSMOtherm (BIOVIA COSMOtherm 2021, version 21.0), EPI Suite

147 (the U.S. EPA's Estimation Programs Interface Suite, version 4.11), OPERA (OPEn
148 structure-activity/property relationship application, version 2.9), and SPARC (SPARC
149 Performs Automated Reasoning in Chemistry).[12-15] The modules used in EPI Suite for
150 each parameter are specified in the SI Text S3, and WATERNT was adopted to predict
151 the S_w for its better reliability than WSKOW (Table S9). Only OPERA can directly
152 predict K_{oa}, using a machine learning algorithm based on the weighted k-nearest
153 neighbor (KNN) model.[15] The EPI Suite KOAWIN model can output K_{oa}; however, this
154 is not a direct calculation but follows the same principle as the other two software
155 applications. K_{oa} is calculated based on K_{aw} and K_{ow} as shown in equations 1-2 (Eqs 1-
156 2). H and R represent the Henry's law constant and the gas constant, respectively. T is
157 the temperature.

158 $$K_{aw} = \frac{H}{RT} \quad (1)$$

159 $$K_{oa} = K_{ow} / K_{aw} \quad (2)$$

160 EPI Suite v4.11 and OPERA are based on the principle of quantitative structure-
161 activity relationship (QSAR) models. QSAR models predict based on mathematical
162 relationships between chemical structures and their physicochemical properties or
163 biological activities. These models are trained on existing datasets to develop regression
164 models that can predict properties of unmeasured compounds. OPERA utilizes a
165 machine learning model based on weighted k-nearest neighbors (KNN) algorithm,
166 incorporating molecular descriptors from the Pharmaceutical Data Exploration
167 Laboratory (PaDEL). Descriptor selection is performed using genetic algorithms, and
168 OPERA is trained and validated on curated PHYSPROP datasets.[15, 38] EPI Suite

primarily relies on fragment and bond contribution methods, which predicts compounds' overall properties by summing the properties of chemical fragments.[13] COSMOtherm, on the other hand, is based on the COSMO-RS (conductor-like screening model for real solvents) theory. This theory simulates surface charge density distribution on molecules to calculate intermolecular interactions, thereby predicting solubility, stability of coordination compounds, and other thermodynamic properties. SPARC utilizes computational algorithms grounded in the fundamental chemical structure theory to estimate a broad spectrum of physicochemical properties directly from molecular structure, enabling the prediction across diverse organic compounds and spanning chemical family boundaries.

The accuracy of each software application is sorted by the root-mean-square error (RMSE) and mean error (ME) calculated by the experimental values and the predicted values of physicochemical parameters of the chemicals in Group (1). The calculation methods of RMSE and ME are described in SI Text S4. Moreover, the applicability domains (ADs) of the *in silico* tools were considered to further evaluate the reliability of predictions, which could provide quantified criteria for identifying the optimal estimation strategy based on chemical structure and MWs.[39, 40] The ADs of the four *in silico* tools was introduced in the SI Text S5. Only the two QSAR models in our study have limited ADs. OPERA provides specific AD values for each substance to evaluate the prediction reliability, while EPI Suite provides specific training set database with applicable ranges of chemical MWs, structural fragment limits and measurements of parameters for individual modules (Tables S25-S27).[15, 39, 41-45] Chemicals falling

outside the EPI Suite estimation domain range were marked as "outside AD", and considered low reliability (Table S10). Reliability scores were used to unify the different ADs forms provided by the two software applications (Tables S11-S12, S26-S27).[15, 45] COSMOtherm and SPARC have a broader chemical space, especially COSMOtherm, having an infinite AD.[45] It has to be clarified that the evaluation of predictions by all of RMSE, ME and ADs replies on experimental data, which have uncertainties due to operation and instrumentation errors, environmental influences or the systematic error and so on, even if standard laboratory methods are followed. This affects the prediction validation and probably will introduce uncertainties into our final estimation strategies. However, best attempts have been made to reduce the uncertainty in this study, and thus better qualified empirical data is still the best priority for prediction validation to date.

Uncertainty evaluation.

Precise knowledge of chemical multimedia distribution supports a better understanding of their environmental exposure routes, while persistence occupies a principal role in environmental risk assessment frameworks as the first criterion of the chemical PBT feature under the Stockholm Convention on Persistent Organic Pollutants.[46] Therefore, the SESAMe v3.4 model (Sino Evaluative Simplebox-MAMI model) was adopted to calculate the chemical multimedia distribution and overall persistence (P_{OV}) of the 46 NOPEs, using different sets of physicochemical parameters predicted by the four *in silico* tools. How uncertainties in these parameters affect the environmental fate of the chemicals was assessed. SESAMe v3.4 is a well-developed

multimedia model having demonstrating good performance on organics with a broad range of properties, covering air, freshwater and sediment, seawater and sediment, soil (classified as natural, agricultural, and urban soil), and vegetation (natural and agricultural vegetation) compartments.[10, 47-51] The average values of environmental variables covering diverse climate zones in China were taken as input to the model for the simulation. The proportion of land use types included in the model was as follows: natural soil (80.2%), agricultural soil (17.1%), urban soil (1.7%), and freshwater (1.0%) (Table S13). The degradation half-lives input to the SESAMe v3.4 model were estimated by EPI Suite (Table S14). The modules used in EPI Suite for calculating degradation half-lives are specified in the SI Text S3. Uncertainties in half-life estimate are present, as (1) some substances are outside the training set of the EPI Suite BIOWIN and AOPWIN models, mainly used to estimate degradation half-lives in individual compartments, which reduces the reliability; and (2) the half-lives in soil and sediment are converted from the estimate for water by 1:2:9 ratio, which is a preliminary estimation method and also reduces the reliability of the estimation.[13] However, this is among the very few datasets/methods available to us, and will not affect the main research purpose of the study.

A theoretical emission was applied to all individual target compounds identically to compare the patterns of their multimedia distribution and P_{OV}. The proportion of emission to air, freshwater, and urban soil (86:11:3) was taken from the study by Chen et al. (2023), as limited information is available from existing studies.[10] The P_{OV} is the average time that a chemical resides in multiple environmental compartments, which

was calculated by the SESAMe v3.4 model based on Eq 3.[52, 53]

$$P_{OV} = \frac{M_{total}}{E} \quad (3)$$

where E and M_{total} represent the emission rate (mol/day) and the total steady-state amount of a chemical (mol) in the system, respectively.[54]

Results and discussion

Comprehensive evaluation of software performance.

The collected measured values of traditional FRs range from -1.38 to 3.87 (log mg/L) for $\log S_w$, -6.33 to 2.15 (log Pa) for $\log V_p$, -0.65 to 9.97 for $\log K_{ow}$, and 6.80 to 14.4 for $\log K_{oa}$. Performance of the four software applications was mainly evaluated by comparing predictions (Tables S15-S18) with experimental data of the traditional FRs as stated above (Figures 2 and S2). The RMSE and ME reveal that none of the four software applications could consistently provide the most accurate estimates for all physicochemical properties of each substance (Figures 2 and S2). The same conclusion was also reached by Rodgers et al. (2021).[21] Overall, EPI Suite demonstrates better performance in predicting V_p, with a lower RMSE compared to the other three *in silico* tools (Figure 2a, b). COSMOtherm, EPI Suite and OPERA exhibit similar accuracy in predicting K_{ow} with RMSE values close to 1.13 log units. Additionally, OPERA provides predictions of S_w and K_{oa} with lower RMSE (0.96, 0.86 log units) and ME values (-0.10, 0.25 log units) than other *in silico* tools (Figure 2c, d).

However, the validation of OPERA predictions for K_{oa} of high-MW substances (e.g., 565 - 959 g/mol) may involve greater uncertainties compared to those for lower-MW substances. The predicted and measured $\log K_{oa}$ values of PBDE 209 (MW, 959

g/mol) are 11.7 and 14.4, respectively, indicating a large discrepancy. Meanwhile, for the other high-MW substance with MWs rapidly increasing from 565 to 723 g/mol, both precited and measured K_{oa} values fall within an extremely narrow range from 11.7 to 12.0 and from 11.4 to 11.7, respectively (Tables S8 and S18). A similar issue is observed in predictions for NOPEs. This is unreasonable, indicating failure of accurate prediction and measurement. High-MW chemicals, with extremely low volatility (log K_{oa} >12), will have a negligible amount partitioning into the air phase during equilibrium partitioning experiments, which more likely result in inaccurate measurement of air concentrations and consequently the K_{oa}, probably due to proximity to the instrument's detection limit. For less volatile chemicals, K_{oa} is also hardly reliably measured, as kinetic limitations may inhibit the phase equilibrium being reached.[35] For predictions, only 78% of the high-MW substances are inside the ADs of OPERA for K_{oa} prediction, which is lower than the percentage of 96% for lower-MW substances (Table S27). The reliability of predictions decreases for chemicals outside the ADs. Meanwhile, OPERA gives similar predictions when targeted chemicals have similar structures. Many high-MW FRs are PBDE congeners with a similar structure, which may result in the close K_{oa} predictions.

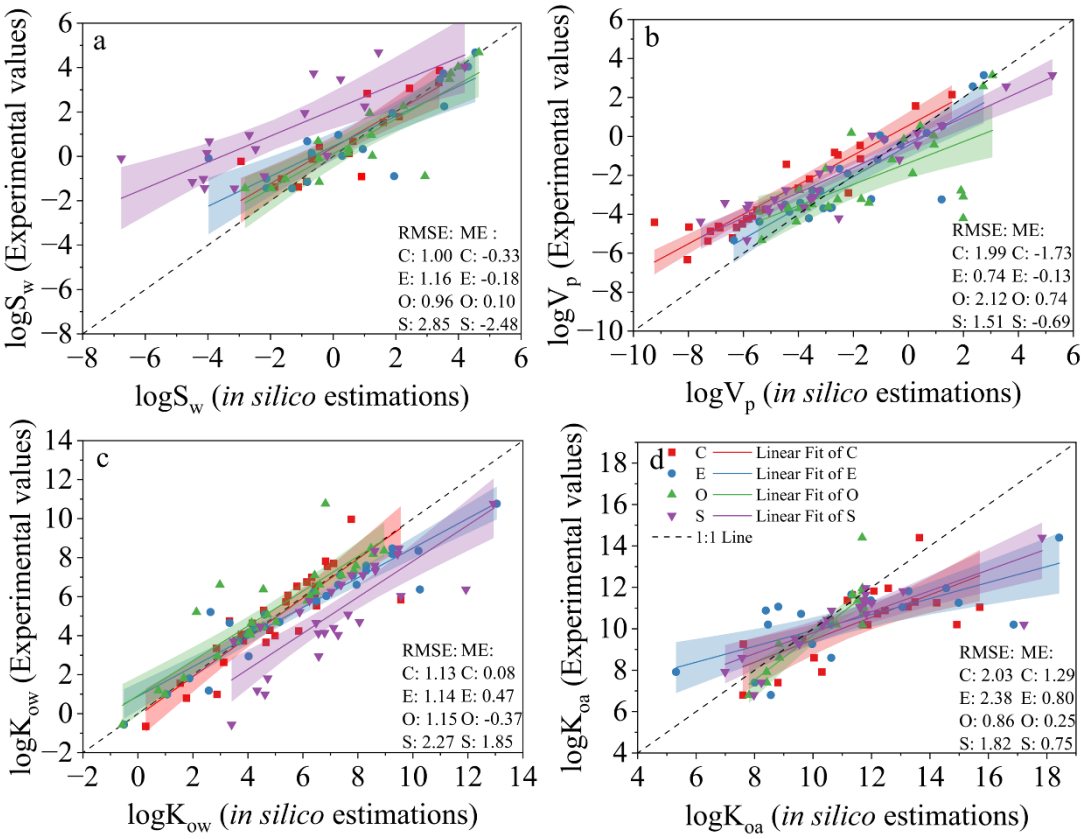


Figure 2. Comparison of the *in silico* estimates and experimental values of the traditional FRs to evaluate the performance of different *in silico* tools. The dashed lines represent the 1:1 agreement; the solid lines show the regressions between the measured and the simulated data; the shaded areas represent the 95% confidence interval of the regression. C, E, O and S represent the COSMOtherm, EPI Suite, OPERA and SPARC model, respectively. RMSE indicates root-mean-square error, ME indicates mean error. RMSE and ME are in logarithmic units.

More specifically, for individual categories of BFRs, alkyl-OPEs, aryl-OPEs, and halogenated-OPEs, the four *in silico* tools show varying uncertainties. COSMOtherm and OPERA perform significantly better (RMSE, 0.27−2.29 log units) than the other two software applications (RMSE, 0.36−4.50 log units) in predicting S_w for all categories (Figure S2a-d). Meanwhile, COSMOtherm also demonstrates the best

287 performance in predicting K_{ow} of BFRs; and it generally has a strong performance on
288 K_{ow} for all OPEs. Although it does not rank the top two on most occasions, the
289 difference from the superior software was minimal. EPI Suite has the best performance
290 on V_p for all categories, S_w for halogenated-OPEs and K_{ow} for alkyl-OPEs; and it is one
291 of the top two *in silico* tools in estimating S_w and K_{ow} of BFRs (inferior to
292 COSMOtherm) and aryl-OPEs (inferior to OPERA). OPERA exhibits the best accuracy
293 for predicting K_{ow} for aryl-OPEs and halogenated-OPEs, K_{oa} for alkyl-OPEs and
294 halogenated-FRs, including BFRs and halogenated-OPEs, and V_p for BFRs. It ranks
295 second in predicting K_{ow} of alkyl-OPEs (inferior to EPI Suite) and K_{oa} of aryl-OPEs
296 (inferior to SPARC), but the difference from the top-ranked software is not significant.
297 SPARC was the best for predicting K_{oa} of aryl-OPEs (Figure S2) and ranked the second
298 for predicting V_p of OPEs. The rank of software performance based on RMSE is
299 provided in Table S19.

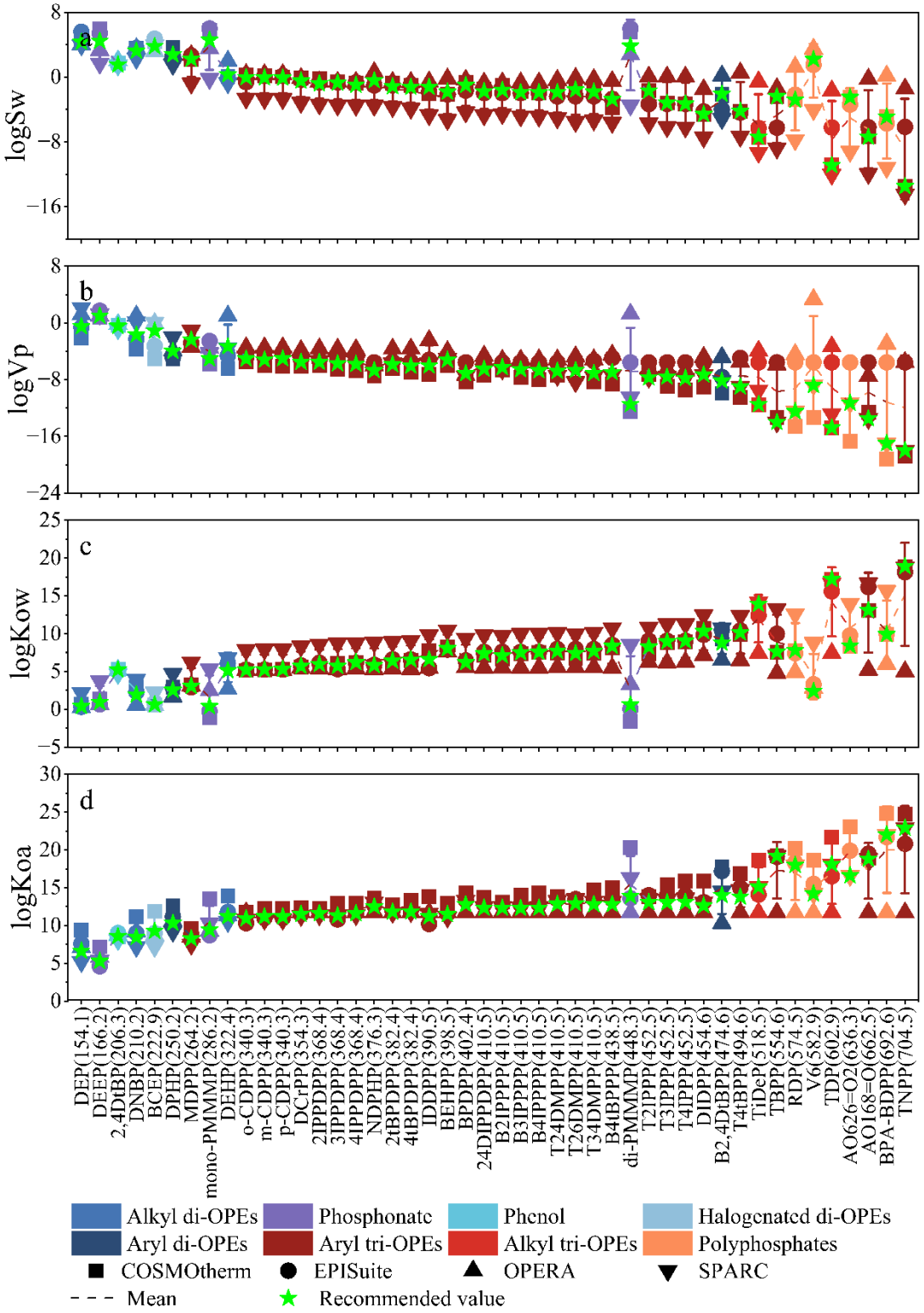
300 However, the uncertainty of predictions increases with growing MWs and
301 structure complexity for all software applications, although to varying extents. As
302 shown in Figure S2, the predictions exhibit a greater discrepancy from the measurement
303 for chemicals with higher MWs across nearly all four parameters, which is illustrated
304 by the increasing deviation of data points from the 1:1 line for higher-MW chemicals.
305 As the MW or structural complexity of a substance increases, it typically comprises a
306 greater number of or more intricate structural fragments. A higher percentage of these
307 substances are outside the ADs or training set of the two QSAR models, compared to
308 those with lower MWs. Even if they fall within the ADs, they are generally associated

with “Caution advised” or “unreliable” reliability scores (Tables S26 and S27). By comparing RMSEs, the QSAR model’s predictions for substances outside the training set have reduced accuracy compared to the substances within the training set (Tables S28 and S29). For example, the RMSEs of the $\log V_p$ for substances inside and outside the training set of the MPBVP module of the EPI Suite model are 0.75 and 1.77 log units, respectively. This was also observed by Zhang et al. (2016), who reported larger discrepancies in model estimates for novel FRs with MWs exceeding 800 g/mol, which were largely attributed to the models’ reliance on training sets composed of lower MW compounds.[55] Wang et al. (2017) also found similar trends when predicting gas-aqueous partitioning for volatile organic compounds with different number of functional groups and MWs.[56]

Estimated physicochemical properties of NOPEs.

Although each *in silico* tool shows advantages in predicting at least one variable for traditional FRs, EPI Suite and OPERA presented limitations on NOPEs owing to their operational principle based on empirical data. Figure 3 shows that when MWs of tri-OPEs and polyphosphate esters increased from 453 to 705 g/mol, OPERA generates very narrow ranges from -6.5 to -5.6 (log Pa) for $\log V_p$, from -1.52 to 0.53 (log mg/L) for $\log S_w$, and from 4.8 to 7.14 for log K_{ow}, while the predicted $\log K_{oa}$ is even stable at around 11.7. Similarly, the estimation of $\log V_p$ by EPI Suite ranges only from -5.6 to -4.9 (log Pa). The same phenomenon was observed in a previous study when estimating $\log K_{aw}$ of PFAS by OPERA.[18] This is due to the use of the nearest neighbor algorithm and the lack of experimental data of chemicals with high MWs in the training set. There

is no measurement data used in the training set due to the challenge of accurately measuring phase distributions when the chemicals predominantly distribute in one phase. For instance, if a chemical has a K_{oa} value exceeding 10^{12}, its transfer from the octanol phase to the air phase would become too small to be measured. As a fragment-based model, the bias of EPI Suite increases with the enlarged chemical structural complexity, due to the aggregate of small errors induced by growing numbers and complexity of bond fragments.[57] Ebert et al. noticed the same phenomenon for K_{aw} prediction using EPI Suite HENRYWIN.[57] Consequently, these two *in silico* tools may face inherent limitations in providing reliable estimates of the four parameters, especially for the NOPEs with larger MWs and complex structures that fall outside existing database of the software.



342

343 **Figure 3.** Predicted values, mean, standard deviation and recommended values of $logV_p$,

344 $logS_w$, $logK_{ow}$ and $logK_{oa}$ for NOPEs. The values in parentheses represent the molecular

345 weights of the substances.

In contrast, COSMOtherm and SPARC are based on quantum chemical principles and fundamental chemical structure theory, respectively. Predictions of chemicals with high MWs using the two approaches show a more reasonable changing pattern.[58-60] As the complexity of the molecular structure increases, both COSMOtherm and SPARC can better predict significant changes in physicochemical parameters, more accurately reflecting the large differences between compounds with higher and lower MWs, as has been observed in previous studies.[55] For example, for tri-OPEs and polyphosphate esters with MWs ranging from 453 to 705 g/mol, the estimates of $\log V_p$, $\log S_w$, $\log K_{ow}$ and $\log K_{oa}$ span from -7 to -19 (log Pa), -1.7 to -14.5 (log mg/L), 7.39 to 18.9, and 13.1 to 24.8, respectively, exhibiting considerably larger variations compared to those predicted by EPI Suite and OPERA (Figure 3 and Tables S20-S23). Therefore, significant discrepancies in predictions over 3 log units are observed among the four *in silico* tools when MWs exceeded 450 g/mol for S_w and V_p, and 500 g/mol for K_{ow} and K_{oa}. Zhang et al. also found large discrepancies among model predictions with rising MWs, which could even reach 12 log units for chemicals having MWs >800 g/mol.[55] Meanwhile, based on reliability score, it is also found that 450 g/mol and 500 g/mol are proper thresholds, as a greater number of NOPEs with MWs higher than the two thresholds fall outside the ADs for individual parameters than NOPEs with MWs lower than the thresholds. This indicates reduced reliability of the two QSAR models in prediction for the high-MWs substances. This reflects that making predictions for high-molecular-weight compounds should be done extremely cautiously. However, some predictions still need to be made—even though they have high uncertainty—to fulfill

368 the research objectives on these compounds.

369 In this case, by comprehensively considering the rank obtained by the software
370 evaluation on traditional FRs, software performance on NOPEs and operational
371 principles, this study recommends varying optimal *in silico* tools to predict the four
372 physicochemical parameters of novel substances with different MWs, using 450 g/mol
373 and 500 g/mol as thresholds. Given the small difference between the *in silico* tools when
374 MWs are below these thresholds, the average value is taken from the top three *in silico*
375 tools for all parameters. The software is ranked based on RMSE derived from validation
376 on traditional FRs (Table S19). For NOPEs with MWs above the thresholds, the first-
377 ranked non-QSAR software is used for individual parameters, considering the reduced
378 reliability of QSAR models on these high-MW chemicals as discussed above. For
379 compounds with MWs greater than 450 g/mol, COSMOtherm and SPARC are preferred
380 for S_w and V_p respectively for all substance categories. For all compounds with MWs
381 exceeding 500g/mol, COSMOtherm and SPARC are recommended to predict K_{ow} and
382 K_{oa}, respectively (Table 1). The recommended values of the four physicochemical
383 parameters for NOPEs are provided in Table S1. However, limited empirical data may
384 introduce uncertainties into the verification and thus the recommendation, which
385 requires more experimental data for confirmation or calibration. But the recommended
386 values have been the most reasonable and reliable based on current available tools and
387 methods.

388 **Table 1.** Recommended software predictions for OPEs with molecular weights
389 (MWs) >450/500 g/mol.

Compounds	S_w (>450 g/mol)	V_p (>450 g/mol)	K_{ow} (>500 g/mol)	K_{oa} (>500 g/mol)
Alkyl-OPEs	COSMOtherm	SPARC	COSMOtherm	SPARC
Aryl-OPEs	COSMOtherm	SPARC	COSMOtherm	SPARC
Halogenated-OPEs	COSMOtherm	SPARC	COSMOtherm	SPARC

For all categories, $\log S_w$ and $\log V_p$ generally decrease when the number of carbon atoms increases, with $\log S_w$ ranging from -13.5 to 5.1 (log mg/L) and $\log V_p$ from -18.0 to 1.3 (log Pa). In contrast, $\log K_{ow}$ and $\log K_{oa}$ exhibit an increasing trend with growing carbon atoms, with $\log K_{ow}$ ranging from 0.4 to 18.9 and $\log K_{oa}$ from 5.3 to 22.9 (Figure 3 and Table S1). This suggests that the hydrophobicity of the compounds increases with the addition of carbon atoms, while the volatility progressively decreases. For chemicals with the same number of carbon atoms, different functional groups effectively influence chemical properties. The alkyl-OPEs, having the same number of carbon atoms as aryl-OPEs, exhibit lower water solubility and higher log K_{ow} than corresponding aryl-OPEs, indicating stronger lipophilicity. For example, TiDeP and TDP, as alkyl-OPEs, have a $\log S_w$ at -7.4 and -10.9 (log mg/L), respectively, and $\log K_{ow}$ at 13.9 and 17.2 (Table S1); while T4tBPP and TBPP, as two aryl-OPEs having the same number of carbon atoms with the above two alkyl-OPEs, present a higher $\log S_w$ at -4.2 and -2.4 (log mg/L), and lower $\log K_{ow}$ values at 8.9 and 7.6, respectively (Table S1). Moreover, an increase in phosphate groups enhances molecular polarity by facilitating hydrogen bond formation when the number of carbon atoms remains the same, thereby increasing the water solubility of the chemicals.[61] For instance, both as aryl-OPEs

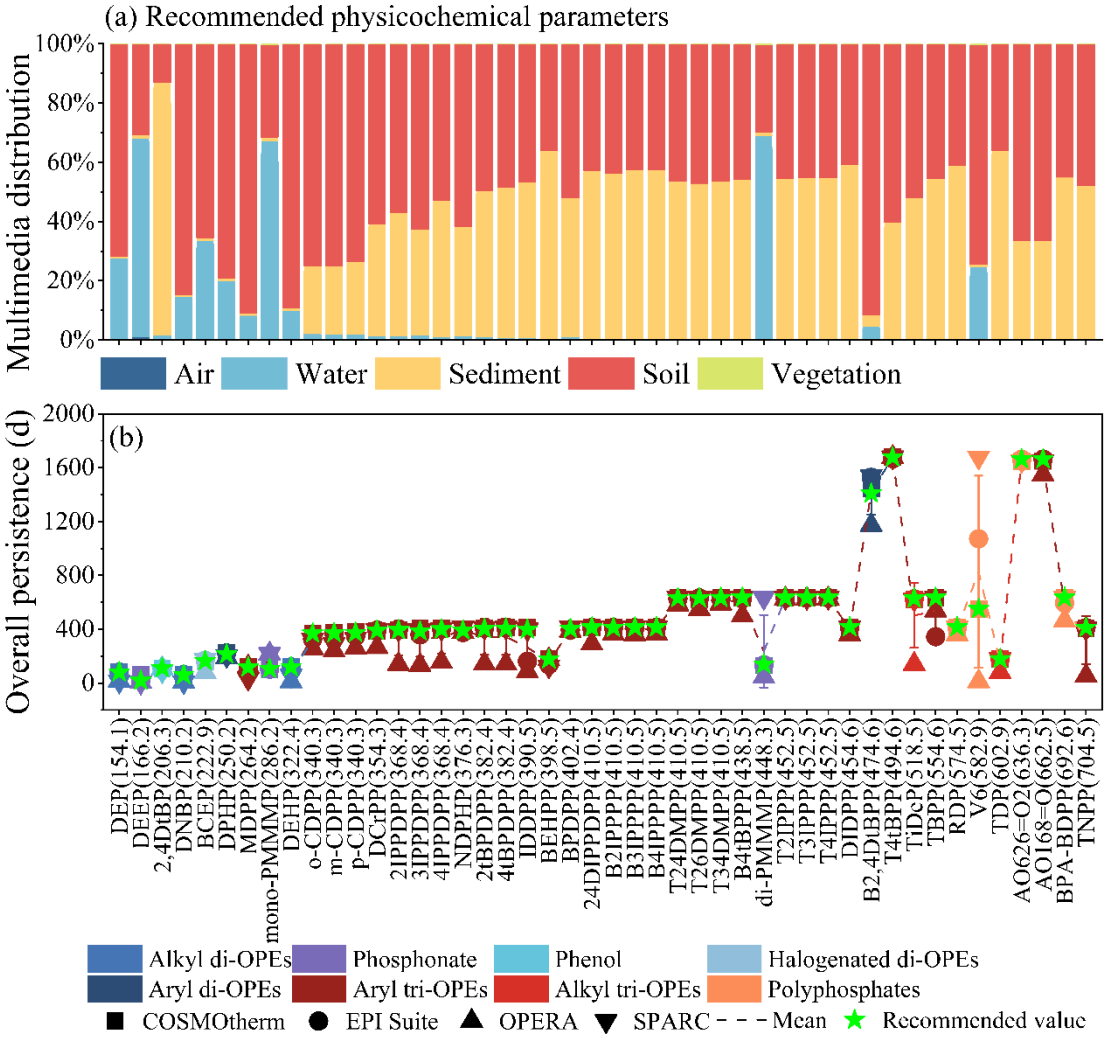
having an identical carbon atom count, RDP possessing an additional phosphate group than T4tBPP has a greater $logS_w$ (-2.8 versus -4.2, log mg/L) and a lower logKow (7.8 versus 8.9) than T4tBPP. Chlorine substituents also play a critical role in enhancing the water solubility of chemicals by amplifying polar interactions. For example, V6, which contains six chlorine atoms, shows higher water solubility with a $logS_w$ value of 2.2 (log mg/L). In addition to chlorination, oxidation reactions—particularly the addition of hydroxyl groups—can greatly increase the hydrophilicity of transformation products compared to their parent tri-OPEs. This rise in polarity is evidenced by higher $logS_w$ values and lower $logK_{ow}$ values (Figure S1). A typical example is BCEP, a transformation product of TCEP, which forms through photooxidation reactions, as reported by Liu et al (2021).[8] BCEP exhibits a lower $logK_{ow}$ of 0.64 compared to 1.57 for TCEP (Figure S1), further highlighting how oxidative processes enhance both solubility and polarity.

Impact of uncertainty in physicochemical parameters on predicting multimedia distribution.

All four *in silico* tools provide physicochemical parameters indicating negligible distribution in the air for all NOPEs under both equilibrium conditions (Figure S3) and the steady state (Figure 4a). The difference of the two conditions is described in SI Text S6. However, equilibrium partitioning illustrated by K_{ow} and K_{aw} in Figure S3 shows that predictions by COSMOtherm and EPI Suite lead to greater variability in chemical partitioning between water and solid phases (sediment/soil) at equilibrium than those from the other software applications. The prediction by SPARC also suggests a

relatively variable partitioning but indicates a greater tendency for substances to distribute in air compared to the other software applications. In contrast, OPERA generates the most consistent partitioning of the NOPEs at equilibrium, with slightly higher partitioning in the water phase compared to other phases (Figure S3).

Steady-state distributions predicted by the SESAMe v3.4 model reveal that, with the input of recommended physicochemical parameters, despite 86% of emissions entering the atmosphere, novel tri-OPEs and polyphosphate esters are primarily distributed in sediments (23−64%) and soils (36−75%), together accounting for over 98% of the total mass remaining in the multimedia system (Figure 4a). As MWs increase, the distribution shifts, showing an increasing proportion in sediment compared to soil. Furthermore, V6, a halogenated-polyphosphate ester, shows the highest percentage in the water compartment (25%) among all polyphosphate esters, as a result of its highest water solubility and long half-life in water (Figure 4a and Table S14).



443

444 **Figure 4.** (a) Multimedia environmental distribution of NOPEs under steady state

445 predicted by the SESAMe v3.4 model with inputs of recommended physicochemical

446 parameter values; (b) Overall persistence predicted by the SESAMe v3.4 model, with

447 input of recommended physicochemical parameter values and predictions by individual

448 *in silico* tools. The values in parentheses represent the molecular weights of the

449 substances.

450 Novel di-OPEs are generally more hydrophilic than their parent compounds, as

451 indicated by higher S_w and lower K_{ow} values, causing a greater distribution in the water

452 compartment (Figure 4a). As a special case, the phenolic transformation product of

AO168=O, i.e., 2,4DtBP, is mainly distributed in the sediment (85%)—the highest distribution in the sediment compartment of all NOPEs. This is because of its moderate water solubility ($logS_w$ at 1.77, log mg/L) and $logK_{ow}$ (5.23), which makes 2,4-DtBP more likely to partition in the soluble phase of the soil pore water than other NOPEs, and be transported to freshwater systems by surface land runoff. Meanwhile, 2,4-DtBP is more easily distributed in the sediment than other novel di-OPEs. The three phosphonate esters, namely DEEP, mono-PMMMP and di-PMMMP, are primarily distributed in water and soil compartments, with 67%, 67% and 69% in water, and 31%, 31% and 30% in soil, respectively (Figure 4a). DEEP exhibits a slightly higher distribution in air (1.2%) compared to other phosphonate esters because of its much higher $logV_p$ and lower $logK_{oa}$ (Figure 4a), while other phosphonate esters and even all other novel tri-OPEs have a $logV_p$ <0 and $logK_{oa}$ >5.5.

Uncertainties are shown in the multimedia distribution simulated by taking the non-recommended values for physicochemical parameters (i.e. S_w, V_p, and K_{ow}) provided by the different software applications, compared to those using recommended values. The most pronounced uncertainties concentrate on soil and sediment distribution for low-molecular-weight (<360 g/mol) and high-molecular-weight OPEs (>600 g/mol), reaching up to 83%. COSMOtherm and EPI Suite provide a closer distribution pattern in sediment (0.5−86%), soil (12−92%), and water (0.1−75%) to that calculated by the recommended values for chemicals with MWs below 450 g/mol (Figures 4a and S4a, b). However, when using SPARC, the deviation from using recommended values is the greatest for the low-molecular-weight chemicals (<360

g/mol) (Figures 4a and S4d). Specifically, adopting SPARC tends to result in a deviation in water distribution by up to 65% (mono-PMMMP), and in sediment and soil distribution by 83% and 77%, respectively (MDPP). For chemicals with MWs between 400 and 450 g/mol, OPERA would underestimate the distribution in sediments by up to 22% and overestimate the distribution in soil by up to 37% (Figures 4a and S4c). For chemicals with MWs >600 g/mol, particularly AO168=O, AO626=O_2, and TNPP, the physicochemical parameters predicted by single software alone cannot obtain reasonable multimedia distributions, with the sediment and soil distributions differing from the recommendations by up to 47%. Furthermore, only using the physicochemical parameters predicted by recommended software like COSMOtherm or SPARC also result in significant inaccuracies in multimedia distributions, e.g. the sediment and soil distribution deviation of TiDeP (MW >500 g/mol) is up to 43%.

Impacts of uncertainty in physicochemical parameters on predicting overall persistence.

High persistence indicates the potential for prolonged environmental and human exposure to a substance, which is difficult to control or remove. As indicated above, it has been suggested as a highly concerned chemical inherent feature on its own. By inputting the recommended values of physicochemical parameters, T4tBPP, AO168=O, AO626=O_2 and B2,4DtBPP display the highest P_{OV} of the tested chemicals, ranging from ~4−4.6 years, longer than the P_{OV} of the other target NOPEs (18 days ~ 1.7 years) (Figure 4b). This is primarily attributed to their predominant distribution in soil (60−91%) and sediment (4−40%), accounting for over 95% of the total remaining mass,

497 and their long half-lives in soil (360 days) and sediment (4.4 years) (Table S14). These
498 four substances could be classified as “very persistent” (vP) under the EU REACH
499 criterion, which requires a residence time exceeding 60 days in water, 180 days in soil,
500 and 540 days in sediment.[62]

501 P_{OV} of other novel tri-OPEs and polyphosphate esters (from 118 days to 1.7 years)
502 is less than that of the aforementioned four substances, with MDPP, BEHPP and TDP
503 exhibiting the lowest P_{OV} (<179 days). Under steady state, MDPP is mainly distributed
504 in soil (91%), with additional 8% in water and 1% in sediment, while both BEHPP and
505 TDP are predominantly distributed in sediment (64%) and soil (36%). These half-lives
506 of these three chemicals in the key environmental compartments are the lowest among
507 all the target compounds (Table S14). It is the reason why they have the lowest P_{OV}.
508 Novel di-OPEs generally have lower P_{OV} (61−214 days, except for B2,4DtBPP) than
509 the selected novel tri-OPEs (Figure 4b). However, most of them (DPHP, MDPP, DEP,
510 DNBP and BCEP) demonstrate P_{OV} 2 to 16 times higher than their precursors (Figure
511 S5a, c-f). Only B2,4DtBPP and 2,4DtBP are less persistent than their precursor
512 AO168=O (Figure S5b). The P_{OV} of the three phosphonate esters ranges from 18 to 140
513 days (Figure 4b).

514 If using only one software tool, COSMOtherm could yield P_{OV} estimates closer to
515 the recommended values for novel tri-OPEs and polyphosphate esters, while SPARC
516 may overestimate the P_{OV} of these substances by up to 3 years (V6) and other two tools
517 may underestimate the Pov by up to 1.5 years. For novel di-OPEs and phosphonate
518 esters, the deviations caused by using OPERA and SPARC can exceed 1 year. Overall,

using single software can result in Pov deviations ranging 0−350%, and generates the highest uncertainty on P_{OV} estimates for di-PMMMP, TiDeP, TNPP and V6, with deviations of 1 to 3 years. The overestimation by SPARC for di-PMMMP is primarily because it overpredicted the diffusive mass flux from water to sediment and sedimentation, combined with an underestimation of soil runoff and sediment resuspension. This increases the mass retained in sediment and soil, consequently, overall persistence. In contrast, OPERA underestimates TiDeP and TNPP for the opposite reason. Using single software could result in the largest deviation of retained mass in soils and sediments at the steady state for V6, compared to other substances, resulting in the high deviation of P_{OV} for V6.

Perspectives.

This research provides the first and most reliable reference values for essential physicochemical parameters of 46 NOPEs, covering a wide variety of MWs and structures. These reference values are valuable because the property ranges of the NOPEs made direct measurement of the physicochemical properties infeasible. The discussion on impact of uncertainties in physicochemical parameters on predicting multimedia distribution and P_{OV} has revealed potential deviations when using different software to predict these parameters and highlighted the characteristics of environmental behavior of these novel compounds. Reliable knowledge of chemical properties is the premise of accurate understanding of environmental behaviors, exposure routes and risks of chemicals, as well as efficient management of chemical use and contamination. This study has filled data gaps and provided solutions regarding

defects of the existing software in predicting physicochemical properties, which could be further applied to other emerging contaminants. However, it should be noted that the MW thresholds for prediction were established based on the evaluation of the targeted FRs and NOPEs in this study. Further evaluation across broader chemical categories may be needed to validate these thresholds in future work when expanding the strategy application on other chemicals.

Future research should fully leverage the powerful computational capabilities of modern computers, particularly machine learning and artificial intelligence technologies, to further develop methods based on COSMO-RS and fundamental chemical structure theory.[63, 64] By integrating machine learning models with existing algorithms for molecular surface charge density distribution and molecular structure deduction, deeper insights can be achieved into the interactions between atoms, electrons, and functional groups, thereby improving the accuracy of chemical property predictions. These methods hold even greater potential in cases where prior data is lacking or for complex structures which are not present in databases. Meanwhile, more attention should focus on building larger training datasets and developing multi-scale computational models to more accurately predict the physicochemical properties of organic compounds with more complex structure under various conditions, thereby advancing fields such as materials science, chemical engineering, environmental science and public health etc. As global concerns about chemical safety grow, this study underscores the critical role of advanced computational tools in filling data gaps and mitigating the risks associated with emerging contaminants.

Associated content

Supporting Information

The Supporting Information is available at https://pubs.acs.org

The information, structure, application and recommended physicochemical properties of the target NOPEs; the information of traditional FRs; SMILES strings for NOPEs and traditional FRs; data for experimental and estimated physicochemical parameters of traditional FRs; EPI Suite model selection; plot of comparison of *in silico* estimates and experimental data of physicochemical parameters for traditional FRs; software performance rankings; application domains and reliability evaluations for QSAR models; RMSE and ME for various validated substances; estimates of physicochemical parameters of NOPEs by the four *in silico* tools; parameters input into the SESAMe v3.4 model; degradation half-lives for NOPEs; chemical structure and overall persistence for OPEs and corresponding transformation products; plots of equilibrium partitioning characteristics and multimedia environmental distributions based on estimated physicochemical parameters of NOPEs.

Acknowledgement

This study is supported by the by the National Natural Science Foundation of China (42377362, 423B1003 and 72088101), the Research Start-up Funding from Shanghai Jiao Tong University and the Shanghai Oriental Talent Program - Youth Project.

Reference

1. van der Veen I, de Boer J. Phosphorus flame retardants: Properties, production, environmental occurrence, toxicity and analysis. *Chemosphere*, **2012**, 88 (10):

1119-1153.
2. Xie Z, Wang P, Wang X, Castro-Jiménez J, Kallenborn R, Liao C, Mi W, Lohmann R, Vila-Costa M, Dachs J. Organophosphate ester pollution in the oceans. *Nature Reviews Earth & Environment*, **2022**, 3 (5): 309-322.
3. Zhu H, Al-Bazi M M, Kumosani T A, Kannan K. Occurrence and Profiles of Organophosphate Esters in Infant Clothing and Raw Textiles Collected from the United States. *Environmental Science & Technology Letters*, **2020**, 7 (6): 415-420.
4. Sühring R, Diamond M L, Bernstein S, Adams J K, Schuster J K, Fernie K, Elliott K, Stern G, Jantunen L M. Organophosphate Esters in the Canadian Arctic Ocean. *Environmental Science & Technology*, **2021**, 55 (1): 304-312.
5. Gu L, Hu B, Fu Y, Zhou W, Li X, Huang K, Zhang Q, Fu J, Zhang H, Zhang A, Fu J, Jiang G. Occurrence and risk assessment of organophosphate esters in global aquatic products. *Water Research*, **2023**, 240: 120083.
6. Tian Y X, Chen H Y, Ma J, Liu Q Y, Qu Y J, Zhao W H. A critical review on sources and environmental behavior of organophosphorus flame retardants in the soil: Current knowledge and future perspectives. *Journal of Hazardous Materials*, **2023**, 452: 131161.
7. Blum A, Behl M, Birnbaum L S, Diamond M L, Phillips A, Singla V, Sipes N S, Stapleton H M, Venier M. Organophosphate Ester Flame Retardants: Are They a Regrettable Substitution for Polybrominated Diphenyl Ethers? *Environmental Science & Technology Letters*, **2019**, 6 (11): 638-649.
8. Liu Q, Li L, Zhang X, Saini A, Li W, Hung H, Hao C, Li K, Lee P, Wentzell J J B, Huo C, Li S M, Harner T, Liggio J. Uncovering global-scale risks from commercial chemicals in air. *Nature*, **2021**, 600 (7889): 456-461.
9. Liu Q, Liu R, Zhang X, Li W, Harner T, Saini A, Liu H, Yue F, Zeng L, Zhu Y, Xing C, Li L, Lee P, Tong S, Wang W, Ge M, Wang J, Wu X, Johannessen C, Liggio J, Li S-M, Hung H, Xie Z, Mabury S A, Abbatt J P D. Oxidation of commercial antioxidants is driving increasing atmospheric abundance of organophosphate esters: Implication for global regulation. *One Earth*, **2023**, 6 (9): 1202-1212.
10. Chen R, Xing C, Shen G, Jones K C, Zhu Y. Indirect Emissions from Organophosphite Antioxidants Result in Significant Organophosphate Ester Contamination in China. *Environmental Science & Technology*, **2023**, 57 (48): 20304-20314.
11. Sühring R, Wolschke H, Diamond M L, Jantunen L M, Scheringer M. Distribution of Organophosphate Esters between the Gas and Particle Phase–Model Predictions vs Measured Data. *Environmental Science & Technology*, **2016**, 50 (13): 6644-6651.
12. BIOVIA, BIOVIA COSMOtherm 2021, version 21.0. Dassault Systèmes: 2021. https://www.3ds.com/products/biovia/cosmo-rs/cosmotherm (accessed 2023-07-15).
13. U.S. EPA, Estimation Programs Interface Suite for Microsoft Windows (EPI Suite v4.11). U.S. Environmental Protection Agency: 2023. (accessed 2023-08-10).
14. ArChemCalc, SPARC Performs Automated Reasoning in Chemistry (SPARC).

ArChemCalc: 2024. http://archemcalc.com/sparc-web/calc (accessed 2023-08-02).
15. Mansouri K, Grulke C M, Judson R S, Williams A J. OPERA models for predicting physicochemical properties and environmental fate endpoints. *Journal of Cheminformatics*, **2018**, 10 (1): 10.
16. Stenzel A, Goss K-U, Endo S. Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOtherm, ABSOLV, and SPARC. *Environmental Toxicology and Chemistry*, **2014**, 33 (7): 1537-1543.
17. Endo S. Intermolecular Interactions, Solute Descriptors, and Partition Properties of Neutral Per- and Polyfluoroalkyl Substances (PFAS). *Environmental Science & Technology*, **2023**, 57 (45): 17534-17541.
18. Endo S, Hammer J, Matsuzawa S. Experimental Determination of Air/Water Partition Coefficients for 21 Per- and Polyfluoroalkyl Substances Reveals Variable Performance of Property Prediction Models. *Environmental Science & Technology*, **2023**, 57 (22): 8406-8413.
19. Gomis M I, Wang Z, Scheringer M, Cousins I T. A modeling assessment of the physicochemical properties and environmental fate of emerging and novel per- and polyfluoroalkyl substances. *Science of The Total Environment*, **2015**, 505: 981-991.
20. Wang Z, MacLeod M, Cousins I T, Scheringer M, Hungerbühler K. Using COSMOtherm to predict physicochemical properties of poly- and perfluorinated alkyl substances (PFASs). *Environmental Chemistry*, **2011**, 8 (4): 389-398.
21. Rodgers T F M, Okeme J O, Parnis J M, Girdhari K, Bidleman T F, Wan Y, Jantunen L M, Diamond M L. Novel Bayesian Method to Derive Final Adjusted Values of Physicochemical Properties: Application to 74 Compounds. *Environmental Science & Technology*, **2021**, 55 (18): 12302-12316.
22. Hammer J, Endo S. Volatility and Nonspecific van der Waals Interaction Properties of Per- and Polyfluoroalkyl Substances (PFAS): Evaluation Using Hexadecane/Air Partition Coefficients. *Environmental Science & Technology*, **2022**, 56 (22): 15737-15745.
23. Kuramochi H, Takigami H, Scheringer M, Sakai S-i. Estimation of physicochemical properties of 52 non-PBDE brominated flame retardants and evaluation of their overall persistence and long-range transport potential. *Science of The Total Environment*, **2014**, 491-492: 108-117.
24. Zhang X, Brown T N, Wania F, Heimstad E S, Goss K-U. Assessment of chemical screening outcomes based on different partitioning property estimation methods. *Environment International*, **2010**, 36 (6): 514-520.
25. Tan H, Chen D, Peng C, Liu X, Wu Y, Li X, Du R, Wang B, Guo Y, Zeng E Y. Novel and Traditional Organophosphate Esters in House Dust from South China: Association with Hand Wipes and Exposure Estimation. *Environmental Science & Technology*, **2018**, 52 (19): 11017-11026.
26. Liu R, Mabury S A. Organophosphite Antioxidants in Indoor Dust Represent an Indirect Source of Organophosphate Esters. *Environmental Science & Technology*, **2019**, 53 (4): 1805-1811.
27. Kutarna S, Chen W, Xiong Y, Liu R, Gong Y, Peng H. Screening of Indoor Transformation Products of Organophosphates and Organophosphites with an in

Silico Spectral Database. *ACS Measurement Science Au*, **2023**, 3 (6): 469-478.
28. Hu J, Lyu Y, Li M, Wang L, Jiang Y, Sun W. Discovering Novel Organophosphorus Compounds in Wastewater Treatment Plant Effluents through Suspect Screening and Nontarget Analysis. *Environmental Science & Technology*, **2024**, 58 (14): 6402-6414.
29. Liu R, Mabury S A. Unexpectedly High Concentrations of a Newly Identified Organophosphate Ester, Tris(2,4-di-tert-butylphenyl) Phosphate, in Indoor Dust from Canada. *Environmental Science & Technology*, **2018**, 52 (17): 9677-9683.
30. Meng W, Li J, Shen J, Deng Y, Letcher R J, Su G. Functional Group-Dependent Screening of Organophosphate Esters (OPEs) and Discovery of an Abundant OPE Bis-(2-ethylhexyl)-phenyl Phosphate in Indoor Dust. *Environmental Science & Technology*, **2020**, 54 (7): 4455-4464.
31. Li J, Zhang Y, Bi R, Ye L, Su G. High-Resolution Mass Spectrometry Screening of Emerging Organophosphate Esters (OPEs) in Wild Fish: Occurrence, Species-Specific Difference, and Tissue-Specific Distribution. *Environmental Science & Technology*, **2022**, 56 (1): 302-312.
32. Xie Z, Zhang X, Xie Y, Liu F, Sun B, Liu W, Wu J, Wu Y. Bioaccumulation and Potential Endocrine Disruption Risk of Legacy and Emerging Organophosphate Esters in Cetaceans from the Northern South China Sea. *Environmental Science & Technology*, **2024**, 58 (9): 4368-4380.
33. Muir D C G, Howard P H. Are There Other Persistent Organic Pollutants? A Challenge for Environmental Chemists. *Environmental Science & Technology*, **2006**, 40 (23): 7157-7166.
34. Ye L, Li J, Gong S, Herczegh S M, Zhang Q, Letcher R J, Su G. Established and emerging organophosphate esters (OPEs) and the expansion of an environmental contamination issue: A review and future directions. *Journal of Hazardous Materials*, **2023**, 459: 132095.
35. Li L, Zhang Z, Men Y, Baskaran S, Sangion A, Wang S, Arnot J A, Wania F. Retrieval, Selection, and Evaluation of Chemical Property Data for Assessments of Chemical Emissions, Fate, Hazard, Exposure, and Risks. *ACS Environmental Au*, **2022**, 2 (5): 376-395.
36. Baskaran S, Lei Y D, Wania F. A Database of Experimentally Derived and Estimated Octanol–Air Partition Ratios (KOA). *Journal of Physical and Chemical Reference Data*, **2021**, 50 (4).
37. Bahadur N P, Shiu W-Y, Boocock D G B, Mackay D. Temperature Dependence of Octanol−Water Partition Coefficient for Selected Chlorobenzenes. *Journal of Chemical & Engineering Data*, **1997**, 42 (4): 685-688.
38. Kamel M, Williams A, *QMRF - Title: KOA model for the octanol/air partition coefficient prediction from OPERA models*. QMRF Report Q17-18-0018, ResearchGate, 2017. http://10.13140/RG.2.2.14409.54883/1.
39. Zhang Z, Sangion A, Wang S, Gouin T, Brown T, Arnot J A, Li L. Chemical Space Covered by Applicability Domains of Quantitative Structure–Property Relationships and Semiempirical Relationships in Chemical Assessments. *Environmental Science & Technology*, **2024**, 58 (7): 3386-3398.

40. OECD *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*; No. 69; Paris, 2014.
41. U.S. EPA WATERNT™ User's Guide. https://episuite.dev/EpiWebSuite/#/help/waternt (accessed 2024-12-10).
42. U.S. EPA MPBPVP™ User's Guide. https://episuite.dev/EpiWebSuite/#/help/mpbpvp (accessed 2024-12-10).
43. U.S. EPA KOWWIN™ User's Guide. https://episuite.dev/EpiWebSuite/#/help/kowwin (accessed 2024-12-10).
44. U.S. EPA KOAWIN™ User's Guide. https://episuite.dev/EpiWebSuite/#/help/koawin (accessed 2024-12-10).
45. Baskaran S, Lei Y D, Wania F. Reliable Prediction of the Octanol–Air Partition Ratio. *Environmental Toxicology and Chemistry*, **2021**, 40 (11): 3166-3180.
46. Cousins I T, Ng C A, Wang Z, Scheringer M. Why is high persistence alone a major cause of concern? *Environmental Science: Processes & Impacts*, **2019**, 21 (5): 781-792.
47. Zhu Y, Tao S, Price O R, Shen H Z, Jones K C, Sweetman A J. Environmental Distributions of Benzo[a]pyrene in China: Current and Future Emission Reduction Scenarios Explored Using a Spatially Explicit Multimedia Fate Model. *Environmental Science & Technology*, **2015**, 49 (23): 13868-13877.
48. Li S, Zhu Y, Zhong G, Huang Y, Jones K C. Comprehensive Assessment of Environmental Emissions, Fate, and Risks of Veterinary Antibiotics in China: An Environmental Fate Modeling Approach. *Environmental Science & Technology*, **2024**, 58 (12): 5534-5547.
49. Zhu Y, Tao S, Price O R, Shen H, Jones K C, Sweetman A J. Environmental Distributions of Benzo[a]pyrene in China: Current and Future Emission Reduction Scenarios Explored Using a Spatially Explicit Multimedia Fate Model. *Environmental Science & Technology*, **2015**, 49 (23): 13868-13877.
50. Zhu Y, Price O R, Kilgallon J, Rendal C, Tao S, Jones K C, Sweetman A J. A Multimedia Fate Model to Support Chemical Management in China: A Case Study for Selected Trace Organics. *Environmental Science & Technology*, **2016**, 50 (13): 7001-7009.
51. Zhu Y, Tao S, Sun J, Wang X, Li X, Tsang D C W, Zhu L, Shen G, Huang H, Cai C, Liu W. Multimedia modeling of the PAH concentration and distribution in the Yangtze River Delta and human health risk assessment. *Science of The Total Environment*, **2019**, 647: 962-972.
52. Webster E, Mackay D, Wania F. Evaluating environmental persistence. *Environmental Toxicology and Chemistry*, **1998**, 17 (11): 2148-2158.
53. Klasmeier J, Matthies M, Macleod M, Fenner K, Scheringer M, Stroebe M, Le Gall A C, McKone T, Van De Meent D, Wania F. Application of Multimedia Models for Screening Assessment of Long-Range Transport Potential and Overall Persistence. *Environmental Science & Technology*, **2006**, 40 (1): 53-60.
54. Zhu Y, Price O R, Tao S, Jones K C, Sweetman A J. A new multimedia contaminant fate model for China: How important are environmental parameters in influencing chemical persistence and long-range transport potential? *Environment*

International, **2014**, 69: 18-27.
55. Zhang X, Sühring R, Serodio D, Bonnell M, Sundin N, Diamond M L. Novel flame retardants: Estimating the physical–chemical properties and environmental fate of 94 halogenated and organophosphate PBDE replacements. *Chemosphere*, **2016**, 144: 2401-2407.
56. Wang C, Yuan T, Wood S A, Goss K U, Li J, Ying Q, Wania F. Uncertain Henry's law constants compromise equilibrium partitioning calculations of atmospheric oxidation products. *Atmos. Chem. Phys.*, **2017**, 17 (12): 7529-7540.
57. Ebert R-U, Kühne R, Schüürmann G. Henry's Law Constant─A General-Purpose Fragment Model to Predict Log Kaw from Molecular Structure. *Environmental Science & Technology*, **2023**, 57 (1): 160-167.
58. Nedyalkova M A, Madurga S, Tobiszewski M, Simeonov V. Calculating the Partition Coefficients of Organic Solvents in Octanol/Water and Octanol/Air. *Journal of Chemical Information and Modeling*, **2019**, 59 (5): 2257-2263.
59. Niederer C, Goss K-U. Effect of ortho-chlorine substitution on the partition behavior of chlorophenols. *Chemosphere*, **2008**, 71 (4): 697-702.
60. Wittekindt C, Goss K-U. Screening the partition behavior of a large number of chemicals with a quantum-chemical software. *Chemosphere*, **2009**, 76 (4): 460-464.
61. A. Fleckenstein C, Optimization Strategies Based on the Structure of the Analytes. In *Optimization in HPLC*, 2021; pp 141-164.
62. Matthies M, Beulke S. Considerations of temperature in the context of the persistence classification in the EU. *Environmental Sciences Europe*, **2017**, 29 (1): 15.
63. Mahdi W A, Obaidullah A J. Combination of machine learning and COSMO-RS thermodynamic model in predicting solubility parameters of coformers in production of cocrystals for enhanced drug solubility. *Chemometrics and Intelligent Laboratory Systems*, **2024**, 253: 105219.
64. Salthammer T, Grimme S, Stahn M, Hohm U, Palm W-U. Quantum Chemical Calculation and Evaluation of Partition Coefficients for Classical and Emerging Environmentally Relevant Organic Compounds. *Environmental Science & Technology*, **2022**, 56 (1): 379-391.