

DOI: 10.7524/AJE.1673-5897.20220616001

蒋岚, 徐悦, 张晓宇, 等. 全氟化合物在鱼类肝脏中的生物富集因子预测与影响因素分析[J]. 生态毒理学报, 2023, 18(2): 249-258

Jiang L, Xu Y, Zhang X Y, et al. Prediction of bioconcentration factor and analysis of influencing factors of perfluorinated compounds in fish liver [J]. Asian Journal of Ecotoxicology, 2023, 18(2): 249-258 (in Chinese)

全氟化合物在鱼类肝脏中的生物富集因子预测与影响因素分析

蒋岚¹, 徐悦^{1,*}, 张晓宇^{2,#}, 徐冰峰¹, 徐西蒙¹, 马懿星³

1. 昆明理工大学建筑工程学院, 昆明 650500

2. 云南省生态环境科学研究院, 昆明 650034

3. 昆明理工大学环境科学与工程学院, 昆明 650500

收稿日期: 2022-06-16

录用日期: 2022-08-24

摘要: 全氟化合物(perfluorinated compounds, PFCs)作为一种新兴有机污染物,因其环境持久性及生物富集性,对生态环境健康存在潜在风险。鉴于实验测定 PFCs 生物富集作用的局限性,为实现对 PFCs 生物富集因子(bioconcentration factor, BCF)的合理预测及其影响因素分析,本研究基于前人报道的 log BCF 实验数据,采用多元线性逐步回归方法建立了 PFCs 对鱼类肝脏生物富集作用的定量结构-活性关系(quantitative structure-activity relationship, QSAR)模型,并对该模型进行了全面验证与评估及机理解释。结果表明,所建 QSAR 模型的决定系数($R^2=0.902$)、内部验证指标($Q_{\text{LoO}}^2=0.852$)、外部验证指标($Q_{\text{Fi}}^2=0.855$ 、 $Q_{\text{F2}}^2=0.850$ 、 $Q_{\text{F3}}^2=0.814$)等统计参数均符合建模标准,模型具备良好的拟合优度、稳健性和预测能力;Williams 图显示所建模型具有较强的泛化能力。PFCs 的分子体积(V)和分子表面电位的最小值($V_{\text{s,min}}$)是影响其在鱼类肝脏中富集的重要原因,且分子体积起主导作用;PFCs 的生物富集效应是疏水与氢键相互作用机制共同影响的结果。

关键词: 全氟化合物;鱼类肝脏;生物富集因子;QSAR

文章编号: 1673-5897(2023)2-249-10

中图分类号: X171.5

文献标识码: A

Prediction of Bioconcentration Factor and Analysis of Influencing Factors of Perfluorinated Compounds in Fish Liver

Jiang Lan¹, Xu Yue^{1,*}, Zhang Xiaoyu^{2,#}, Xu Bingfeng¹, Xu Ximeng¹, Ma Yixing³

1. Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650500, China

2. Yunnan Research Academy of Eco-environmental Sciences, Kunming 650034, China

3. Faculty of Environmental Science and Engineering, Kunming University of Science and Technology, Kunming 650500, China

Received 16 June 2022

accepted 24 August 2022

Abstract: As an emerging organic pollutant, perfluorinated compounds (PFCs) are considered to be a potential threat to the ecological environment due to their persistent, bioaccumulative, and toxic properties. Based on the previously reported experimental log bioconcentration factor (BCF) data, the multiple linear stepwise regression meth-

基金项目: 国家自然科学基金面上项目(42177464)

第一作者: 蒋岚(1994—), 女, 硕士研究生, 研究方向为水处理理论与技术, E-mail: 514551718@qq.com

* 通信作者(Corresponding author), E-mail: xvyue1029@hotmail.com

共同通信作者(Co-corresponding author), E-mail: 32754543@qq.com

od was introduced to establish a quantitative structure-activity relationship (QSAR) model to predict the BCF of PFCs on fish liver and its impact mechanism, avoiding the limitations of experimental determination of PFCs bioaccumulation. Subsequently, the proposed model was validated and evaluated, and the identified impact mechanism was explained. The coefficient of determination ($R^2=0.902$), internal validation metric ($Q_{\text{Loo}}^2=0.852$), and external validation metrics ($Q_{\text{F1}}^2=0.855$, $Q_{\text{F2}}^2=0.850$, $Q_{\text{F3}}^2=0.814$) demonstrated the good performance of the proposed model in terms of goodness of fit, robustness and prediction accuracy. Moreover, the Williams graph confirmed the strong generalization ability of the proposed model. Furthermore, molecular volume (V) and the minimum value of molecular surface potential ($V_{s, \min}$) were identified as the key influencing factors affecting the enrichment of PFCs in fish liver, and the former played a dominant role. The bioaccumulation effect of PFCs was found to be the result of the interaction and co-influence of hydrophobic and hydrogen bonds.

Keywords: PFCs; fish liver; bioconcentration factor; QSAR

随着工业化进程的不断加速,环境中的新兴污染物层出不穷,其风险防范工作已日益得到广泛关注。全氟化合物(perfluorinated compounds, PFCs)是化合物分子中与碳原子连接的氢原子全部被氟原子所取代的一类典型新兴持久性有机污染物^[1]。PFCs 因具有疏水疏油、表面活性等特性,被普遍应用于化工及日常消费品的生产中^[2-3]。PFCs 的生产使用及其前驱物质的转化,使其遍及各类水体^[4]、空气^[5]和土壤^[6]等环境介质。PFCs 稳定性强,可通过环境介质进入生物体内并富集,达到一定阈值时将会产生相应的毒效应,损害生物体神经^[7]、生殖^[8]和肝脏^[9]系统等,严重危害人体健康。

生物富集因子(bioconcentration factor, BCF)是描述化合物在生物体内富集作用大小的重要参数。现今,PFCs 及其衍生物层出不穷,但有关 PFCs 的生物富集性研究却极为有限。究其原因,主要受以下 2 个方面的制约:(1) BCF 主要通过实验测定,但测定成本高、过程复杂、耗时长、可操作性低,且在时间上具有滞后性;(2) BCF 的测定涉及动物试验,测定过程中的生态伦理问题不容忽视。及时、全面地完善和补充 PFCs 的生物富集效应理论对科学建立 PFCs 环境标准,准确评估 PFCs 的环境生态风险具有重要意义。

定量结构-活性关系(quantitative structure-activity relationship, QSAR)模型作为一种发展迅速、应用范围较广的理论预测工具,将化合物的分子结构与其性质之间建立函数关系,从而实现对化合物性质的有效预测^[10]。利用 QSAR 模型可快速、准确地预测多种有机污染物的生物活性,例如芳香族化合物对厌氧菌群的毒性预测^[11],含氧多环芳烃对斑马鱼胚胎的毒性预测等^[12]。目前,关于 PFCs 生物富集

性的预测模型还鲜有报道。

因此,鉴于实验研究的局限性及对 PFCs 生物富集效应理论研究的迫切需求,本研究在遵循国际经济合作与发展组织(Organization for Economic Co-operation and Development, OECD)提出的 QSAR 模型构建准则^[13]的基础上建立 PFCs 在鱼类肝脏中生物富集的 QSAR 模型,并对该模型进行了全面验证与评估,以实现对鱼类肝脏中 PFCs 的 log BCF 的合理预测,为评估 PFCs 进入生物体内可能产生的损害奠定理论基础。同时,通过对 QSAR 模型的解释,为从分子角度阐述影响 PFCs 生物富集的主要因素提供了见解,对理解 PFCs 在生态系统中的迁移转化途径及规律具有一定的参考意义。

1 材料与方法 (Materials and methods)

1.1 数据集的收集与处理

从 PFCs 对白鲢、罗非鱼和乌鳢肝脏的生物富集作用研究文献中收集了 log BCF 值^[14-15],包括 11 种全氟烷基羧酸(perfluorinated carboxylic acids, PF-CAs)、5 种全氟烷基磺酸(perfluoroalkyl sulfonic acids, PFSAAs)和 1 种全氟烷基磺酰胺(perfluoroalkane sulfonamides, FOSAs),涵盖了日常生活和工业领域中常见的 PFCs。其中,针对白鲢的 PFCs 生物富集性研究以武汉汤逊湖为取样地^[14],而对罗非鱼和乌鳢开展的相关实验以珠江三角洲为研究区域^[15]。二者的实验方法相似,且均对不同鱼类的生活水体及其肝脏的 PFCs 浓度进行了测定。具体实验方法为:地表水样经滤膜过滤、固相萃取后进行 PFCs 本底浓度的测定;鱼体肝脏经研磨、冷冻和液液萃取后,测定其 PFCs 浓度。经预处理后,水体及鱼类肝脏中 PFCs 浓度的测定均采用高效液相色谱-串联质谱法,进而计算鱼体肝脏的 log BCF 值。

鱼类肝脏中 PFCs 的生物富集性与其生活水体中化合物的本底浓度有关^[15]。经比较发现,当地表水中 PFCs 浓度相近时,罗非鱼和乌鳢肝脏对 PFCs 的富集差异不大,而 PFCs 更易在白鲢肝脏中形成富集,例如两地水体中全氟癸酸(perfluorodecanoic acid, PFDA)的环境浓度接近,但白鲢肝脏的 log BCF 值更大。

因建模数据来源于不同实验条件和实验环境,为减少分析的不确定性,按以下规则对原始数据进行处理:对所收集的同一 PFCs 的多个 log BCF 实验值,先剔除明显偏离整体数据样本的离群值,使得实验数据的变异系数 $\leq 15\%$,再取其平均值建立 QSAR 模型。原始数据经处理后,同一 PFCs 的 log BCF 实验值的变异系数均 $< 15\%$,以其平均值为建模数据具有足够的稳健性,可支持 QSAR 分析^[16-17]。17 种 PFCs 的 log BCF 数值范围为 1.82 ~ 5.29,跨度范围为 3.47,平均值(mean)为 3.51,所对应的标准偏差(standard deviation, SD)为 1.10。所有数据均落在(mean-3SD, mean+3SD)区间范围内,满足建模的数据要求^[18]。使用 Excel 软件随机选择数据集中 80% 的数据为训练集(14 种 PFCs),以建立 QSAR 模型;其余 20% 的数据为验证集(3 种 PFCs),用于对 QSAR 模型进行外部验证。

1.2 分子描述符的计算与筛选

量子化学描述符所代表的物理化学意义明确,有利于解释化合物对生物作用的反应机理。本研究运用 Gaussian 09 软件在 B3LYP/6-31G* 基组水平下优化 PFCs 的分子结构,以获得能量最低的 PFCs 稳定分子构型。采用 Multiwfn 程序^[19]对优化后的 PF-Cs 分子结构进行计算,得到 32 个量子化学描述符,包括但不限于以下描述符:(1)与分子结构特征相关的描述符,如分子量、分子体积和分子表面积等;(2)与分子静电势相关的描述符,如分子静电势正值区域的平均值、分子表面静电势的平均偏差等;(3)与分子极性相关的描述符,如偶极矩、分子极性指数等。

为简化 QSAR 模型的构建过程,对所有描述符进行相关性分析,对相关系数 0.95 以上的描述符,只保留一个与 log BCF 之间相关系数较高的分子描述符。经筛选后,保留 16 个 PFCs 的分子描述符进入 QSAR 模型构建环节。PFCs 的 16 个分子描述符的具体信息如表 1 所示。

1.3 QSAR 模型的建立与验证

基于 SPSS 26 软件,以经筛选后所保留的 16 个

分子描述符为自变量,log BCF 为因变量进行逐步线性回归,得到包含不同个数分子描述符的 QSAR 模型。从模型的拟合优度、稳健性和预测能力 3 个方面考量模型性能,经比较分析后,确定本研究所建立的最终模型。

采用决定系数(R^2)对 QSAR 模型的拟合优度进行评估;以内部验证指标(Q_{Loo}^2)评估 QSAR 模型的稳健性;利用验证集对 QSAR 模型进行外部验证,以外部验证指标(Q_{F1}^2 、 Q_{F2}^2 和 Q_{F3}^2)评估 QSAR 模型的预测能力;对由 QSAR 模型所得的预测值进行残差分析,评估模型是否存在系统误差;由 QSAR 模型中所含分子描述符的显著性指标(P)和方差膨胀系数(VIF),判断各分子描述符是否具有显著性,以及描述符之间是否存在多重共线性。 Q_{Loo}^2 、 Q_{F1}^2 、 Q_{F2}^2 和 Q_{F3}^2 参照文献资料中的通用方法进行计算^[21],其余统计参数由 SPSS 26 软件计算而得。

1.4 QSAR 模型的应用域分析

以 Williams 图^[22]表征 QSAR 模型的应用域,评估其适用范围,判断建模样本中是否存在异常值。Williams 图由标准化残差(δ)与杠杆值(h)组成,当 δ 的绝对值 > 3 时,表明该样本的预测值为异常值;当 h 超过警告杠杆值(h^*)时,表明该样本的结构与训练集中样本的结构有较大差异^[22]。 δ 与 h 采用文献资料中通用方法进行计算^[23]。

2 结果 (Results)

2.1 QSAR 模型的建立与验证

经逐步线性回归后得到了包含不同个数分子描述符的 QSAR 模型,如式(1)~(2)所示,其统计参数如表 2 所示。分子描述符包括分子体积(V)与分子表面电位的最小值($V_{\text{s,min}}$)。

$$\log \text{BCF} = 0.001641 V + 0.007 \quad (1)$$

$$\log \text{BCF} = 0.001577 V - 14.55 V_{\text{s,min}} - 14.90 \quad (2)$$

由表 2 可知,2 个 QSAR 模型均具备显著性。对比两模型的训练集统计参数,模型(2)的拟合优度与稳健性较好,优于模型(1);测试集范畴内,模型(1)的外部预测能力更佳。

两模型的建立均符合样本数/变量数 ≥ 5 的建模经验规则^[24]。由表 2 可知,两模型的 $R^2 > 0.8$, $Q_{\text{Loo}}^2 > 0.5$, Q_{F1}^2 、 Q_{F2}^2 和 $Q_{\text{F3}}^2 > 0.5$,根据模型应用的评估标准^[21],表明两模型均具备良好的拟合优度、稳健性和预测能力,符合 QSAR 模型构建准则的要求。两模型的 $R^2 - Q_{\text{Loo}}^2 < 0.3$,即两模型均不存在过度拟合现象。

表 1 16 个分子描述符的具体信息
Table 1 Specific information of 16 molecular descriptors

序号 No.	分子描述符名称 Molecular descriptor name	含义 Description	计算的基组水平 The basis set level of the calculation
1	分子体积(V) Molecular volume (V)	与分子的大小有关 Related to the size of the molecule	B3LYP/def2-TZVP
2	分子表面电位的最小值($V_{s,min}$) The minimum value of molecular surface potential ($V_{s,min}$)	与分子接受质子形成氢键的能力有关 Related to the ability of molecules to accept protons to form hydrogen bonds	B3LYP/def2-TZVP
3	分子表面静电势的平均偏差(π) Average deviation of electrostatic potential on molecular surfaces (π)	与分子的电荷分离有关 Related to charge separation of molecules	B3LYP/def2-TZVP
4	最高占据分子轨道能量与 最低未占据分子轨道能量的差值(E_{gap}) The energy difference between the highest occupied molecular orbital and the lowest unoccupied molecular orbital (E_{gap})	与分子电离的点位有关 Related to the point of molecular ionization	B3LYP/def2-TZVP
5	以碳原子为探针原子的分子 范德华势 ^[20] 的全局极小值(V^{dw}) Global minima of molecular van der Waals potential ^[20] with carbon atoms as probe atoms (V^{dw})	与分子间范德华相互作用有关 Related to intermolecular van der Waals interactions	B3LYP/def2-TZVPD
6	分子静电势负值区域的表面积(S^-) The surface area of the region where the molecular electrostatic potential is negative (S^-)	与分子的静电势有关 Related to the electrostatic potential of the molecule	B3LYP/def2-TZVP
7	分子静电势正值区域的平均值(V_{mean}^+) The average value of the positive region of the molecular electrostatic potential (V_{mean}^+)	与分子的静电势有关 Related to the electrostatic potential of the molecule	B3LYP/def2-TZVP
8	分子表面平均局部离子化能的最大值($ALIE_{max}$) The maximum average local ionization energy on the molecular ($ALIE_{max}$)	与分子在空间任意点电离电子所需的平均能量有关 Required to ionize electrons with molecules at any point in space average energy	B3LYP/def2-TZVP
9	分子表面平均局部离子化能的最小值($ALIE_{min}$) The minimum average local ionization energy on the molecular ($ALIE_{min}$)	与分子在空间任意点电离电子所需的平均能量有关 Required to ionize electrons with molecules at any point in space average energy	B3LYP/def2-TZVP
10	分子极性指数(MPI) Molecular polarity index (MPI)	与分子的极性有关 Related to the polarity of the molecule	B3LYP/def2-TZVP
11	偶极矩(μ) Dipole moment (μ)	与分子的极性有关 Related to the polarity of the molecule	B3LYP/def2-TZVPD
12	第一超极化率(β) Magnitude of first hyperpolarizability (β)	与分子的极性有关 Related to the polarity of the molecule	B3LYP/def2-TZVPD
13	垂直电离能(VIP) Vertical ionization potential (VIP)	与分子得失电子的能力有关 Related to the ability of molecules to gain and lose electrons	B3LYP/def2-TZVP
14	Mulliken 电负性(X_M) Mulliken electronegativity (X_M)	与分子得失电子的能力有关 Related to the ability of molecules to gain and lose electrons	B3LYP/def2-TZVP
15	电子硬度(η) Electronic hardness (η)	与分子得失电子的能力有关 Related to the ability of molecules to gain and lose electrons	B3LYP/def2-TZVP
16	亲电指数(ω_{cubic}) Electrophilic index (ω_{cubic})	与分子得失电子的能力有关 Related to the ability of molecules to gain and lose electrons	B3LYP/def2-TZVP

表 2 不同 QSAR 模型的统计参数
Table 2 Statistical parameters of QSAR model

模型 Model	训练集 Training set						测试集 Validating set				
	<i>n</i>	<i>R</i> ²	<i>Q</i> _{Lo0} ²	RMSE	<i>F</i>	<i>P</i>	<i>n</i>	<i>Q</i> _{F1} ²	<i>Q</i> _{F2} ²	<i>Q</i> _{F3} ²	RMSE
(1)	14	0.847	0.800	0.442	66.539	0.000	3	0.954	0.952	0.941	0.254
(2)	14	0.902	0.852	0.369	50.784	0.000	3	0.855	0.850	0.814	0.451

注:*n* 为样本集个数;RMSE 为标准误差;*F* 为方差比;*P* 为显著性指标,当 $P<0.05$ 时,表明模型具有显著性。
Notes:*n* means the number of data set; RMSE means root mean square error; *F* means variance ratio; *P* means significance level, when $P<0.05$, the model is significant.

象^[25]。在 QSAR 模型构建过程中,应考虑尽可能多的分子描述符,以提高其拟合优度与稳健性^[26]。同时,当测试集中的样本数据相对较少,分析的不确定性将会有所提高^[24]。另一方面,相比于单因子模型,双因子模型的解释能力更强。具有相同官能团的同系 PFCs,随着碳链长度(CF₂)的增加,其分子体积随之增加,log BCF 值呈递增趋势。例如,全氟壬酸(perfluorononanoic acid, PFNA)与 PFDA 的分子表面电位的最小值相近,但因 PFDA 的分子体积较大,其 log BCF 也更大。非同系 PFCs 不仅存在碳链长度(CF₂)的差异,其末端所带官能团也不同。通过对 PFDA 和全氟辛基磺酰胺(perfluorooctane sulfonamide, PFOSA)进行比较分析发现,PFOSA 的分子体积小于 PFDA,但其 log BCF 却更大,多因 PFOSA 的分子表面电位的最小值更小所致。双因子模型在对非同系 PFCs 的生物富集性作出定性机理解释的同时,还具有一定的定量分析意义。基于上述原因,本研究选择模型(2)为最终 QSAR 模型。

表 3 中给出了使用 QSAR 模型预测 PFCs 的 log BCF 值的具体结果。log BCF 预测值与实验值之间的相关性如图 1 所示,所有数据点均分布在 45° 线附近,模型对 PFCs 的 log BCF 值预测精度较高。log BCF 的残差分布如图 2 所示,所有残差均随机分布在基线两侧,无明显的规律性,模型不存在系统误差。

表 4 列出了 QSAR 模型中所含分子描述符的 *P* 值和 VIF 值。当分子描述符的 VIF 值均<10 时,表明分子描述符间不存在多重共线性^[27]。由表 4 可知,QSAR 模型中各分子描述符均为关键描述符且彼此之间不存在多重共线性。

2.2 QSAR 模型的应用域分析

通过 Williams 图划定本文所建 QSAR 模型的适用范围,如图 3 所示(图中杠杆值 $h<0.64$ 与标准

残差的绝对值 $|\delta|<3$ 的界定范围内)。测试集与验证集中所有数据样本均在 δ 与 h 的界定范围内,表明该 QSAR 模型具有较强的预测能力及泛化能力。

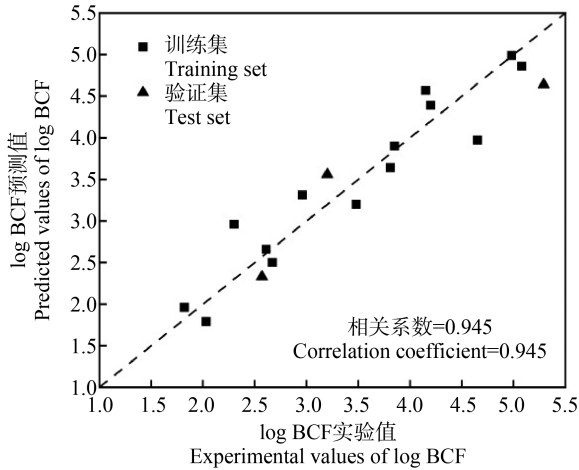


图 1 实验值与 QSAR 模型预测值的相关性

Fig. 1 Experimental and predicted values of the QSAR model

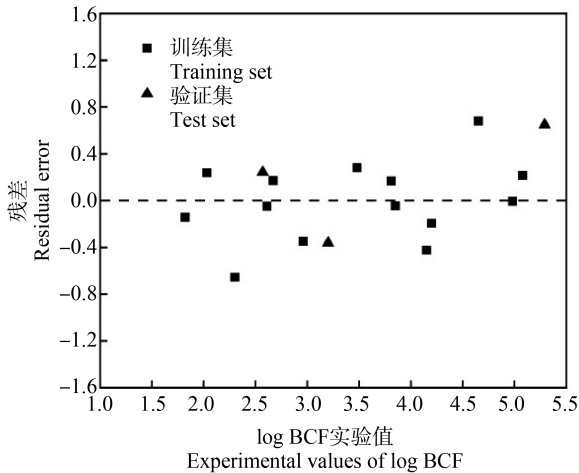
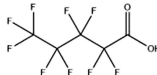
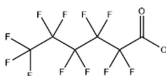
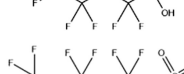


图 2 QSAR 模型的残差分布图

Fig. 2 Residual diagram of the QSAR models

表 3 QSAR 模型参数与模型预测结果
Table 3 Model parameters and model predicting results

全氟化合物(PFCs) Perfluorinated compounds (PFCs)	CAS	分子结构 Molecular structures	分子描述符 Molecular descriptors		生物富集因子 log BCF		
			V/Bohr ³	V _{s,min} /eV	实验值 Exp.	预测值 Pred.	残差 Res.
全氟丁酸 Perfluorobutyric	375-22-4		1 067.45	-1.0315	2.03	1.79	0.24
全氟戊酸 perfluorovaleric acid	2706-90-3		1 280.86	-1.0202	1.82	1.96	-0.14
全氟己酸* Perfluorohexanoic acid*	307-24-4		1 494.00	-1.0221	2.57	2.33	0.24
全氟庚酸 Perfluoroheptanoic acid	375-85-9		1 707.31	-1.0218	2.61	2.66	-0.05
全氟辛酸 Perfluorooctanoic acid	335-67-1		1 920.77	-1.0190	2.30	2.96	-0.66
全氟壬酸 Perfluorononanoic acid	375-95-1		2 133.68	-1.0203	2.96	3.31	0.35
全氟癸酸 Perfluorodecanoic acid	335-76-2		2 346.86	-1.0198	3.81	3.64	0.17
全氟十一酸 Perfluoroundecanoic acid	2058-94-8		2 560.05	-1.0194	4.65	3.97	0.68
全氟十二酸 Perfluorododecanoic acid	307-55-1		2 774.07	-1.0577	5.08	4.86	0.22
全氟十三酸* Perfluorotridecanoic acid*	72629-94-8		2 986.35	-1.0194	5.29	4.64	0.65
全氟十四酸 Perfluorotetradecanoic acid	376-06-7		3 199.66	-1.0199	4.98	4.99	-0.01
全氟丁烷磺酸 Perfluorobutane sulfonic acid	375-73-5		1 420.07	-1.0420	2.67	2.50	0.17
全氟己烷磺酸 Perfluorohexane sulfonic acid	355-46-4		1 846.81	-1.0437	3.48	3.20	0.28
全氟庚烷磺酸* Perfluoroheptane sulfonic acid*	375-92-8		2 060.03	-1.0455	3.20	3.56	-0.36
全氟辛烷磺酸 Perfluorooctane sulfonic acid	1763-23-1		2 273.14	-1.0454	3.85	3.90	-0.05
全氟癸烷磺酸 Perfluorodecane sulfonic acid	335-77-3		2 699.39	-1.0458	4.15	4.57	-0.42
全氟辛基磺酰胺 Perfluorooctane sulfonamide	754-91-6		2 312.04	-1.0755	4.20	4.39	-0.19

注:表中带* 的 PFCs 为验证集数据。

Notes: * Chemicals were in the validation set.

表 4 QSAR 模型中不同分子描述符的统计参数

Table 4 Statistical parameters of different Molecular descriptor		
分子描述符 Molecular descriptor	显著性指标(P) Significance indicator (P)	方差膨胀系数(VIF) Variance inflation factor (VIF)
V	0.000	1.023
V _{s,min}	0.030	1.023

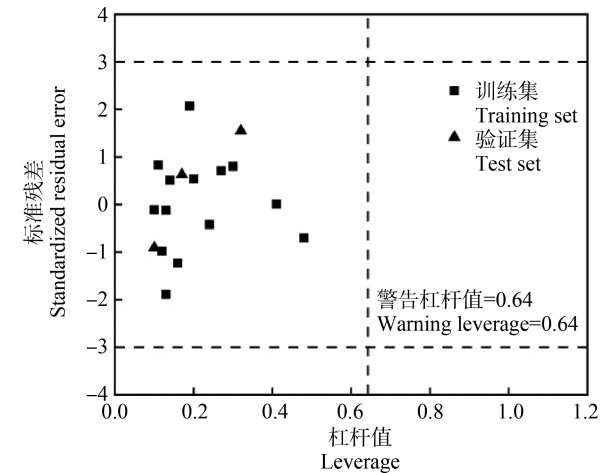


图 3 QSAR 模型的 Williams 图

注:当 δ 的绝对值 >3 时,表明该样本的预测值为异常值;
若 h 超过警告杠杆值 h^* 时,表明该样本的结构
与训练集中样本的结构有较大差异。

Fig. 3 Williams diagram of the QSAR models
Note: When the absolute value of $\delta > 3$, it indicates that the predicted value of the sample is abnormal; if h exceeds the warning leverage value h^* , it indicates that the structure of the sample is quite different from that in the training set.

2.3 log BCF 的影响因素分析

由 QSAR 模型可知,PFCs 的 log BCF 与其 V 和 V_{s,min} 有关,且与 V 呈正相关,与 V_{s,min} 呈负相关。经计算,QSAR 模型中 V 与 V_{s,min} 的标准化回归系数^[28]分别为 0.885、-0.237,表明 PFCs 的 V 是影响其 log BCF 大小的主要因素。

V 反映分子的大小,较大分子尺寸的 PFCs 具有更强的疏水性^[29],而疏水相互作用可驱动吸附质与吸附剂间的吸附^[30]。PFCs 的 V 越大,其分子尺寸越大,疏水性也越强,更易吸附在鱼类的肝脏细胞上,从而表现出的生物富集性也就越大。

V_{s,min} 反映分子接受质子形成氢键的能力(即氢键碱度)^[31],氢键碱度较高的分子更倾向于与水形成氢键^[32]。PFCs 的 V_{s,min} 越大,其氢键碱度越高,更易

与水发生相互作用形成氢键,从而表现出的生物富集性也就越小。

3 讨论 (Discussion)

本文基于 17 种 PFCs 的分子结构特性,建立鱼类肝脏中 PFCs 的 log BCF 的 QSAR 预测模型,与现有相关模型相比,本模型具有可靠性强、应用域广和作用机理清晰的特点。Liu 等^[33]以半最大效应浓度(log EC₅₀)为单分子描述符建立了 PFCs 在青口贻贝中的生物富集性的 QSAR 模型($R^2=0.999$)。该模型具有高 R^2 ,但其建模数据集所含 PFCs 的数量较少(仅包含 4 种 PFCs),模型的建模过程未完全遵循 QSAR 模型构建的五项准则^[13],缺乏对 QSAR 模型外部预测能力的验证及对应用域的定义;此外,以 log EC₅₀ 为分子描述符建立 QSAR 模型,并非纯粹的“结构-性质”关系,该描述符多由实验测定,其实验结果受多因素影响,数据的准确性和可靠性受限,模型不确定性较高。本文建立的 QSAR 模型所采用的数据量有所增加,数据集的覆盖面有所加大;描述符由理论计算获得,数据稳定,可靠性佳,建模过程完全基于 QSAR 模型构建的准则框架^[13],模型适用性较好;且经全面验证表明,本模型具备良好的拟合优度、稳健性和预测能力。Bhatarai 和 Gramatica^[34]分析了 PFCs 在虹鳟鱼体内的 log BCF 值随 PFCs 碳链长度的增加而增加,PFCs 的疏水性越强越容易在鱼体组织中富集。这与本研究模型筛选出的表征疏水性的 V 值与 PFCs 在鱼类肝脏中生物富集性呈正相关的结果一致。本模型研究表明,除与疏水性相关的 PFCs 的分子体积外,与化合物氢键碱度相关的分子表面电位的最小值对 PFCs 的生物富集性也具有影响作用;对于非同系 PFCs,模型可从定性与定量分析的角度予以阐释,丰富了 PFCs 生物富集效应研究理论。

综上所述,在适用范围内,本文所建的 QSAR 模型可预测目前大多数 PFCs 的 log BCF 值,从生物富集性角度,为 PFCs 的生态风险评估提供数据支持。

基于本文的研究结果,未来可侧重于以下几个方向开展研究。

(1)结合多种机器学习方法,对 PFCs 的分子结构与 log BCF 之间进行线性与非线性拟合尝试,进一步提升 QSAR 模型的拟合优度、稳健性与预测能力。虽然大多数非线性模型属于“黑箱”模型,不利于作用机理的解释,但 PFCs 在生物体内的生物富

集行为是一复杂过程,极有可能涉及非线性关系,因此,开展非线性模型的研究有利于完善 PFCs 的生物富集性理论体系。

(2)通过对 QSAR 模型的机理解释,发现 PFCs 的分子体积和分子表面电位的最小值是影响其在鱼类肝脏中富集的重要因子,且分子体积起主导作用。PFCs 的生物富集效应受多种作用机制共同影响,如疏水作用、氢键等。在化工生产中,可根据影响 PFCs 生物富集性的结构因素,合成对生态风险影响较小的新型 PFCs。

(3)现研究大多采用美国环境保护局推荐的非致癌物质健康风险模型^[35]对 PFCs 进行风险评估,该方法主要以 PFCs 在环境中的质量浓度和人体摄入量为基础进行分析,其评估结果具有滞后性。在今后的研究中可尝试建立包含多环境生态风险因子的综合指标体系(如生物富集因子、生物毒性和化合物降解难易程度指标等),从 PFCs 固有的分子结构特性出发,探索评估 PFCs 健康风险的新理论与新方法,对现有或新兴 PFCs 可能产生的生态安全威胁进行客观及时预警。

通信作者简介:徐悦(1982—),女,博士,讲师,主要研究方向为水处理理论与技术。

共同通信作者简介:张晓宇(1982—),男,博士,高级工程师,主要研究方向为生态毒理学。

参考文献 (References):

- [1] 李春梅,岳宁,周杰,等.基于超高效液相色谱-四极杆/静电场轨道阱高分辨质谱技术研究全氟化合物物质谱裂解规律[J].食品安全质量检测学报,2020,11(22): 8380-8386
Li C M, Yue N, Zhou J, et al. Study on the mass spectrometry fragmentation of perfluorinated compounds based on ultra performance liquid chromatography-quadrupole/electrostatic field orbitrap high resolution mass spectrometry [J]. Journal of Food Safety & Quality, 2020, 11(22): 8380-8386 (in Chinese)
- [2] 顾春节,钟哲辉,徐晨烨,等.印染末端废水中全氟化合物的污染特征、影响因素及风险评估[J].环境科学学报,2021,41(5): 1920-1929
Gu C J, Zhong Z H, Xu C Y, et al. Occurrence, influencing factors, and risks assessment of perfluorinated compounds (PFCs) in wastewater from textile and dyeing industry [J]. Acta Scientiae Circumstantiae, 2021, 41 (5): 1920-1929 (in Chinese)
- [3] Ahrens L, Bundschuh M. Fate and effects of poly- and perfluoroalkyl substances in the aquatic environment: A review [J]. Environmental Toxicology and Chemistry, 2014, 33(9): 1921-1929
- [4] Liu Y, Li X, Wang X, et al. Contamination profiles of perfluoroalkyl substances (PFAS) in groundwater in the alluvial-pluvial plain of Hutuo River, China [J]. Water, 2019, 11(11): 1-2316
- [5] 李梦丹.上海市大气中全氟化合物迁移和转化的数值模拟研究[D].上海:上海交通大学,2016: 13-14
Li M D. Numerical simulation of perfluorinated compounds migration and transformation in Shanghai's atmosphere [D]. Shanghai: Shanghai Jiao Tong University, 2016: 13-14 (in Chinese)
- [6] Xiao F, Simcik M F, Halbach T R, et al. Perfluorooctane sulfonate (PFOS) and perfluorooctanoate (PFOA) in soils and groundwater of a U.S. metropolitan area: Migration and implications for human exposure [J]. Water Research, 2015, 72: 64-74
- [7] Mariussen E. Neurotoxic effects of perfluoroalkylated compounds: Mechanisms of action and environmental relevance [J]. Archives of Toxicology, 2012, 86(9): 1349-1367
- [8] 张小梅,宋锦兰,金一和,等.全氟辛烷磺酸对雄性鹌鹑生殖毒性影响[J].生态毒理学报,2011,6(2): 143-148
Zhang X M, Song J L, Jin Y H, et al. Study on reproductive toxicity of perfluorooctane sulfonate (PFOS) in male quail [J]. Asian Journal of Ecotoxicology, 2011, 6(2): 143-148 (in Chinese)
- [9] 谢蕾,章涛,孙红文.全氟烷基化合物在人体肝脏中的富集特征及其与肝损伤的关系[J].环境化学,2020,39(6): 1479-1487
Xie L, Zhang T, Sun H W. Enrichment characteristics of perfluoroalkyl substances (PFASs) in human liver and its association with hepatocyte apoptosis [J]. Environmental Chemistry, 2020, 39(6): 1479-1487 (in Chinese)
- [10] Schultz T W, Cronin M T D, Walker J D, et al. Quantitative structure - activity relationships (QSARs) in toxicology: A historical perspective [J]. Journal of Molecular Structure: THEOCHEM, 2003, 622(1-2): 1-22
- [11] 吴少奇,王黎,夏正海,等. QSAR 模型预测石化废水中芳香族物质对厌氧菌群的综合毒性[J].生态毒理学报,2020,15(6): 167-174
Wu S Q, Wang L, Xia Z H, et al. QSAR modelling for predicting comprehensive toxicity of aromatic substances to anaerobic microflora in petrochemical wastewater [J]. Asian Journal of Ecotoxicology, 2020, 15(6): 167-174 (in Chinese)

- [12] 金玲敏, 徐童, 乔显亮. OPAHs 对斑马鱼胚胎的急性毒性预测[J]. 生态毒理学报, 2021, 16(6): 53-59
Jin L M, Xu T, Qiao X L. Prediction of acute toxicity by OPAHs to zebrafish embryos [J]. Asian Journal of Ecotoxicology, 2021, 16(6): 53-59 (in Chinese)
- [13] Organization for Economic Co-operation and Development (OECD). Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q) SAR] Models [M]// Environment Health and Safety Publications Series on Testing and Assessment. Paris: OECD, 2007, 2: 1-154
- [14] 胡哲. 不同生物介质中全氟化合物的分布特征及富集规律研究[D]. 武汉: 华中农业大学, 2016: 57
Hu Z. Study on the distribution characteristics and enrichment laws of perfluorinated compounds in different biological media [D]. Wuhan: Huazhong Agricultural University, 2016: 57 (in Chinese)
- [15] Pan C G, Zhao J L, Liu Y S, et al. Bioaccumulation and risk assessment of per- and polyfluoroalkyl substances in wild freshwater fish from rivers in the Pearl River Delta region, South China [J]. Ecotoxicology and Environmental Safety, 2014, 107: 192-199
- [16] 百度百科. 变异系数[EB/OL]. (2021-12-31) [2022-07-18]. <http://baike.baidu.com/item/变异系数/6463621?fr=aladdin>
- [17] Brusseau M L. Estimating the relative magnitudes of adsorption to solid-water and air/oil-water interfaces for per- and poly-fluoroalkyl substances [J]. Environmental Pollution, 2019, 254(Pt B): 113102
- [18] Zhang L. Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems [M]. Singapore: Springer, 2016: 497-503
- [19] Lu T, Chen F W. Multiwfn: A multifunctional wavefunction analyzer [J]. Journal of Computational Chemistry, 2012, 33(5): 580-592
- [20] Lu T, Chen Q X. Van der Waals potential: An important complement to molecular electrostatic potential in studying intermolecular interactions [J]. Journal of Molecular Modeling, 2020, 26(11): 315
- [21] Kunal R, Supratik K, Rudra N D. A Primer on QSAR/QSPR Modeling [M]. Berlin: Springer, 2015: 48-52
- [22] Zhu T Y, Wu J, He C D, et al. Development and evaluation of MTLSE and QSAR models for predicting polyethylene-water partition coefficients [J]. Journal of Environmental Management, 2018, 223: 600-606
- [23] 何培, 潘勇, 蒋军成, 等. 芳香族硝基化合物爆速的定量构效关系预测[J]. 中国安全科学学报, 2018, 28(7): 32-37
He P, Pan Y, Jiang J C, et al. Prediction of detonation velocity of nitro aromatic compounds based on quantitative structure-property relationship [J]. China Safety Science Journal, 2018, 28(7): 32-37 (in Chinese)
- [24] 李建凤, 廖立敏. 不饱和烃类化合物溶解度(-lg S_w)的QSPR 研究[J]. 湖南师范大学自然科学学报, 2020, 43(1): 68-74
Li J F, Liao L M. QSPR study on solubility (-lg S_w) of unsaturated hydrocarbons [J]. Journal of Natural Science of Hunan Normal University, 2020, 43(1): 68-74 (in Chinese)
- [25] Kiralj R, Ferreira M M. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application [J]. Journal of the Brazilian Chemical Society, 2009, 20: 770-787
- [26] Li Y, Yu X Y, Chen X, et al. Underlying mechanisms for the impacts of molecular structures and water chemistry on the enrichment of poly/perfluoroalkyl substances in aqueous aerosol [J]. Science of the Total Environment, 2022, 803: 150003
- [27] 吕利平, 李兵, 何树华, 等. 基于定量-构效关系预测含低碳酯二元共沸物的共沸温度[J]. 化学工程, 2019, 47(11): 44-49
Lyu L P, Li B, He S H, et al. Predicting the azeotropic temperature of binary azeotropes containing low carbon esters based on the quantitative structure-property relation theory [J]. Chemical Engineering (China), 2019, 47(11): 44-49 (in Chinese)
- [28] 刘莹, 于影, 石宝友, 等. 供水管网铁释放的影响因素相对重要性分析[J]. 环境科学, 2017, 38(12): 5090-5096
Liu Y, Yu Y, Shi B Y, et al. Relative importance of factors influencing iron release in drinking water distribution systems [J]. Environmental Science, 2017, 38(12): 5090-5096 (in Chinese)
- [29] Rayne S, Forest K, Friesen K J. Estimated bioconcentration factors (BCFs) for the C4 through C8 perfluorinated alkylsulfonic acid (PFSA) and alkylcarboxylic acid (PFCA) congeners [J]. Journal of Environmental Science and Health, Part A, 2009, 44(6): 598-604
- [30] Apul O G, Wang Q L, Shao T, et al. Predictive model development for adsorption of aromatic contaminants by multi-walled carbon nanotubes [J]. Environmental Science & Technology, 2013, 47(5): 2295-2303
- [31] Pedretti A, Villa L, Vistoli G. VEGA: A versatile program to convert, handle and visualize molecular structure on Windows-based PCs [J]. Journal of Molecular Graphics and Modelling, 2002, 21(1): 47-49
- [32] Apul O, Perreault F, Ersan G, et al. Linear solvation ener-

- gy relationship development for adsorption of synthetic organic compounds by carbon nanomaterials: An overview of the last decade [J]. *Environmental Science: Water Research & Technology*, 2020, 6(11): 2949-2957
- [33] Liu C H, Chang V W C, Gin K Y H. Oxidative toxicity of perfluorinated chemicals in green mussel and bioaccumulation factor dependent quantitative structure-activity relationship [J]. *Environmental Toxicology and Chemistry*, 2014, 33(10): 2323-2332
- [34] Bhatarai B, Gramatica P. Prediction of aqueous solubility, vapor pressure and critical micelle concentration for aquatic partitioning of perfluorinated chemicals [J]. *Environmental Science & Technology*, 2011, 45 (19): 8120-8128
- [35] US Environmental Protection Agency. Guidelines for exposure assessment [R]. Washington DC: US Environmental Protection Agency, 1992 ◆