

DOI:10.7524/AJE.1673-5897.20191223003

于海英, 李美萍, 郝俊生. 基于 MLP-ANN 和 SVM 方法的多氯代二苯并呋喃光解半衰期 QSPR 比较研究[J]. 生态毒理学报, 2020, 15(4): 240-247
Yu H Y, Li M P, Hao J S. A comparative QSPR study on photolysis half-lives of polychlorinated dibenzofurans by MLP-ANN and SVM [J]. Asian Journal of Ecotoxicology, 2020, 15(4): 240-247 (in Chinese)

基于 MLP-ANN 和 SVM 方法的多氯代二苯并呋喃光解半衰期 QSPR 比较研究

于海英¹, 李美萍², 郝俊生¹

1. 山西大学化学化工学院, 太原 030006

2. 山西大学生命科学学院, 太原 030006

收稿日期: 2019-12-23 录用日期: 2020-03-30

摘要: 多氯代二苯并呋喃(PCDFs)是全球性污染物之一, 光化学降解是其主要的的环境降解途径。基于分子二维拓扑结构提出的用于表征化合物结构参数的分子电性距离矢量描述子(MEDV), 应用多层感知器神经网络(MLP-ANN)和支持向量机(SVM)对 PCDFs 在云杉针叶和飞灰表面的光解半衰期($t_{1/2}$)进行定量结构-性质相关(QSPR)分析, 并用交互检验和外部样本对所建模型的稳定性进行了检验。旨在为 PCDFs 光解机理的 QSPR 研究提供新思路。结果表明, 所建模型均具有良好的稳定性和预测能力, 尤以 MLP-ANN 模型为佳, 其建模相关系数(R_{cum})、留一法交互检验相关系数(Q_{Loo})以及外部样本检验相关系数(Q_{ext})分别为 0.850、0.816、0.954(云杉针叶表面)和 0.892、0.753、0.897(飞灰表面)。

关键词: 多氯代二苯并呋喃(PCDFs); 分子电性距离矢量(MEDV); 光解半衰期; QSPR; 多层感知器神经网络(MLP-ANN); 支持向量机(SVM)

文章编号: 1673-5897(2020)4-240-08 中图分类号: X171.5 文献标识码: A

A Comparative QSPR Study on Photolysis Half-lives of Polychlorinated Dibenzofurans by MLP-ANN and SVM

Yu Haiying¹, Li Meiping^{2,*}, Hao Junsheng¹

1. College of Chemistry & Chemical Engineering, Shanxi University, Taiyuan 030006, China

2. College of Life Sciences, Shanxi University, Taiyuan 030006, China

Received 23 December 2019 accepted 30 March 2020

Abstract: Environmental pollution by polychlorinated dibenzofurans (PCDFs) is one of the global environment issue. Photolysis is the main pathway of PCDFs degradation in the environment. In this paper, molecular electronegativity distance vector (MEDV) descriptors, based on two-dimensional topological structure, are used to describe the structures of PCDFs compounds. Using multilayer perceptron neural network (MLP-ANN) and support vector machine (SVM), a quantitative structure-property relationship (QSPR) study was performed to correlate the photolysis half-lives ($t_{1/2}$) of PCDFs adsorbed on the spruce needles and fly ash surfaces with their molecular structures. Meanwhile, the stability of the model was tested by cross validation and the external validation of the test samples.

第一作者: 于海英(1970—), 女, 硕士, 研究方向为药物合成及分析, E-mail: haiyingy@sxu.edu.cn

* 通讯作者(Corresponding author), E-mail: lmpmg@sxu.edu.cn

The aim of this study is to provide a new idea or method reference for QSPR study on the photolysis mechanism of PCDFs. The established models, especially the MLP-ANN model showed good stability and predictive ability. The correlation coefficients including cumulative multiple correlation coefficient of training set (R_{cum} , cumulative cross-validated R_{cum} of training set Q_{LOO} and external Q of test set Q_{ext}) of the MLP-ANN model are 0.850, 0.816, 0.954 (spruce needles) and 0.892, 0.753, 0.897 (flying ash surface), respectively.

Keywords: polychlorinated dibenzofuran (PCDFs); molecular electronegativity distance vector (MEDV); photolysis half-life; QSPR; multi-layer perceptron-artificial neural network (MLP-ANN); support vector machine (SVM)

多氯代二苯并呋喃 (polychlorinated dibenzofurans, PCDFs) 属二恶英 (PCDD/Fs) 类化合物, 是一种典型的持久性有机污染物 (POPs), 也是首批列入《斯德哥尔摩公约》的 12 种 POPs 之一, 具有较强的稳定性, 在自然条件下很难发生生物和化学降解, 光降解是其在环境中转化的主要途径^[1-5]。植物表面光解过程对于有机污染物从大气到食物链的迁移有重要的影响。PCDD/Fs 在植物表面上的光解主要发生在植物叶面角质层中, 其中, 光解半衰期 ($t_{1/2}$) 对于考察此类污染物的环境行为并进行环境风险评价有重要的意义。由于氯原子的取代数目和取代位置的不同, PCDFs 共有 135 种分子结构。实验测定 PCDD/Fs 的光解半衰期需要特殊设备, 消耗大量的时间和财力, 因此, 发展 PCDD/Fs 光解行为的定量结构-性质关系模型十分必要。研究表明, 定量结构-性质关系 (quantitative structure-property relationship, QSPR)^[6-9] 在预测化合物生物性质方面已成为化学、环境等学科的一个重要方法。Chen 等^[10-11] 采用量子化学参数分别对 PCDD/Fs 在水和乙腈中的量子产率以及樱桃树叶表面的光降解速率进行了 QSPR 研究; 王蕾等^[12] 采用电性拓扑指数对 PCDFs 在云杉针叶和飞灰表面的光解半衰期进行了 QSPR 模拟分析; Niu 等^[3-4, 13] 和牛军峰等^[4] 先后采用量子化学参数应用遗传算法对 PCDD/Fs 在云杉针叶和飞灰表面的光降解行为进行了 QSPR 研究, 赵继颖等^[15] 采用分子空间坐标参数对 PCDFs 在云杉针叶和飞灰表面的光降解行为进行了 QSPR 研究, 这些分析结果都从不同方面揭示了影响 PCDFs 光解速率的因素, 但存在有些参数较抽象或者所建模型用的参数较多, 本文在课题组前期研究的基础上^[16], 以具有明确物理-化学意义的分子电性距离矢量 (MEDV) 描述子^[17] 为自变量, 应用多层感知器神经网络 (MLP-ANN) 和支持向量机 (SVM) 建模方法对 PCDFs 在云杉针叶和飞灰表面的光解半衰期 ($t_{1/2}$) 进行定量构效关系分析, 建立了具有较强预测能力和稳定性的 QSPR

模型, 有助于对 PCDD/Fs 光解机理的研究。

1 数据来源 (Data source)

2,3,7,8-PCDF 的结构如图 1 所示, 当氯在不同的取代位置时, 就得到不同的 PCDFs。云杉针叶表面和飞灰表面的 48 个 PCDF 的光解半衰期值取自文献^[16], 在此用光解半衰期的对数值 ($\lg t_{1/2}$) 进行 QSPR 建模分析。

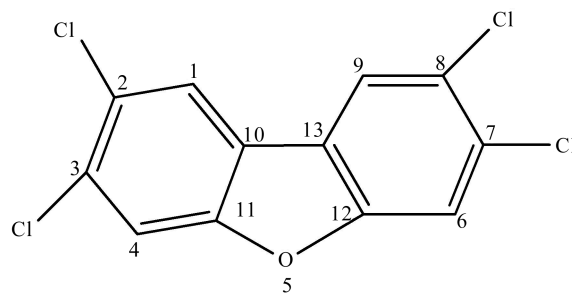


图 1 2,3,7,8-PCDF 结构图

Fig. 1 The generic structure of 2,3,7,8-PCDF

2 研究方法 (Research method)

2.1 MEDV 的计算

48 个 PCDF 的 10 个 MEDV 描述子按文献^[17] 计算, 其原理已有详细介绍, 在此不再赘述。由于 PCDFs 分子中不含有第四类型的原子, 所以 M_{14} 、 M_{24} 、 M_{34} 和 M_{44} 均为零向量, 剩余向量依次为 M_{11} 、 M_{12} 、 M_{13} 、 M_{22} 、 M_{23} 和 M_{33} , 48 个 PCDF 的 MEDV 描述子及其光解半衰期 ($\lg t_{1/2}$) 详见文献^[1]。

为深入研究 PCDFs 的 6 个 MEDV 与光解半衰期的 QSPR 隐含关系, 使用 MLP-ANN 和 SVM 建立模型。

2.2 基于 MEDV 的 QSPR 预测模型

基于 MEDV 计算方法和高维非线性模型人工神经网络 (ANN) 和 SVM, 提出一种基于 MEDV 的 QSPR 预测模型用于描述分析 PCDFs 的 6 个 MEDV 与光解半衰期的 QSPR 的高维非线性关系, 具体步骤如下。

输入:训练集为 $\{(x_i, y_i)\}_{i=1}^{42}$, 式中, $x_i = (m_{11}, m_{12}, m_{13}, m_{22}, m_{23}, m_{33})$, 为输入向量, $y_i = y_1$ 或 $y_i = y_2$ 。 $m_{11}, m_{12}, m_{13}, m_{22}, m_{23}, m_{33}$ 为 PCDFs 的 MEDV 值, y_1 和 y_2 分别为 PCDFs 在云杉针叶和飞灰表面光解半衰期的对数值。测试集 $\{(x_i, y_i)\}_{i=43}^{47}$ 中的输入向量 $x_i, i=43, \dots, 47$ 。

输出: $y_i (i=43, \dots, 47)$ 的估计结果 $f(x_i)$ 。

第一步:使用训练集 $\{(x_i, y_i)\}_{i=1}^{42}$ 训练预测模型 $y_1 = f_1(x)$ 和 $y_2 = f_2(x)$, 其中模型 f_1 为 ANN, f_2 为 SVM。

第二步:依据平均绝对误差(MAE)最小的原则分别最优化模型 $y_1 = f_1(x)$ 和 $y_2 = f_2(x)$ 中的参数。

第三步:将测试集输入向量 $x_i (i=43, \dots, 47)$ 输入以上训练好的模型 $y_1 = f_1(x)$ 和 $y_2 = f_2(x)$, 得到 QSPR 的预测结果 $f(x_i) (i=43, \dots, 47)$ 。

ANN 及 SVM 模型的具体构建过程为:

(1) MLP-ANN

ANN 是一种通过模拟生物神经信号处理系统发展起来的机器学习方法,广泛用于模式识别、函数拟合、过程控制和图形映射等领域。多层感知器(MLP)是一种前向结构的人工神经网络,映射一组输入向量到另一组输出向量。MLP 可以被看做是一个有向图,由多个节点层组成,每一层全连接到下一层。除了输入节点,每个节点都是一个带有非线性激活函数的神经元(或称处理单元)。MLP 已被证明是一种通用的函数近似方法,可以被用来拟合复杂的函数,或解决分类问题^[18-21]。该建模过程的技术框架如图 2 所示,主要用 Weka 完成。

(2) SVM

SVM 是一种以 VC 维(Vapnik-Chervonenkis dimension)理论、SRM(structural risk minimization)准则为基础建立的机器学习方法。主要用途是按监督学习对样本进行分类,适用于非线性、小样本和高维度等问题,可以有效避免“过度协调”。同时,与传统的机器学习方法相比,该方法具有较强的鲁棒性,特别适合分类、时间序列预测等^[22-23]。

在支持向量回归(support vector regression, SVR)中,学习机(learning machine)将处于低维空间的数据投射到高维特征空间,然后,在高维特征空间里,超平面(包含松弛变量)可以在训练数据和输出数据之间形成非线性关系,称为 SVR 函数,也可以表示为凸优化问题:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, y_j) + \epsilon \sum_{i=1}^l (\alpha_i^* + \alpha_i) - \sum_{i=1}^l y_i (\alpha_i^* - \alpha_i) \tag{1}$$

$$\text{s.t.} \sum_{i=1}^l (\alpha_i^* - \alpha_i) = 0, \quad 0 \leq \alpha_i, \alpha_i^* \leq \frac{C}{l}, \quad i=1, 2, \dots, l$$

式中: $K(x_i, y_j)$ 是核函数, $\alpha_i, \alpha_j, \alpha_i^*$ 和 α_j^* 是拉格朗日乘数,表示约束方程梯度线性组合里每个向量的系数, l 表示样本数量, C 表示惩罚系数。得到的回归函数 $f(x)$ 如下:

$$f(x) = \sum_{i=1}^l (\alpha_i^* - \alpha_i)K(x, x_i) + b \tag{2}$$

式中: b 表示超平面截距。

本文分别用线性 Linear、多项式 poly、径向基 RBF、sigmoid 核函数进行 SVM 建模,发现 poly 核函数建模结果较优。

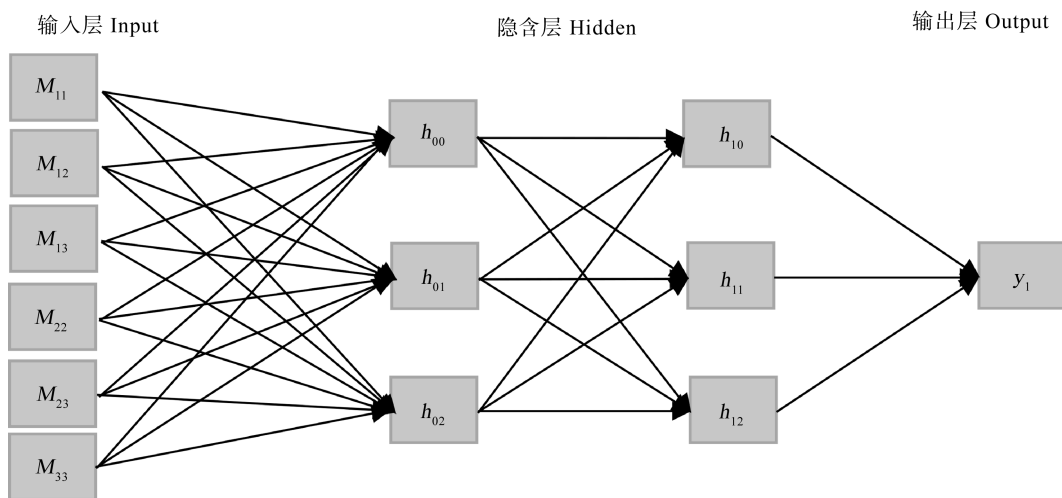


图2 Weka 建模
Fig. 2 Weka modeling

3 结果与讨论 (Results and discussion)

为寻找影响 PCDFs 对光解半衰期与各参数之间的 QSPR 关系,将全部样本分为 2 个组,云杉针叶表面 37 个样本作训练集(training set),飞灰表面 42 个样本作训练集,两者均选取 5 个相同的样本作为测试集(test set),且测试集选取方法相同,即在全部样本的每种取代物中随机选取一个样本作为测试集,选取方法同文献[1]。训练集样本用于建立 QSPR 模型,测试集样本用于检验 QSPR 模型的预测能力。它们的建模结果如下。

3.1 MLP-ANN 模型拟合结果

MLP-ANN 模型对训练集样本拟合结果及对测试集样本预测值的相关统计参数如表 1 和图 3

所示。

由图 3 和表 1 可知,用 MLP-ANN 方法分别对 PCDFs 训练集和测试集进行拟合,得到光解半衰期的 MAE 分别为 0.0290、0.0994(云杉针叶)和 0.0387、0.0464(飞灰表面),其均方根误差(RMSE)分别为 0.0381、0.1199(云杉针叶)和 0.0464、0.0518(飞灰表面),由此说明,MLP-ANN 模型优于文献[15-16]的线性模型,这可能是由于 MLP-ANN 模型所选描述子基本与 PCDFs 光解半衰期呈非线性关系,MLP-ANN 法使其拟合精度进一步提高。

3.2 SVM 模型拟合结果

SVM 模型对训练集样本拟合结果及对测试集样本预测值的相关统计参数如表 1 和图 4 所示。

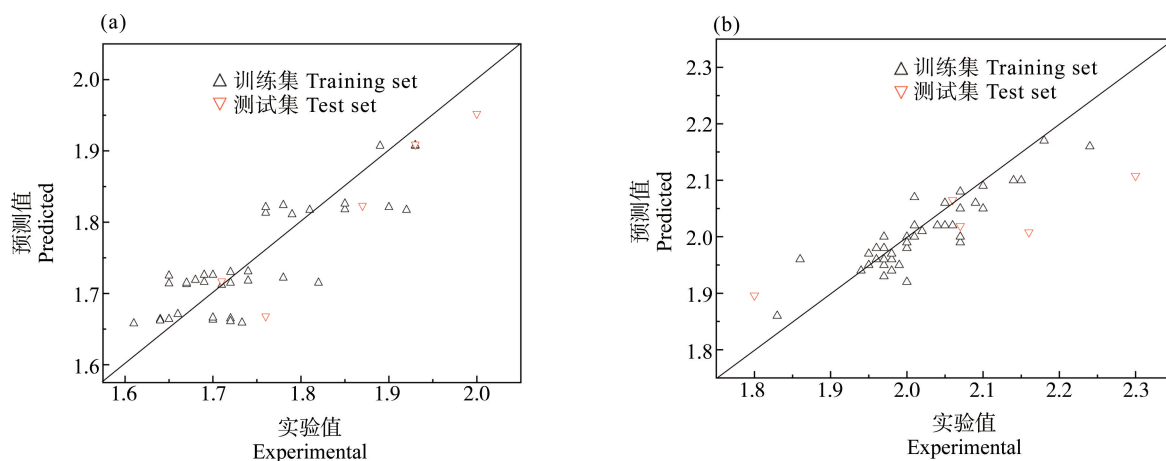


图 3 多层感知器神经网络 (MLP-ANN) 模型中 PCDFs 在云杉针叶 (a) 和飞灰 (b) 表面光解半衰期预测值与实验值的相关性

Fig. 3 Correlation between predicted value and experimental value of PCDFs photolysis half-life on the surface of spruce needles (a) and fly ash (b) by multi-layer perceptron-artificial neural network (MLP-ANN)

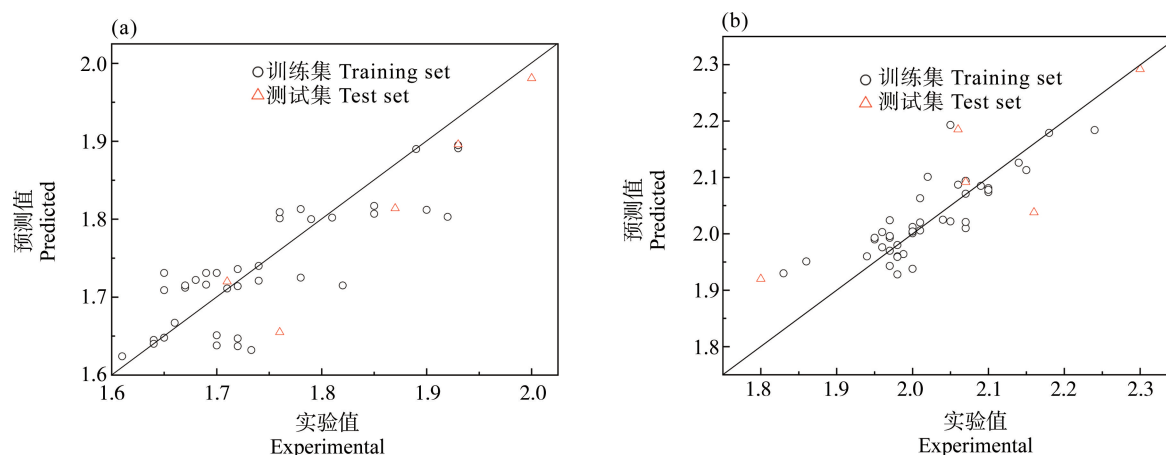


图 4 支持向量机 (SVM) 模型中 PCDFs 在云杉针叶 (a) 和飞灰 (b) 表面光解半衰期预测值与实验值的相关性

Fig. 4 Correlation between predicted value and experimental value of PCDFs photolysis half-life on the surface of spruce needles and fly ash by support vector machine (SVM)

表1 本工作与文献的比较
Table 1 Comparison between this work and some literature

文献 Literature	介质 Medium	方法 Method	训练集 Number of samples of training set (n_t)		测试集 Number of samples of test set (n_b)		描述子个数 (m/LV_s) Number of descriptors (m/LV_s)	训练集样本交互 检验相关系数(Q_{LoO})		测试集样本 外部相关系数 (Q_{Ext})		平均绝对误差 (MAE)		均方根误差 (RMSE)	
			Number of samples of training set (n_t)	Number of samples of test set (n_b)	训练集样本交互 检验相关系数(Q_{LoO})	Cumulative cross-validated R_{cum} of training set (Q_{LoO})		外部相关系数 (Q_{Ext})	外部相关系数 (Q_{Ext})	Mean absolute error (MAE)	Mean absolute error (MAE)	Root mean square error (RMSE)	Root mean square error (RMSE)		
[16]	云杉针叶 Spruce needle	MLR	42	0	0.860	0.839	-	-	-	-	-	-	-	-	-
		PLS	42	0	0.857	0.849	-	-	-	-	-	-	-	-	-
	飞灰 Fly ash	MLR	37	5	0.826	0.797	0.939	-	-	-	-	-	-	-	-
		PLS	37	5	0.802	0.789	0.892	-	-	-	-	-	-	-	-
	云杉针叶 Spruce needle	MLR	47	0	0.836	0.807	-	-	-	-	-	-	-	-	-
		PLS	47	0	0.829	0.807	-	-	-	-	-	-	-	-	-
	飞灰 Fly ash	MLR	42	5	0.824	0.778	0.853	-	-	-	-	-	-	-	-
		PLS	42	5	0.814	0.790	0.833	-	-	-	-	-	-	-	-
	本工作 This study	云杉针叶 Spruce needle	MLP	37	5	0.850	0.816	0.954	0.826	0.806	0.945	0.0290	0.0994	0.0381	0.1199
		飞灰 Fly ash	SVM	37	5	0.826	0.806	0.945	0.892	0.753	0.897	0.0387	0.0430	0.0464	0.0518
飞灰 Fly ash		MLP	42	5	0.826	0.755	0.831	0.826	0.755	0.831	0.0400	0.0531	0.0395	0.0507	
云杉针叶 Spruce needle		BP	42	6	-	-	0.872	-	-	-	-	-	-	-	
[15]	云杉针叶 Spruce needle	MLR	37	5	0.842	-	6	0.842	-	0.941	-	-	-	-	
		BP	42	-	0.9994	-	6	0.9994	-	-	-	-	-	-	
	飞灰 Fly ash	MLR	47	0	0.852	-	7	0.852	-	-	-	-	-	-	
		BP	42	5	0.821	-	7	0.821	-	0.893	-	-	-	-	
[12]	云杉针叶 Spruce needle	PLS	47	0	0.882	0.870	5	0.882	0.870	-	-	-	-	-	
		PLS	47	0	0.857	0.837	4	0.857	0.837	-	-	-	-	-	
[14]	云杉针叶 Spruce needle	GA	70	0	0.849	0.839	4	0.849	0.839	-	-	-	-	-	
		PLS	-	-	-	-	-	-	-	-	-	-	-	-	

注: m 为多元线性回归描述子个数, LVS_s 为偏最小二乘回归主成分数, MLR 为多元线性回归, PLS 为偏最小二乘回归分析, MLP 为多层感知器, SVM 为支持向量机, BP 为反向传播算法, GA 为遗传算法; - 表示无。
 Note: m represents number of descriptors for multiple linear regression (MLR); LVS_s represents number of latent variables for the partial least squares regression (PLS) model; MLP stands for multiple linear regression; PLS stands for partial least squares regression; MLR stands for multiple linear regression; SVM stands for support vector machine; BP stands for back propagation; GA stands for genetic algorithm; - represents "no data."

由图4和表1可知,用SVM方法分别对PCDFs训练集和测试集进行拟合,得到光解半衰期的MAE分别为0.0329、0.0794(云杉针叶)和0.0400、0.0531(飞灰表面),RMSE分别为0.0444、0.0954(云杉针叶)和0.0395、0.0507(飞灰表面),SVM模型优于文献[15-16]的线性模型,原因可能是SVM模型所选描述子MEDV基本与PCDFs光解半衰期呈非线性关系,且SVM建模方法更适合其相关性。

3.3 模型评价

采用MAE、RMSE和决定系数(R^2),以留一法交互检验相关系数(R_{Loo})以及外部样本相关系数(Q_{ext})对各模型的预测精度进行了评价,其相应的计算公式如式(3)、(4)、(5)和(6)所示。

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

$$R^2 = \frac{\text{ESS} \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\text{TSS} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

$$Q_{\text{ext}} = \sqrt{1 - \frac{\sum_{i=1}^{\text{test}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{\text{test}} (y_i - \bar{y}_{\text{tr}})^2}} \quad (6)$$

式中: \bar{y} 是样本*i*实际值的平均值, y_i 为样本*i*的实际值, \hat{y}_i 为样本*i*的预测值, \bar{y}_{tr} 为训练集样本实验的平均值, n 是样本数;ESS为平方和,TSS为总平方和。

MAE和RMSE值越低,表示模型的预测误差越小,预测结果越优。 R^2 的值表示所解释的ESS在TSS中的比例,该值介于0和1之间。 R^2 的值越大,则拟合效果越好, R^2 同时也可作为回归方程拟合优度的度量指标。以交互检验的 R^2 为标准判断所建模型的稳定性和内部预测能力, Q_{ext} 评价模型对外部样本的预测能力, R_{Loo} 和 Q_{ext} 越接近,表明所建模型的预测结果好,亦即说明该模型并非是偶然相关得到的,进一步说明所建模型的泛化能力较好。

3.4 与文献的比较

本研究与文献比较,评价指标更为全面。与文献[16]相比可以看出,同样使用MEDV描述子,本文采用非线性方法建立的模型更加稳定,外部预测能

力也优于文献。与文献[15]相比,所用描述子数量少,建模结果和文献相当。文献[12]和[14]中,虽建模相关系数优于本工作,但未将模型进行外部检验。为便于理解,表1列出了本文与文献模型比较。

综上所述,对PCDFs类化合物采用非线性方法建立其光解半衰期的QSPR模型,所建模型稳健且预测能力强。与线性方法相比,本文采用的非线性方法拟合精度更好,适合处理复杂的非线性问题。但若要提供PCDFs类化合物对光解半衰期的影响因素及其机理,笔者认为线性模型是较好的选择。同时,MEDV不仅能反映分子的大小、形状等整体,蕴含化合物丰富的结构信息,能很好地区分化合物的结构,同时其计算简单,取值完全来自分子本身的结构,不需要加入任何经验性的性质参数或校正参数,较为客观。同时,大量文献也表明,MEDV能用于多种化合物结构参数的表征。运用QSPR方法为环境中POPs光降解行为的研究提供了一种便捷的评价方法。后期可以考虑将本文的非线性方法应用于其他类型化合物的环境中分配行为与其结构之间的关系分析。

通讯作者简介:李美萍(1977—),女,博士,副教授,主要研究方向为化学计量学。

参考文献(References):

- [1] 李美萍. QSAR/QSPR方法在环境、药物和材料化学中的应用[D]. 太原: 山西大学, 2014: 29-58
Li M P. Application of QSAR/QSPR in environment, medicine and material chemistry [D]. Taiyuan: Shanxi University, 2014: 29-58 (in Chinese)
- [2] 夏传海, 徐杰, 吴文忠, 等. 二噁(英)降解方法的研究[J]. 化学进展, 2004, 16(1): 123-130
Xia C H, Xu J, Wu W Z, et al. Advance in degradation of polychlorinated dibenzo-*p*-dioxins/furans [J]. Progress in Chemistry, 2004, 16(1): 123-130 (in Chinese)
- [3] Niu J F, Chen J W, Henkelmann B, et al. Photodegradation of PCDD/Fs adsorbed on spruce (*Picea abies* (L.) Karst.) needles under sunlight irradiation [J]. Chemosphere, 2003, 50(9): 1217-1225
- [4] Niu J F, Chen J W, Martens D, et al. The role of UV-B on the degradation of PCDD/Fs and PAHs sorbed on surfaces of spruce (*Picea abies* (L.) Karst.) needles [J]. Science of the Total Environment, 2004, 322(1-3): 231-241
- [5] 牛军峰, 余刚, 刘希涛. 水相中 POPs 光化学降解研究

- 进展[J]. 化学进展, 2005, 17(5): 938-948
- Niu J F, Yu G, Liu X T. Advances in photolysis of persistent organic pollutants in water [J]. Progress in Chemistry, 2005, 17(5): 938-948 (in Chinese)
- [6] 唐自强, 冯长君. 取代苯酚类化合物抑藻活性的 CoMFA 模型[J]. 生态毒理学报, 2019, 14(4): 192-196
- Tang Z Q, Feng C J. CoMFA model for inhibitory activity of chlorinated phenolic compounds to *Dunaliella salina* [J]. Asian Journal of Ecotoxicology, 2019, 14(4): 192-196 (in Chinese)
- [7] 张文华, 赵继颖, 王琼瑶, 等. 多氯代苯并噻吩分子空间坐标指数与气相色谱保留时间的 QSPR 研究[J]. 生态毒理学报, 2018, 13(3): 220-225
- Zhang W H, Zhao J Y, Wang Q Y, et al. QSPR study for molecular spatial coordinates index of poly-chlorinated dibenzothiophenes and retention time of gas chromatography [J]. Asian Journal of Ecotoxicology, 2018, 13(3): 220-225 (in Chinese)
- [8] 薛建芳, 史雅娟, 王尘辰, 等. 持久性有机污染物的作物吸收及迁移模型研究进展[J]. 生态毒理学报, 2018, 13(1): 75-88
- Xue J F, Shi Y J, Wang C C, et al. Assessment of POPs absorption in agriculture crop using migration models [J]. Asian Journal of Ecotoxicology, 2018, 13(1): 75-88 (in Chinese)
- [9] Fissa R M, Lahiouel Y, Khaouane L, et al. QSPR estimation models of normal boiling point and relative liquid density of pure hydrocarbons using MLR and MLP-ANN methods [J]. Journal of Molecular Graphics and Modeling, 2019, 87(3): 109-120
- [10] Chen J W, Xie Q, Schramm K W, et al. Quantitative structure-property relationships (QSPRs) on direct photolysis of PCDDs [J]. Chemosphere, 2001, 45(2): 151-159
- [11] Chen J W, Quan X, Pei N B, et al. Quantitative structure-property relationships (QSPRs) on direct photolysis quantum yields of PCDDs [J]. Chemosphere, 2001, 43(2): 235-241
- [12] 王蕾, 刘新会, 侯娟, 等. 电性拓扑态指数预测 PCDFs 光解半衰期[J]. 化学学报, 2007, 65(3): 184-190
- Wang L, Liu X H, Hou J, et al. Prediction of photolysis half-lives of PCDFs with the electrotopological state indices [J]. Acta Chimica Sinica, 2007, 65(3): 184-190 (in Chinese)
- [13] Niu J F, Chen J, Yu G, et al. Quantitative structure-property relationships on direct photolysis of PCDD/Fs on surfaces of fly ash [J]. SAR and QSAR in Environmental Research, 2004, 15(4): 265-277
- [14] 牛军峰, 余刚, 韩文亚. 应用遗传算法建立云杉针叶表面 PCDD/Fs 光解半衰期的预测模型[J]. 环境科学, 2005, 26(2): 28-33
- Niu J F, Yu G, Han W Y. Prediction of half-lives of PCDD/Fs adsorbed on spruce needles optimized by genetic algorithm [J]. Environmental Science, 2005, 26(2): 28-33 (in Chinese)
- [15] 赵继颖, 张文华, 王琼瑶, 等. 多氯代二苯并呋喃分子空间坐标与光解半衰期的 QSPR 研究[J]. 化学研究与应用, 2018, 30(10): 1649-1654
- Zhao J Y, Zhang W H, Wang Q Y, et al. QSPR study for the molecular spatial coordinates index and photolysis half-lives of polychlorinated dibenzofurans [J]. Chemical Research and Application, 2018, 30(10): 1649-1654 (in Chinese)
- [16] 李美萍, 张生万, 胡永刚, 等. 分子电性距离矢量用于多氯代二苯并呋喃光解半衰期的 QSPR 研究[J]. 生态毒理学报, 2010, 5(2): 208-214
- Li M P, Zhang S W, Hu Y G, et al. Molecular electronegativity distance vector (MEDV) applied to quantitative structure-property relationship study on photolysis half-lives of PCDFs [J]. Asian Journal of Ecotoxicology, 2010, 5(2): 208-214 (in Chinese)
- [17] 刘树深, 刘堰, 李志良, 等. 一个新的分子电性距离矢量(MEDV)[J]. 化学学报, 2000, 58(11): 1353-1357
- Liu S S, Liu Y, Li Z L, et al. A novel molecular electronegativity-distance vector (MEDV) [J]. Acta Chimica Sinica, 2000, 58(11): 1353-1357 (in Chinese)
- [18] 李双成, 王晓玥, 谢爱丽, 等. 基于多层感知器模型的 MODIS 地表温度降尺度研究[J]. 环境科学研究, 2017, 3(12): 1889-1897
- Li S C, Wang X Y, Xie A L, et al. Downscaling MODIS land surface temperatures using multilayer perceptron model [J]. Research of Environmental Sciences, 2017, 3(12): 1889-1897 (in Chinese)
- [19] 张鹏程, 贾旻旻. 一种基于多层感知器的动态区域联合短时降水预报方法[J]. 计算机应用与软件, 2018, 35(11): 153-158
- Zhang P C, Jia Y Y. A dynamic regional combined short-term precipitation forecasting approach based on multilayer perceptron [J]. Computer Applications and Software, 2018, 35(11): 153-158 (in Chinese)
- [20] 任清, 孙雪剑, 戴晓爱, 等. 全谱段光谱分析的块状商品煤种类鉴别[J]. 光谱学与光谱分析, 2018, 38(2): 352-357

- Ren Y, Sun X J, Dai X A, et al. Variety identification of bulk commercial coal based on full-spectrum spectroscopy analytical technique [J]. *Spectroscopy and Spectral Analysis*, 2018, 38(2): 352-357 (in Chinese)
- [21] 丛嘉昕, 宋江峰, 李大婧, 等. 基于 RSM 和 MLP-ANN 的草莓果浆超声酶解参数优化[J]. *江苏农业学报*, 2018, 34(6): 1354-1361
- Cong J X, Song J F, Li D J, et al. Optimization of ultrasonic enzymolysis of strawberry pulp based on RSM and MLP-ANN [J]. *Jiangsu Journal of Agricultural Sciences*, 2018, 34(6): 1354-1361 (in Chinese)
- [22] 周梦妮, 崔会芳, 曹锐, 等. 基于排列熵和支持向量机的癫痫发作预测研究[J]. *计算机应用研究*, 2019, 36(6): 1696-1699
- Zhou M N, Cui H F, Cao R, et al. Epileptic seizure prediction research based on permutation entropy and support vector machine [J]. *Application Research of Computers*, 2019, 36(6): 1696-1699 (in Chinese)
- [23] 李荟, 赵云敏. GMM-UBM 和 SVM 在说话人识别中的应用[J]. *计算机系统应用*, 2018, 27(1): 225-230
- Li H, Zhao Y M. Application of GMM-UBM and SVM in speaker recognition [J]. *Computer Systems Applications*, 2018, 27(1): 225-230 (in Chinese) ◆