

应用线性溶剂化能参数预测多氯有机物的分配性质

戴家银*, 陈玲, 赵建夫 (同济大学污染控制与资源化研究国家重点实验室, 上海 200092)

摘要: 应用线性溶剂化能方法构建定量结构-性质相关(QSPR)模型. 该模型能较好地预测多氯有机物的正辛醇/水分配系数(K_{ow})和沉积物吸附系数(K_{oc}). K_{ow} 和 K_{oc} 平均误差分别为0.43和0.37个对数单位. 模型的稳健性检验结果表明, 模型中自变量不存在相关关系, 残差呈正态分布; 对 K_{oc} 来说, PCB和PCDD/PCDF在构建模型中起核心作用, 剔除该类化合物导致模型相关系数明显下降, 农药类和氯硝基苯类化合物的存在降低了模型的精度.

关键词: 正辛醇/水分配系数; 沉积物吸附系数; 定量结构-性质相关; 稳健性

中图分类号: X131 文献标识码: A 文章编号: 1000-6923(2002)04-0329-05

Prediction of partition property of polychlorinated organic compounds using linear solvation energy parameter.

DAI Jia-yin, CHEN Ling, ZHAO Jian-fu (State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai 200092, China). *China Environmental Science*. 2002,22(4): 329~333

Abstract: Linear solvation energy technique was used to develop quantitative structure-property relationships (QSPR) model. The model could predict more effectively the octanol/water partition coefficient (K_{ow}) and the sediment sorption coefficient (K_{oc}) of the polychlorinated organic compounds (PCOCs). The mean errors of K_{ow} and K_{oc} of the compounds were 0.43 and 0.37 log units respectively. Results of robustness test of the model demonstrated that the independent variables in the model did not interrelate and the residuals presented normal distribution. For K_{oc} , PCB and PCDD/PCDF are important in developing the model, the interrelation coefficient of the model decreased markedly by deleting these compounds; the accuracy of the model decreased due to existence of pesticides and chlorinated nitrobenzene compounds.

Key words: octanol/water partition coefficient; sediment sorption coefficient; quantitative structure-property relationships; robustness

目前我国每年排放干污泥约为500万t,部分地区已在尝试污泥的土地利用,让污泥回归于自然^[1].美国环境保护局的统计表明,污泥的土地利用已经成为污泥处置的主要发展趋势.现有的研究工作主要集中在重金属迁移富集等方面^[2],但国内外对污泥土地利用的“负效应”研究较少.污泥中含有一些难降解有机物,如多环芳烃类(PAH)以及多氯有机物(PCB)等^[3].尽管这些物质含量较低,但它们的脂溶性强(具有大的 K_{ow} 值),且有一定的毒性,容易被土壤中有有机碳吸附,从而通过食物链的放大作用,直接或间接对人类健康产生不良影响.

对污泥中多氯有机物进行环境风险评价,首先要了解多氯有机物的环境参数,正辛醇/水分配系数(K_{ow})和沉积物吸附系数(K_{oc})是环境风险评价的基本数据之一.测定该类化合物 K_{ow} 和 K_{oc}

既困难又费时.线性溶剂化能相关(LSER)被公认是一种十分有效的定量结构-性质相关(QSPR)模型.通过LSER方程式可精确预测化学品的 K_{ow} 、 K_{oc} 和对各种生物的毒性^[4].作者以线性溶剂化能参数建立QSPR模型,预测多氯有机物的分配性质,对这类化合物的暴露评价具有重要意义.

1 研究方法

K_{ow} 采用摇瓶法、慢慢搅动法以及产生柱法等实验测定的数据; K_{oc} 收集最近报道的实验数据,如同一化合物有几个 K_{oc} 值时,则取中间值,数据取自文献[5-10].

收稿日期: 2001-10-24

基金项目: 国家自然科学基金资助项目(20007002)

* 通讯联系人

模型的优劣由两方面来评价:回归分析中所有自变量显著水平取 $\alpha=0.05$;用变异膨胀因子(VIF)来评价各自变量的相关程度,VIF定义为: $VIF=1/(1-r^2)$, r 为方程中某一自变量与其他变量间多元回归的相关系数(经自由度校正). $VIF=1.0$,表示各变量间无自相关; $VIF=1.0\sim 5.0$,表示相

关方程可以接受; $VIF>10$,表示回归方程不稳定,需重新检验.

70 种多氯有机物的线性溶剂化能参数(V_i , π^* , β_m , α_m)按文献[4]方法估算(表 1).并与分子连接性指数法(MCIs)作比较.

表 1 70 种多氯有机物的分子结构描述符及分配系数
Table 1 Molecular structure descriptors and partition properties for 70 PCOCs

化合物	LSER				MCIs χ^1	K_{oc}			K_{ow}		
	$V_i/100$	π^*	β_m	α_m		实验值	计算值	残差	实验值	计算值	残差
氯丹	1.648	1.07	0.55	0.52	8.114	5.15	5.04	0.11	5.80	5.95	-0.15
艾氏剂	1.674	1.04	0.45	0.43	8.276	4.69	5.27	-0.58	6.50	6.29	0.21
α -硫丹	1.656	1.64	1.37	0.38	7.686	4.13	3.76	0.37	3.83	3.95	-0.12
七氯	1.558	1.02	0.50	0.51	7.703	-	4.86	-	6.10	5.76	0.34
α -六六六	1.031	0.70	0.00	0.00	5.464	3.25	4.12	-0.87	3.80	5.18	-1.38
β -六六六	1.031	0.70	0.00	0.00	5.464	3.36	4.12	-0.76	3.78	5.18	-1.40
γ -六六六	1.146	0.68	0.24	0.29	5.464	3.00	4.07	-1.07	3.76	4.99	-1.23
p,p' -滴滴涕	1.632	1.90	0.24	0.00	8.874	5.31	5.49	-0.18	6.91	6.66	0.25
p,p' -敌敌畏	1.516	1.90	0.39	0.05	8.575	4.82	4.91	-0.09	6.96	5.89	1.07
六氯苯	1.031	0.70	0.00	0.00	5.464	4.45	4.12	0.33	6.18	5.18	1.00
西玛津	1.219	0.89	1.48	0.00	6.258	2.10	2.31	-0.21	2.18	2.18	0.00
阿拉特津	1.307	0.89	1.48	0.34	6.613	2.24	2.57	-0.33	2.63	2.48	0.15
氰乙酰胺	1.406	1.09	1.85	0.34	7.465	2.28	2.27	0.01	2.22	1.91	0.31
麦草畏	0.965	0.89	0.43	0.54	6.075	2.20	3.24	-1.04	2.21	3.89	-1.68
甲草胺	1.470	1.11	1.26	0.66	8.855	3.09	3.39	-0.30	3.59	3.58	0.01
二苯基甲醇	1.641	1.93	0.54	0.33	9.248	-	5.03	-	5.02	5.95	-0.93
2,2'-二氯联苯	1.1	0.60	0.21	0.00	6.788	3.92	3.99	-0.07	4.97	4.90	0.07
2,4'-二氯联苯	1.1	0.60	0.21	0.00	6.771	4.14	3.99	0.15	5.10	4.90	0.20
2,4,4'-三氯联苯	1.19	0.70	0.21	0.00	7.165	4.62	4.25	0.37	5.62	5.21	0.41
2,2',5-三氯联苯	1.19	0.55	0.18	0.00	7.182	4.23	4.29	-0.06	5.55	5.29	0.26
2,2',4-三氯联苯	1.19	0.70	0.21	0.00	7.182	4.84	4.25	0.59	5.60	5.21	0.39
2,2',6,6'-四氯联苯	1.28	0.70	0.18	0.00	7.609	5.01	4.56	0.45	5.94	5.60	0.34
2,5,3',4'-四氯联苯	1.28	0.65	0.14	0.00	7.575	4.84	4.62	0.22	6.23	5.69	0.54
2,2',5,5'-四氯联苯	1.28	0.50	0.14	0.00	7.575	4.73	4.62	0.11	6.09	5.69	0.40
3,3',4,4'-四氯联苯	1.28	0.80	0.14	0.00	7.575	4.41	4.62	-0.21	6.21	5.69	0.52
2,2',3,4,5'-五氯联苯	1.37	0.70	0.14	0.00	8.003	4.62	4.88	-0.26	6.37	6.00	0.37
2,2',4,5,5'-五氯联苯	1.37	0.55	0.11	0.00	7.986	4.63	4.93	-0.30	6.50	6.08	0.42
2,2',4,4',5-五氯联苯	1.37	0.65	0.11	0.00	7.486	5.68	4.93	0.75	6.41	6.08	0.33
2,2',3,3',4,4'-六氯联苯	1.46	0.90	0.11	0.00	8.430	5.05	5.19	-0.14	6.93	6.39	0.54
2,2',4,4',5,5'-六氯联苯	1.46	0.60	0.08	0.00	8.397	5.65	5.24	0.41	7.16	6.46	0.70
3,3',4,4',5,5'-六氯联苯	1.46	0.90	0.08	0.00	7.700	6.60	5.24	1.36	7.41	6.46	0.95
2,2',3,4,5,5',6-七氯联苯	1.55	0.60	0.11	0.00	8.841	5.95	5.45	0.50	7.00	6.70	0.30
1,2,3,4-四氯二苯并二噁噻	1.976	0.70	0.48	0.00	8.017	-	6.10	-	7.18	7.25	-0.07
1,2,3,7-四氯二苯并二噁噻	1.976	0.80	0.48	0.00	8.559	5.98	6.10	-0.12	6.91	7.25	-0.34
1,3,6,8-四氯二苯并二噁噻	1.976	0.70	0.48	0.00	8.541	5.98	6.10	-0.12	7.20	7.25	-0.05
2,3,7,8-四氯二苯并二噁噻	1.976	1.05	0.48	0.00	8.541	6.30	6.10	0.20	7.02	7.25	-0.23
2,8-二氯二苯并呋喃	1.76	0.80	0.24	0.00	7.237	-	5.86	-	5.44	7.10	-1.66

续表1

化合物	LSER				MCIs	K_{oc}			K_{ow}		
	$V_i/100$	π^*	β_m	α_m		实验值	计算值	残差	实验值	计算值	残差
2,3,7,8-四氯二苯并呋喃	1.941	1.14	0.18	0.00	8.059	5.20	6.48	-1.28	6.53	7.87	-1.34
1,2-二氯苯	0.671	0.79	0.07	0.00	3.805	2.50	2.96	-0.46	3.44	3.77	-0.33
1,3-二氯苯	0.671	0.84	0.07	0.00	3.788	2.48	2.96	-0.48	3.53	3.77	-0.24
1,4-二氯苯	0.671	0.79	0.07	0.00	3.788	2.44	2.96	-0.52	3.44	3.77	-0.33
1,2,3-三氯苯	0.761	0.84	0.04	0.00	4.215	3.29	3.27	0.02	4.14	4.15	-0.01
1,2,4-三氯苯	0.761	0.79	0.04	0.00	4.198	3.15	3.27	-0.12	4.05	4.15	-0.10
1,3,5-三氯苯	0.761	0.79	0.04	0.00	4.182	2.85	3.27	-0.42	4.19	4.15	0.04
1,2,3,4-四氯苯	0.851	0.79	0.04	0.00	4.626	3.48	3.53	-0.05	4.60	4.46	0.14
1,2,3,5-四氯苯	0.851	0.79	0.04	0.00	4.609	3.20	3.53	-0.33	4.66	4.46	0.20
1,2,4,5-四氯苯	0.851	0.64	0.04	0.00	4.609	3.25	3.53	-0.28	4.70	4.46	0.24
五氯苯	0.951	0.74	0.04	0.00	5.037	4.11	3.82	0.29	5.18	4.81	0.37
2,4-二氯苯胺	0.751	0.87	0.45	0.00	4.198	2.72	2.59	0.13	2.91	3.11	-0.20
2,6-二氯苯胺	0.751	0.87	0.45	0.26	4.215	3.25	2.59	0.66	2.82	3.11	-0.29
3,5-二氯苯胺	0.751	0.77	0.45	0.26	4.182	2.49	2.59	-0.10	2.90	3.11	-0.21
3,4-二氯苯胺	0.751	0.92	0.45	0.26	4.198	2.29	2.59	-0.30	2.69	3.11	-0.42
2,3,4-三氯苯胺	0.841	0.97	0.45	0.26	4.626	2.60	2.85	-0.25	3.68	3.42	0.26
2,3,4,5-四氯苯胺	0.931	0.92	0.45	0.26	5.037	3.03	3.11	-0.08	4.27	3.73	0.54
2,3,5,6-四氯苯胺	0.931	0.82	0.45	0.26	5.037	3.94	3.11	0.83	4.10	3.73	0.37
2,3,4,5,6-五氯苯胺	1.021	0.87	0.45	0.26	5.464	4.62	3.37	1.25	4.82	4.04	0.78
2,3,4,5-四氯硝基苯	0.991	0.89	0.29	0.26	5.947	4.23	3.54	0.69	3.93	4.33	-0.40
2,3,5,6-四氯硝基苯	0.991	0.79	0.29	0.16	5.947	4.05	3.54	0.51	3.89	4.33	-0.44
2,3,4,5,6-五氯硝基苯	1.081	0.84	0.29	0.16	6.375	4.36	3.80	0.56	4.64	4.64	0.00
2,3-二氯酚	0.716	0.52	0.26	0.16	4.215	2.65	2.79	-0.14	2.84	3.45	-0.61
2,4-二氯酚	0.716	0.67	0.26	0.60	4.198	2.75	2.79	-0.04	3.06	3.45	-0.39
2,4,6-三氯酚	0.806	0.67	0.23	0.60	4.609	3.02	3.10	-0.08	3.69	3.84	-0.15
2,4,5-三氯酚	0.806	0.82	0.23	0.60	4.609	3.36	3.10	0.26	3.72	3.84	-0.12
3,4,5-三氯酚	0.806	0.97	0.23	0.60	4.609	3.56	3.10	0.46	4.01	3.84	0.17
2,3,5-三氯酚	0.806	0.87	0.23	0.60	4.069	3.36	3.10	0.26	4.21	3.84	0.37
2,3,6-三氯酚	0.806	0.97	0.23	0.60	4.626	-	3.10	-	3.77	3.84	-0.07
五氯酚	0.986	0.87	0.23	0.60	5.464	3.73	3.62	0.11	5.12	4.46	0.66
4,5,6-三氯愈创木酚	0.949	1.07	0.45	0.66	5.575	3.11	3.16	-0.05	3.83	3.79	0.04
四氯愈创木酚	1.039	1.02	0.45	0.66	6.002	3.15	3.43	-0.28	4.42	4.10	0.32

注:-为数据未得到, $V_i/100$ 为分子的本征体积, π^* 为分子偶极/极化率电性, β_m 为氢键受体碱性, α_m 为氢键受体酸性, χ^1 为分子连接性指数

2 结果与讨论

2.1 预测模型的建立

- 采用逐步回归法获得如下方程:
- $$\log K_{ow}=1.63+3.45V_i/100-2.47\beta_m \quad n=70$$
$$r^2=0.838 \quad SE=0.59 \quad F=179.44 \quad (1)$$
- $$\log K_{oc}=1.12+2.92V_i/100-1.60\beta_m \quad n=65$$
$$r^2=0.812 \quad SE=0.499 \quad F=139.27 \quad (2)$$
- $$\log K_{oc}=0.46+0.72\log K_{ow} \quad n=65$$
$$r^2=0.859 \quad SE=0.435 \quad F=385.70 \quad (3)$$

- $$\log K_{oc}=0.641+0.526\chi^1 \quad n=65$$
$$r^2=0.572 \quad SE=0.759 \quad F=84.07 \quad (4)$$
- $$\log K_{ow}=0.855+0.632\chi^1 \quad n=70$$
$$r^2=0.518 \quad SE=1.026 \quad F=73.17 \quad (5)$$
- 式中: n 为化合物的数目; r^2 为经自由度校正的相关系数; SE 为标准误差; F 为 F -检验值.
- 考虑到70种化合物的母体结构差别较大,物
- 理化学性质差异也较大,因此可以认为方程(1),
- 方程(2)具有高 r^2 ,较小的 SE ;方程(4),方程(5)与方
- 程(1),方程(2)相比, SE 较大,这表明应用线性溶剂

化能方法建立的模型能更好地预测该类化合物的 K_{ow} 和 K_{oc} 值.用方程(1),方程(2)计算的预测值和残差见表1.观察值和实验值吻合得较好. K_{ow} 和 K_{oc} 平均估算误差分别为0.43个对数单位和0.37个对数单位.

2.2 模型检验

表 2 显示方程(1),方程(2)中自变量间不存在相关关系, VIF 值为 1.07,表示方程(1),方程(2)具有统计意义,可以应用于分配机理解释.对 70 种化合物 K_{ow} 和 K_{oc} 的残差值进行检验,结果表明残差呈正态分布.

表 2 方程(1),方程(2)中自变量相关系数矩阵及 VIF 值

Table 2 Correlation coefficient matrix for variables and VIF for equations (1) and (2)				
项目	$V_i/100$	β_m	$R^2(\text{adj.})$	(VIF)
$V_i/100$	1.00		0.0667	1.07
β_m	-0.28	1.00		

注: $R^2(\text{adj.})$ 为 Jackknife 系数

2.3 稳健性检验

应用修正的“Jackknife”检验法,检验方程是否有异常值以及方程(1),方程(2)的稳健性.随机删除 n 个观察值进行 m 次回归的结果表明,方程(1),方程(2)是稳健的,不是随机相关的(表3).

表 3 随机删除 n 个观察值进行 m 次回归的结果

Table 3 Summary of results of random deletion test					
删除样本数 ($<10\%$ of N)	随机回 归次数	K_{ow}		K_{oc}	
		校正的 r^2 平均值	平均 SE	校正的 r^2 平均值	平均 SE
3	14	0.840	0.58	0.817	0.493
5	10	0.839	0.59	0.812	0.496
		0.840*	0.59*	0.814*	0.495*

注: N 为总样本数, * 为总平均值

将化合物分成 7 类,每次剔除 1 类,对余下 6 类进行多元回归,jackknife 系数的变化见图 1.从图 1 可以看出,分子结构相对复杂的农药和 PCDD/PCDF 等 2 类化合物对模型稳健性有较大的影响.对 K_{oc} 来说,PCB 和 PCDD/PCDF 在模型中起核

心作用,剔除该类化合物的结果导致模型相关系数明显下降.农药类、氯硝基苯类及氯苯胺类化合物的存在降低了模型的精度.

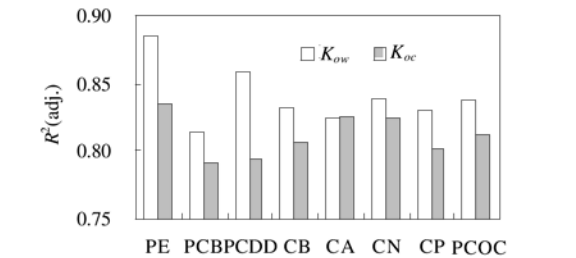


图 1 删除不同类别化合物的jackknife系数分布
(方程1和方程2)

Fig.1 Distribution of jackknife R^2 by eliminating various series of compounds [eqs. (1) and (2)]

PE 农药 PCB 多氯联苯 PCDD 二噁英 CB 氯苯
CA 氯代苯胺 CN 氯代硝基苯 CP 氯酚
PCOC 多氯有机物

方程(1),方程(2)中变量 $V_i/100$ 和 β_m 的 t -检验值分别为 17.40 和-12.10;15.94 和-9.21.表明对化合物的 K_{ow} 和 K_{oc} 来说,分子体积 V_i 与氢键受体碱性 β_m 参数对其影响显著.其中,分子体积是最重要的影响因素, V_i 提供的是有关溶质分子大小的信息.

综合方程(1),方程(2)可以看出,分子体积越大,需要在溶剂中形成的窝穴更大,为克服分子间的内聚作用所需的能量也更大.由于分子在水溶液中的排斥作用,使分子更容易进入非极性的有机相如正辛醇或有机碳,因此化合物的 $\log K_{ow}$ 和 $\log K_{oc}$ 增大.化合物氢键受体碱性参数对 $\log K_{ow}$ 和 $\log K_{oc}$ 值的影响说明了正辛醇和沉积物有机相与水酸性的不同,水的氢键给体酸性比正辛醇和有机相要强,化合物氢键受体碱性更有利于化合物在水中的溶解,因此 β_m 值越大, $\log K_{ow}$ 和 $\log K_{oc}$ 越小.

3 结语

基于线性溶剂化能参数建立的模型能较好地预测多氯有机物的 K_{ow} 和 K_{oc} , K_{ow} 和 K_{oc} 平均误

差分别为0.43和0.37个对数单位.模型中自变量不存在相关关系,且线性溶剂化能参数具有一定的理化意义,有利于从机理上解释分配性质.

参考文献:

- [1] 蒋成爱,黄国锋,吴启堂.城市污水污泥处理利用研究进展 [J]. 农业环境与发展, 1999,16(1):13-17.
- [2] Wu Q T, Nyriandge P, Mo C H. Removal of heavy metals from sewage by lower costing chemical method and recycling in agriculture [J]. Journal of Environmental Sciences, 1998, 10(1): 122-128.
- [3] Disse G, Weber H, Hamann R, *et al.* Comparison of PCDD and PCDF concentrations after aerobic and anaerobic digestion of sewage sludge [J]. Chemosphere, 1995, 31(7): 3617-3625.
- [4] Kamlet M J, Abboud J L M. Solubility properties in polymers and biological media. 8. An analysis of the factors that influence toxicities of organic nonelectrolytes to the Golden orfe fish (*leuciscosidus melanotus*) [J]. Environ. Sci. Technol., 1987, 21(2): 149-155.
- [5] Mackay D, Shiu W Y, Ma K C. Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals [M]. London: Boca Raton Inc., 1992. 1, 2, 4.
- [6] Meylan W M, Howard P H, Boethling R S. Improved method for estimation water solubility from octanol/water partition coefficient [J]. Environ. Toxicol. Chem., 1996, 15(1): 100-106.
- [7] Meylan W, Howard P H, Boethling R S. Molecular topology/fragment contribution method for predicting soil sorption coefficients [J]. Environ. Sci. Technol., 1992, 26(8): 1560-1567.
- [8] Nevalainen T, Kolehmainen E. New QSAR models for polyhalogenated aromatics [J]. Environ. Toxicol. Chem., 1994, 13(10): 1699-1706.
- [9] Reddy K N, Locke M A. Relationships between molecular properties and $\log P$ and soil sorption (K_{oc}) of substituted phenyl-lureas: QSAR models [J]. Chemosphere, 1994, 28(11):1929-1941.
- [10] Sabljic A, Gusten H, Verhaar H, *et al.* QSAR modeling of soil sorption improvements and systematics of $\log K_{oc}$ vs. $\log K_{ow}$ correlations [J]. Chemosphere, 1995, 31(11-12): 4489-4514.

作者简介: 戴家银(1965-),男,安徽怀宁人,博士.现为同济大学环境科学与工程学院污染控制与资源化研究国家重点实验室在站博士后,主要从事有机污染物的构效关系以及生态毒理研究.发表论文 17 篇.