

有机污染物分配系数(正辛醇/水)预测软件比较研究

商照聪^{1,2}, 贺少鹏¹

1. 上海化工研究院, 上海 200062
2. 中国科学院沈阳应用生态研究所, 沈阳 110016

摘要 介绍了虚拟化学实验室网站提供的 10 种常用分配系数(正辛醇/水)预测软件(ALOGPs、AC logP、AB/LogP、COSMOFraq、miLogP、ALOGP、MLOGP、KOWWIN、XLOGP 2、XLOGP 3)。以美国环境保护局《国家推荐水质标准》中的 105 种有机污染物作为测试集, 使用成对 t 检验和线性回归的方法比较了 10 种软件的预测值与实验值的差异性。成对 t 检验结果表明, ALOGPs、AC logP、AB/LogP 和 KOWWIN 软件的预测值和实验值没有显著的差别, 对应 p -水平分别为 0.26517、0.08149、0.43717 和 0.89145。线性回归的结果表明, ALOGPs 和 KOWWIN 软件的预测值和实验值相关性最好, 相关系数分别为 0.9919 和 0.9884, 拟合曲线的斜率分别为 0.9765 和 0.9982, 截距分别为 0.0564 和 0.0102。比较而言, ALOGPs 软件界面友好, 操作方便, 输入方式多样, 通过 Internet 免费运行并且同时提供多种软件预测值作为参考等优点, 因此其具有更大的优势。

关键词 分配系数(正辛醇/水); 预测软件; 有机污染物; 分子疏水性

中图分类号 X131

文献标识码 A

doi 10.3981/j.issn.1000-7857.2012.19.010

Comparison of Softwares Used for Predicting Partition Coefficient (n -Octanol/Water) of Organic Contaminants

SHANG Zhaocong^{1,2}, HE Shaopeng¹

1. Shanghai Research Institute of Chemical Industry, Shanghai 200062, China
2. Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China

Abstract Ten kinds of general partition coefficient (n -octanol/water) prediction software (ALOGPs, AC logP, AB/LogP, COSMOFraq, miLogP, ALOGP, MLOGP, KOWWIN, XLOGP 2, and XLOGP 3) provided by Virtual Computational Chemistry Laboratory is introduced. Using 105 kinds of organic pollutants listing in the National Recommended Water Quality Criteria of USEPA as a test set, paired t -test and liner regression method are performed on the predictive values and experimental values. The paired t -test results show that there is no significant difference between predictive values and experimental values for ALOGPs, AC logP, AB/LogP, and KOWWIN. The corresponding p -values are 0.26517, 0.08149, 0.43717, and 0.89145, respectively. The results of liner regression method indicate that the ALOGPs and KOWWIN are more accurate predictors than other software. For ALOGPs, the correlation coefficient between predictive values and experimental values is 0.9919. The slope and the intercept of the fitting curve are 0.9765 and 0.0564, respectively. For KOWWIN, the correlation coefficient is 0.9884. The slope and the intercept of the fitting curve are 0.9982 and 0.0102, respectively. Since the ALOGPs has a more user friendly interface, more convenient operation, more methods for entering input data, is able to free online perform, and provides the average $\lg P$ value of 10 kinds of software, therefore the ALOGPs has advantage over KOWWIN.

Keywords partition coefficient (n -octanol/water); prediction software; organic contaminants; molecular hydrophobicity

0 引言

化合物的分配系数(正辛醇/水)是指其在正辛醇相和水

相中的平衡浓度之比。用于表示化合物的亲脂性或疏水性, 是有机化合物环境归趋研究中一项非常重要的性质, 通常用

收稿日期: 2012-06-12; 修回日期: 2012-07-02

基金项目: “十二五”国家科技支撑计划项目(2012BAK03B01)

作者简介: 商照聪, 教授级高级工程师, 研究方向为化学品危险性分类鉴定、储运安全保障技术及化肥与分子筛质量评价, 电子信箱: szcsmads@gmail.com

lgP 表示^[1]。

化合物的 lgP 可以通过实验或者软件预测方法得到。现有的摇瓶法^[2-3]、高效液相色谱法^[4-5]、产生柱法^[6]、缓慢搅拌法^[7-9]和针对离子化合物的 pH 电极法^[9]等试验方法获得的数据虽然可靠,但这些方法需要得到纯度较高的物质,并且耗时长、花费较大,难以满足对化学品危险性识别和风险评估的需求;预测方法得到的数据虽然精度略差,但具有速度快、不需要获得物质、成本低等众多优点,受到很多研究者的欢迎^[10-13],并且也逐渐得到欧盟和世界经济合作与发展组织(OECD)等机构的认可。尤其是近年来借助于计算机科学和互联网技术的发展,诞生了大量可用于预测 lgP 的软件,更显示出预测方法的发展优势。

然而,这些大量的软件都是根据不同的算法建立起来的,使用不同的软件预测同一化合物的 lgP 往往会出现较大的差异^[14-15],使人们无法确定该相信哪个软件的预测值。为了解决这一问题,过去曾有人以药物分子作为测试集比较了部分 lgP 预测软件^[16-17],也有学者以普通化学品为测试集比较了部分 lgP 预测软件^[18-19]。但是近年来又开发了很多 lgP 预测软件,并且这些软件对于常见有机污染物预测的准确性还未见评价,因此本文选择 105 种有机污染物作为测试集,对目前常用的 10 种 lgP 预测软件进行对比,以评价不同软件对有机污染物 lgP 预测的准确性。

1 测试集选择和实验数据来源

1.1 测试集选择

选择 105 种有机污染物作为测试集,均来自美国环境保护局 (USEPA) 国家推荐水质标准 (National Recommended Water Quality Criteria, 2006) 中的优先毒性污染物。

由于测试集中有机污染物的数量较多,因此其具体名称和 CAS 编号等信息这里省略。105 种有机污染物根据其化学结构大致可以分为以下几类:(1) 卤代(烷、烯)烃类;(2) 苯系物;(3) 氯代苯类;(4) 酚类;(5) 硝基苯类;(6) 苯胺类;(7) 多环芳烃类;(8) 酞酸酯类;(9) 农药及其代谢物;(10) 丙烯腈;(11) 亚硝胺类;(12) 卤代醚类;(13) 氯代苯酚。此外还包括丙烯醛、2-氯萘、3,3'-二氯联苯胺、异佛尔酮等物质。

1.2 实验数据来源

选择的有机污染物 lgP 实验值来源于美国 SRC 公司的 PhysProp 数据库。

著名物理化学性质数据库 PhysProp 的在线版包括超过 25000 种化合物的化学结构、名称和物理性质^[20]。PhysProp 数据库中的数据均经过资深科学家的评价,且在不断更新。目前,已有很多研究者使用该数据库进行定量结构-性质/活性相关 (quantitative structure-activity/property relationships, QSPR/QSAR) 研究^[21]。

1.3 预测值计算方法和相应软件简介

在虚拟计算化学实验室网站 (<http://www.vccclab.org>) 计算

105 种有机污染物的 lgP, 该网站集成的 10 种可用于预测 lgP 的软件分别是: ALOGPs、AC logP、AB/LogP、COSMOFraq、miLogP、ALOGP、MLOGP、KOWWIN、XLOGP 2 和 XLOGP 3。虚拟计算化学实验室网站可以通过 CAS 编号、SMILES 式和结构式输入化合物,进而使用上述软件计算其 lgP、水溶解度和解离常数(对于可解离有机污染物而言),此外还提供了部分有机污染物的 lgP 实验值和水溶解度实验值作为参考。在这里选择 105 种有机污染物的 CAS 编号作为输入项(少数没有预测结果的物质使用 SMILES 式作为输入项)。下面对 10 种预测软件作简要介绍。

(1) ALOGPs

ALOGPs 软件^[22-26]由 I. V. Tetko 等开发,通过神经网络分析 12908 种物质,然后预测出相应物质的 lgP。该软件使用原子和化学键型的电子拓扑状态 (E-state) 指数、氢原子的数量、非氢原子的数量等指数描述分子结构。该软件使用多元线性回归分析法选择指数,共选出 75 种参数。一些参数组合了具有相似物理化学性质的原子和化学键指数。该软件使用逐一剔除法测试软件的预测能力。

(2) AC logP

AC logP 软件^[21]是由 Actelion 公司开发的 OSIRIS 性质研究程序中的一部分。AC logP 根据原子/碎片贡献法预测 lgP,它根据分子中原子及邻近原子的性质(键型、芳香性等)将原子分为 368 种类型。然后使用超过 5000 种化合物的实验 lgP 值训练计算出 368 种原子类型的贡献值。预测未知化合物的 lgP 时,将相应的原子/碎片求和。

(3) AB/LogP

AB/LogP 软件^[18,21]是由 Pharma Algorithms 开发的预测 lgP 的软件,该软件使用碎片法预测,并通过可靠性指数值定量给出了预测结果的可靠程度。

(4) COSMOFraq

COSMOFraq 软件使用线性溶剂化模型开发。它先根据超过 40000 种化合物计算出分子的 σ 碎片,在预测未知化合物的 lgP 时将相应的 σ 碎片组合计算得到。

(5) miLogP

miLogP 软件^[17-18]是 Molinspiration Chemoinformatics 开发的化学信息学软件工具中的一个模块。基本原理是基于基团贡献的计算方式。它通过超过 12000 个分子化合物的实验数值作为训练集预测恰当的 lgP 值。该软件通过 35 个小的基本碎片、185 个大碎片描述分子内氢键作用对 lgP 的贡献和电荷相互作用。miLogP 软件计算 lgP 的方法非常强大,可以预测几乎全部有机分子和大多数有机金属分子的 lgP。

(6) ALOGP

ALOGP 软件^[18]的算法是 Ghose-Crippen 法的原子加和法。该方法是纯粹的原子贡献加和,没有引入任何修正项。ALOGP 最初 (1989 年) 使用 Talete Srl 的 Dragon 1.4 软件运行。该软件的修订版是 ALOGP 98。

(7) MLOGP

MLOGP 软件^[18]使用疏水性和亲水性作为两类基本描述符。其中,疏水性描述符是指碳原子数和卤族原子数的加权后求和;亲水性描述符是指 N 原子和 O 原子数之和。这两类描述符解释了 1230 种化合物实验 $\lg P$ 值 73% 的变化。此外还引入了 11 种描述以提高模型的预测精度。这些描述符分别是:PRX 描述 N/O 原子邻近效应;UB 描述不饱和键数量;POL 描述极性芳香取代基数量;RNG 描述是否存在环状结构;NO₂ 描述硝基数量;HB 描述分子内氢键数量;AMP 描述化合物两性性质;ALK 描述化合物是否为烷烃、烯烃、环烷烃或环烯烃;QN 描述 4 价 N 原子;NCS 为描述异硫氰基;BLM 描述是否存在 β -内酰胺。由于 MLOGP 模型比较简单实用,较广泛用于计算 $\lg P$ 。

(8) KOWWIN

KOWWIN 软件^[19,27-28]根据化学结构式计算 $\lg P$ 。在 KOWWIN 中结构式以 SMILES 式的形式输入。以非 SMILES 式的其他形式的结构式也可以直接输入 KOWWIN 软件。KOWWIN 软件的预测方法由 SRC 公司开发。KOWWIN 软件使用“碎片常数法”预测 $\lg P$ 。在“碎片常数法”中,结构式被分成若干碎片(包括原子或者官能团),而每个碎片或官能团的系数值加起来就是分子的 $\lg P$ 预测值。KOWWIN 中各碎片和基团的系数值是根据 2400 种可靠的实验 $\lg P$ 值由多元回归方法得到的。

(9) XLOGP 2

XLOGP 2 软件^[18,28-29]是由北京大学物理化学研究所开发。XLOGP 2 软件本质上是一个原子加和模型,只是增加了一项校正因子。XLOGP 2 软件使用共 90 种原子类型来对中性有机化合物中的原子进行分类。其对原子分类的依据为:(1) 元素类型;(2) 杂化状态;(3) 溶剂可及性;(4) 相邻原子性质;(5) 与 π -体系接近程度。对于属于同一原子类型的原子一般具有相近的溶剂可及表面积和电荷密度。根据这种方法,得到某一特定原子类型的原子对 $\lg P$ 具有特定的贡献。XLOGP 2 软件引入了 10 种校正因子用于修正分子内的基团间相互作用(例如疏水性碳原子、分子内氢键等)。

(10) XLOGP 3

XLOGP 3^[18,28,30]是由中国科学院上海有机化学研究所开发,基本原理也是附加有修正项的原子加和模型。该软件使用最优化分类方法得到 87 种原子/基团类型以及用于校正分子内氢键和氨基酸类化合物的校正因子。它与 XLOGP 2 的区别在于:可以使用已知 $\lg P$ 值的化合物作为参照物质,在此基础上计算相应化合物的 $\lg P$ 以提高预测精度。该软件首先计算出要求化合物与参照化合物的差异,然后按照加和模型计算出对应的 $\lg P$ 。XLOGP 3 软件在其训练集中寻找与要预测的分子二维结构相似的分子作为参照物质,然后根据模型计算出要求的 $\lg P$ 。

2 数据处理方法

2.1 平均绝对偏差

平均绝对偏差(AAE)用于评价不同软件对该测试集的总体预测效果,根据下式计算各个软件的平均绝对偏差^[19]:

$$AAE = \frac{\sum_{i=1}^n |\lg P(\text{exp}) - \lg P(\text{pre})|}{n} \quad (1)$$

式中, $\lg P(\text{exp})$ 为分配系数 $\lg P$ 的实验值; $\lg P(\text{pre})$ 为分配系数 $\lg P$ 的软件预测值。

2.2 成对结果 t 检验

成对 t 检验(paired t -test)通常用于比较两种方法(或两个实验室)对一组样品进行分析,且这一组样品的被检测组分的含量不尽相同。

该方法中每一样品都来自两种方法的分析结果,构成“对子”。如果“对子”之间差值很小(平均值等于零或接近零),可以认为这两种方法的分析结果的差值 d 来自于平均值 $\mu_d = 0$ 的总体。为假设这 n 对结果是否存在显著性差异,即是否来自同一总体,假设 $H_0: \mu_d = 0$, 计算 t 为

$$t = \frac{\bar{d} \cdot \sqrt{n}}{S_d} \quad (2)$$

式中, \bar{d} 和 S_d 分别为配对结果的差别 d 的平均值和标准误差。 t 的自由度为 $n-1$ 。如果计算的 t 值小于临界值,零假设被接受,即两种方法不存在显著差异;否则说明两种方法之间存在显著差异^[31-32]。

置信区间:两种方法出现差异的 95% 置信区间为

$$\mu = \bar{d} \pm \frac{t_{0.05} \cdot S_d}{\sqrt{n}} \quad (3)$$

这里,以 105 种有机污染的分配系数实验值与一种软件的预测值组成“对子”,如果“对子”间的差值很小,则说明该种软件的预测精度很好,否则说明软件的预测精度比较差。

2.3 线性回归

一元线性回归所得的标准曲线也可用于比较两种方法的差异。这里以 105 种有机污染物的实验分配系数值为横轴,不同软件的估算值为纵轴,标准曲线上的每一个点表示一种化合物的实验分配系数值和软件估算值,通过对 105 种有机污染物的实验值和一种软件估算值进行线性回归得到回归曲线。除了以回归方程作为评价软件估算值与实验值的一致性程度外,还可以计算相关系数、标准偏差等指标^[33]。

线性回归的最终方程形式为

$$y = A + Bx$$

其中 A 为截距, $A = \bar{y} - B\bar{x}$ (4)
 B 为斜率,

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (5)$$

相关系数 R 定义为

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

拟合的标准偏差 SD 为

$$SD = \sqrt{\frac{\sum_{i=1}^n (y_i - (A + Bx_i))^2}{n-2}} \quad (7)$$

这里使用 OriginLab 公司的 OriginPro 7.5 软件进行成对的假设 t 检验 (Two Sample t -Test (Paired test)) 和线性回归运算。

3 结果与讨论

以 105 种选定的有机污染物作为测试集,使用上述 10 种预测软件分别计算 105 种化合物的 $\lg P$, 并与实验值进行比较。图 1 所示为 105 种有机污染物的分配系数实验值的分布情况。可以看出,105 种有机污染物的 $\lg P$ 分布在 -1—9 之间,绝大多数化合物的分配系数 $\lg P$ 值分布在 1—7 之间,其中 $\lg P$ 在 1—4 之间的化合物最多,超过化合物总数的 1/2。测试集的这一分布情况与环境实际存在的有机污染物的 $\lg P$ 的分布情况类似,因此可以认为这一测试集具有明显的代表性。

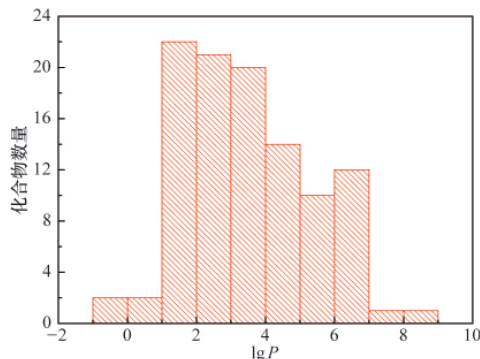


图 1 105 种有机污染物的分配系数 (正辛醇/水) 分布

Fig. 1 Distribution of $\lg P$ values for 105 organic pollutants

3.1 成对 t -检验结果和讨论

表 1 为 10 种软件预测值与实验值的成对 t 检验结果。可以看到,在 95% 置信度上,软件 KOWWIN、ALOGPs、AB/LogP 和 AC logP 都和实验值没有显著的差别,其中软件 KOWWIN、AB/LogP 和 ALOGPs 具有最高的 p -水平,分别为 0.89145、0.43717 和 0.26517。这说明对所选的 105 种有机污染物而言,KOWWIN 具有最好的准确性,其次分别为 AB/LogP 和 ALOGPs,而 AC logP 的预测值虽然与实验值没有显著差别,但精度比前面 3 个软件低。

表 1 不同预测软件的预测值与实验值的成对 t 检验结果

Table 1 Paired t -test results between predicted values of different softwares and experimented values

编号	软件名称	$t(t_{0.05}=1.98304)$	p -水平	预测值与实验值的差异性 (95% 置信度)	预测值与实验值差值的 95% 置信区间
1	ALOGPs	1.12029	0.26517	无	-0.01995—0.07176
2	AC logP	-1.75919	0.08149	无	-0.22954—0.01373
3	AB/LogP	-0.77999	0.43717	无	-0.09885—0.04304
4	COSMOFraq	2.0556	0.04233	有	0.00382—0.21275
5	miLogP	2.37793	0.01924	有	0.01351—0.14916
6	ALOGP	7.88647	0.00000	有	0.37207—0.62203
7	MLOGP	5.31271	0.00000	有	0.17901—0.39223
8	KOWWIN	-0.13681	0.89145	无	-0.05903—0.05141
9	XLOGP 2	2.51088	0.01358	有	0.03938—0.33528
10	XLOGP 3	2.25973	0.02592	有	0.01242—0.19044

表 2 具体分析了对于 105 种有机污染物,不同软件的预测值与实验值的绝对偏差分布情况及平均绝对偏差。从平均绝对偏差 AAE 看,ALOGPs 软件预测的 AAE 最小,仅为 0.18343,其次为 KOWWIN、XLOGP 3 和 AB/LogP 软件;ALOGP 软件的 AAE 最大,为 0.61286。从各个软件的绝对偏差分布情况可以看出,对于所选择的 105 种有机污染物,软件 ALOGPs 的预测值有 86 个绝对偏差在 0—0.3,19 个在 0.3—0.9,没有绝对偏差大于 0.9 的化合物,在所有软件中的精度最好。其次分别为 KOWWIN、AB/LogP 和 XLOGP 3 软件。综合平均绝对偏差和绝对偏差分布可以看出,ALOGPs、

KOWWIN 和 AB/LogP 软件的预测效果都比较好。XLOGP 3 软件虽然 AAE 较小,但是其与分配系数实验值的成对 t 检验却表明它与实验值有显著差异,进一步考查各个软件的预测值与实验值的绝对偏差发现,对于研究的 105 种有机污染物,XLOGP 3 预测值的绝对偏差有近 1/2 (50 个) 都等于零,而其他软件预测值的绝对偏差为零的化合物数量很少 (均 < 10 个)。Benfenati 等^[34]在比较不同软件的分配系数预测值时认为,如果对于一组化合物某个软件的预测值与实验值的绝对偏差为零,很大可能是软件的“预测值”并不是真正意义的预测值,而是与所选择实验值来源相同的“实验值”。在这里,考

表 2 不同软件的预测值与实验值的绝对偏差分布和平均绝对偏差
Table 2 AAE between predicted values and experimented values

编号	软件名称	$0 \leq \Delta < 0.3$	$0.3 \leq \Delta < 0.9$	$0.9 \leq \Delta < 1.5$	$1.5 \leq \Delta$	AAE
1	ALOGPs	86	19	0	0	0.18343
2	AC logP	61	34	5	5	0.41857
3	AB/LogP	76	26	3	0	0.25514
4	COSMOFraq	51	45	8	1	0.41114
5	miLogP	64	39	2	0	0.27695
6	ALOGP	36	42	17	10	0.61286
7	MLOGP	54	40	8	3	0.41933
8	KOWWIN	79	25	1	0	0.20819
9	XLOGP 2	67	25	4	9	0.43933
10	XLOGP 3	83	14	3	5	0.22105

虑到所选择的 105 种有机污染物都是比较常见的化合物,很大可能是 XLOGP 3 软件使用了部分物质的实验值。如果这样,平均绝对偏差就不适于评价 XLOGP 3 的预测准确性,也就是说对于更广泛的有机污染物而言,XLOGP 3 的预测精度更差。

3.2 线性回归结果和讨论

为了更深入地比较不同软件的预测精度,以 105 种有机污染物的分配系数实验值 $\lg P(\text{exp})$ 为横轴,相应软件的预测值 $\lg P(\text{pre})$ 为纵轴作图(图 2),并进行线性拟合。图中的虚线

为 $y=x$,实线为拟合曲线。得到的拟合方程及相关系数等信息见表 3。

从图 2 和表 3 可以看出,ALOGPs 和 KOWWIN 软件的预测值与实验值相关性最好,而且拟合曲线的截距(拟合曲线与直线 $y=x$ 基本重合)很小,说明在研究的 10 种软件当中,ALOGPs 和 KOWWIN 软件的预测精度最高。

ALOGPs 软件使用神经网络的方法建立模型,神经网络是一种抽象的数学模型,它可以在很大程度上提高模型预测的精度。此外,ALOGPs 软件使用了原子拓扑电子状态等 75

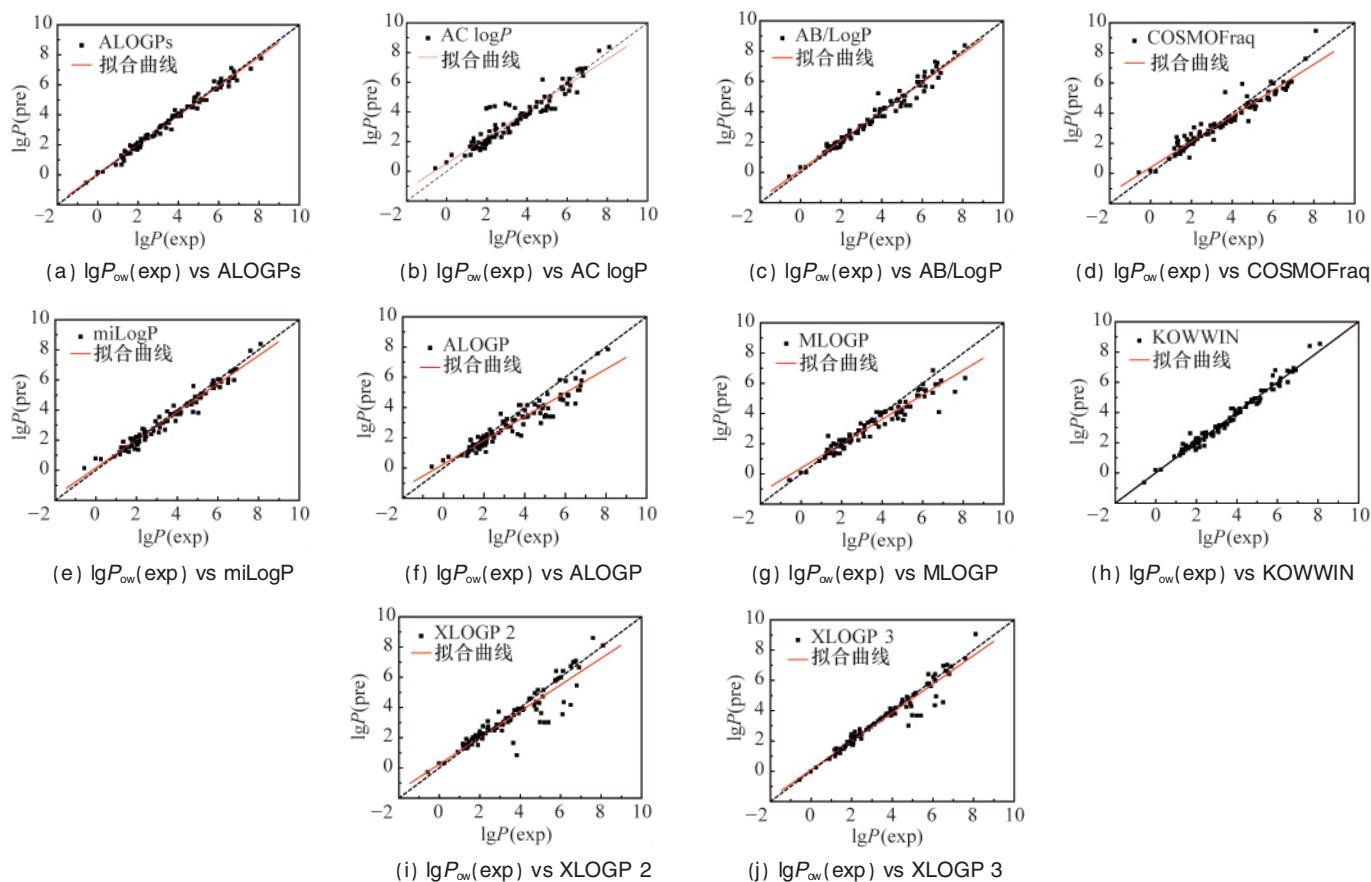


图 2 实验值对软件预测值

Fig. 2 Experimental and predicted $\lg P$ values

表 3 lgP 实验值与预测值的回归方程

Table 3 Regression equation between predicted and experimental lgP values

序号	软件名称	拟合方程	相关系数	标准偏差
1	ALOGPs	$y=0.0564+0.9765x$	0.9919	0.2340
2	AC logP	$y=0.5374+0.8773x$	0.9412	0.5885
3	AB/LogP	$y=0.1655+0.9607x$	0.9804	0.3610
4	COSMOFraq	$y=0.3970+0.8557x$	0.9593	0.4706
5	miLogP	$y=0.1540+0.9328x$	0.9827	0.3290
6	ALOGP	$y=0.2537+0.7855x$	0.9446	0.5103
7	MLOGP	$y=0.3823+0.8092x$	0.9630	0.4234
8	KOWWIN	$y=0.0102+0.9982x$	0.9884	0.2867
9	XLOGP 2	$y=0.2496+0.8752x$	0.9128	0.7319
10	XLOGP 3	$y=0.0964+0.9435x$	0.9690	0.4499

种描述符,能够比其他软件更全面地描述分子结构,这也极大地提高了其预测精度。Pyka 等^[6]以 193 种药物作为测试集,比较了 ALOGPs、IAllogP、KOWWIN 等 7 种分配系数预测软件,结果表明 ALOGPs 和 KOWWIN 的预测值与实验值的相关性最好。

KOWWIN 软件使用原子/碎片贡献法计算分配系数。各个原子/碎片的贡献与校正项的值分别根据多元线性回归的方法求得。Benfenati 等^[34]比较了 Pallas、KOWWIN 等 4 种单机版软件,结果表明 KOWWIN 的预测精度较高。

4 结论

本文通过选取的 105 种有机污染物比较了 10 种常见分配系数预测软件的精度,这 105 种物质包括了卤代烃类、苯系物、硝基苯类等常见有机污染物类型。105 种有机污染物的 lgP 范围为-1—9。成对 *t* 检验和一元线性回归的结果均表明,ALOGPs 和 KOWWIN 两种软件的预测结果更准确一些。

所比较的软件中,有基于结构的方法(包括碎片法和原子加和法),也有基于性质的方法。预测结果较好的 ALOGPs 和 KOWWIN 软件分别为基于性质的方法和碎片法。从各种软件的发展前景来看,KOWWIN 软件是在传统的碎片加和的基础上开发的,虽然目前可以嵌套在部分网站上运行,但是本质上属于单机版软件。ALOGPs 是完全通过 Internet 运行的,并且同时提供了化合物的实验值(如果有)和其他软件的预测值作为参考,这给使用者带来了极大方便。

总之,分配系数的预测值虽然在短期内不能完全取代实验值,但它在缩短预备实验的时间以及化学品的风险评估方面具有很大的优势,需要进一步深入研究。

参考文献 (References)

- [1] Leo A, Hansch C, Elkins D. Partition coefficients and their uses[J]. *Chemical Reviews*, 1971, 71(6): 525-616.
- [2] United States Environmental Protection Agency. Product properties test

- guidelines: OPPTS 830.7550 partition coefficient (*n*-octanol/water), shake flask method[M]. EPA 712-C-96-038, Washington, DC: USEPA, 1996.
- [3] Organisation for Economic Co-operation and Development. Test no 107: Partition coefficient (*n*-octanol/water), shake flask method[M]//OECD Guidelines for the Testing of Chemicals. Paris: OECD, 1995.
- [4] Organisation for Economic Co-operation and Development. Test no 117: Partition coefficient (*n*-octanol/water), HPLC method[M]//OECD Guidelines for the Testing of Chemicals. Paris: OECD, 2004.
- [5] United States Environmental Protection Agency. Product properties test guidelines: OPPTS 830.7570 partition coefficient (*n*-octanol/water), estimation by liquid chromatography [M]. EPA 712-C-96-040, Washington, DC: USEPA, 1996.
- [6] United States Environmental Protection Agency. Product properties test guidelines: OPPTS 830.7560 partition coefficient (*n*-octanol/water), generator column method[M]. EPA 712-C-96-039, Washington, DC: USEPA, 1996.
- [7] Braekvelt E, Tittlemier S A, Tomy G T. Direct measurement of octanol-water partition coefficients of some environmentally relevant brominated diphenyl ether congeners[J]. *Chemosphere*, 2003, 51(7): 563-567.
- [8] Fisk A T, Rosenberg B, Cymbalisky C D, et al. Octanol/water partition coefficients of toxaphene congeners determined by the "slow-stirring" method[J]. *Chemosphere*, 1999, 39(14): 2549-2562.
- [9] Organisation for Economic Co-operation and Development. Test no 123: Partition coefficient (1-octanol/water), slow-stirring method[M]//OECD Guidelines for the Testing of Chemicals. Paris: OECD, 2006.
- [10] Breindl A, Beck B, Clark T, et al. Prediction of the *n*-octanol/water partition coefficient, logp, using a combination of semiempirical molecular calculations and a neural network [J]. *Journal of Molecular Modeling*, 1997, 3(3): 142-155.
- [11] Li L, Wang L, Han S, et al. Comparison of four methods of predicting newly measured octanol/water coefficients (LogKow) for heterocyclic nitrogen compounds and the partition mechanism [J]. *Environmental Toxicology and Chemistry*, 1999, 18(12): 2723-2728.
- [12] Livingstone D J, Ford M G, Huuskonen J J, et al. Simultaneous prediction of aqueous solubility and octanol/water partition coefficient based on descriptors derived from molecular structure [J]. *Journal of Computer-Aided Molecular Design*, 2001, 15(8): 741-752.
- [13] Sârbu C, Onisor C, Posa M, et al. Modeling and prediction (correction) of partition coefficients of bile acids and their derivatives by multivariate regression methods [J]. *Talanta Special Section: Remote Sensing*, 2008, 75(3): 651-657.
- [14] Eros D, Kevosdi I, Orfi L, et al. Reliability of logP predictions based on calculated molecular descriptors: A critical review[J]. *Current Medicinal Chemistry*, 2002, 9(20): 1819.
- [15] Organisation for Economic Co-operation and Development. Test no. 117: Partition coefficient (*n*-octanol/water), High Performance Liquid Chromatograph (HPLC) method[M]//OECD Guidelines for the Testing of Chemicals. Paris: OECD, 2004.
- [16] Pyka A, Babu K A, Zachariasz M. A comparison of theoretical methods of calculation of partition coefficients for selected drugs [J]. *Acta Poloniae Pharmaceutica Drug Research*, 2006, 63(3): 159-167.
- [17] 苏笠, 杨劲, 王友群, 等. 化合物脂水分系数计算软件及比较研究 [J]. 中国药科大学学报, 2008, 39(2): 178-182.
Su Li, Yang Jin, Wang Youqun, et al. *Journal of China Pharmaceutical University*, 2008, 39(2): 178-182.
- [18] Mannhold R, Poda G I, Ostermann C, et al. Calculation of molecular

- lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds [J]. *Journal of Pharmaceutical Sciences*, 2009, 98(3): 861–893.
- [19] Machatha S G, Yalkowsky S H. Comparison of the octanol/water partition coefficients calculated by ClogP, ACDlogP and KowWin to experimentally determined values [J]. *International Journal of Pharmaceutics*, 2005, 294(1–2): 185–192.
- [20] SRC. Interactive PhysProp database demo [EB/OL]. <http://www.srcinc.com/what-we-do/databaseforms.aspx?id=386>, 2012.
- [21] Tetko I V. The WWW as a tool to obtain molecular parameters [J]. *Mini-Reviews in Medicinal Chemistry*, 2003, 3(8): 809–820.
- [22] Igor V T. Application of ALOGPS to predict 1-octanol/water distribution coefficients, log P , and log D , of AstraZeneca in-house database [J]. *Journal of Pharmaceutical Sciences*, 2004, 93(12): 3103–3110.
- [23] Tetko I V. Computing chemistry on the web [J]. *Drug Discovery Today*, 2005, 10(22): 1497–1500.
- [24] VCCLAB. Virtual computational chemistry laboratory [EB/OL]. <http://www.vcclab.org/>, 2005.
- [25] Kujawski J, Bernard M K, Janusz A, *et al.* Prediction of log P : ALOGPS application in medicinal chemistry education [J]. *Journal of Chemical Education*, 2011, 89(1): 64–67.
- [26] Tetko I, Gasteiger J, Todeschini R, *et al.* Virtual computational chemistry laboratory: Design and description [J]. *Journal of Computer-Aided Molecular Design*, 2005, 19(6): 453.
- [27] United States Environmental Protection Agency. Estimation Program Interface (EPI) suite [EB/OL]. v1.68 ed. <http://www.epa.gov/oppt/exposure/pubs/episuite.htm>, 2010.
- [28] 王艳玲, 李婕, 王任小. 有机化合物脂水分配系数和溶解度的计算方法[J]. *物理化学学报*, 2010, 26(7): 1742–1754.
Wang Yanling, Li Jie, Wang Renxiao. Computational approaches to estimating octanol-water partition coefficient and water solubility of organic compounds [J]. *Acta Physico-Chimica Sinica*, 2010, 26 (7): 1742–1754.
- [29] Institute of Physical Chemistry, Peking University. XLOGP2[EB/OL]. <http://www.ics.uci.edu/~dock/manuals/xlogp2.1/download.html>, 2012.
- [30] State Key Laboratory Of Bioorganic Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy Of Sciences. XLOGP3[EB/OL]. <http://www.sioc-ccbg.ac.cn/?p=42&software=xlogp3&action=download>, 2012.
- [31] 许禄, 邵学广. 化学计量学方法[M]. 2版. 北京: 科学出版社, 2004.
Xu Lu, Shao Xueguang. *Methods of chemometrics* [M]. 2nd ed. Beijing: Science Press, 2004.
- [32] 倪永年. 化学计量学在分析化学中的应用 [M]. 北京: 科学出版社, 2004.
Ni Yongnian. *Chemometrics in analytical chemistry*[M]. Beijing: Science Press, 2004.
- [33] 漆德瑶, 肖明耀, 吴芯芯. 理化分析数据处理手册 [M]. 北京: 中国计量出版社, 1990.
Qi Deyao, Xiao Mingyao, Wu Xinxin. *Data processing manual of physical and chemical analysis*[M]. Beijing: China Metrology Publishing House, 1990.
- [34] Benfenati E, Gini G, Piclin N, *et al.* Predicting log P of pesticides using different software[J]. *Chemosphere*, 2003, 53(9): 1155–1164.

(责任编辑 朱宇)

· 学术动态 ·

“第九届全国微波化学暨第四届样品制备学术会议”征稿

“第九届全国微波化学暨第四届样品制备学术会议”由中国电子学会微波分会微波化学专业委员会,中国分析仪器学会样品制备分会主办,大会定于2012年11月9—11日在昆明市召开。

征稿范围:(1)微波化学研究及应用领域:①微波化学理论及应用研究进展;②微波技术在化学、化工、食品科学、环境等领域中的研究和应用;③微波在制药工业中应用及基本理论研究;④微波在医药检测方面应用及研发;⑤微波化学反应设备研制、开发和应用;⑥微波化学及技术与绿色制备。(2)样品制备及技术领域:①样品前处理理论及其技术的新进展;②样品前处理专用仪器设备的研制、开发和应用;③样品前处理技术和仪器在工业监测及分析中的应用;④工业产品的样品前处理技术和仪器。(3)微波冶金及微波材料处理领域:①微波在冶金工业中应用及基本理论研究;②微波冶金设备研发及工业推广;③微波在材料制备及研发领域应用研究。(4)微波机理研究及微波设备研发:①微波腔体设计及计算;②微波场模拟及温度测量;③微波产业研发及应用研究。

全文截稿日期:2012年8月30日。

联系电话:microwave2012@163.com。

电子信箱:microwave2012@163.com。

大会网站:<http://um.kmust.edu.cn>。