

根据分子连接性指数估算非极性有机污染物的吸附系数*

卢晓霞 陶澍 李杭 朴海善

(北京大学城市与环境学系, 北京 100871)

摘要 根据分子拓扑理论,研究了分子连接性指数与有机污染物吸附系数的关系。通过对 300 种有机化合物的 $^1\chi$ 与 $\log K_{oc}$ 进行回归分析,发现 $^1\chi$ 与非极性化合物 K_{oc} 的对数有很好的线性关系。当在模型中引入 $^1\chi$ 、 $^2\chi$ 和 $^3\chi$ 时,计算的准确度进一步提高。在研究的 99 种非极性化合物中,55%的估算误差在 0.2 个对数单位之内。据此建立了根据分子连接性指数估算非极性有机污染物 K_{oc} 的定量模型。

关键词 分子连接性指数,有机污染物,非极性化合物, K_{oc}

中图分类号 X592

农药的施用、工业有机溶剂的排放、以及石油组分的泄漏等均使土壤受到有机物的污染。有机污染直接或间接危害人体健康,因而受到当前世界上许多国家的关注。有机污染物进入土壤后,与土壤中的固、液和气相物质发生一系列化学、物理化学和生物学反应。土壤对有机污染物的吸附是影响其迁移和归宿的重要过程,吸附作用使有机污染物残留于土壤中,从而影响其移动性和生物毒性^[1]。就吸附机理而言,土壤对有机和无机污染物的吸附有较大的差异。后者主要是表面络合作用和离子交换吸附,吸附等温方程式大多表现为非线性形式。而前者主要为憎水吸附,吸附等温方程式大多表现为线性形式。这种相似相溶的吸附作用的强度与土壤中的有机质含量紧密相关^[2]。因此,在定量描述土壤或沉积物对有机污染物的吸附作用时常常可以忽略非有机吸附剂的作用,而仅仅以有机化合物在纯有机质或纯有机碳上的吸附系数 K_{om} 或 K_{oc} 来描述其吸附平衡。在大多数情况下,可以直接根据 K_{oc} 值和土壤有机质含量计算某种化合物在土壤中的吸附系数。

K_{oc} 被定义为达到吸附平衡时,有机化合物被土壤或沉积物中有机质(以有机碳计)吸附的浓度与其在体系中的水相浓度之比,它是用于评价有机污染物环境行为的一个重要参数。 K_{oc} 的获取有两种方式,即实测和计算。实测 K_{oc} 具有准确、可靠和实用性强的优点,但由于有机化合物种类繁多,难以满足实际工作的需要。而通过计算获得 K_{oc} ,尽管不如实测准确,但其方法简便、成本低,并且可在短时间内获得大量有机化合物的 K_{oc} ,有利于实际工作的进行。计算 K_{oc} 的传统方法是建立有机化合物 K_{oc} 与其它理化性质的统计关系式,其中最常用的是 K_{oc} 与 K_{ow} (辛醇/水分配系数),以及 K_{oc} 与 S (水相溶解度)的

* 国家自然科学基金重点项目(批准号49632060)及国家杰出青年基金(49525102)资助

关系式^[3]。但由于 K_{ow} 中的辛醇相与天然有机吸附剂在本质上有很大差异,而 S 中并不包含对有机相的描述,因此这些计算方法有较大局限性。近年来,许多学者开始研究根据有机化合物的分子结构直接计算 K_{oc} , 分子连接性指数法便是其中重要的一种。

分子连接性指数 (MCI) 是由 Kier 和 Hall 等人根据拓扑理论,在分子分枝指数基础上提出和发展起来的,至今只有二十几年的历史^[4]。MCI 包含有分子大小、分枝多少、原子种类和数目等结构信息。已有研究表明, MCI 与有机化合物的许多理化性质,例如沸点、水溶解度、以及分配系数等有很好的相关性,它可以用来预测一些有机化合物的未知性质^[5]。本研究的目的是通过研究 MCI 与非极性有机化合物 K_{oc} 的定量关系,建立根据化合物分子结构直接计算这些化合物 K_{oc} 的定量模型。

1 研究方法

1.1 分子连接性指数计算

分子连接性指数 (MCI) 是从有机化合物分子的隐氢骨架中产生的拓扑指数。对于某一给定的分子结构,可以计算出不同类型和阶项的 MCI。本研究用到的 MCI 主要有 ${}^1\chi$ 、 ${}^1\chi^v$ 、 ${}^2\chi$ 和 ${}^2\chi^v$, 计算公式如下:

$${}^1\chi = \sum (\delta_i \times \delta_j)^{-1/2} \quad (1)$$

$${}^1\chi^v = \sum (\delta_i^v \times \delta_i^v)^{-1/2} \quad (2)$$

$${}^2\chi = \sum (\delta_i \times \delta_i \times \delta_k)^{-1/2} \quad (3)$$

$${}^2\chi^v = \sum (\delta_i^v \times \delta_j^v \times \delta_k^v)^{-1/2} \quad (4)$$

式中, δ 为原子的点价, δ^v 为原子的价点价, i, j, k 为分子中依次排列的非氢原子。

1.2 数据来源

本研究收集了 330 种有机化合物的 K_{oc} 实测数据。数据来源包括 G.G. Briggs^[6,7], E. E. Kenaga^[8], S. W. Karickhoff^[9,10,11], A. G. Douglas^[12], P. D. Vowles^[13], J. J. Hassett^[14,15] 和 C. T. Chiou^[16,17] 等学者发表或整理的文献。对于有多个实测数据的化合物,取其中值。

所研究化合物包括脂肪烃、芳香烃、卤代烃、苯胺、乙酰苯胺、苯脲、酸、醇、酚、硝基化合物、含氮杂环化合物、以及各类酯等。研究时将 330 种化合物随机地分成两组: A 组和 B 组。A 组为训练组,共有 300 种化合物, $\log K_{oc}$ 范围为 $-0.40 \sim 6.22$ 。B 组为检验组,共有 30 种化合物, $\log K_{oc}$ 范围为 $0.98 \sim 6.07$ 。

1.3 研究方案

为了探讨各类有机化合物 MCI 与 K_{oc} 的一般关系,首先对 A 组所有化合物的 ${}^1\chi$ 与 K_{oc} 进行回归分析。回归结果显示极性化合物和非极性化合物在 ${}^1\chi \sim \log K_{oc}$ 关系图上的分布不同,其中非极性化合物的 ${}^1\chi$ 与 $\log K_{oc}$ 表现出较好的线性关系,因此下面单独对非极性化合物进行研究。研究步骤为:

(1) 研究 ${}^1\chi$ 与 K_{oc} 关系; (2) 研究多项 MCI 与 K_{oc} 关系; (3) 验证模型; (4) 以所有非极性化合物为基础建立最终模型。本研究使用 Statistica (version 4.5) 统计软件实现所有计算。

2 结果与讨论

2.1 各类化合物 $^1\chi$ 与 K_{oc} 关系

计算 A 组各类化合物的 $^1\chi$, 对 $^1\chi$ 与 $\log K_{oc}$ 的关系进行线性回归, 得到回归方程:

$$\log K_{oc} = 0.407 + 0.380^1\chi, \quad n = 300, \quad R^2 = 0.403, \quad s = 0.988 \quad (5)$$

式中, n 为化合物总数, R^2 为可决系数, s 为估计的标准差。 $^1\chi$ 与 $\log K_{oc}$ 的关系如图 1 所示。

从图 1 可以看出, 尽管 $^1\chi$ 与 $\log K_{oc}$ 显著相关(在 0.05 水平), 但大部分数据偏离回归直线较远, 不能直接用该回归模型计算各类化合物的 K_{oc} 。如果仔细分析图中数据点, 可以发现分布在回归直线之上的化合物表现出一定的规律, 即位于回归直线之上的化合物大多为不含极性基团的化合物, 而位于回归直线之下的化合物大多为含有极性基团的化合物。事实上, 100% 的非极性化合物都表现为对回归方程正偏离, 78% 的极性化合物负向偏离回归方程。由此可见, 极性与非极性化合物在 $^1\chi \sim \log K_{oc}$ 定量关系上表现出明显不同的特性。

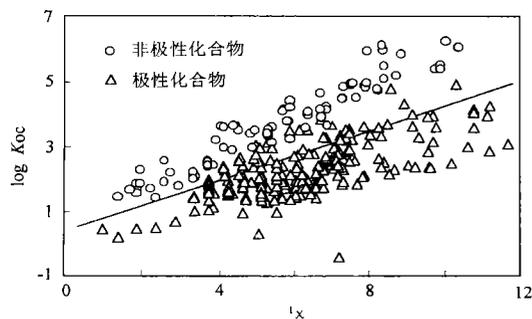


图1 各类化合物 $^1\chi \sim \log K_{oc}$ 回归直线

Fig.1 $^1\chi \sim \log K_{oc}$ regression line of all compounds

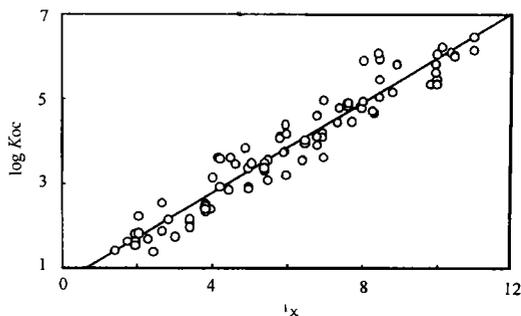


图2 训练组非极性化合物 $^1\chi \sim \log K_{oc}$ 回归直线

Fig.2 $^1\chi \sim \log K_{oc}$ regression line of nonpolar compounds in the training set

2.2 非极性化合物 $^1\chi$ 与 K_{oc} 关系

图 1 中非极性化合物比极性化合物表现出更好的 K_{oc} 与 $^1\chi$ 的关系。因此仅取非极性化合物的 K_{oc} 与 $^1\chi$ 值(共 89 种化合物), 得到以下回归方程:

$$\log K_{oc} = 0.689 + 0.529^1\chi, \quad n = 89, \quad R^2 = 0.925, \quad s = 0.387 \quad (6)$$

$^1\chi$ 与 $\log K_{oc}$ 的关系如图 2 所示。

从图 2 可以看出, 非极性化合物的 $^1\chi$ 与 $\log K_{oc}$ 有很好的线性关系。将所研究化合物的 $^1\chi$ 代入方程(6)计算 K_{oc} , 得到的 K_{oc} 计算值和实测值的平均误差为 0.31 个对数单位, 45% 的绝对误差低于 0.2 个对数单位。根据方程 (6) 计算 K_{oc} 的误差分布如图 3 所示。

2.3 非极性化合物多项 MCI 与 K_{oc} 关系

为充分利用 MCI 所包含的信息特征, 在回归方程中引入 $^1\chi^v$, $^2\chi$ 和 $^2\chi^v$, 这三个指数包

含了不饱和键、杂原子、以及分枝的信息。对 ${}^1\chi$ 、 ${}^1\chi^v$ 、 ${}^2\chi$ 和 ${}^2\chi^v$ 与 $\log K_{oc}$ 的关系进行多元线性回归,得到以下回归方程:

$$\begin{aligned} \log K_{oc} = & 0.79 - 0.159{}^1\chi + 0.408{}^1\chi^v \\ & + 0.593{}^2\chi - 0.291{}^2\chi^v, \quad n = 89, \\ R^2 = & 0.943, \quad s = 0.344 \end{aligned} \quad (7)$$

回归方程可决系数 R^2 由(6)式的0.925提高到0.943,估计的标准差 s 由0.387降低到0.344,表明回归方程对数据的拟合质量有所提高。由此可见,非极性化合物的 K_{oc} 主要取决于分子表面积大小(${}^1\chi$ 可视为分子中非氢骨架所占分子表面积大小),同时还与分子的电子特征有关。对方程变量进行相关分析,发现 ${}^1\chi$ 与 ${}^2\chi$ 显著相关($P < 0.05$),表明这两个指数包含的信息特征重复或部分重复。最后选定 ${}^1\chi^v$ 、 ${}^2\chi$ 和 ${}^2\chi^v$ 作为参数,建立模型:

$$\log K_{oc} = 0.757 + 0.300{}^1\chi^v + 0.453{}^2\chi - 0.198{}^2\chi^v, \quad n = 89, \quad R^2 = 0.941, \quad s = 0.345 \quad (8)$$

根据方程(8)计算所研究化合物的 K_{oc} ,得到的 $\log K_{oc}$ 计算值与实测值关系如图4所示。计算的平均误差为0.27个对数单位,其中48%的误差低于0.2个对数单位,误差分布见图5。

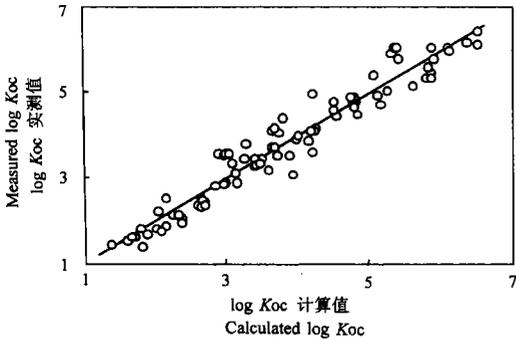


图4 训练组非极性化合物 $\log K_{oc}$ 的计算值—实测值关系

Fig.4 Relationship between the calculated and measured values of $\log K_{oc}$ of nonpolar compounds in the training set

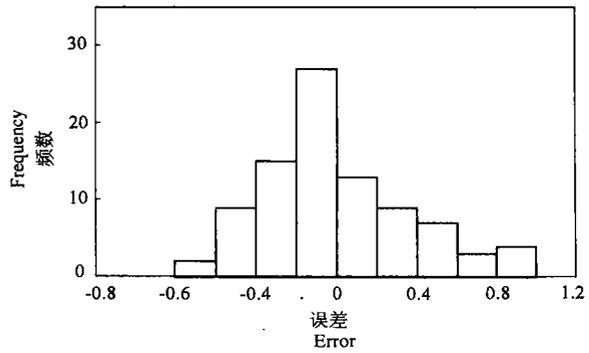


图3 训练组非极性化合物计算 K_{oc} 的误差分布(对数单位)

Fig.3 Error distribution of calculating K_{oc} of nonpolar compounds in the training set (log unit)

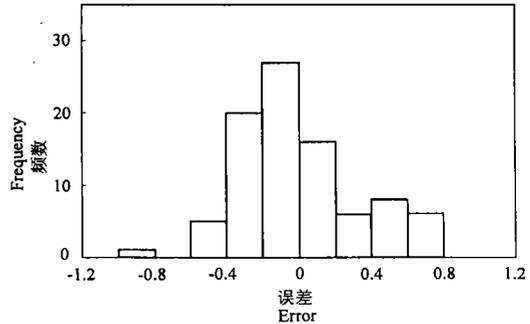


图5 训练组非极性化合物计算 K_{oc} 的误差分布(对数单位)

Fig.5 Error distribution of the calculated K_{oc} of nonpolar compounds in the training set (log unit)

2.4 模型验证

为进一步验证该模型的可靠性,用(8)式计算了检验组中10种非极性化合物的 K_{oc} 值,图6和图7分别给出了 $\log K_{oc}$ 计算值与实测值的关系及计算 K_{oc} 的误差分布。

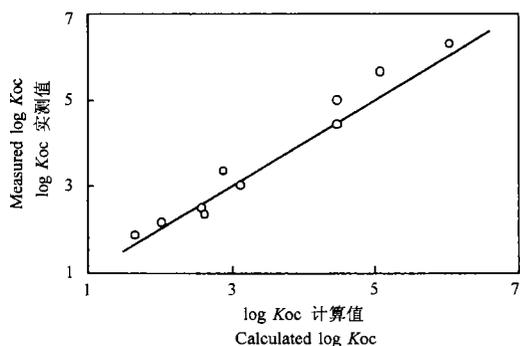


图6 检验组非极性化合物Koc的计算值—实测值关系

Fig.6 Relationship between the calculated and measured values of logKoc of nonpolar compounds in the test set

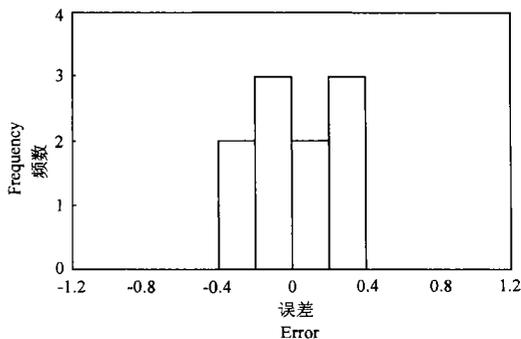


图7 检验组非极性化合物计算Koc的误差分布(对数单位)

Fig.7 Error distribution of the calculated Koc of nonpolar compounds in the test set (log unit)

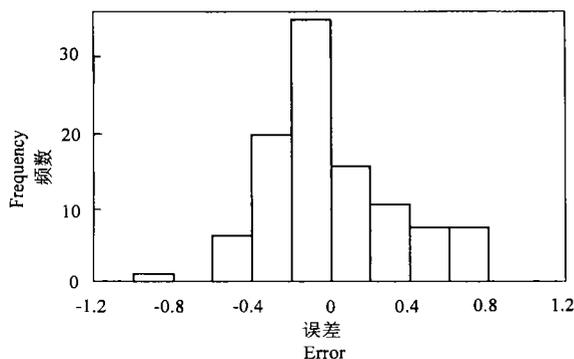


图8 所有非极性化合物计算Koc的误差分布(对数单位)

Fig.8 Error distribution of the calculated Koc of all nonpolar compounds (log unit)

算非极性化合物 Koc 的定量模型。该模型对所有 99 种非极性化合物 Koc 的估算误差如图 8 所示。其平均误差为 0.26 个对数单位,其中 55% 的误差小于 0.2 个对数单位。

3 结论

非极性有机化合物分子连接性指数与它们的吸附系数(Koc)有很好的线性关系,可以根据分子连接性指数较准确地计算非极性化合物的 Koc。应用多项分子连接性指数,能够提高对 Koc 的估算精度。对大多数非极性化合物而言,选择 ${}^1\chi^v$ 、 ${}^2\chi$ 和 ${}^2\chi^v$ 即可较准确地计算 Koc。根据 99 种非极性化合物实测 Koc 数据建立的估算模型的平均误差为 0.26 个对数单位,其中 55% 小于 0.2 个对数单位。

检验组 Koc 计算值与实测值的平均误差仅为 0.19 个对数单位,由于这些化合物的实测数据并未用于模型建立过程中,从而证实了所建模型的可靠性。

2.5 最终模型确定

为充分利用现有数据信息,改善模型参数的准确性,以两组中所有非极性化合物实测 Koc 值为基础,按方程 (8) 方式重新建立回归模型,得:

$$\begin{aligned} \log Koc = & 0.728 + 0.302^1\chi^v + 0.457^2\chi \\ & - 0.198^2\chi^v, n = 89, R^2 = 0.942, \\ & s = 0.338 \end{aligned} \quad (9)$$

该式即为根据分子连接性指数计

参 考 文 献

1. Chiou C T. Theoretical considerations of the partition uptake of nonionic organic compounds by soil organic matter. In: *Reactions and Movement of Organic Chemicals in Soils*. SSSA, WI: Madison, 1989, 1~29
2. Chiou C T, Reters L J, Freed V H. Partition equilibria of nonionic organic compounds between soil organic matter and water. *Environ. Sci. & Technol.* 1983, 17: 227~231
3. Meylan W, Howard P H, Boething R S. Molecular topology/fragment contribution method for predicting soil sorption coefficients. *Environ. Sci. & Technol.* 1992, 26: 1560~1567
4. Kier L B, Hall L H. *Molecular connectivity in structure activity analysis*, Chichester: Research Studies Press, 1986
5. 王连生.关于分子连接性指数与理化参数相关性的探讨. *环境科学丛刊*, 1989, 10(3): 1~12
6. Briggs G G. A simple relationship between soil sorption of organic chemicals and their octanol-water partitioning behaviour. *Proc 7th British insecticide and fungicide conf.*, 1973, 11: 475~478
7. Briggs G G. Theoretical and experimental relationships between soil adsorption, octanol-water partition coefficients, water solubilities, bioconcentration factors, and the parachor. *J. Agric. Food Chem.*, 1981, 29: 1050~1059
8. Kenaga E E, Goring C A. Relationship between water solubility, soil sorption, octanol-water partitioning, and concentration of chemicals in biota In: Eaton J G, Parish P P, Hendricks A C. eds. *Aquatic Toxicology. Proceedings of the Third Annual Symposium on Aquatic Toxicology*. ASTM, STP 707, Philadelphia, 1980. 78~115
9. Karickhoff S W, Brown D S, Scott T A. Sorption of hydrophobic pollutants on natural sediments. *Water Res.*, 1979, 13: 241~248
10. Karickhoff S W. Semiempirical estimation of sorption of hydrophobic pollutants on natural sediments and soils. *Chemosphere*, 1981, 10(8): 833~846
11. Karickhoff S W, Morris K R. Sorption dynamics of hydrophobic pollutants in sediment suspensions. *Environ. Toxicol. Chem.*, 1985, 4(4): 469~479
12. Douglas A G. *Proceedings of 9th international meeting on organic geochemistry*. Newcastle upon Tyne, 1979
13. Vowles P D, Mantoura R F C. Sediment-water partition coefficients and HPLC-retention factors of aromatic hydrocarbons. *Chemosphere*, 1987, 16: 109~116
14. Hassett J J. Sorption properties of sediments and energy-related pollutants. EPA-600/3-80-041, 1980, 23~67
15. Hassett J J. Sorption of Alpha-naphthol: Implications concerning the limits of hydrophobic sorption. *Soil Sci. Soc. Am. J.*, 1981, 45(1): 38~42
16. Chiou C T, Peters L J, Freed V H. A physical concept of soil-water equilibria for nonionic organic compounds. *Science*, 1979, 206: 831~832
17. Chiou C T, Porter P E, Schmeddi D W. Partition equilibria of nonionic organic compounds between soil organic matter and water. *Environ. Sci. & Technol.*, 1983, 17(4): 227~231

ESTIMATION OF ADSORPTION COEFFICIENT VALUES OF NONPOLAR ORGANIC COMPOUNDS BASED ON MOLECULAR CONNECTIVITY INDICES

Lu Xiao-xia Tao Shu Li Hang Piao Hai-shan

(*Department of Urban and Environmental Sciences, Peking University, Beijing 100871*)

Summary

Relationship between the molecular connectivity indices and adsorption coefficient (K_{oc}) of nonpolar organic compounds was investigated based on molecular topology. Close linear relationship between ${}^1\chi$ and $\log K_{oc}$ of nonpolar compounds was revealed by the regressive analysis of ${}^1\chi$ and $\log K_{oc}$ of 300 organic compounds. Calculation accuracy would be further improved as ${}^1\chi^v$, ${}^2\chi$ and ${}^2\chi^v$ were introduced into the model. Fifty five percent of the estimated error was within 0.2 log unit for 99 nonpolar compounds studied. A quantitative model was developed for estimating K_{oc} of nonpolar organic pollutant based on molecular connectivity indices.

Key words Molecular connectivity indices, Nonpolar organic compounds, K_{oc}