

DOI: 10.7524/j.issn.0254-6108.2015.10.2015022701

杨静, 王建兵, 王亚华, 等. 酚类物质臭氧氧化降解的定量构效关系[J]. 环境化学, 2015, 34(10): 1932-1939

YANG Jing, WANG Jianbing, WANG Yahua, et al. Quantitative structure-activity relationship for the ozonation of phenols[J]. Environmental Chemistry, 2015, 34(10): 1932-1939

## 酚类物质臭氧氧化降解的定量构效关系\*

杨 静<sup>1</sup> 王建兵<sup>2\*\*</sup> 王亚华<sup>2</sup> 张峰源<sup>2</sup> 何绪文<sup>2</sup>

(1. 国家自然科学基金委, 北京, 100085; 2. 中国矿业大学(北京)化学与环境工程学院, 北京, 100083)

**摘 要** 测定了 23 种酚的臭氧氧化速率, 分别采用遗传算法(GA)结合偏最小二乘法(PLS)、遗传算法结合人工神经网络(ANN)建立了酚类物质臭氧氧化速率的定量构效关系(QSAR)模型. 研究表明, 臭氧氧化酚的速率可用伪一级反应速率模型描述, 苯环上取代基得失电子的能力对酚的氧化速率影响较大. 基于 GA-PLS 算法建立的 QSAR 模型为  $\lg k = 3.439 - 0.206 \lg P$  (辛醇-水分配系数对数值)  $+ 0.122 \times pK_a$  (解离常数)  $+ 0.346^4 \chi_{pc}$  (四阶路径/簇分子连接性指数)  $- 0.0236qC^-$  (碳原子所带最大负电荷). 基于 GA-ANN 算法建立的 QSAR 模型含有参数  $\lg P$ 、 $\chi_{pc}$ 、 $pK_a$  和  $\alpha$  (平均分子极化率). 留一法交叉验证结果表明, 基于 GA-ANN 算法建立的模型比基于 GA-PLS 算法建立的模型具有更好的稳健性. QSAR 研究表明, 酚的臭氧氧化速率与电子云分布以及苯环上取代基的性质密切相关, 另外, 水的溶剂化作用对酚的氧化速率也有显著影响.

**关键词** 臭氧氧化, 酚, 遗传算法, 偏最小二乘法, 人工神经网络.

## Quantitative structure-activity relationship for the ozonation of phenols

YANG Jing<sup>1</sup> WANG Jianbing<sup>2\*\*</sup> WANG Yahua<sup>2</sup> ZHANG Fengyuan<sup>2</sup> HE Xuwen<sup>2</sup>

(1. National Science Foundation of China, Beijing, 100085, China; 2. School of Chemical and Environmental Engineering, Beijing Campus, China University of Mining and Technology, Beijing, 100083, China)

**Abstract:** Ozonation rates of twenty-three phenols were measured. Their Quantitative Structure Activity Relationship (QSAR) models were developed by the method of genetic algorithm (GA) combining with Partial Least Squares (PLS) and Artificial Neural Networks (ANN), respectively. The degradation rate of phenols can be described by the pseudo-first-order reaction rate model. The capacity of releasing or taking electron of the substitution group in the ring has obvious effect on the ozonation rate of the phenols. The QSAR model developed by GA-PLS is  $\lg k = 3.439 - 0.206 \lg P$  (the logarithm of octanol-water partition coefficients)  $+ 0.122 \times pK_a$  (dissociation constant)  $- 0.346^4 \chi_{pc}$  (four order path/cluster molecular connectivity index)  $- 0.0236qC^-$  (the maximum negative charge of carbon atom). The QSAR model developed by GA-ANN model has the descriptors of  $\lg P$ ,  $\chi_{pc}$ ,  $pK_a$  and  $\alpha$  (molecular average polarizability). Based on leave-one-out cross validation, the QSAR model constructed by GA-ANN has better robustness than that by GA-PLS. The study of QSAR shows that the ozonation rate of phenols has a close relationship with electron cloud distribution and the properties of substitution groups in benzene ring. It also shows that the solvent effect of water obviously influences the ozonation rate of phenols.

**Keywords:** ozonation, phenol, genetic algorithm, least square regression, artificial neural network.

2015 年 2 月 27 日收稿.

\* 国家自然科学基金(20907072)资助.

\*\* 通讯联系人, Tel: 010-62339087; E-mail: wangjb@cumt.edu.cn

酚类化合物普遍存在于工业废水(如煤化工废水)二级生化处理系统出水中,由于具有较高的毒性,并且常规污水深度处理工艺(混凝、沉淀、过滤等)很难将其浓度降低到排放标准规定的限值以下,必须采取有效的工艺加以去除.臭氧氧化是一种典型的高级氧化工艺,被广泛用于水中低浓度酚的去除<sup>[1-3]</sup>.水中有机物的臭氧氧化包含传质和化学反应两个过程,化学反应又包括酚与臭氧、酚与自由基、中间产物与臭氧、中间产物和自由基之间的反应,过程复杂,影响因素较多<sup>[4-7]</sup>.由于水中臭氧溶解度小,臭氧氧化酚的速率较快,现有研究一般在快速均匀混合体系中用过量的臭氧氧化水溶液中微量酚,通过精确分析微量酚的浓度变化,采用伪一级反应速率模型计算臭氧与酚之间的直接反应速率常数,实验难度较大.而且,自然界存在有 2000 多种酚类化合物,逐一测定它们的臭氧氧化速率必然耗费大量的人力物力.

近年来,定量构效关系(Quantitative structure activity relationship, QSAR)模型被广泛用来研究有机物光解、光催化氧化、臭氧氧化的速率<sup>[8-12]</sup>.水中酚的臭氧氧化降解 QSAR 模型研究也受到关注, Liu<sup>[13]</sup>等测试了 26 种酚的臭氧氧化速率,选用 12 种性质参数,通过多元线性回归建立了臭氧氧化降解酚的 QSAR 模型,最优的模型包含分子最低未占据轨道能、平均分子极化率、分子中电负性最大原子的净电荷、标准熵等结构参数. Sudhakaran 和 Amy<sup>[8]</sup>建立了 40 种有机物(含 6 种酚)的臭氧氧化降解 QSAR 模型,建模所采用的速率常数摘自不同的文献.然而,不同文献中的实验条件不同,获取速率常数过程中的误差也会不同,这必然会影响到模型的准确性.相对于酚的总数,已研究的酚类物质数目仍是非常少的,现有的研究还难以充分揭示臭氧氧化酚的构效关系,继续开展臭氧氧化降解酚的 QSAR 模型研究仍然具有非常重要的意义.

本文测试了 23 种酚的臭氧氧化速率,选用 27 种性质参数,分别采用偏最小二乘(Partial least squares, PLS)和人工神经网络(Artificial neural net, ANN)两种算法,建立了臭氧氧化降解酚的 QSAR 模型,在此基础上,探讨了臭氧与酚之间的直接反应速率与酚结构的关系.由于选用的结构参数较多,本文采用遗传算法(Genetic algorithm, GA)来优选变量,遗传算法(GA)是一种良好的变量选择方法,采用 GA-PLS 能利用 GA 排除特别不适合的变量,有利于用较低维数建立具有预言能力的模型. GA-ANN 较之单纯的 BP 神经网络是全局最优的算法,并且在高维空间能以较快速度逼近全局最优点<sup>[14-18]</sup>.

## 1 材料与方 法

### 1.1 臭氧氧化降解酚的实验方法

臭氧氧化降解酚的实验在玻璃反应器(高 250 mm,直径 40 mm)中进行.先向反应器中注入 490 mL 去离子水,在剧烈搅拌条件下( $500 \text{ r} \cdot \text{min}^{-1}$ ),将臭氧发生器(EXT120, Longevity Resources Inc., Canada)产生的臭氧从底部鼓入反应器,待臭氧浓度饱和后(约 120 min),注入 10 mL 的酚溶液( $10 \text{ mg} \cdot \text{L}^{-1}$ ),整个实验过程中不停地通入臭氧,保持反应体系中溶解性的臭氧相对于酚是过量的,并在  $\text{pH}=7.0$  和碳酸氢盐浓度为  $100 \text{ } \mu\text{mol} \cdot \text{L}^{-1}$  的条件下进行,每隔一定时间取样分析水中臭氧和酚的浓度.

### 1.2 分析方法

臭氧浓度采用碘量法分析<sup>[19-20]</sup>.酚浓度采用高效液相色谱 HPLC (Shimadzu, LC-10AD, Japan)分析,色谱柱为 Inertsil ODS-3 ( $5 \text{ } \mu\text{m}$ ,  $250 \text{ mm} \times 4.6 \text{ mm i.d.}$ ),流动相为甲醇和 1% 的冰醋酸混合溶液 ( $V/V=50:50$ ),流速  $1.0 \text{ mL} \cdot \text{min}^{-1}$ ,柱温  $40 \text{ } ^\circ\text{C}$ ,检测波长根据酚类物质在紫外可见分光光度计(Hach, DR/4000U, USA)的最大吸收波长确定.

### 1.3 QSAR 建模方法

共选用 23 种酚,初步选取酚的 27 种性质参数(表 1):辛醇-水分配系数对数值  $\lg P$ (无量纲)、分子折射系数  $\text{MR}(\text{m}^3 \cdot \text{mol}^{-1})$ 、解离常数  $\text{p}K_{\text{a}}$ (无量纲)、偶极矩  $\mu$ (Debye)、平均分子极化率  $\alpha$ (a.u.)、氧原子最大净电荷  $q\text{O}^-$ (a.c.u.),分子最高占据轨道能  $E_{\text{HOMO}}$ (eV)、分子最低未占据轨道能  $E_{\text{LUMO}}$ (eV)、分子总核排斥能  $\text{CCR}(\text{eV})$ 、电子能  $\text{EE}(\text{eV})$ 、分子总能量  $\text{TE}(\text{eV})$ 、分子中 C 原子所带最大负电荷  $q\text{C}^-$ (a.c.u.)、分子生成热  $\text{HOF}(\text{eV})$ 、分子离子化势  $\text{IP}$ (无量纲)、分子量  $\text{Mw}$ (无量纲)、分子长度  $\text{Dimen}(\text{nm})$  以及分子

连接性指数.分子连接性指数包括路径连接性指数( ${}^0\chi_p, {}^1\chi_p, \dots, {}^6\chi_p$ ) (无量纲)和路径/簇分子连接性指数( ${}^3\chi_c, {}^4\chi_{pc}, {}^5\chi_{pc}, {}^6\chi_{pc}$ ) (无量纲)<sup>[21-22]</sup>.  $\lg P$ 、MR、 $\mu$ 、 $\alpha$ 、 $qO^-$ 、 $E_{\text{homo}}$ 、 $E_{\text{lumo}}$ 、CCR、EE、TE、 $qC^-$ 、HOF、IP、Mw、Dimen 等参数的值通过 ChemOffice2004 软件采用 PM3 算法计算,分子连接性指数采用 Matlab 编程计算<sup>[8,23-24]</sup>.

表 1 臭氧氧化降解酚 QSAR 模型的描述算符

Table 1 Descriptors of QSAR model for the ozonation of phenols

| 酚         | $\lg P$ | CMR  | HOF    | TE       | EE        | CCR     | IP   | Mw     | Dimen | $E_{\text{homo}}$ | $E_{\text{lumo}}$ | $\mu$ | $qO^-$ | $qC^-$ |
|-----------|---------|------|--------|----------|-----------|---------|------|--------|-------|-------------------|-------------------|-------|--------|--------|
| 2,4-二氯酚   | 2.97    | 3.82 | -147.4 | -1891.10 | -6992.13  | 5101.03 | 9.27 | 163.00 | 6.28  | -0.10             | 0.82              | 0.40  | -0.24  | -0.14  |
| 对苯二酚      | 0.81    | 2.99 | -275.1 | -1491.51 | -5747.48  | 4255.97 | 8.73 | 110.11 | 6.22  | -8.73             | 0.23              | 2.27  | -0.25  | -0.18  |
| 对硝基苯酚     | 1.85    | 3.45 | -82.37 | -2001.84 | -8166.18  | 6164.34 | 10.1 | 139.11 | 6.83  | -0.42             | 0.52              | 5.27  | -0.36  | -0.23  |
| 对溴苯酚      | 2.63    | 3.62 | -74.00 | -1510.56 | -5574.19  | 4063.63 | 9.19 | 173.01 | 6.43  | -9.19             | 0.02              | 1.59  | -0.25  | -0.22  |
| 邻氨基苯酚     | 0.62    | 3.21 | -96.25 | -1391.94 | -5722.90  | 4330.96 | 8.64 | 109.13 | 5.81  | -8.64             | 0.41              | 1.33  | -0.27  | -0.17  |
| 对氨基苯酚     | 0.25    | 3.21 | -96.39 | -1391.94 | -5646.74  | 4254.80 | 8.27 | 109.13 | 6.57  | -8.27             | 0.44              | 2.10  | -0.25  | -0.18  |
| 对氯苯酚      | 2.48    | 3.33 | -123.0 | -1531.05 | -5640.58  | 4109.54 | 9.12 | 128.56 | 6.25  | -9.12             | 0.09              | 1.48  | -0.25  | -0.21  |
| 邻硝基苯酚     | 1.85    | 3.45 | -66.62 | -2001.68 | -8355.48  | 6353.80 | 9.95 | 139.11 | 6.08  | -0.34             | 0.66              | 6.34  | -0.38  | -0.23  |
| 间氨基苯酚     | 0.25    | 3.21 | -101.0 | -1391.99 | -5662.85  | 4270.87 | 8.53 | 109.13 | 5.83  | -8.53             | 0.52              | 2.46  | -0.25  | -0.28  |
| 2,3-二甲基苯酚 | 2.37    | 3.77 | -149.7 | -1482.60 | -6901.19  | 5418.59 | 8.93 | 122.17 | 5.93  | -8.93             | 0.37              | 1.70  | -0.26  | -0.18  |
| 间甲基苯酚     | 1.97    | 3.31 | -125.2 | -1326.81 | -5573.55  | 4246.74 | 9.03 | 108.14 | 5.90  | -9.03             | 0.39              | 1.53  | -0.25  | -0.22  |
| 邻甲基苯酚     | 1.92    | 3.31 | -123.6 | -1326.79 | -5634.86  | 4308.07 | 9.00 | 108.14 | 5.88  | -9.00             | 0.37              | 1.42  | -0.25  | -0.16  |
| 2-甲氧基苯酚   | 1.32    | 3.46 | -251.3 | -1646.80 | -7092.01  | 5445.21 | 8.79 | 124.14 | 7.11  | -8.79             | 0.39              | 2.41  | -0.25  | -0.16  |
| 1-萘酚      | 2.65    | 4.53 | -9.602 | -1710.29 | -8344.51  | 6634.22 | 8.49 | 144.17 | 7.23  | -0.37             | 0.01              | 1.46  | -0.25  | -0.17  |
| 邻苯二酚      | 0.88    | 2.99 | -277.9 | -1492.54 | -5829.94  | 4337.41 | 8.89 | 110.11 | 5.68  | -8.89             | 0.30              | 2.13  | -0.27  | -0.17  |
| 1,2,4-苯三酚 | 0.21    | 3.15 | -459.7 | -1812.10 | -7223.42  | 5411.32 | 8.62 | 126.11 | 6.23  | -8.62             | 0.13              | 2.80  | -0.25  | -0.26  |
| 五氯苯酚      | 4.71    | 5.30 | -187.7 | -2971.92 | -11706.51 | 8735.59 | 9.57 | 266.34 | 6.28  | -0.75             | 0.22              | 1.24  | -0.22  | -0.15  |
| 对叔丁基苯酚    | 3.56    | 4.70 | -248.7 | -1646.77 | -6972.34  | 5325.56 | 8.65 | 124.14 | 7.68  | -8.65             | 0.31              | 2.41  | -0.25  | -0.18  |
| 苯酚        | 1.48    | 2.84 | -93.43 | -1170.94 | -4456.22  | 3285.28 | 9.11 | 94.11  | 5.66  | -9.11             | 0.40              | 1.23  | -0.38  | -0.21  |
| 对羟基苯甲酸    | 1.56    | 3.49 | -206.9 | -1794.27 | -9180.74  | 7386.47 | 8.91 | 150.22 | 10.25 | -8.91             | 0.43              | 1.30  | -0.25  | -0.21  |
| 4-甲氧基苯酚   | 1.57    | 3.46 | -248.7 | -1646.77 | -6972.34  | 5325.56 | 8.65 | 124.14 | 7.68  | -8.65             | 0.31              | 2.41  | -0.25  | -0.25  |
| 对甲基苯酚     | 1.97    | 3.31 | -125.0 | -1326.81 | -5560.10  | 4233.29 | 8.88 | 108.14 | 6.67  | -8.88             | 0.43              | 1.36  | -0.25  | -0.17  |
| 2-萘酚      | 2.65    | 4.53 | -15.9  | -1710.36 | -8253.07  | 6542.71 | 8.64 | 144.17 | 7.75  | -0.35             | 0.14              | 0.98  | -0.25  | -0.22  |

| 酚         | $a$   | $pK_a$ | ${}^0\chi_p$ | ${}^1\chi_p$ | ${}^2\chi_p$ | ${}^3\chi_p$ | ${}^4\chi_p$ | ${}^5\chi_p$ | ${}^6\chi_p$ | ${}^3\chi_c$ | ${}^4\chi_{pc}$ | ${}^5\chi_{pc}$ | ${}^6\chi_{pc}$ |
|-----------|-------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------|-----------------|-----------------|
| 2,4-二氯酚   | 68.74 | 7.85   | 5.94         | 3.09         | 2.44         | 1.45         | 0.89         | 0.46         | 0.18         | 0.42         | 0.60            | 0.68            | 0.68            |
| 对苯二酚      | 59.49 | 9.91   | 4.20         | 2.27         | 1.52         | 0.85         | 0.45         | 0.29         | 0.05         | 0.15         | 0.17            | 0.14            | 0.15            |
| 对硝基苯酚     | 71.59 | 7.15   | 4.98         | 2.58         | 1.73         | 1.03         | 0.56         | 0.34         | 0.10         | 0.18         | 0.26            | 0.23            | 0.22            |
| 对溴苯酚      | 64.29 | 9.34   | 5.72         | 3.03         | 2.39         | 1.35         | 0.71         | 0.54         | 0.13         | 0.40         | 0.46            | 0.37            | 0.46            |
| 邻氨基苯酚     | 62.58 | 9.28   | 4.33         | 2.34         | 1.56         | 0.94         | 0.49         | 0.27         | 0.06         | 0.15         | 0.25            | 0.21            | 0.18            |
| 对氨基苯酚     | 64.21 | 8.50   | 4.33         | 2.33         | 1.59         | 0.89         | 0.47         | 0.31         | 0.06         | 0.17         | 0.20            | 0.16            | 0.17            |
| 对氯苯酚      | 61.51 | 9.43   | 4.89         | 2.61         | 1.91         | 1.08         | 0.57         | 0.40         | 0.09         | 0.26         | 0.30            | 0.24            | 0.29            |
| 邻硝基苯酚     | 70.63 | 7.22   | 4.98         | 2.59         | 1.70         | 1.04         | 0.60         | 0.32         | 0.08         | 0.16         | 0.28            | 0.31            | 0.27            |
| 间氨基苯酚     | 63.25 | 9.83   | 4.33         | 2.33         | 1.60         | 0.87         | 0.52         | 0.27         | 0.07         | 0.17         | 0.18            | 0.20            | 0.16            |
| 2,3-二甲基苯酚 | 72.53 | 10.5   | 5.68         | 2.97         | 2.22         | 1.58         | 0.74         | 0.37         | 0.14         | 0.33         | 0.76            | 0.77            | 0.67            |
| 间甲基苯酚     | 63.49 | 10.0   | 4.76         | 2.55         | 1.84         | 1.00         | 0.63         | 0.31         | 0.10         | 0.24         | 0.26            | 0.29            | 0.24            |
| 邻甲基苯酚     | 63.21 | 10.3   | 4.76         | 2.55         | 1.79         | 1.12         | 0.56         | 0.32         | 0.08         | 0.21         | 0.37            | 0.32            | 0.27            |
| 2-甲氧基苯酚   | 69.17 | 9.99   | 5.16         | 2.66         | 1.67         | 1.09         | 0.63         | 0.33         | 0.09         | 0.12         | 0.25            | 0.29            | 0.26            |
| 1-萘酚      | 94.64 | 9.30   | 5.99         | 3.55         | 2.50         | 1.77         | 1.21         | 0.78         | 0.33         | 0.22         | 0.49            | 0.69            | 0.82            |
| 邻苯二酚      | 58.74 | 9.36   | 4.20         | 2.27         | 1.49         | 0.88         | 0.46         | 0.26         | 0.05         | 0.13         | 0.21            | 0.18            | 0.15            |
| 1,2,4-苯三酚 | 64.39 | 11.82  | 4.57         | 2.41         | 1.67         | 0.96         | 0.51         | 0.29         | 0.07         | 0.20         | 0.28            | 0.26            | 0.25            |
| 五氯苯酚      | 92.32 | 4.75   | 9.11         | 4.55         | 3.80         | 3.44         | 1.72         | 0.86         | 0.38         | 0.76         | 2.30            | 3.25            | 3.91            |
| 对叔丁基苯酚    | 70.16 | 10.39  | 6.88         | 4.11         | 2.77         | 1.76         | 1.04         | 0.74         | 0.33         | 0.19         | 0.31            | 0.33            | 0.42            |
| 苯酚        | 53.88 | 10     | 3.83         | 2.13         | 1.34         | 0.76         | 0.43         | 0.24         | 0.03         | 0.07         | 0.09            | 0.07            | 0.06            |
| 对羟基苯甲酸    | 88.94 | 4.75   | 5.11         | 2.72         | 1.85         | 1.11         | 0.60         | 0.37         | 0.13         | 0.20         | 0.31            | 0.27            | 0.27            |
| 4-甲氧基苯酚   | 70.16 | 10.4   | 5.16         | 2.66         | 1.70         | 1.07         | 0.58         | 0.35         | 0.12         | 0.14         | 0.23            | 0.21            | 0.21            |
| 对甲基苯酚     | 63.99 | 10.26  | 4.76         | 2.55         | 1.84         | 1.03         | 0.54         | 0.38         | 0.08         | 0.24         | 0.28            | 0.22            | 0.26            |
| 2-萘酚      | 95.36 | 9.51   | 5.99         | 3.54         | 2.53         | 1.73         | 1.18         | 0.80         | 0.34         | 0.24         | 0.46            | 0.61            | 0.74            |

分别采用 GA-PLS 和 GA-ANN 两种方法建立 QSAR 模型, GA 的参数为:初始种群数 80, 交叉系数 0.5, 变异系数 0.1, 迭代次数 500. PLS 程序中的交互验证系数大于 0.5 即可, 该值越大说明模型的预测能力和稳定性越高. ANN 采用 BP 神经网络, 输入层节点数由 GA 筛选的因子个数决定, 输出层节点数为 1, 隐含层节点数由程序确定, 输入函数采用 tansig, 输出函数采用 pureline, 训练函数采用 traingdm.

## 2 结果与讨论

### 2.1 臭氧与酚直接反应速率

在臭氧氧化降解酚的实验中, 分析水中臭氧浓度, 发现其浓度基本维持不变, 而酚的浓度迅速降低. 实验中关掉臭氧发生器, 通入纯氧气开展空白对照实验(以 1-萘酚、2,4-二氯酚、对溴酚、2,3-二甲基酚为对象), 发现酚的浓度几乎没有变化, 说明吹脱对水中酚去除的影响几乎可以忽略不计. 同时, 反应在中性 pH 和非常低的碳酸盐浓度下进行, 自由基氧化对水中有机物去除的贡献非常小. 因此, 水溶液中的酚主要是通过臭氧与酚的直接氧化反应去除的.

由于反应过程中臭氧浓度保持不变, 臭氧与酚的直接氧化反应速率可用伪一级反应速率方程表示<sup>[3]</sup>. 作酚浓度对数值与时间关系图可获得一条直线, 斜率即为臭氧氧化降解酚的表观速率常数<sup>[3]</sup>. 采用这种方法获得了 23 种酚臭氧氧化降解的表观速率常数, 通过臭氧氧化降解酚的表观速率常数和水中溶解性臭氧浓度(基本维持不变), 可以计算臭氧与酚直接氧化反应的绝对速率常数, 其对数值见表 2.

表 2 QSAR 模型的  $\lg k$  预测结果比较

Table 2 Comparison of predicted values of  $\lg k$  by QSAR models

| 酚          | 实验值<br>( $\lg k$ ) | 模型预测   |       |        |       | 留一法交叉验证 |       |        |       |
|------------|--------------------|--------|-------|--------|-------|---------|-------|--------|-------|
|            |                    | GA-PLS |       | GA-ANN |       | GA-PLS  |       | GA-ANN |       |
|            |                    | 预测值    | 残差    | 预测值    | 残差    | 预测值     | 残差    | 预测值    | 残差    |
| 2,4-二氯酚    | 4.02               | 3.99   | -0.03 | 3.91   | 0.11  | 3.96    | -0.06 | 3.87   | 0.15  |
| 对苯二酚       | 4.21               | 4.54   | 0.33  | 4.02   | 0.19  | 4.47    | 0.26  | 4.31   | -0.10 |
| 对硝基苯酚      | 3.84               | 4.02   | 0.18  | 3.91   | -0.07 | 4.03    | 0.19  | 3.84   | 0.00  |
| 对溴苯酚       | 3.91               | 4.20   | 0.29  | 3.90   | 0.01  | 4.25    | 0.34  | 4.02   | -0.11 |
| 邻氨基苯酚      | 4.22               | 4.53   | 0.31  | 4.05   | 0.17  | 4.47    | 0.25  | 4.30   | -0.08 |
| 对氨基苯酚      | 4.23               | 4.50   | 0.27  | 4.07   | 0.16  | 4.43    | 0.20  | 4.25   | -0.02 |
| 对氯苯酚       | 3.96               | 4.19   | 0.23  | 3.82   | 0.14  | 4.16    | 0.20  | 4.00   | -0.04 |
| 邻硝基苯酚      | 3.81               | 4.04   | 0.23  | 3.95   | -0.14 | 4.04    | 0.23  | 3.85   | -0.04 |
| 间氨基苯酚      | 4.23               | 4.66   | 0.43  | 4.09   | 0.14  | 4.61    | 0.38  | 4.44   | -0.21 |
| 2,3-二甲基苯酚* | 4.39               | 4.50   | 0.11  | 4.28   | 0.11  | 4.54    | 0.15  | 4.36   | 0.03  |
| 间甲基苯酚*     | 4.12               | 4.35   | 0.23  | 3.94   | 0.18  | 4.26    | 0.14  | 4.08   | 0.04  |
| 邻甲基苯酚*     | 4.23               | 4.43   | 0.20  | 4.00   | 0.23  | 4.33    | 0.10  | 4.16   | 0.07  |
| 2-甲氧基苯酚*   | 4.24               | 4.48   | 0.24  | 4.17   | 0.07  | 4.40    | 0.16  | 4.24   | 0.00  |
| 1-萘酚       | 3.97               | 4.20   | 0.23  | 4.13   | -0.16 | 4.19    | 0.22  | 4.13   | -0.16 |
| 邻苯二酚       | 4.09               | 4.48   | 0.39  | 3.89   | 0.20  | 4.40    | 0.31  | 4.23   | -0.14 |
| 1,2,4-苯三酚* | 4.82               | 4.94   | 0.12  | 4.67   | 0.15  | 5.26    | 0.44  | 5.01   | -0.19 |
| 五氯苯酚*      | 4.87               | 3.85   | -1.02 | 4.68   | 0.19  | 3.91    | -0.96 | 4.21   | 0.66  |
| 对叔丁基苯酚     | 3.61               | 4.08   | 0.47  | 3.78   | -0.17 | 4.61    | 1.00  | 4.08   | -0.47 |
| 苯酚         | 3.99               | 4.39   | 0.40  | 3.80   | 0.19  | 4.35    | 0.36  | 4.19   | -0.20 |
| 对羟基苯甲酸     | 3.87               | 3.81   | -0.06 | 4.10   | -0.23 | 3.85    | -0.02 | 3.55   | 0.32  |
| 4-甲氧基苯酚    | 4.25               | 4.47   | 0.22  | 4.20   | 0.05  | 4.37    | 0.12  | 4.20   | 0.05  |
| 对甲基苯酚      | 4.09               | 4.38   | 0.29  | 3.97   | 0.12  | 4.30    | 0.21  | 4.13   | -0.04 |
| 2-萘酚       | 4.40               | 4.22   | -0.18 | 4.56   | -0.16 | 4.11    | -0.29 | 3.85   | 0.55  |

注: \* 表示溶解性臭氧浓度维持在  $6.67 \mu\text{mol} \cdot \text{L}^{-1}$ , 其余实验溶解性臭氧浓度维持在  $20 \mu\text{mol} \cdot \text{L}^{-1}$ .

对比臭氧与酚直接氧化反应的绝对速率常数, 可以发现苯环上的酚羟基越多, 直接氧化反应的速

率越快,即反应活性顺序为 1,2,4-苯三酚>对苯二酚>苯酚.另外,含斥电子基团的酚臭氧氧化速率高于含吸电子基团的酚,即对于含有不同取代基的酚,其臭氧氧化速率顺序为:对氨基苯酚>对苯二酚>对甲苯酚>对氯苯酚>对溴苯酚>对硝基苯酚.可见,取代基的性质对臭氧与酚之间的直接反应速率有较大影响,这主要是因为:芳香化合物的臭氧氧化主要是通过亲电取代反应进行,反应的第一步是形成碳正离子,第二步是碱性基团从亲核取代位上获得一个质子,第一步的反应速率较慢,是取代反应速率的限速步骤<sup>[3]</sup>;碳正离子的稳定性越高,亲电加成反应的速率越大;—R、—OH 和—NH<sub>2</sub> 倾向于释放电子,提高碳正离子的稳定性,从而提高苯环反应活性;相反,—Cl、—Br 和—NO<sub>2</sub> 倾向于获得电子,降低碳正离子的稳定性,从而降低苯环反应活性;由于—OH 具有较强的给电子倾向,因此尽管苯环很稳定,臭氧氧化酚的速率仍较快,而且—OH 数目越多,氧化速率越快.因此,苯环上的电子云分布对酚的臭氧氧化速率有较大的影响.

## 2.2 基于 GA-PLS 算法的模型

选取 23 种酚,以酚与臭氧直接氧化反应的速率常数对数值为因变量,以 23 种性质参数为自变量,采用 GA-PLS 算法建立臭氧与酚之间直接氧化反应速率常数的 QSAR 模型.由于多种酚的  $qO^-$  值相差非常小,因此在零值检验中剔除变量  $qO^-$ .基于 GA-PLS 算法建立的 QSAR 模型为: $\lg k = 3.439 - 0.206 \lg P + 0.122 pK_a + 0.346 \chi_{pc} - 0.0236 qC^-$ , 累积有效性系数  $Q^2_{cm} = 0.587$ , 大于 PLS 要求的  $Q^2$  最小值(0.5).模型对 23 种酚的  $\lg k$  预测值列于表 2.计算得到 GA-PLS 模型的  $R^2$ 、SSE 和 RMSE 分别为 0.904、2.67 和 0.348,结果表明预测值与实测值相关性显著,模型预测能力较强.

## 2.3 基于 GA-ANN 算法的模型

选取 23 种酚,以酚与臭氧直接反应的速率常数对数值为因变量,以 23 种性质参数为自变量,采用 GA-ANN 算法建立臭氧与酚之间直接氧化反应速率常数的 QSAR 模型.跟 GA-PLA 建模类似,在零值检验过程中剔除变量  $qO^-$ .然后进行自变量间的相关性分析,得到酚的性质参数之间的相关性如表 3,相关度大于 0.9 的两个变量剔除其中之一,共剔除了  $EE$ 、 $\chi_p$ 、 ${}^3\chi_p$ 、 ${}^4\chi_p$ 、 ${}^6\chi_p$ 、 ${}^6\chi_{pc}$ .神经网络模型是非线性模型,不能获得明确的数学表达式,基于 GA-ANN 算法建立的 QSAR 模型含有  $\lg P$ 、 ${}^4\chi_{pc}$ 、 $pK_a$  和  $\alpha$  等 4 个性质参数,输入层节点数为 4,隐含层节点数也为 4,输出层节点数为 1,阈值和权重如表 4.模型对 23 种酚的  $\lg k$  预测值列于表 2.计算得到 GA-ANN 模型的  $R^2$ 、SSE 和 RMSE 分别为 0.955、2.18 和 0.315,结果表明预测值与实测值相关性显著,模型预测能力较强.

## 2.4 模型验证和解析

本研究采用留一法(leave-one-out)交叉验证来检验和比较 QSAR 模型的稳健性.分别利用 GA-PLS 和 GA-ANN 算法对每一种酚的  $\lg k$  值进行预测,交叉验证的预测结果列于表 2 中,并绘制预测值与实测值的相关图,如图 1 所示.图 1(a)是由 GA-PLS 建立模型的  $\lg k$  预测值与实测值的相关图,该图显示除了反应速率较快和较慢的酚,其他酚的臭氧氧化速率常数预测值与实测值间的相关性显著,误差较小.同时该图还显示模型的预测值总体偏高.图 1(b)是由 GA-ANN 建立模型的  $\lg k$  预测值与实测值的相关图,该图显示所有酚的臭氧氧化速率常数预测值与实测值间的相关性显著,误差较小.

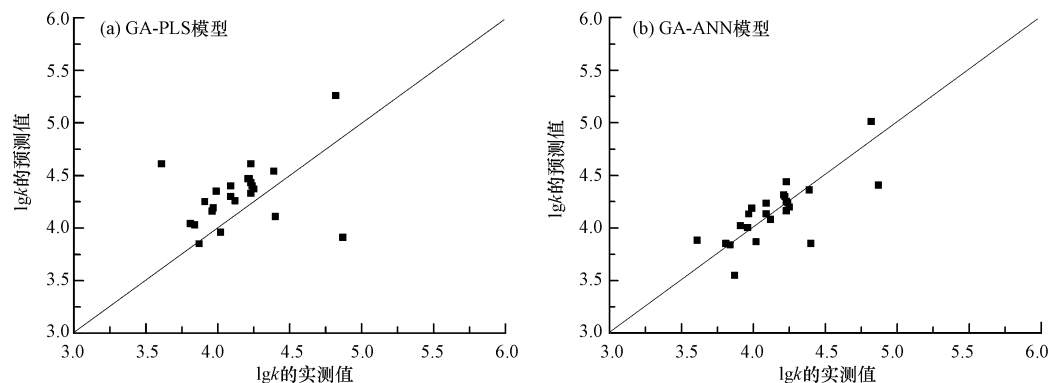


图 1 速率常数实验值与模型预测值之间的关系

Fig.1 Comparison of the measured rate constants and the predicted rate constants from the QSAR models



表 4 基于 GA-ANN 算法建立的 QSAR 模型的参数

Table 4 Parameters for the standardized QSAR model based on GA-ANN method

|         | 阈值    | 权重      |          |                        |                        |
|---------|-------|---------|----------|------------------------|------------------------|
|         |       | $\lg P$ | $\alpha$ | $\text{p}K_{\text{a}}$ | ${}^6\chi_{\text{pc}}$ |
| 输入层到隐含层 | -1.98 | 0.01    | -1.56    | 1.08                   | 0.55                   |
|         | 0.66  | -0.11   | 1.59     | 1.18                   | 0.07                   |
|         | -0.66 | -1.28   | 0.06     | 0.64                   | 1.37                   |
|         | 1.97  | -1.37   | -1.19    | 0.51                   | -1.4                   |
| 隐含层到输出层 | -0.83 | 0.60    | -0.09    | -0.14                  | 0.65                   |

对比图 1 (a)和图 1 (b)可以看出,由 GA-ANN 建立的 QSAR 模型实测值和预测值相关性更好.另外,GA-ANN 模型的决定系数、SSE 和 RMSE 也比 GA-PLS 更优,可见 GA-ANN 模型的稳健性相对较好.GA-ANN 通过不断提取有用信息、调整不同层与节点的权重和阈值,获得更好的拟合效果,是一种非线性方法<sup>[25]</sup>.通常,一种性质参数会与酚的多种性质有关联,不同性质对酚的降解影响不同,因而一些性质参数与反应速率之间的关系是非线性的,因此,GA-ANN 模型的稳健性强于 GA-PLS 模型,但 GA-ANN 模型不能给出明确的回归关系式,且收敛慢<sup>[26]</sup>.

Hu<sup>[12]</sup>等人提出臭氧氧化有机物的速率与  $E_{\text{homo}}$  和电负性 (Absolute electronegativity,  $\chi$ ) 密切相关.Liu<sup>[13]</sup>等人提出臭氧氧化酚的速率与  $E_{\text{lumo}}$ 、 $\alpha$ 、 $q^-$  (The most negative atomic partial charge in molecule) 和标准熵 (Standard entropy,  $S^{\theta}$ ) 密切相关.Sudhakaran 和 Amy<sup>[8]</sup>提出臭氧氧化有机物的速率与双键当量数 (Double bond equivalence, DBE)、IP 和电子亲和力 (Electron-affinity, EA) 密切相关.这些文献中提出的模型参数主要包含两类,一类与电子云分布相关 (如  $\chi$ 、 $\alpha$ 、 $q^-$ 、DBE 和 IP), 另一类与分子轨道能相关 (如  $E_{\text{homo}}$  和  $E_{\text{lumo}}$ ).电子云分布影响亲电加成反应第一步生成的碳正离子的稳定性,可反映分子得失电子的能力,分子轨道能也可表征酚类物质得失电子的能力,而酚与臭氧直接氧化反应的典型反应机制就是亲电加成反应,因此,这两类模型参数必然与酚的臭氧氧化速率密切相关.在本研究中,基于 GA-PLS 算法建立的 QSAR 模型含有算符  $qC^-$ , 基于 GA-ANN 算法建立的 QSAR 模型含有算符  $\alpha$ , 这两个参数均与电子云分布直接相关,因此它们与酚的臭氧氧化速率密切相关.

在本研究中,基于 GA-PLS 算法建立的模型和基于 GA-ANN 算法建立的模型均含有同样的算符  ${}^4\chi_{\text{pc}}$ 、 $\text{p}K_{\text{a}}$  和  $\lg P$ .  ${}^4\chi_{\text{pc}}$  为四阶路径/簇分子连接性指数,与苯环取代基的位置、长度和类型有关,前面的讨论中也证实了苯环上取代基的性质对臭氧氧化酚的速率有着明显的影响.Sudhakaran 和 Amy 研究的 40 种有机物 (除了含有 6 种酚,还有卤代烃、醇、酮、羧酸等有机物),种类较多,有机物与臭氧氧化反应速率的变化范围较宽;本研究所涉及的化合物都是酚,有机物种类较单一,酚与臭氧之间的反应速率变化范围较窄<sup>[8]</sup>.因此,Sudhakaran 和 Amy 的研究中,官能团的不同 (例如双键当量数) 是导致有机污染物与臭氧反应速率存在较大不同的主要原因;而本研究中,苯环取代基种类、位置和性质不同是反应速率常数存在差别的主要原因.这也是本研究最终建模中含有参数  ${}^4\chi_{\text{pc}}$  的原因.

$\text{p}K_{\text{a}}$  与物质在水溶液中的电离程度有关.与文献 [8] 和 [12] 中研究的大多数有机物明显不同的是,酚在水溶液具有较大电离度,不少文献中均报导有机物在水中电离会影响其与臭氧之间的反应速率<sup>[4,26]</sup>, Qiu 等<sup>[26]</sup> 在研究完全混合式搅拌反应器中臭氧吸收和臭氧氧化二氯苯酚模型时,提出水中二氯苯酚与臭氧之间的反应速率为二氯苯酚分子的氧化速率与二氯苯酚离子的氧化速率之和,二氯苯酚离子的臭氧氧化不可忽略.酚电离后会形成负离子,苯环上电子云分布发生明显改变,亲电反应特性也会发生明显改变,这种改变的实质是酚溶解后,在水化作用下发生电子云分布改变,其得失电子的能力也随着改变,必然会导致臭氧氧化速率发生改变.因此,  $\text{p}K_{\text{a}}$  对臭氧氧化酚的速率存在显著影响,建立的 QSAR 模型中含有参数  $\text{p}K_{\text{a}}$  是合理的.

$\lg P$  为辛醇-水分配系数对数值,辛醇-水分配系数与溶解度有关,而溶解度受溶剂 (水) 和溶质 (酚) 之间的相互作用影响.Sudhakaran 和 Amy 建立的臭氧氧化有机物 QSAR 模型中虽然没有表征溶剂对溶质作用的参数,但是,他们建立的羟基自由基氧化有机物 QSAR 模型显示溶剂可及表面积弱极性组分与反应速率密切相关<sup>[14]</sup>.可见,溶剂化作用会对臭氧氧化酚的速率产生显著影响.

### 3 结论

(1) 在臭氧浓度过量的条件下, 酚的臭氧氧化速率较快, 含斥电子基团的酚与臭氧之间反应的速率高于含吸电子基团的酚。

(2) 臭氧氧化降解酚的 QSAR 模型显示, 反应速率不仅与电子云分布以及苯环上取代基性质密切相关, 水溶液中溶剂化作用也会对臭氧氧化降解酚的速率产生显著影响。

(3) 基于 GA-PLS 算法建立的 QSAR 模型具有直观的方程式, 基于 GA-ANN 算法建立的 QSAR 模型的稳健性更好。

### 参 考 文 献

- [ 1 ] FOCKEDEY E, VAN LIERDE A. Coupling of anodic and cathodic reactions for phenol electro-oxidation using three-dimensional electrodes [J]. *Water Research*, 2002, 36(16):4169-4175
- [ 2 ] HAN W, ZHU W, ZHANG P, et al. Photocatalytic degradation of phenols in aqueous solution under irradiation of 254 and 185nm UV light [J]. *Catalysis Today*, 2004, 90(3-4):319-324
- [ 3 ] FERNANDO J B, BELTRÁN F J. Ozone reaction kinetics for water and wastewater systems[M]. New York: Lewis Publishers, 2004:6-12
- [ 4 ] HOIGNÉ J, BADER H. Rate constants of reactions of ozone with organic and inorganic compounds in water-II: Dissociating organic compounds[J]. *Water Research*, 1983, 17(2):185-194
- [ 5 ] HOIGNÉ J, BADER H. Rate constants of reactions of ozone with organic and inorganic compounds in water-I: Non-dissociating organic compounds[J]. *Water Research*, 1983, 17(2):173-183
- [ 6 ] HU S T, YU Y H. Preozonation of chlorophenolic wastewater for subsequent biological treatment[J]. *Ozone Science & Engineering*, 1994, 16(1):13-28
- [ 7 ] KUO C H, HUANG C H. Aqueous phase ozonation of chlorophenols[J]. *Journal of Hazardous Materials*, 1995, 41(1):31-45
- [ 8 ] SUDHAKARAN S, AMY G L. QSAR models for oxidation of organic micropollutants in water based on ozone and hydroxyl radical rate constants and their chemical classification[J]. *Water Research*, 2013, 47(3):1111-1122
- [ 9 ] HUANG J, YU G, YANG X, et al. Predicting physico-chemical properties of polychlorinated diphenyl ethers (PCDEs): Potential organic pollutants (POPs)[J]. *Journal of Environmental Science*, 2004, 16(2):204-207
- [ 10 ] KUŠIĆ H, RASULEV B, LESZCZYŃSKA D, et al. Prediction of rate constants for radical degradation of aromatic pollutants in water matrix: A QSAR study[J]. *Chemosphere*, 2009, 75(8):128-134
- [ 11 ] LEI H, SNYDER S A. 3D QSPR models for the removal of trace organic contaminants by ozone and free chlorine[J]. *Water Research*, 2007, 41(18):4051-4060
- [ 12 ] HU J Y, MORITA, MAGARA Y, et al. Evaluation of reactivity of pesticides with ozone in water using the energies of frontier molecular orbitals[J]. *Water Research*, 2000, 34(8):2215-2222
- [ 13 ] LIU H, TAN J, YU H X, et al. Determination of the apparent reaction rate constants for ozone degradation of substituted phenols and QSPR/QSAR analysis[J]. *International Journal of Environmental Research*, 2010 4(3):507-512
- [ 14 ] LI T, MEI H, CONG P. Combining nonlinear PLS with the numeric genetic algorithm for QSAR [J]. *Chemometrics and Intelligent Laboratory Systems*, 1999, 45(1/2):177-184
- [ 15 ] DAREN Z. QSPR studies of PCBs by the combination of genetic algorithms and PLS analysis[J]. *Computers & Chemistry*, 2001, 25(2):197-204
- [ 16 ] GUPTA V K, KHANI H, AHMADI-ROUDI B, et al. Prediction of capillary gas chromatographic retention times of fatty acid methyl esters in human blood using MLR, PLS and back-propagation artificial neural networks[J]. *Talanta*, 2001, 83(3):1014-1022
- [ 17 ] SAVORY N, ABE K, SODE K, et al. Selection of DNA aptamer against prostate specific antigen using a genetic algorithm and application to sensing[J]. *Biosensors & Bioelectronics*, 2010, 26(4):1386-1391
- [ 18 ] TONG D L, SCHIERZ A C. Hybrid genetic algorithm-neural network: Feature extraction for unprocessed microarray data[J]. *Artificial Intelligence Review*, 2011, 53(1):47-56
- [ 19 ] WANG J, ZHOU Y, ZHU W, et al. Catalytic ozonation of dimethyl phthalate and chlorination disinfection by-product precursors over Ru/AC[J]. *Journal of Hazardous Materials*, 2009, 166(1):502-507
- [ 20 ] BADER H, Hoigné J. Determination of ozone in water by the indigo method[J]. *Water Research*, 1981, 15(4):449-456
- [ 21 ] BAHNICK D A, DOUCETTE W J. Use of molecular connectivity indices to estimate soil sorption coefficients for organic chemicals[J]. *Chemosphere*, 1988, 17(9):1703-1715
- [ 22 ] KIER L B, MURRAY W J, Hall L H. Molecular connectivity. 4. Relations to biological activities[J], *Journal of Medicinal Chemistry*, 1975, 18(12):1272-1274
- [ 23 ] NIU J, YU G. Molecular structural characteristics governing biocatalytic chlorination of PAHs by chloroperoxidase from *Caldariomyces fumago*[J]. *SAR and QSAR in Environmental Research*, 2004, 15(3):159-167
- [ 24 ] NIU J, HUANG L, CHEN J, et al. Quantitative structure-property relationships on photolysis of PCDD/Fs adsorbed to spruce (*Picea abies* (L) Karst) needle surfaces under sunlight irradiation[J]. *Chemosphere*, 2005, 58(7):917-924
- [ 25 ] YAZDANMEHR M, ANJIDAN S H, BAHRAMI A. Using GA-ANN algorithm to optimize soft magnetic properties of nanocrystalline mechanically alloyed Fe-Si powders[J]. *Computational Materials Science*, 2009, 44(4):1218-1221
- [ 26 ] QIU Y Q, KUO C H, ZAPPI M E. Performance and simulation of ozone absorption and reactions in a stirred-tank reactor [J]. *Environmental Science & Technology*, 2001, 35(1):209-215