

BP 人工神经网络法研究 PBDEs 类化合物的 RRT

曹红翠,孙海霞,保英莲

(青海大学 化工学院,西宁 810016)

摘要:运用 BP 人工神经网络方法对 PBDEs 的相对保留时间(RRT)进行了 QSPR 研究.所建的 BP 人工神经网络对 PBDEs 的 RRT 预测准确度非常高,网络训练误差几乎为 0,网络回判 MSE 误差为 0.003 9,明显低于逐步回归分析结果,独立检测集 MSE 误差为 0.000 4,也很低,说明 BP 人工神经网络具有较好的泛化能力.此方法得到的模型预测能力要优于逐步回归模型.

关键词:多溴联苯醚(PBDEs);BP 人工神经网络;RRT

中图分类号:O641.12

文献标志码:A

多溴联苯醚(Polybrominated Diphenyl Ethers, PBDEs)在环境中难以进行化学、生物以及光的降解,滞留时间长,具有强亲脂性和憎水性,可以通过生物链进行富集和放大,给人类的健康带来严重威胁,也可通过“蚱蜢跳效应”广域迁移,导致全球污染^[1-4]. PBDEs 与物质的结合不是通过化学键,故很容易在产品的生产、使用以及废物处理的过程中释放出来,造成污染环境并且危害人体健康.目前 PBDEs 几乎无处不在,即使在遥远北极的哺乳动物体内也已检测出 PBDEs^[5-10].

PBDEs 的通式为 $C_{12}H_{10}X_oBr_x$ ($x=1-10$),根据溴原子的取代数目和位置的不同, PBDEs 存在 209 种可能的结构,但目前只对部分 PBDEs 的保留时间(RRT)进行了测定和研究.因此对其进行分子结构性质关系(QSPR)的研究和预测对于评价 PBDEs 的环境行为和归宿有着非常重要的意义.金皓、刘宗峰、孟范平、许惠英等作者主要以拓扑法、分子表面静电势和溴原子取代个数对其建模研究的.本文选取 PBDEs 类化合物的保留时间(RRT)利用量子化学参数进行定量结构性质相关(QSPR)研究,为这种持久性有机污染物在环境中的风险评价提供基础数据支持.

1 人工神经网络原理与方法

1.1 人工神经网络原理

人工神经网络技术是以生物神经网络为基础^[11],模拟人脑行为的一种信息处理方法,是一个由人工神经元连接而成,以有向图为拓扑结构的动态系统,并通过对输入信息做出的反应而完成信息处理.在预测领域中应用最广泛的是 BP 神经网络^[12],BP 网络的学习是一种误差反向传播式的网络权值训练方法,多层神经网络模型是它的理论基础.

1.2 BP 网络变量的选择和网络训练及预测仿真

本文研究所用的 PBDEs 部分 RRT 的实验数据均取自文献^[13].采用 ChemOffice 8.0 中 MOPAC-AM1 量子化学法,优化计算了 124 个 PBDEs 分子的 12 个量子化学参数.利用逐步回归分析进行建模,在进行逐步回归分析中只有 Elce 量子化学参数进入方程.其余的量子化学参数对 RRT 都不具有显著性,予以剔除.

收稿日期:2014-09-28;修回日期:2014-12-23.

基金项目:教育部“春晖计划”(Z2012085)

作者简介(通信作者):曹红翠(1975-),女,江苏南京人,青海大学副教授,研究方向为计算化学,E-mail:caohongcui_1975@163.com.

采用上述变量参数进行建模的方程见表 1,

表 1 PCBES 类化合物 RRT 的 QSPR 方程及拟合参数

性质	回归方程 (样品数 $n=124$)	拟合参数			
		R	SE	F	Sig
RRT	$RRT = -0.727 - 6.1 \times 10^{-5} Elce$	0.972	0.048	2 553.231	0.000

这个方程具有较高的 R 值、 F 值、较小的 SE 和置信水平 Sig , 但自变量除了分子总能 (TE) 其余参数均没有通过 t 检验, 由此说明自变量与应变量之间不具有典型的线性关系, 采用逐步回归分析方法是不合适的, 因此改用 BP 神经网络来研究 RRT 的影响因素。

输入变量的选择是神经网络建模的前提和基础, 要求各输入变量间互不相关或是相关性很小, 在数据分析前, 对 12 个参数进行筛选, 筛选原则是使预测集的预测误差最小, 并且在相同的网络训练参数下进行, 经筛选, 确定偶极矩 (μ)、分子生成热 (Hof)、分子总能 (TE)、最负的氧原子净电荷 (q_o)、最正的氢原子净电荷 (q_H)、最正的溴原子净电荷 (q_B)、排斥能 (NRE) 等为 BP 网络的输入变量。

隐含层的结点选取是神经网络中的一个重要问题, 隐含层结点过少, 则网络中权重不充分; 隐含层结点过多, 则会产生过拟合。本文以选取的量化参数作为输入变量, 训练集样本 115, 独立检测集样本 9, 最终确定网络结构为 7-10-1, 隐含层传递函数为 \logsig 函数; 输出层传递函数为 $purelin$ 函数; 学习训练算法为 L-M 优化算法, 网络训练参数: 最大迭代次数为 6 000 次, 网络误差参数为 $goal=0.001$ 。BP 神经网络方法通过 Matlab7.1 函数编程完成^[14-15]。

2 BP 网络结果分析与讨论

BP 网络变量筛选结果表明: 偶极矩 (μ)、分子生成热 (Hof)、分子总能 (TE)、最负的氧原电荷 (q_o)、最正的氢原子净电荷 (q_H)、最正的溴原子净电荷 (q_B)、排斥能 (NRE) 对多溴联苯醚类化合物的保留时间 (RRT) 影响显著, 训练集样本的 MSE 为 0.003 9, 保留时间 (RRT) 预测集的实验值和预测值 (表 2), BP 神经网络训练预测值及实验值相关图 (图 1)、残差频率分布图 (图 2)。

表 2 9 个独立预测集 PBDEs 化合物 RRT 的实验值和预测值

序号	取代形式	RRT		
		<i>exp</i>	<i>cal</i>	<i>Res</i>
1	3,3',5,5'-tetrabro	0.289	0.260	0.029
2	3,4,4',5-tetrabro	0.350	0.332	0.018
3	2,2',3,4,4'-pentabro	0.486	0.467	0.019
4	2,2',3,4,5-pentabro	0.462	0.432	0.030
5	2,2',3,4,5'-pentabro	0.460	0.493	-0.033
6	2,2',3,4,6-pentabro	0.423	0.418	0.005
7	2,2',3,4',5'-pentabro	0.457	0.454	0.003
8	2,2',3,4',6'-pentabro	0.414	0.398	0.016
9	2,2',4,4',5-pentabro	0.433	0.436	-0.003

从图中可以看出, BP 神经网络对 PBDEs 化合物的 RRT 具有很强的预测能力, 网络训练误差几乎为 0, 网络回判 MSE 误差为 0.003 9, 明显低于逐步回归分析结果, 独立检测集 MSE 误差为 0.000 4, 也很低, 说明 BP 神经网络具有较好的泛化能力, 预测准确度高。

3 结 论

运用 BP 神经网络对 PBDEs 的保留时间 (RRT) 进行了 QSPR 研究, 所建 BP 神经网络对 PBDEs 的 RRT 预测准确度非常高, 网络训练好后, 回判预测为 0.003 9, 独立检测集误差为 0.000 4, 也很低。预测精度明显高于逐步回归分析结果。

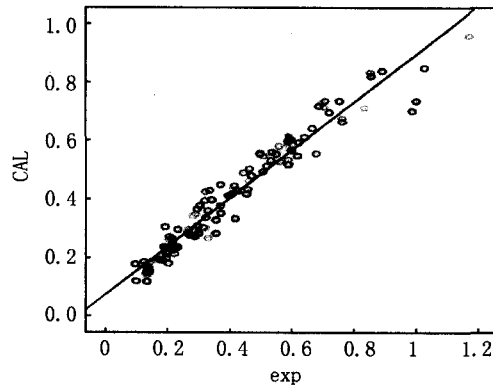


图1 BP神经网络训练预测值及实验值相关图

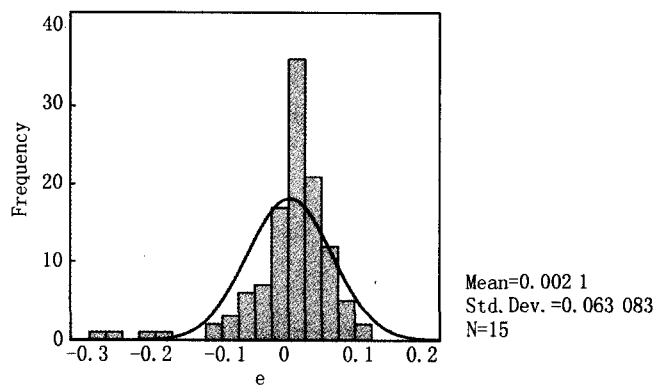


图2 RRT模型的预测残差频率分布图

参 考 文 献

- [1] Muir D C G, Backus S, Derocher A E. Brominated flame retardants in polar bears (*Ursus maritimus*) from Alaska, the Canadian Arctic, East Greenland and Svalbard[J]. *Environmental Science & Technology*, 2006, 40(2): 449-455.
- [2] DE Wit C. An overview of brominated flame retardants in the environment[J]. *Chemosphere*, 2002, 46(5): 583-624.
- [3] Rahman F, Langford K H, Scrimshaw M D. Polybrominated diphenyl ether (PBDE) flame retardants[J]. *Sci Total Environ*, 2001, 275(1/3): 1-17.
- [4] Gouin T, Hamer T. Modelling the environmental fate of the polybrominated diphenyl ethers [J]. *Environ Int*, 2003, 29(6): 717-724.
- [5] Martin M, Paul K S, Richardson B J. An Asian quandary: where have all of the PBDEs gone [J]. *Marine Pollution Bulletin*, 2004, 49(5-6): 375-382.
- [6] 王连生. 有机污染物化学[M]. 北京: 科学出版社, 1991.
- [7] Cynthia A W, Mehaan A, Derek C G. Levels and trends of brominated flame retardants in the Arctic [J]. *Chemosphere*, 2006, 64(2): 209-233.
- [8] DE Boer J, G Wester P, Der Host. Polybrominated diphenyl ethers in influents, suspended particulate matter, sediments, sewage treatment plant and effluents and biota from the Netherlands[J]. *Environmental Pollution*, 2003, 122 (1): 63-74.
- [9] Sellstregard A, DE Wit C. Polybrominated diphenyl ethers and hexabromo-cyclododecane in sediment and fish from a Swedish river [J]. *Environmental Toxicology and Chemistry*, 1998, 17(6): 1065-1072.
- [10] Takumi T, Kurunthachalamr, Hiroaki T. Impact of fermented brown rice with *Aspergillus oryzae* (FEBRA) intake and concentration of polybrominated diphenylethers (PBDEs) in blood of humans from Japan [J]. *Chemosphere*, 2004, 57(8): 795-811.
- [11] 许 禄, 邵学广. 化学计量学方法[M]. 北京: 科学出版社, 2004
- [12] 任国宾, 王静康, 尹秋响. 神经网络在半水盐酸帕罗西汀溶解度预测中的应用[J]. *化工学报*, 2006, 57(4): 853-860.
- [13] 许惠英, 张建英, 王艳花, 等. 多溴联苯醚定量结构-性质关系的分子表面静电势应用研究[J]. *环境科学*, 2008, 29(2): 398-408.
- [14] 许国根, 许萍萍, 谭宪林. MATLAB在化学中的应用[M]. 西安: 西安交通大学出版社, 2005.
- [15] 飞思科技产品研发中心. Matlab6.5辅助神经网络分析与设计[M]. 北京: 电子工业出版社, 2003.

BP Artificial Neural Network Method For QSPR Study on RRT of PBDE_s

CAO Hongcui, SUN haixia, BAO Yinglian

(School of Chemical Engineering, Qinghai University, Xining 810016, China)

Abstract: QSPR studies on RRT of PBDEs were carried out by BP artificial neural network method. The prediction accuracy of the BP artificial neural network on PBDEs RRT was very high. The network training error was almost 0. The network predates MSE error was 0.003 9 which was significantly lower than the results of stepwise regression analysis. Detection of MSE in the independent predictive error was 0.000 4, and it was very low. The results showed the generalization ability by BP artificial neural network was better, and the model constructed by this method had better prediction ability than the model of stepwise regression model.

Keywords: Polybrominated diphenyl ethers (PBDEs); BP artificial neural network; RRT