

GANN模型在十溴联苯醚分散液液微萃取中的应用

王婷¹, 胡艳¹, 翦英红^{2,3}, 李鱼^{1*}

(1.华北电力大学能源与环境研究中心,北京 102206; 2.吉林大学环境与资源学院,吉林 长春 130012;
3.吉林化工学院环境与生物学院,吉林 吉林 132022)

摘要:文章研究目的在于用遗传神经网络模型(GANN模型)快速优化水样中十溴联苯醚分散液液微萃取的萃取条件。以水样中十溴联苯醚分散液液微萃取的正交试验为训练样本,建立十溴联苯醚分散液液微萃取条件的遗传神经网络模型。比较遗传神经网络模型和BP神经网络模型的学习速度、学习精度及网络泛化能力。采用Matlab遗传算法工具箱运用遗传神经网络模型对影响萃取回收率的因素进行优化求解,获得了水样中十溴联苯醚分散液液微萃取优化后的萃取条件,并进行实验验证。文章建立的遗传神经网络模型得到的预测值与实验值平均偏差为14.41%, R^2 为0.8887;最佳DLLME萃取条件为10 μ L四氯乙烯、0.71mL丙酮、pH=5、离子强度为20%NaCl、萃取时间10min;优化后十溴联苯醚分散液液微萃取的萃取回收率和富集因子比优化前分别提高了54%和580。

关键词:分散液液微萃取; 十溴联苯醚; 遗传神经网络

中图分类号:X132 **文献标志码:**A **doi:**10.3969/j.issn.1003-6504.2010.10.004 **文章编号:**1003-6504(2010)10-0015-04

Application of Genetic Algorithm Neural Network Model in Dispersive Liquid-Liquid Microextraction Condition Optimization of Decabrominated Diphenyl Ether

WANG Ting¹, HU Yan¹, JIAN Ying-hong^{2,3}, LI Yu^{1*}

(1.Energy and Environmental Research Centre, North China Electric Power University, Beijing 102206, China;

2.School of Environment and Resource, Jilin University, Changchun 130012, China;

3.School of Environment and Biology, Jilin Institute of Chemistry and Technology, Jilin 132022, China)

Abstract: The study aims to optimize dispersive liquid-liquid microextraction (DLLME) of decabrominated diphenyl ether (BDE-209) in water samples rapidly through genetic algorithm neural network (GANN) model. A GANN model was established based on an orthogonalized experiment of BDE-209 DLLME. Convergence speed, learning precision and generalization were compared between BP neural network model and GANN model. Matlab genetic algorithm toolbox was applied to seek a solution of DLLME optimization, and the optimum DLLME conditions were validated by experiments. Average errors between predicted and experimental values of extraction recovery were less than 14.41%, and R^2 was 0.8887. The optimum DLLME condition of BDE-209 was 10 μ L of tetrachloroethylene, 0.71mL of acetone, pH 5, 20% of NaCl and extraction time of 10min. Results showed that the extraction recovery and enrichment factor of optimum condition had an increase of 54% and 580 respectively through GANN model.

Key words: dispersive liquid-liquid microextraction (DLLME); decabrominated diphenyl ether; genetic algorithm neural network (GANN)

十溴联苯醚(BDE-209)是商用十溴联苯醚的主要组成部分(98%)。BDE-209虽毒性小于其低溴同系物(如五溴联苯醚、四溴联苯醚等),但其在各相介质中可光降解或生物降解为毒性更强的已列为POPs的低溴代联苯醚^[1-3]。分散液液微萃取(DLLME)是近年来建立的新型液液微萃取技术,由于其可以快速、简

单、高效的萃取水样中的有机物,已被用于水样中多种有机污染物的测定。在DLLME的萃取过程中可能影响DLLME萃取效率及富集因子的因素有很多,如萃取剂体积、分散剂体积等,有必要对DLLME的萃取条件进行优化。萃取条件的优化方法主要有单因素试验、正交试验设计、BP神经网络和遗传神经网络

收稿日期:2009-06-22;修回 2009-08-10

基金项目:973 国家重点基础研究发展规划资助项目(2004CB3418501)

作者简介:王婷(1985-),女,博士研究生,研究方向为环境污染与控制化学,(电话)010-51971241(电子信箱)wangting_btt@163.com;* 通讯作者,男,教授,博士,博士生导师,研究方向为环境污染与控制化学、环境污染物形态、环境规划与评价研究,(电话)010-51971241(电子信箱)liyuxx@jlu.edu.cn.

(GANN)模型等^[4-9]。GANN 模型是采用遗传算法与 BP 神经网络学习相结合的网络预测模型,该方法先利用遗传算法在解空间进行全局搜索,然后在遗传算法搜索到的最优区间内用神经网络学习求出最优解^[10]。GANN 模型可以模拟不同水平间距的正交试验,使多因素、多位级的试验效率大大增强,较少的实验次数就能将实验条件或结果优化。相比 BP 神经网络模型,GANN 模型在训练速度及预测精度上有所提高^[10-13],在萃取条件优化方面已有应用^[9]。

本文以液液微萃取水样中十溴联苯醚(BDE-209)正交试验数据为训练样本,建立 BDE-209DLLME 萃取条件的 GANN 模型,并比较 BP 神经网络和 GANN 模型的性能,同时采用遗传算法对影响萃取回收率(ER)的因素进行优化,以期获得水样中 BDE-209 的最佳 DLLME 萃取条件。

1 实验部分

1.1 试剂与仪器

BDE-209 标准品购于 Dr. Ehrenstorfer (Augsburg, 德国)。色谱纯丙酮、乙腈、甲醇、氯苯、四氯乙烯和四氯化碳均购于 J.T. Baker(美国)。氯化钠(分析纯)购于 Aldrich(Sigma-Aldrich, 德国)。去离子水采用 Millipore Direct-Q UV 系统制备(Bedford, MA, 美国)。离心机(SIGMA 2-16PK, Sigma)。10mL 带盖锥形玻璃离心管。

色谱条件:安捷伦 1200 双泵高效液相色谱仪,配备紫外检测器;流动相:甲醇/水(97/3, V/V);流速:1.0mL/min;Zorbax Eclipse XDB-C18 色谱柱(5 μ m, 150mm \times 4.6mm, Agilent);检测波长 226nm。该色谱条件下 BDE-209 的保留时间为 15.9min。

1.2 十溴联苯醚分散液液微萃取

混合分散剂与萃取剂,用 2.00mL 注射器将分散剂与萃取剂的混合液快速注入 5.0mL 含有 100ng/mL BDE-209 的 10mL 锥形玻璃离心管,形成乳浊液。放

置 10min 后,乳浊液离心(6000r/min)10min。用 25 μ L 微型注射器移取离心管中的有机沉积相,注入 5 μ L 至液相色谱中分析 BDE-209 的浓度。

1.3 分散液液微萃取条件的 GANN 模型建立及优化

采用 GANN 模型,以分散液液微萃取影响因素:萃取剂体积(A)、分散剂体积(B)、萃取剂类型(C)、分散剂类型(D)、pH(E)、离子强度(F)及萃取时间(G)为模型的输入条件,以 ER 为输出条件,对 BDE-209 的分散液液微萃取过程建立模型,通过误差分析以及相关性分析验证网络模型的可靠性,并利用 Matlab 遗传算法工具箱对萃取条件进行优化。优化过程中选取分散剂类型为丙酮,将显著性水平相对较低的 pH 和萃取时间设定为固定值,对另外 4 个影响因素进行优化求解。从应用角度确定萃取回收率(Y)取值最高为优化目标,即寻求萃取回收率最高时(即为最优情况下),4 个影响因素萃取剂体积(A)、分散剂体积(B)、萃取剂类型(C)、离子强度(F)最优取值。

2 结果与讨论

2.1 分散液液微萃取正交试验设计

以 BDE-209 DLLME 的富集因子(EF)为指标,选取 DLLME 所涉及的分散剂、分散剂体积、萃取剂、萃取剂体积、pH、离子强度、萃取时间 7 个条件为考察因素,每个因素选取 3 个水平(表 1),同时考察萃取剂体积与分散剂体积的交互作用(A*B)。用 $L_{27}(3^3)$ 正交表进行 DLLME 萃取条件实验,对正交试验结果进行极值分析,分析得出的萃取条件为:20 μ L 四氯化碳、1.0mL 乙腈、溶液 pH=7、离子强度为 5% NaCl、萃取时间 5min。在此萃取条件下,BDE-209 的 ER 为 50.45%,EF 为 226。对正交试验结果进行方差分析,萃取剂体积(A)对 DLLME 的 EF 影响显著,离子强度(F)对 DLLME 的 ER 有显著性影响,而萃取剂体积与分散剂体积的交互作用(A*B)、分散剂(D)、pH(E)和萃取时间(G)对 EF 及 ER 都没有显著性影响。

表 1 BDE-209 分散液液微萃取正交试验因素水平
Table 1 Orthogonal experimental factors and levels of DLLME BDE-209

因素	水平
A	萃取剂体积(μ L)
B	分散剂体积(mL)
C	萃取剂
D	分散剂
E	pH
F	NaCl 含量(%)
G	萃取时间(min)

2.2 分散液液微萃取条件的 GANN 模型建立

以正交设计试验数据为 GANN 模型的训练样本,采用 LM 贝叶斯正则化算法对模型进行训练,设

定目标误差为 0.005,建立 GANN 模型。模型设置输入和输出节点个数分别为 7 和 1,其中输入条件:萃取剂体积(A)、分散剂体积(B)、萃取剂类型(C)、分散剂

类型(D)、pH(E)、离子强度(F)及萃取时间(G)。

EF 和 ER 关系如式(1)所示:

$$ER = \frac{M_{sed}}{M_0} \times 100 = \frac{C_{sed} \times V_{sed}}{C_0 \times V_0} \times 100 = \left(\frac{V_{sed}}{V_0} \right) EF \times 100 \quad (1)$$

式(1)中, M_{sed} 和 M_0 分别为有机沉积相和水样中 BDE-209 含量、 C_{sed} 和 C_0 分别为有机沉积相和水样中 BDE-209 浓度、 V_{sed} 和 V_0 分别为有机沉积相和水样体积。由式(1)可见, EF 和 ER 存在相关性, 因此模型输出条件确定为 ER。隐含层神经元节点数的确定采用动态构造法, 最终确定隐含层节点数选为 13, 即本模型最终采用 $7 \times 13 \times 1$ 结构。

为提高网络泛化能力, 对输入数据进行归一化处理, 选取 LM 贝叶斯正则化算法对模型进行训练, 设置最大训练步数 epochs=100, 训练目标误差 goal=0.005, 学习效率 lr=0.01, 输入层、隐含层和输出层节点激励函数分别为 tansig、tansig 正切 S 型函数和 purelin 线性函数, 以上过程通过 Matlab 软件编程实现, 当训练误差达到预先设定的目标误差时训练停止, 如图 1(b)所示。

对验证样本的模型预测输出值和实验值进行相关性分析, 计算得出 GANN 模型检验样本预测输出值和实验值的决定系数(R^2)为 0.8887(>0.80)^[14], 模型预测值与实验值的平均偏差为 14.41%。

对本文所得 BDE-209 DLLME 萃取条件的 GANN 模型与 BDE-209 DLLME 萃取条件的 BP 神经网络模型进行比较, 两模型优化随训练步数的变化曲线如图 1(a)、(b)所示。由图 1 可以看出, GANN 模型在训练步数为 8 步时达到设定的目标误差(0.005), 学习速度远大于 BP 神经网络(经 2544 步训练后训练误差为 0.492854)。GANN 模型预测值与实验值的平均偏差为 14.41%, 小于 BP 神经网络预测值与实验值的平均偏差 17.24%。GANN 模型输入值与实验值的决定系数(R^2)为 0.8887, 略大于 BP 神经网络模型输入值与实验值的决定系数(R^2)0.8734。因此, GANN 模型和 BP 神经网络模型相比学习速度更快, 学习精度更高, 网络泛化能力较好。

2.3 基于 GANN 模型的分散液液微萃取回收率的优化

基于 GANN 模型的分散液液微萃取回收率优化模型目标函数可表示为:

$$Y = net(A, B, C, F) \quad (2)$$

式(2)中 net 表示神经网络函数, A 、 B 、 C 、 F 分别为萃取剂体积、分散剂体积、萃取剂类型、离子强度。

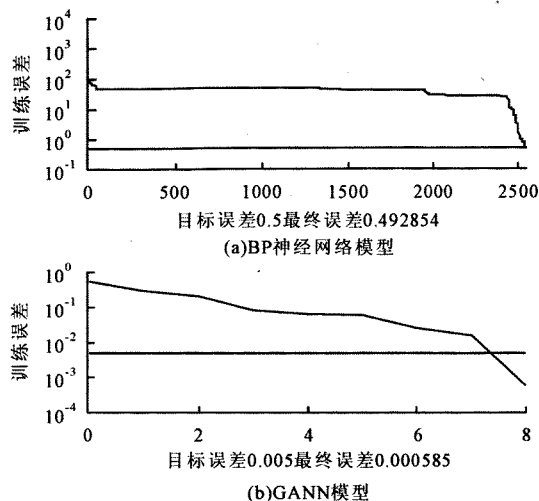


图1 训练误差随训练步数变化曲线
Fig. 1 Training errors and training epochs based on BP neural network model and GANN model

目标函数的约束条件为 4 个主要影响因素的取值范围, 按实验条件设置为 $10 \leq A \leq 30$ (μL), $0.7 \leq B \leq 1.1$ (mL), $1 \leq C \leq 2$ (mg/mL) 以及 $0 \leq F \leq 3$ (%). 利用遗传算法工具箱对 GANN 模型进行优化求解。由此得到萃取回收率最大时所对应的萃取条件, 并利用该条件进行验证实验。萃取条件、优化值和验证实验值如表 2 所示, 萃取回收率优化值和实验值的范围分别为 80.6%~108.99% 和 78.54%~104.75%, 优化值和实验值之间有一定差异, 可能因为实验过程中实际采用的萃取剂密度和优化出的萃取剂密度有一定差异。优化后实验值的 ER (78.54%~104.75%) 和 EF (648~9371) 明显高于优化前的 ER (50.45%) 和 EF (226)。而在 4 号萃取条件下, 有机沉积相体积 $\leq 0.5 \mu\text{L}$, 不利于 BDE-209 在液相色谱中的测定, 因此最佳 DLLME 萃取条件为 $10 \mu\text{L}$ 四氯乙烯、 0.71 mL 丙酮、 $\text{pH}=5$ 、 $\text{NaCl}=20\%$ 、萃取时间 10min, 此萃取条件下, BDE-209 的 EF 和 ER 分别为 104.75 和 806, 和优化前相比分别提高了 54% 和 580。

表 2 基于 GANN 模型的 DLLME 萃取条件的遗传算法优化及验证

Table 2 Optimization and validation of extraction conditions of DLLME based on the GANN model through Matlab software

编号	pH	分散剂	A(μL)	B(mL)	C	F(%)	G(min)	ER		验证实验值 EF
								Matlab 优化值	验证实验值	
1	9	丙酮	11.1	0.9	四氯乙烯	2.50	10	104.67	90.65	648
2	5	丙酮	10.2	0.71	四氯乙烯	2.00	5	88.26	78.54	654
3	5	丙酮	10	0.71	四氯乙烯	2.00	10	80.16	104.75	806
4	9	丙酮	10.1	1.09	氯苯	2.99	5	108.99	93.71	9371

3 结论

本文利用 Matlab 软件建立了水样中十溴联苯醚 DLLME 的 GANN 模型,可有效预测水样中十溴联苯醚 DLLME 萃取回收率和富集因子,且优化后十溴联苯醚 DLLME 的萃取回收率和富集因子可分别提高了 2 倍和 3.6 倍。

[参考文献]

- [1] Söderström G, Sellström U, De Wit C A, et al. Photolytic debromination of decabromodiphenyl ether (BDE 209)[J]. *Environmental Science & Technology*, 2004, 38(1): 127-132.
- [2] He J Z, Robrock K R, Alvarez-Cohen L. Microbial reductive debromination of polybrominated diphenyl ethers (PB-DEs)[J]. *Environmental Science & Technology*, 2006, 40(14): 4429-4434.
- [3] Gerecke A C, Hartmann P C, Heeb N V, et al. Anaerobic degradation of decabromodiphenyl ether[J]. *Environmental Science & Technology*, 2005, 39(4): 1078-1083.
- [4] Li Y Y, Wei G H, Hu J, et al. Dispersive liquid-liquid microextraction followed by reversed phase-high performance liquid chromatography for the determination of polybrominated diphenyl ethers at trace levels in landfill leachate and environmental water samples[J]. *Analytica Chimica Acta*, 2008, 615: 96-103.
- [5] Zhou R, Wang J Y, Li S F, et al. Supercritical fluid extraction of monoamine oxidase inhibitor from antler velvet[J]. *Separation and Purification Technology*, 2009, 65(3): 275-281.
- [6] Sun Y X, Li Y J, Li M Q, et al. Optimization of extraction technology of the anemone raddeana polysaccharides (ARP) by orthogonal test design and evaluation of its anti-tumor activity[J]. *Carbohydrate Polymers*, 2009, 75: 575-579.
- [7] Pourmortazavi S M, Hajimirsadeghi S S, Kohsari I, et al. Orthogonal array design for the optimization of supercritical carbon dioxide extraction of different metals from a solid matrix with cyanex 301 as a ligand[J]. *Journal of Chemical & Engineering Data*, 2004, 49: 1530-1534.
- [8] 汤卫东, 朱海涛. 遗传 BP 神经网络在正交试验最优化中的应用[J]. *信息技术与信息化*, 2004, 29(6): 44-46, 59.
- Tang Wei-dong, Zhu Hai-tao. The application of genetic algorithms and BP neural network in orthogonal experimental method optimization[J]. *Information Technology and Informatization*, 2004, 29(6): 44-46, 59. (in Chinese)
- [9] 付敏. 蚕砂低甲氧基果胶提取工艺的研究[D]. 成都: 四川大学, 2004.
- Fu Min. Study on Extraction Technique of Low Methoxy Pectin from Silkworm Excrement[D]. Chengdu: Sichuan University, 2004. (in Chinese)
- [10] 李目, 何怡刚, 周少武, 等. 混沌时间序列的混合遗传神经网络预测方法[J]. *系统仿真学报*, 2008, 20(21): 5825-5828.
- Li Mu, He Yi-gang, Zhou Shao-wu, et al. Hybrid genetic neural network method for predicting chaotic time series[J]. *Journal of System Simulation*, 2008, 20(21): 5825-5828. (in Chinese)
- [11] 欧阳玉梅, 马志强, 方若森. 基于 MATLAB 的遗传神经网络的设计与实现[J]. *信息技术*, 2008, 6: 73-76, 80.
- Ouyang Yu-mei, Ma Zhi-qiang, Fang Ruo-sen. Design and realization of genetic-neural network based on MATLAB[J]. *Information Technology*, 2008, 6: 73-76, 80. (in Chinese)
- [12] 陈海英, 郭巧, 徐立. 基于混合遗传神经网络的百米跑成绩预测方法[J]. *计算机仿真*, 2004, 21(2): 89-91.
- Chen Hai-ying, Guo Qiao, Xu Li. Prediction method of 100 sprint performance based on hybrid genetic neural network[J]. *Computer Simulation*, 2004, 21(2): 89-91. (in Chinese)
- [13] 韩飞, 莫建华, 龚攀. 基于遗传神经网络的数字化渐进成形回弹预测[J]. *华中科技大学学报(自然科学版)*, 2008, 36(1): 121-124.
- Han Fei, Mo Jian-hua, Gong Pan. Incremental sheet NC forming springback prediction using genetic neural network[J]. *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, 2008, 36(1): 121-124. (in Chinese)
- [14] 苏保林, 王建平, 贾海峰, 等. 密云水库流域非点源模型系统[J]. *清华大学学报(自然科学版)*, 2006, 46(3): 355-359.
- Su Bao-lin, Wang Jian-ping, Jia Hai-feng, et al. Non-point source modeling system of Miyun Reservoir watershed[J]. *Journal of Tsinghua University (Science and Technology)*, 2006, 46(3): 355-359. (in Chinese)