

文章编号: 1006-9941 (2016)07-0644-07

基于遗传算法的人工神经网络预测多硝基化合物撞击感度

钱博文, 陈利平, 陈网桦

(南京理工大学化工学院, 江苏 南京 210094)

摘要: 为获得精度较高、适用范围较大的多硝基化合物撞击感度预测模型, 根据定量构效关系 (QSPR) 原理, 用遗传算法 (GA) 进行变量筛选, 建立了多元线性回归 (GA-MLR) 模型和人工神经网络 (GA-ANN) 模型, 用于 149 种多硝基化合物撞击感度的定量构效关系的研究。两个模型的相关系数分别为 0.854 和 0.974, 均方根误差分别为 0.195 和 0.071, 平均绝对误差分别为 0.157 和 0.051。通过模型比较, 结果表明所得模型的预测精度均高于已有的 QSPR 模型, 且 GA-ANN 模型要明显优于 GA-MLR 模型。

关键词: 多硝基化合物; 撞击感度; 预测; 遗传算法 (GA); 人工神经网络 (ANN); 定量构效关系 (QSPR)

中图分类号: TJ55; O64

文献标志码: A

DOI: 10.11943/j.issn.1006-9941.2016.07.004

1 引言

撞击感度是衡量含能材料安全性能的一个重要参数^[1], 多硝基化合物在民用生产和军事上应用广泛, 因此获得多硝基化合物的撞击感度数据具有一定实用价值。由于实验测试费时费力, 且测试结果易受环境、操作人员、设备等因素的影响, 此外, 对于正处于探索阶段尚未合成出的物质而言, 也无法通过实验获得测试值, 因此有必要对理论预测撞击感度的方法进行研究。定量结构-性质相关性研究 (Quantitative Structure Property Relationship, QSPR) 是一种能根据分子结构实现有机物理化性质预测的有效方法, 已广泛应用于有机物的各类理化性质及生物活性的预测中^[2]。

近年来, 国内外学者已经对撞击感度与其分子结构间的相关关系进行了广泛的探讨。赵俊^[3]等人对 36 种硝基化合物和 33 种硝胺化合物的撞击感度和其分子特征量之间的关联顺序进行了研究。王睿^[4-5]等人应用电性拓扑状态指数对 41 种多硝基苯类化合物和 46 种多硝基脂肪族化合物分别进行了撞击感度的定量构效关系研究。葛素红^[6]等人计算并分析了 18 种三硝基芳香族炸药分子的原子化能以及原子化能与分子的电子结构能的比值与撞击感度的关联关系。Guillaume

Fayet^[7]等人用最佳多元线性回归 (B-MLR) 研究 60 个硝胺类化合物的撞击感度与 Codessa 计算所得 64 个分子描述符间的线性关系。N. R. Badders^[8]等人研究了 22 个多硝基化合物的撞击感度与 6 个量子化学参数间的相关性。这些研究只专注于特定种类的样本集, 所用样本较少, 所得模型的适用范围有很大的局限性。因此, 也有学者针对较大样本集对硝基化合物的撞击感度进行了进一步研究。袁方强^[9]等人根据分子内氢键、分子结构、对称性、氧平衡 OB_{100} 、活性指数 F 、“拥挤性”等分子结构参数与对 156 个硝基类炸药的撞击感度进行了研究。房伟^[10]等人选取氧系数、对称性、 $-COOR$ 、氧杂环、苯环、 $\alpha-H$ 、 $\alpha-OH$ 、 $\alpha-CH$ 、硝基 ($-NO_2$)、氨基 ($-NH_2$) 等作分子结构描述符, 用多元线性回归计算了 123 个硝基化合物的撞击感度。而这些研究是凭借分子结构与撞击感度之间的定性关联, 人为地发掘数个对撞击感度有影响的结构参数, 因此往往构建的模型所用变量较多, 或者可能存在选用参数不足以很好得表征研究的性质的问题。因此, 有必要在较大样本集条件下, 从分子结构角度出发对结构参数进行筛选, 从而建立多硝基化合物撞击感度和其分子结构间的定量关系模型。

为了实现上述目标, 这里根据 QSPR 基本原理, 采用遗传算法^[11] (Genetic Algorithm, GA) 对计算所得的描述符进行筛选, 分别建立多元线性回归 (MLR) 模型和人工神经网络^[12] (ANN) 模型对 149 种多硝基化合物的撞击感度定量构效关系进行研究, 并进行模型比较。

收稿日期: 2015-10-19; 修回日期: 2015-12-10

作者简介: 钱博文 (1992-), 男, 硕士生, 主要从事热安全研究。

e-mail: 897189968@qq.com

通信联系人: 陈利平 (1981-), 女, 讲师, 主要从事热安全研究。

e-mail: 30629372@qq.com

2 样本与方法

2.1 试验样本

实验样本的数量和准确度直接影响撞击模型的适用性和预测精度。这里选取了 149 种多硝基化合物的撞击敏感度^[13]作为实验样本,其中包含 47 个芳香族类化合物、70 个脂肪族化合物以及 32 个硝胺类化合物。按照 4:1 的比例随机选择 120 种作为训练集用于建立 QSPR 模型,剩余的 29 种作为测试集用于对所建模型进行外部评估。 H_{50} 为 2.5 kg 落锤实验中测得的样品爆炸几率为 50% 时的落高,所涉及的 149 种多硝基化合物的撞击敏感度实验值($\lg H_{50}$)及 QSPR 计算得到的预测值见附表(附表可查阅本刊网站)。由于篇幅原因,表 1 只给出了部分多硝基化合物的撞击敏感度实验值和预测值。

2.2 研究方法

首先,借助化学软件 HyperChem7.5^[14]进行分子结构的输入。在采用分子力学 MM+初步优化的基础上,采用量子化学半经验方法 PM3 对分子结构进一步优化以获得能量低的稳定构型。所有计算均限制在 Hartree-Fock 能级,采用 Ploak-Ribiere 方法直至均方根误差梯度达到 0.001 kJ/mol。然后将优化的分子结构导入 Dragon2.1 软件^[15],计算获得 1481 种分子描述符,将近似常数或者相关系数大于 0.95 的描述符删除,初步筛选获得 561 个分子描述符。采用 Material Studio 7.0 软件^[16]实现遗传算法对分子描述符进行进一步筛选。继而分别采用 MLR 和 ANN 研究所得描述符与撞击敏感度间的定量关系,多元线性回归由 SPSS 软件^[17]实现,人工神经网络建模由 MATLAB R2014a^[18]实现。试验步骤流程如图 1 所示。

表 1 部分多硝基化合物的撞击敏感度值 ($\lg H_{50}$)

Table 1 Impact sensitivity values of part of polynitro compounds ($\lg H_{50}$)

No.	compound	experiment value	GA-MLR		GA-ANN	
			predicted value	deviation	predicted value	deviation
1	1-methoxy-3,5-dichloro-2,4,6-trinitrobenzene	1.875	1.873	0.002	1.896	-0.021
2	1,3,5-triamino-2,4,6-trinitrobenzene	2.690	2.498	0.192	2.587	0.103
3	2,4,6-trinitrophenol	1.431	2.100	-0.669	1.473	-0.042
4	3,3'-dihydroxy-2,2',4,4',6,6'-hexanitrobiphenyl	1.602	1.774	-0.172	1.619	-0.017
5	2,4,6-Trinitrobenzyl chloride	1.643	1.899	-0.256	1.682	-0.039
6	2,4,6-Trinitrobenzyl alcohol	1.716	2.149	-0.433	1.797	-0.081
7	1-hydroxyethyl-2,4,6-trinitrobenzene	1.833	2.291	-0.458	1.835	-0.002
8	2,4,6-Trinitrobenzoic acid	2.037	1.911	0.126	2.037	0.000
9	1-ethoxy-2,4,6-trinitrobenzene	2.279	2.367	-0.088	2.213	0.066
10	hexanitro benzene	1.079	0.822	0.257	1.078	0.001
11	2',2',2'-trinitroethyl-2,4,6-trinitrobenzoate	1.380	1.410	-0.030	1.321	0.059

最终,将采用相关系数(R^2)、均方根误差(RMSE)和平均绝对误差(MAE)等统计参数对所建立的模型性能进行全面的评估,并采用 Williams 图对模型的应用域进行分析。

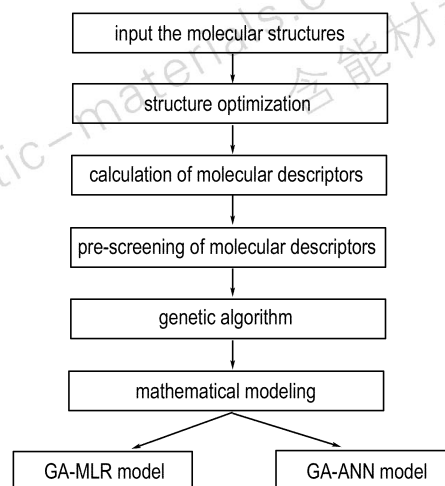


图 1 QSPR 研究步骤流程图

Fig. 1 The flow chart of the QSPR research steps

3 结果与讨论

3.1 GA 筛选结果

基于 Accelrys 公司的 Material Studio 7.0 软件,采用遗传算法(GA)对 561 个分子描述符进行筛选。设置方程式的初始长度 6,最大方程式长度为 10,种群数为 50,最大代数为 500,变异概率为 0.1。选用 Friedman 提出的拟合缺失分数(Lack-Of-Fit, LOF)作为其适应度函数,比例控制 LOF 的平滑参数 α 为 0.5。在遗传过程中,选择在整个种群中出现次数最多的变量为适应度函数的最优解。最终确定的 6 个分子描述符的名称、类型和定义见表 2。

表 2 GA 筛选出的描述符
Table 2 Descriptors obtained by genetic algorithm

descriptor	type	definition
IC1	topological descriptors	information content index(neighborhood symmetry of 1 order)
ATS1e	2D autocorrelations	Broto-Moreau autocorrelation of a topological structure-lag 1 /weighted by atomic Sanderson electronegativities
MATS1e	2D autocorrelations	Moreau autocorrelation- lag 1 /weighted by atomic Sanderson electronegativities
RDF050v	RDF descriptors	radial distribution function-5.0 /weighted by atomic van der Waals volumes
RTu	GETAWAY descriptors	R total index / unweighted
R6m ₊	GETAWAY descriptors	R maximal autocorrelation of lag 6 / weighted by atomic masses

3.2 GA-MLR 模型的建立

首先,运用 MLR 研究 GA 筛选出的描述符和 129 种多硝基化合物撞击感度之间可能存在的线性关系。基于训练集所建立的撞击感度多元线性回归模型如下:
 $\lg H_{50}=16.268-0.285\times IC1-10.656\times ATS1e-$

$$1.578\times MATS1e+0.031\times RDF050v-0.037\times RTu+0.326\times R6m_+$$
$$n=129,R^2=0.854,Q^2_{LOO}=0.808, RMSE=0.195, S_0=0.200, F=110.514, P<0.001 \quad (1)$$

式中, IC1、ATS1e、MATS1e、RDF050v、RTu 和 R6m₊ 是由 GA 筛选获得的描述符,具体含义见表 2; n 为样本数; R² 为相关系数; Q²_{LOO} 为“留一法”交互验证的相关系数; RMSE 为均方根误差; S₀ 为标准偏差; F 为 F 检验值; P 为模型显著性概率。

该模型有较高的相关系数和交互验证系数,均方根误差和标准偏差均较小,表明该模型是可靠的。模型的 F 检验值远大于 F_{理论} (95% 置信度) = 2.198,因此可以认为该回归方程及所选变量的影响均是显著的。表 3 所示是模型的显著性检验结果,标准系数可以确定不同分子描述符对模型的影响程度。从表 3 可以看出,6 个描述符对撞击感度的影响大小排序为 ATS1e>MATS1e>RTu>RDF050v>IC1>R6m₊,所有描述符与撞击感度 lgH₅₀ 均呈负相关性。

表 3 GA-MLR 模型系数检验结果
Table 3 Test results of the coefficients of GA-MLR model

descriptor	regression coefficient	standard error	standardized coefficient	t-value
constant	16.268	0.700	-	23.252
IC1	-0.285	0.080	-0.157	-3.567
ATS1e	-10.656	0.477	-0.933	-22.355
MATS1e	-1.578	0.191	-0.402	-8.252
RDF050v	0.031	0.006	-0.251	5.554
RTu	-0.037	0.005	-0.366	-7.785
R6m ₊	0.326	0.139	-0.095	2.349

为检验模型的外部预测能力,应用测试集对所建模型进行外部验证。计算得到,训练集的相关系数 R² 为 0.854,均方根误差 RMSE 为 0.195,平均绝对误差 MAE 为 0.157; 测试集的相关系数 R²_{EXT} 为 0.790,均方根误差 RMSE_{EXT} 为 0.215,平均绝对误差 MAE 为 0.183。模型的主要性能参数见表 4,可以看出,无论对于训练集还是测试集,模型均有较小的均方根误差和平均绝对误差,说明该线性模型有着较好的内部稳定性。

表 4 GA-MLR、GA-ANN 撞击感度预测模型的主要性能参数
Table 4 Main performance parameters of GA-MLR model and GA-ANN model

parameter	model	
	GA-MLR	GA-ANN
R ²	0.854	0.974
RMSE	0.195	0.071
MAE	0.157	0.051
R ² _{EXE}	0.790	0.740
RMSE _{EXT}	0.215	0.248
MAE _{EXT}	0.183	0.197
MAPE/%	11.12	7.98
Q ² _{LOO}	0.808	0.975
Q ² _{F1}	0.796	0.746
Q ² _{F2}	0.791	0.740
Q ² _{F3}	0.731	0.666

训练集和测试集撞击感度预测值与实验值的比较及残差分布分别见图 2 和图 3。如图 2 所示,所有的数据点均随机分布在对角线附近,表明此线性模型具有一定的预测能力,但是模型的精度不高,预测结果不是很好。由此猜测,多硝基类化合物的撞击感度和分子结构间可能存在非线性关系,因此将尝试建立 GA-ANN 非线性模型进行研究。而由图 3 可知,模型的预测残差均匀且随机分布于基准线的两侧。由此可以认为,GA-MLR 模型在建立过程中未产生系统误差。

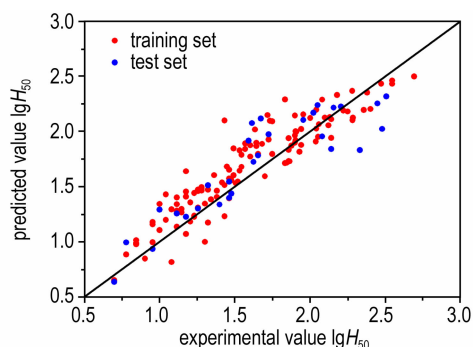


图2 GA-MLR 模型所得撞击感度的预测值与实验值的比较

Fig. 2 Comparison of the predicted values of impact sensitivity obtained by GA-MLR model and the experimental ones

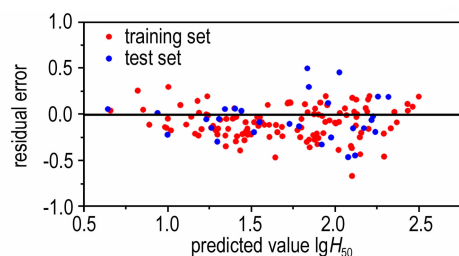


图3 GA-MLR 模型撞击感度预测值和残差的关系图

Fig. 3 Relation of the predicted values of impact sensitivity obtained by GA-MLR model and the residual

3.3 GA-ANN 模型的建立

在用神经网络研究化学结构性质的领域,三层反向传播的前馈神经网络应用最为广泛。理论上已经证明:具有偏差和至少一个 S 型隐含层加上一个线性输出层的网络,能够逼近任何有理函数^[12]。因此,这里建立包含输入层、输出层和一层隐含层的三层 BP 神经网络用于非线性模型研究。BP 神经网络的构建在 MATLAB R2014a 环境中实现,经过多次训练试验,最终确定 6-8-1 的神经网络结构的精度最好。具体参数设定如下:输入层节点数为 6,隐含层节点数为 8,输出层节点数为 1,选择最小训练速率 0.1,最大迭代次数为 1000,隐含层激活转移函数采用双曲正切 S 型函数 tansig,输出层激活转移函数采用线性函数 purelin,训练函数采用 trainlm,网络学习函数采用带动量项的 BP 学习规则 learngdm。

模型的预测结果见附表,相关模型性能参数见表 4。计算得到,该模型的相关系数为 0.974,交叉验证系数为 0.975,训练集和测试集的均方根误差分别为 0.071 和 0.248,平均绝对误差分别为 0.051 和 0.197。从相关系数、交叉验证系数和均方根误差来看,GA-ANN 模型的精度明显高于 GA-MLR 模型。

如图 4 所示,所有多硝基化合物撞击感度预测值和实验值的比较均未偏离对角线很远,整体基本紧挨对角线,这说明该模型比 MLR 建立的模型精度高。图 5 给出了 GA-ANN 模型的残差图,预测残差均匀且随机分布于基准线,该模型在建立过程中未产生系统误差。

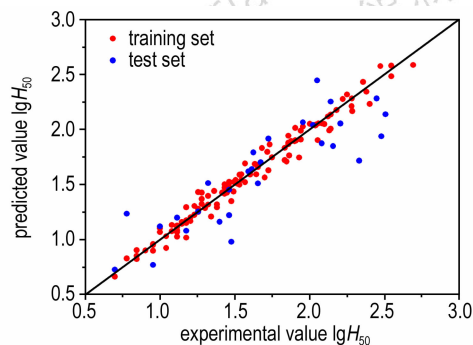


图4 GA-ANN 模型所得撞击感度预测值与实验值的比较

Fig. 4 Comparison of the predicted values of impact sensitivity obtained by GA-ANN model and the experimental ones

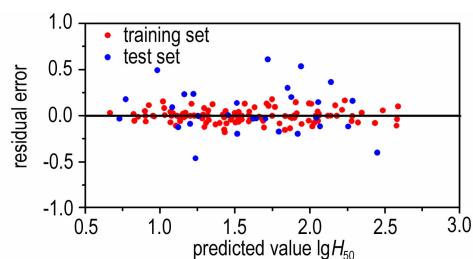


图5 GA-ANN 模型撞击感度预测值和残差的关系图

Fig. 5 Relation of the predicted values of impact sensitivity obtained by GA-ANN model and the residual

3.4 模型比较

附表给出了两个模型对所有 149 种多硝基化合物的预测值。为便于比较分析 GA-MLR 和 GA-ANN 模型,计算得到两个模型的主要性能参数见表 4。比较两个模型的相关系数 R^2 、均方根误差 RMSE、平均绝对误差 MAE、平均绝对误差百分比 MAPE 和交叉验证系数 Q_{LOO}^2 可以发现,GA-ANN 模型的性能参数都要优于 GA-MLR 模型,这表明非线性模型的精度要明显高于线性模型。然而,比较两个模型的测试集相关系数 R_{EXT}^2 和外部验证系数 Q_{F1}^2 、 Q_{F2}^2 和 Q_{F3}^2 ,GA-MLR 模型均略优于 GA-ANN 模型,这表明非线性模型的外部预测能力相比线性模型并没有得到预期的提升。即,所建的两个模型虽然具有一定的预测能力,但是预测结果并没有取得理想的预期效果。这可能由于两个原因:(1)样本集的数据存在问题;(2)建模所用的参数并不能很好表征多硝基化合物的撞击感度。

这里采用 Williams 图^[19]进一步分析比较两个模型的应用域。横坐标为臂比值,纵坐标为标准残差。标准残差落在(-3,+3)以外的化合物,认为其实验值是离群点;当化合物的臂比值 h_i 大于警戒值 h^* 时,认为该物质显著影响模型的回归效果^[14]。臂比值的计算式如下:

$$h_i = x_i^T (X \cdot X^T)^{-1} x_i \quad (i=1, \cdots, n)$$
 (2)

警戒值 h^* 的表达式如下:

$$h^* = 3 \times (m+1) / n$$
 (3)

式中, m 为建立模型所用变量的个数。

如图 6、图 7 所示的 Williams 图中,GA-MLR 模型和 GA-ANN 模型分别有 2 个物质和 5 个物质的标准残差落在了基准线以外,其相应的化合物和撞击感度值见表 5 和表 6。

观察表 5、表 6 发现,出现异常值的物质其撞击感度的预测残差均较大,这表明,建模所用的参数没能很好的表征其撞击感度。已有研究表明^[20],除了分子结构外,含能材料的感度还与炸药晶体特性密切相关,如颗粒内部空洞的微观结构、颗粒表面缺陷、颗粒大小及分布、颗粒形状和表面光滑度对感度影响非常显著。因此,仅用分子结构参数来表征撞击感度还不够完善。而 2',2'-二硝基丙基-2,4,6-三硝基苯甲酸酯的标准残差在两个模型中同时出现异常,可认为主要原因是其实验值数据本身不够准确。在表 7 中,六硝基苯的臂比值超出了警戒值,这表明,六硝基苯显著影响了模

型的回归效果,可以认为其结构相对于整个样本集而言是比较特殊的。

将所建的两个模型与几种已有的多硝基化合物撞击感度 QSPR 模型进行比较,相关统计参数见表 8。

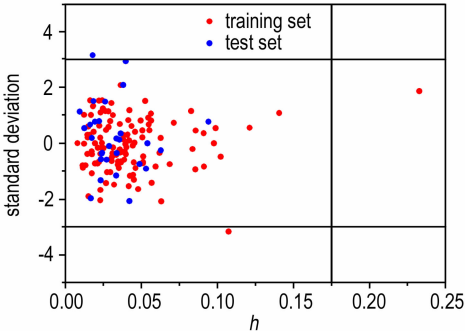


图 6 GA-MLR 撞击感度模型的 Williams 图
Fig. 6 Williams plot of GA-MLR model for impact sensitivity

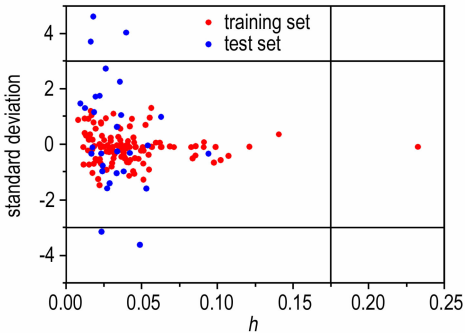


图 7 GA-ANN 撞击感度模型的 Williams 图
Fig. 7 Williams plot of GA-ANN model for impact sensitivity

表 5 根据残差判定的 GA-MLR 模型中的异常值

Table 5 Outliers of GA-MLR model determined according to the residual

compound	standardized residual	experimental value	predicted value	residual error
2,4,6-trinitrophenol	-3.168	1.431	2.100	-0.669
2',2'-dinitropropyl-2,4,6-trinitrobenzoate	3.166	2.330	1.833	0.497

表 6 根据残差判定的 GA-ANN 模型中的异常值

Table 6 Outliers of GA-ANN model determined according to the residual

compound	standardized residual	experimental value	predicted value	residual error
2',2'-dinitropropyl-2,4,6-trinitrobenzoate	4.611	2.330	1.718	0.612
trinitroethyl-bis-(trinitroethoxy)-acetate	-3.619	0.778	1.237	-0.459
N-(2-propyl)-trinitroacetamide	-3.149	2.049	2.447	-0.398
bis-(2,2,2-trinitroethyl)-succinate	3.711	1.477	0.982	0.495
bis-(2,2-dinitropropyl)-carbonate	4.039	2.477	1.940	0.537

表 7 根据臂比值判定的撞击感度预测模型中的异常值

Table 7 Outliers of predicted model of impact sensitivity determined according to the arm ratio

compound	h^*	h	experimental value	predicted value	residual error
hexanitro benzene	0.175	0.232	1.079	1.078	0.001

表 8 所建模型与已有模型的比较

Table 8 Comparison between the established models with the existing models

author	model	data set	R^2	Q_{LOO}^2	$RMSE_{\text{EXT}}$
this paper	GA-MLR	149	0.854	0.808	0.215
	GA-ANN(6-8-1)	149	0.974	0.975	0.248
Nefati ^[21] (1995)	ANN(13-2-1)	204	0.794	0.795	0.257
Wang Rui ^[22] (2008)	MLR	156	0.771	0.593	0.251
	ANN(16-12-1)	156	0.816	—	0.247
Guillaume Fayet ^[23] (2012)	MLR	161	0.816	0.793	0.230

本文建模采用的是由 GA 筛选获得分子描述符,因此能更系统全面地表征所研究的性质,且模型中涉及的变量个数远少于已有模型。比较各模型的相关系数、交互验证系数和外部验证的均方根误差,可以认为这里获得的 GA-MLR 模型和 GA-ANN 模型明显优于已有模型。

4 结 论

(1) 运用遗传算法筛选出 6 个分子描述符,分别构建了 149 种多硝基化合物撞击感度的 GA-MLR 模型和 GA-ANN 模型,结果表明两个模型均是稳定可靠的;

(2) 所得 GA-ANN 模型明显优于 GA-MLR 模型,说明多硝基化合物的撞击敏感度与分子结构间存在较强的非线性关系,且两个模型的预测精度均高于已有的 QSPR 模型;

(3) 由于落锤实验受诸如晶体特性、环境等其他因素的影响,因此所建的两个模型虽然具有一定的预测能力,但是预测效果仍有继续完善的空间。

参考文献:

- [1] 金韶华,王伟,松全才,等. 含能材料机械撞击敏感度判据的认识和发展[J]. 爆破器材, 2006, 35(6): 11-14.
JIN Shao-hua, WANG Wei, SONG Quan-cai, et al. Understanding and development of criteria of impact sensitivity of energetic materials[J]. *Explosive Materials*, 2006, 35(6): 11-14.
- [2] Katritzky A R, Lobanov V S, Karelson M. QSPR: The correlation and quantitative prediction of chemical and physical properties from structure[J]. *Chemical Society Reviews*, 1995, 24(4): 279-287.
- [3] 赵俊,何碧,程新路,等. 多硝基炸药撞击敏感度与分子特征量关联度的 BP 神经网络方法研究[J]. 四川大学学报(自然科学版), 2006, 17(5): 501-507.
ZHAO Jun, HE Bi, CHENG Xin-lu, et al. Neural networks study on the correlation between impact sensitivity and molecular structures for nitramines[J]. *Atmospheric Chemistry & Physics*, 2006, 17(5): 501-507.
- [4] 王睿,蒋军成,潘勇,等. 电性拓扑态指数预测硝基类含能材料撞击敏感度[J]. 固体火箭技术, 2008, 31(6): 657-662.
WANG Rui, JIANG Jun-cheng, PAN Yong, et al. Prediction on impact sensitivity of nitro energetic materials by means of electro-topological state indices[J]. *Journal of Solid Rocket Technology*, 2008, 31(6): 657-662.
- [5] 王睿,蒋军成,潘勇. 脂肪族硝基含能化合物撞击感度的 QSPR 研究[J]. 南京工业大学学报(自然科学版), 2011, 33(3): 15-21
WANG Rui, JIANG Jun-cheng, PAN Yong. Impact sensitivity analysis on aliphatic nitro energetic compounds by QSPR[J]. *Journal of Nanjing University of Technology*, 2011, 33(3): 15-21.
- [6] 葛素红,孙桂华,董光兴. 硝基芳香族炸药的分子结构和撞击敏感度关系研究[J]. 四川师范大学学报, 2014(5): 684-690.
GE Su-hong, SUN Gui-hua, DONG Guang-xing. Molecular structure, atomization energy and impact sensitivity of eighteen trinitro-aromatic-explosives[J]. *Journal of Sichuan Normal University*, 2014(5): 684-690.
- [7] Fayet G, Rotureau P, Joubert L, et al. Development of a QSPR model for predicting thermal stabilities of nitroaromatic compounds taking into account their decomposition mechanisms[J]. *Journal of Molecular Modeling*, 2011, 17(10): 2443-2453.
- [8] Badders N R, Wei C, Aldeeb A A, et al. Predicting the impact sensitivities of polynitro compounds using quantum chemical descriptors[J]. *Journal of Energetic Materials*, 2006, 24(1): 17-33.
- [9] 袁方强,蔡从中,赵帅. 用结构参数预测硝基类炸药的撞击敏感度[J]. 爆炸与冲击, 2013, 33(1): 79-84.
YUAN Fang-qiang, CAI Cong-zhong, ZHAO Shuai. Prediction of impact sensitivity of nitro energetic compounds by using structural parameters[J]. *Explosion & Shock Waves*, 2013, 33(1): 79-84.
- [10] 房伟,王建华,刘玉存,等. 多因素预测硝基化合物的撞击敏感度[J]. 火炸药学报, 2014, 37(5): 42-46.
FANG Wei, WANG Jian-hua, LIU Yu-cun, et al. Multivariate prediction of impact sensitivity of nitro compounds[J]. *Chinese Journal of Explosives & Propellants*, 2014, 37(5): 42-46.
- [11] 陈国良,王煦法,庄镇泉,等. 遗传算法及其应用[M]. 北京: 人民邮电出版社, 1999: 2-4.
CHEN Guo-liang, WANG Xu-fa, ZHUANG Zhen-quan, et al. Genetic algorithm and its application[M]. Beijing: Posts & Telecom Press, 1999: 2-4.
- [12] 丛爽. 面向 MATLAB 工具箱的神经网络理论与应用第三版[M]. 合肥: 中国科学技术大学出版社, 2009: 72-81.
CONG Shuang. The neural network theory and application facing the toolbox of MATLAB of the third edition [M]. Hefei: Press of University of Science and Technology of China, 2009: 72-81.

- [13] Storm C B, Stine J R, Kramer J F. Sensitivity relationships in energetic materials[J]. *Chemistry and Physics of Energetic Materials*, 1990: 605–639.
- [14] SHI Jing-jie, Chen L P, Chen W H. QSPR models of compound viscosity based on iterative self-organizing data analysis technique and ant colony algorithm[J]. *Acta Physico-chimica Sinica*, 2014, 30(5): 803–810.
- [15] Tebby C, Mombelli E, Pandard P. et al. Exploring an eco toxicity database with the OECD (Q) SAR toolbox and DRAGON descriptors in order to prioritize testing on algae, daphnis, and fish[J]. *Science of the Total Environment*, 2011, 409: 3334–3343.
- [16] SHI Jing-jie. Prediction of auto-ignition temperatures of organic compounds by SVM based on the genetic algorithm[J]. *Journal of China Safety Science*, 2011, 21(7): 125–129.
- [17] 薛薇. 统计分析与 SPSS 的应用(第三版)[M]. 北京: 中国人民大学出版社, 2011.
XUE Wei. Statistical analysis and the application of SPSS (3rd edition)[M]. Beijing: China Renmin University Press, 2011.
- [18] 王小川. MATLAB 神经网络 43 个案例分析[M]. 北京: 北京航空航天大学出版社, 2013.
WANG Xiao-chuan. 43 cases analysis of MATLAB neural network[M]. Beijing: University Press, 2013.
- [19] Gramatica P. Principles of QSAR models validation: internal and external[J]. *QSAR and Combinatorial Science*, 2007, 26(5): 694–701.
- [20] 李洪珍, 康彬, 李金山, 等. RDX 晶体特性对冲击感度的影响规律[J]. 含能材料, 2010, 18(5): 487–491.
LI Hong-zhen, KANG Bin, LI Jin-shan, et al. Effects of RDX crystal characteristics on shock sensitivities [J]. *Chinese Journal of Energetic Materials (Hanneng Caifiao)*, 2010, 18(5): 487–491.
- [21] Nefati H, Cense J M, Legendre J J. Prediction of the impact sensitivity by neural networks[J]. *Journal of Chemical Information & Modeling*, 1996, 36(4): 804–810.
- [22] Wang Rui, Jiang Jun-cheng, Pan Yong, et al. Prediction of impact sensitivity of nitro energetic compounds by neural network based on electro topological-state indices[J]. *Hazard. Mater*, 2009, 166: 155–186.
- [23] Fayet G, Rotureau P, Prana V, et al. Global and local quantitative structure-property relationship models to predict the impact sensitivity of nitro compounds[J]. *AIChE*, 2012, 31(3): 291–303.

Prediction of Impact Sensitivity of Polynitro Compounds by Artificial Neural Network Based on the Genetic Algorithm

QIAN Bo-wen, CHEN Li-ping, CHEN Wang-hua

(School of Chemical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: To obtain the prediction model with higher precision and larger suitable scope for predicting impact sensitivity of polynitro compounds, based on the quantitative structure-property relationship (QSPR) principle, genetic algorithm (GA) was adopted in variable selection and GA-multiple linear regression (MLR) model and GA-artificial neural networks(ANN) model were established and employed in research for figuring out the relationship between impact sensitivity and molecular structure of 149 kinds of polynitro compounds. Two model's correlation coefficient is 0.854 and 0.974, respectively, the root-mean-square error is 0.195 and 0.071, respectively and the mean-absolute-error is 0.157 and 0.051, respectively. By comparing the models, the results show that the prediction precision of two models are higher than the existing QSPR model, and GA-ANN model is obviously superior to GA-MLR model.

Key words: polynitro compounds; impact sensitivity; predict; genetic algorithm(GA); artificial neural network(ANN); quantitative structure-property relationship (QSPR)

CLC number: TJ55; O64

Document code: A

DOI: 10.11943/j.issn.1006-9941.2016.07.004