

用分子拓扑参数预估多硝基芳香族化合物的撞击感度

杜军良^{1,2}, 舒远杰², 周 阳², 罗娅君¹, 胡晓黎¹, 边清泉¹

(1. 绵阳师范学院化学与化学工程系, 四川 绵阳 621000;

2. 中国工程物理研究院化工材料研究所, 四川 绵阳 621900)

摘要:引入主成分回归研究了36种多硝基芳香族化合物(PNAC)分子的撞击感度。在DFT-B3LYP/6-311+G(d,p)水平上对这些炸药分子进行了分子优化和频率振动分析。结合炸药分子的拓扑结构参数和所得的量化结构参数,最终确定硝基个数、氨基个数、芳香性NICS(1)、最长C-NO₂键键长、HOMO以及 α -C-H(0,1)6个参数与撞击感度($\lg H_{50}$)具有较好的相关性。以这6个参数作为主成分回归的输入参数构建模型,得到测试集的决定系数 R^2 和交叉验证系数 R_{adj}^2 分别为0.97和0.89,优于由氧平衡指数确定的传统模型的0.91和0.33。得出主成分回归是一种研究撞击感度的有效方法。

关键词:材料科学;主成分回归;撞击感度;多硝基芳香族化合物;量子化学

中图分类号:TJ55;O64

文献标志码:A

文章编号:1007-7812(2010)06-0005-06

Predication on Impact Sensitivity of Polynitroaromatic Compounds using Principal Component Regression

DU Jun-liang¹, SHU Yuan-jie², ZHOU Yang², LUO Ya-jun¹, HU Xiao-li¹, BIAN Qing-quan¹

(1. Department of Chemistry and Chemical Engineering, Mianyang Normal University, Mianyang Sichuan 621000, China; 2. Institute of Chemical Materials, CAEP, Mianyang Sichuan 621900, China)

Abstract: A method was introduced for predicting the impact sensitivity of thirty-six polynitroaromatic compounds (PNACs) by principal component regression (PCR). Geometry optimizations and frequency calculations were carried out for 36 PNACs molecules by means of density functional theory (DFT) at the DFT-6-311+G(d,p) level. Combining with the topological parameters and the quantum-chemical parameters, six molecular descriptors close related to $\lg H_{50}$, including the numbers of nitro group and amido group, aromaticity index NICS (1), the longest C-NO₂ bond length, HOMO (Highest Occupied Molecular Orbital), indicator of $-\text{CH}$ in α (0 or 1) were selected. The PCR with these descriptors was established to predict the impact sensitivity of PNACs. The predicted data of the PCR were compared with experimental results and those of traditional models established by the oxygen balance index (OB100). Results show that the coefficient of determination (R^2) and Cross Validation of PCR are 0.97 and 0.89, respectively, which of the traditional models is 0.91 and 0.33. The PCR is an efficient tool for predicting impact sensitivity of PNACs.

Key words: material science; principal component regression; impact sensitivity; PNAC; quantum chemical

引言

炸药撞击感度与分子结构关系的研究一直是炸药理论的一个重要课题。Kamlet和Adolph^[1]等

发现,对于分解机理相同的炸药分子,其撞击感度的对数和氧平衡之间存在近似的线性关系。Politzer^[2]等人研究了C-NO₂键中点的静电势极值与该键的离解能之间的关系,发现它们之间存在近似的反比例关系:C-NO₂键中点的静电势极值越大,C-NO₂

收稿日期:2010-09-13; 修回日期:2010-11-20

基金项目:国家自然科学基金—中物院NSAF联合基金重点项目(No. 10576030)和中俄国际合作基金(No. 10610194)资助。

作者简介:杜军良(1979—),男,助教,硕士研究生,从事含能材料的制备与性能研究。

键的离解能越低,其撞击感度越高。Zeman^[3]等用量子力学方法研究撞击感度与一系列硝基芳香化合物近似爆轰热、溶化热及升华热间的关系。通过计算芳香族炸药芳香性指标 NICS(核独立化学位移)的方法来评定一些炸药分子的稳定性,发现 NICS 值越小,化合物越敏感^[4-5]。尽管上述研究找到了炸药撞击感度和分子特征量间的某些关系,但这些关联只涉及一个或两个特征量。因此为了进一步研究撞击感度与分子特征量之间的关系,引入了主成分回归方法来研究多个不同的分子特征量与撞击感度的关系。

主成分回归是一种多变量回归方法,首先对数据矩阵进行主成分分析(PCA),将得到的变量转换成新的变量,再对得到的新变量采用多元回归进行建模。主成份回归已被应用于化学、经济学、光谱研究及医学等领域^[6]。本研究运用主成分回归法来分析炸药撞击感度和炸药分子特征量间的关联关系。

1 理论方法

1.1 计算方法

用 Gaussian98 程序^[7]中的密度泛函方法,对文中选取的 37 个炸药分子(见表 1)在 B3LYP/6-311++G(d,p)水平下进行几何结构全优化^[8-10],并求得在各自平衡构型下的电子结构的各特征量。包括核独立化学位移(NICS),最高占据轨道 HOMO,最长 C-NO₂ 键键长等。其中 NICS 是在 B3LYP/6-311++G(d,p)水平下用标准 GIAO 程序计算的,这是由于 Stanger^[11]证明准确计算 NICS 值的最小基组为 B3LYP/6-311++G(d,p)。硝基个数,氨基个数, α -C-H(0/1)为分子结构的拓扑参数。

1.2 参数的选择

本研究计算了 36 个分子的量化参数,结合分子拓扑参数,总共 22 种参数,在 SPSS 软件中利用逐步向后法(backward)进行分析,最终选用 6 个对感度有重要影响的参数(见表 1):硝基个数,氨基个数, NICS,最长 C-NO₂ 键键长, HOMO 以及 α -C-H(0/1)作为主成分回归的输入参数。

其中 α -C-H(0/1)是指示变量,如果炸药分子含 α -C-H 基团,则取 1,反之取 0。NICS 是芳香性指标,对于芳香族化合物,由于存在离域 π 电子,在垂直于环分子平面的外磁场 H^0 的作用下,离域 π 电子会产生环电流,从而产生一个垂直于分子平面的感应磁场 H^1 ,这个感应磁场在环外是增加磁场的,在环内是抵抗外磁场的,所以在环内是屏蔽区(+),在环外是去屏蔽区(-),见图 1^[12]。NICS 值本质是环电流的反映,它在一定程度上反映了芳香族化合物的稳定性。NICS(1)为距苯环中心 1 Å 处的 NICS 值。

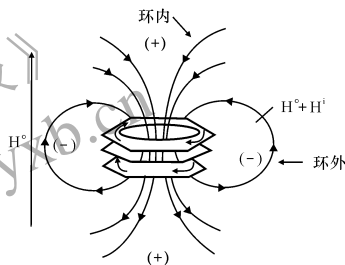


图 1 苯环的环电流和感应磁场

Fig. 1 The ring electric current and inductive magnetic field of benzene ring

硝基具有致爆性,而氨基具有钝感的作用^[13]。C-NO₂ 键长曾被用作表征 C-NO₂ 键的强度。在一定程度上与炸药的撞击感度相关。

表 1 所选炸药相关参数

Table 1 Molecular descriptors of PNACs selected

| 编号 | 化合物 | NICS(1) (X_1) | 最长 C-NO ₂ 键键长(X_2) | HOMO (X_3) | 硝基个数 (X_4) | 氨基个数 (X_5) | α -C-H (X_6) | lg H_{50} (exp.) |
|----|----------------------|----------------------|--------------------------------------|-------------------|-------------------|-------------------|----------------------------|--------------------|
| 1 | 2-氨基-3,4,5,6-四硝基甲苯 | -8.4021 | 1.49575 | -0.292 | 4 | 1 | 1 | 1.55630 |
| 2 | 3,5-二甲基-2,4,6-三硝基苯酚 | -9.1305 | 1.48570 | -0.291 | 3 | 0 | 1 | 1.88649 |
| 3 | 4,6-二硝基间苯二酚 | -7.9156 | 1.45132 | -0.298 | 2 | 0 | 0 | 2.54407 |
| 4 | 2,3,4-三硝基甲苯 | -10.1254 | 1.48862 | -0.317 | 3 | 0 | 1 | 1.74819 |
| 5 | 2,4,6-三硝基甲苯 | -7.6798 | 1.47580 | -0.307 | 3 | 0 | 0 | 1.63347 |
| 6 | 1,3,5-三羟基-2,4,6-三硝基苯 | -6.1905 | 1.45157 | -0.313 | 3 | 0 | 0 | 1.43136 |
| 7 | 2,4,6-三硝基苯胺 | -7.6990 | 1.47173 | -0.293 | 3 | 1 | 0 | 2.24797 |
| 8 | 3,4,5-三硝基甲苯 | -10.0550 | 1.48892 | -0.313 | 3 | 0 | 1 | 2.02938 |
| 9 | 2,3,4,5-四硝基甲苯 | -10.1482 | 1.48926 | -0.329 | 4 | 0 | 1 | 1.11394 |
| 10 | 2,3,4,6-四硝基苯 | -10.5495 | 1.49205 | -0.344 | 4 | 0 | 0 | 1.51851 |
| 11 | 2,3,4,6-四硝基苯胺 | -8.0974 | 1.49679 | -0.305 | 4 | 1 | 0 | 1.61278 |
| 12 | 2,3,4,6-四硝基甲苯 | -10.2602 | 1.49012 | -0.333 | 4 | 0 | 1 | 1.27875 |

续表 1

| 编号 | 化合物 | NICS(1) (X_1) | 最长C-NO ₂ 键键长(X_2) | HOMO (X_3) | 硝基个数 (X_4) | 氨基个数 (X_5) | α -C-H (X_6) | $\lg H_{50}(\text{exp.})$ |
|----|----------------------|----------------------|-------------------------------------|-------------------|-------------------|-------------------|----------------------------|---------------------------|
| 13 | 2,3,5,6-四硝基苯 e | -10.2328 | 1.48805 | -0.336 | 4 | 0 | 1 | 1.39794 |
| 14 | HMX | -3.5750 | 1.40085 | -0.321 | 4 | 0 | 0 | 1.41497 |
| 15 | 均三硝基苯 | -10.7244 | 1.48813 | -0.345 | 3 | 0 | 0 | 2 |
| 16 | 2,4,6-三硝基苯酚 | -10.2080 | 1.48043 | -0.318 | 3 | 0 | 0 | 2.27875 |
| 17 | 2,4,6-三硝基苯甲醚 | -10.2578 | 1.48081 | -0.321 | 3 | 0 | 0 | 2.2833 |
| 18 | 1,3,5-三氨基-2,4,6-三硝基苯 | -4.0127 | 1.44023 | -0.281 | 3 | 3 | 0 | 2.6902 |
| 19 | 五硝基甲苯 | -9.9302 | 1.49067 | -0.344 | 5 | 0 | 1 | 1.25527 |
| 20 | 均三硝基甲苯 | -10.3429 | 1.48533 | -0.324 | 3 | 0 | 1 | 2.20412 |
| 21 | 六硝基苯 | -10.0400 | 1.49429 | -0.362 | 6 | 0 | 0 | 1.07918 |
| 22 | 2,4,6-三硝基苯乙醇 | -10.2244 | 1.49099 | -0.310 | 3 | 0 | 1 | 1.83251 |
| 23 | 2,4,6-三硝基氯甲苯 | -10.4068 | 1.49060 | -0.329 | 3 | 0 | 1 | 1.64345 |
| 24 | 2,4,6-三硝基氯苯 | -10.1374 | 1.48540 | -0.329 | 3 | 0 | 0 | 1.89763 |
| 25 | 2,4,6-三硝基腈苯 | -10.5131 | 1.48801 | -0.345 | 3 | 0 | 0 | 2.14613 |
| 26 | 3-氨基-2,4,6-三硝基腈苯 | -8.1747 | 1.47768 | -0.303 | 3 | 1 | 0 | 2.14613 |
| 27 | 2,4,6-三硝基乙酰苯 | -10.4685 | 1.48766 | -0.303 | 3 | 0 | 0 | 1.89763 |
| 28 | 2,4,6-三硝基苯甲酸甲酯 | -10.6298 | 1.48789 | -0.323 | 3 | 0 | 0 | 1.95424 |
| 29 | 3,5-二氨基-2,4,6-三硝基苯酚 | -4.7969 | 1.46407 | -0.279 | 3 | 2 | 0 | 2.04922 |
| 30 | 2,4,6-三硝基苯酸 | -10.5982 | 1.48828 | -0.330 | 3 | 0 | 0 | 2.03743 |
| 31 | 2,4-二硝基间苯二酚 | -7.5448 | 1.45584 | -0.286 | 2 | 0 | 0 | 2.47129 |
| 32 | 苦味酸(PA) | -9.1195 | 1.48079 | -0.316 | 3 | 0 | 0 | 1.93952 |
| 33 | 五硝基苯胺 | -8.1731 | 1.49814 | -0.315 | 5 | 1 | 0 | 1.17609 |
| 34 | 2,4,6-三硝基苯乙醇 | -9.6478 | 1.48530 | -0.323 | 3 | 0 | 1 | 1.716 |
| 35 | 3-氨基-2,4,6-三硝基苯酚 | -6.9203 | 1.46873 | -0.292 | 3 | 1 | 0 | 2.13988 |
| 36 | 1,3-二氨基-2,4,6-三硝基苯 | -5.7051 | 1.45923 | -0.280 | 3 | 2 | 0 | 2.50515 |

注:化合物1~30为训练集,31~36为测试集。

2 结果与讨论

2.1 使用 Spss Linear Regression 过程的向后逐步法筛选变量和进行多重共线性诊断

在 Linear Regression 对话框,把应变量 Y 放入 Dependent 栏,全部自变量放入 Independent 栏,在 Method 下拉菜单中,选择 Backward 法,筛选有统计学意义 ($p < 0.05$) 的自变量。在 Statistics 对话框,选择 Descriptives、Covariance matrix 和 Collinearity Diagnostic,其他为系统默认值。

运行 Linear Regression 过程后,确定了与 Y 密切相关的自变量为: X_1 、 X_2 、 X_4 、 X_5 和 X_6 ,应变量和自变量的均数±标准差依次为 1.8270 ± 0.41048 、 -9.0509 ± 2.03137 、 1.4796 ± 0.02022 、 3.3667 ± 0.76489 、 0.3000 ± 0.70221 、 0.3667 ± 0.49013 。其他结果见表 2 和表 3。

2.2 使用 Spss Factor Analysis 过程对自变量 X_1 、 X_2 、 X_4 、 X_5 和 X_6 进行主成分分析

在 Factor Analysis 对话框,把自变量 X_1 、 X_2 、

X_4 、 X_5 和 X_6 放入 Variables 栏。在 Extraction 对话框,选择 Number of factors 栏,并在该栏内输入 5,其他为系统默认值。

表 2 回归系数及其共线性统计量

Table 2 Linear regression coefficients and their statistics

| 模型 | 偏回归系数 | t | p | 共线性统计量 | |
|-------|---------|--------|-------|--------|--------|
| | | | | 容差 | VIF |
| 常数 | 20.604 | 3.490 | 0.002 | | |
| X_1 | -0.208 | -3.933 | 0.001 | 0.098 | 10.251 |
| X_2 | -13.257 | -3.063 | 0.005 | 0.147 | 6.792 |
| X_4 | -0.336 | -7.127 | 0.000 | 0.867 | 1.153 |
| X_5 | 0.460 | 5.336 | 0.000 | 0.308 | 3.244 |
| X_6 | -0.145 | -1.921 | 0.067 | 0.825 | 1.212 |

表 3 共线性诊断指标

Table 3 Collinearity diagnostic criteria

| 维数 | 特征值 | 条件指数 | 方差/% | | | | | |
|----|----------|---------|------|-------|-------|-------|-------|-------|
| | | | 常数 | X_1 | X_2 | X_4 | X_5 | X_6 |
| 1 | 4.525 | 1.000 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2 | 0.954 | 2.178 | 0 | 0 | 0 | 0 | 22 | 10 |
| 3 | 0.468 | 3.108 | 0 | 0 | 0 | 0 | 8 | 78 |
| 4 | 0.039 | 10.830 | 0 | 3 | 0 | 77 | 2 | 0 |
| 5 | 0.013 | 18.390 | 0 | 17 | 0 | 13 | 24 | 5 |
| 6 | 1.43E-05 | 561.715 | 1 | 80 | 1 | 10 | 11 | 5 |

经 Factor Analysis 过程运算后结果见表 4。

表 4 各主成分的特征值、方差及其系数

Table 4 Eigenvalue of each principal component, the variance and coefficient

| 主成分 | 特征值 | 方差/% | 累计方差/% | 标化自变量 | | | | |
|-----|-------|--------|--------|--------|-------|--------|--------|--------|
| | | | | X_1 | X_2 | X_4 | X_5 | X_6 |
| 1 | 2.489 | 49.786 | 49.786 | -0.936 | 0.854 | 0.276 | -0.695 | 0.569 |
| 2 | 1.022 | 20.442 | 70.228 | 0.247 | 0.019 | 0.862 | 0.315 | 0.344 |
| 3 | 0.764 | 15.282 | 85.510 | 0.088 | 0.094 | -0.424 | 0.368 | 0.658 |
| 4 | 0.670 | 13.407 | 98.917 | -0.159 | 0.494 | -0.016 | 0.525 | -0.353 |
| 5 | 0.054 | 1.083 | 100.00 | 0.171 | 0.134 | -0.019 | -0.081 | -0.009 |

2.3 使用 SPSS Compute Variable 过程把应变变量 Y 和自变量 X_1 、 X_2 、 X_4 、 X_5 和 X_6 标化并求主成分 C_1 、 C_2 、 C_3 、 C_4 和 C_5 的值

在 Compute Variables 对话框, 把 betay 输入 Target 栏作为标化应变变量名, 在 Numeric Expression 栏输入 $(Y-1.8270)/0.41048$ 。标化自变量 X_i 时, 在 Target 栏输入标化自变量名分别为 $\text{beta}X_1$ 、 $\text{beta}X_2$ 、 $\text{beta}X_4$ 、 $\text{beta}X_5$ 和 $\text{beta}X_6$, 在 Numeric Expression 栏输入数学表达式分别为 $(X_1+9.0509)/2.03137$ 、 $(X_2-1.4796)/0.02022$ 、 $(X_4-3.3667)/0.76489$ 、 $(X_5-0.3000)/0.70221$ 、 $(X_6-0.3667)/0.49013$ 。计算主成分 C_i 时, 在 Target 栏输入主成分变量名分别为 C_1 、 C_2 、 C_3 、 C_4 和 C_5 , 在 Numeric Expression 栏输入数学表达式分别为 $-0.936\text{beta}X_1+0.854\text{beta}X_2+0.276\text{beta}X_4-0.695\text{beta}X_5+0.569\text{beta}X_6$ 、 $0.247\text{beta}X_1+0.019\text{beta}X_2+0.862\text{beta}X_4+0.315\text{beta}X_5+0.344\text{beta}X_6$ 、 $0.088\text{beta}X_1+0.094\text{beta}X_2-0.424\text{beta}X_4+0.368\text{beta}X_5+0.658\text{beta}X_6$ 、 $-0.159\text{beta}X_1+0.494\text{beta}X_2-0.016\text{beta}X_4+0.525\text{beta}X_5-0.353\text{beta}X_6$ 、 $0.171\text{beta}X_1+0.134\text{beta}X_2-0.019\text{beta}X_4-0.081\text{beta}X_5-0.009\text{beta}X_6$ 。不断运行 Compute Variables 过程, 在工作数据表生成标化应变变量、各标化自变量和主成分变量。

2.4 使用 SPSS Linear Regression 过程进行主成分回归分析, Compute Variable 过程的计算残差和残差绝对值

从累计方差大于 85% 所包括的主成分开始建立标准化主成分回归方程。从表 4 中可以看出, 首先选择主成分为 3 的时候开始建模, 在 Linear Regression 对话框, 选择 betay 放入 Dependent 栏, 选择 C_1 、 C_2 、 C_3 放入 Independent 栏。同理, 拟合和的操作仅仅是在 Independent 栏放入的主成分变量有所增加, 前者是放入 C_1 和 C_2 、 C_3 和 C_4 , 后者是 C_1 和 C_2 、 C_3 、 C_4 和 C_5 。拟合 3 个主成分回归方程: $y' = -0.411C_1 - 0.619C_2 + 0.289C_3$; $y' = -0.411C_1 - 0.619C_2 + 0.289C_3 + 0.325C_4$; $y' = -0.410C_1 - 0.620C_2 + 0.289C_3 + 0.325C_4 - 0.314C_5$ 。3 个方程中的标化偏

回归系数的 P 值均 < 0.05 。

Compute Variable 对话框, 把 E_1 输入 Target 栏作为残差变量, 在 Numeric Expression 栏输入 $\text{betay} - (-0.411C_1 - 0.619C_2 + 0.289C_3)$ 。同理, 用表达式 $\text{betay} - (-0.411C_1 - 0.619C_2 + 0.289C_3 + 0.325C_4)$ 和 $\text{betay} - (-0.410C_1 - 0.620C_2 + 0.289C_3 + 0.325C_4 - 0.314C_5)$ 可产生残差变量 E_2 和 E_3 。

在 Compute Variable 对话框, 把 AE_1 输入 Target 栏作为残差绝对值变量, 在 Function 下拉菜单中, 选择 ABS(Numeric Expression), 把它放入 Numeric Expression 栏, 并在括号内输入 E_1 。同理, 可产生残差绝对值变量 AE_2 和 AE_3 。在 Descriptives 对话框, 把残差绝对值变量 AE_1 、 AE_2 和 AE_3 放入 Variables 栏, 求出各标化主成分回归方程的残差绝对值均数和标准差, 见表 5。

表 5 标化主成分回归方程、残差绝对值均数和标准差

Table 5 Standardized principal component regression equation, residuals mean and standard deviation of the absolute value

| 标化主成分回归方程 | 残差绝对值均数±标准差 |
|--|----------------------|
| $y' = -0.411C_1 - 0.619C_2 + 0.289C_3$ | 0.5889 ± 0.62101 |
| $y' = -0.411C_1 - 0.619C_2 + 0.289C_3 + 0.325C_4$ | 0.5913 ± 0.53694 |
| $y' = -0.410C_1 - 0.620C_2 + 0.289C_3 + 0.325C_4 - 0.314C_5$ | 0.5855 ± 0.53104 |

2.5 使用 SPSS Bivariate 过程计算 L_{yy} 和 L_{xixi}

在 Bivariate Correlations 对话框, 选择应变变量 Y 和自变量 X_1 、 X_2 、 X_4 、 X_5 和 X_6 放入 Variables 栏。在 Options 对话框, 选择 Crossproduct deviation and covariance, 运行 Bivariate 过程产生 $L_{yy} = 4.886$, $L_{x_1x_1} = 119.667$, $L_{x_2x_2} = 0.012$, $L_{x_4x_4} = 16.967$, $L_{x_5x_5} = 14.300$, $L_{x_6x_6} = 6.967$ 。其中 $L_{Y Y}$ 为 Y 的离均差平方和, $L_{X_i X_i}$ 为 X_i 的离均差平方和。

2.6 把标化主成分回归方程转换为标化线性回归方程然后再转换为一般线性回归方程

按较小平均绝对残差和较大累计方差百分比的原则, 确定最佳标化主成分回归方程为: $y = -0.410C_1 - 0.620C_2 + 0.289C_3 + 0.325C_4 - 0.314C_5$ 。从 3.3 可得: $C_1 = -0.936\text{beta}X_1 + 0.854\text{beta}X_2 + 0.276$

$\beta_4 X_4 - 0.695\beta_5 X_5 + 0.569\beta_6 X_6, C_2 = 0.247\beta_1 X_1 + 0.019\beta_2 X_2 + 0.862\beta_4 X_4 + 0.315\beta_5 X_5 + 0.344\beta_6 X_6, C_3 = 0.088\beta_1 X_1 + 0.094\beta_2 X_2 - 0.424\beta_4 X_4 + 0.368\beta_5 X_5 + 0.658\beta_6 X_6, C_4 = -0.159\beta_1 X_1 + 0.494\beta_2 X_2 - 0.016\beta_4 X_4 + 0.525\beta_5 X_5 - 0.353\beta_6 X_6, C_5 = 0.171\beta_1 X_1 + 0.134\beta_2 X_2 - 0.019\beta_4 X_4 - 0.081\beta_5 X_5 - 0.009\beta_6 X_6$, 并代入最佳标化主成分回归方程, 整理后得标化线性回归方程: $Y' = 0.151X'_1 - 0.216X'_2 - 0.769X'_4 + 0.392X'_5 - 0.368X'_6$ 。

用标化偏回归系数 $b'_1 = 0.151, b'_2 = -0.216, b'_3 = -0.769, b'_4 = 0.392$ 和 $b'_5 = -0.368$ 按(1)式计算 b_i , 并按(2)式计算 $b_0, b_1 = 0.0305, b_2 = -4.359, b_3 =$

$-0.413, b_4 = 0.229, b_5 = -0.308$ 和 $b_6 = 9.985$ 。得一般线性回归方程: $Y = 9.985 + 0.0305X_1 - 4.359X_2 - 0.413X_4 + 0.229X_5 - 0.308X_6$ 。

$$b_i = b'_i \cdot \sqrt{\frac{L_{yy}}{L_{xixi}}} \quad (1)$$

$$b_0 = \bar{Y} - \sum b_i \bar{x}_i \quad (2)$$

式中: b_i 为一般线性回归方程的第 i 个偏回归系数, b'_i 为标化回归方程的第 i 个标化偏回归系数, b_0 为一般线性回归方程的常数。

2.7 用一般线性回归方程预测6种炸药

用以上得到的一般线性回归方程预测6种炸药的撞击感度, 并与实验值比较, 结果见表6。

表6 测试组化合物撞击感度的实验值与预测值的比较

Table 6 Comparison of impact sensitivity between experimentation and prediction for test set compounds

| 编号 | 化合物 | OB_{100} | $\lg H_{50}$ (实验值) | $\lg H_{50}$ | | | $\lg H_{50}$ | | |
|-------|--------------------|------------|-----------------------|--------------|--------|---------|--------------------------|--------|---------|
| | | | | 主成份回归 | 误差 | 相对误差 | 文献预测值 ^[14-15] | 误差 | 相对误差 |
| 31 | 2,4-二硝基间苯二酚 | -2.000 | 2.471 | 2.583 | 0.112 | 0.0453 | 2.334 | -0.137 | -0.0554 |
| 32 | 苦味酸(PA) | -0.440 | 1.940 | 2.013 | 0.073 | 0.0376 | 1.850 | -0.090 | -0.0464 |
| 33 | 五硝基苯胺 | 1.880 | 1.176 | 1.369 | 0.193 | 0.1641 | 0.789 | -0.387 | -0.3291 |
| 34 | 2,4,6-三硝基苯乙醇 | -2.060 | 1.716 | 1.669 | -0.047 | -0.0273 | 1.856 | 0.140 | 0.0816 |
| 35 | 3-氨基-2,4,6-三硝基苯酚 | -0.810 | 2.140 | 2.362 | 0.222 | 0.1037 | 1.965 | -0.175 | -0.0818 |
| 36 | 1,3-二氨基-2,4,6-三硝基苯 | -1.950 | 2.505 | 2.669 | 0.164 | 0.0655 | 2.319 | -0.186 | -0.0743 |
| 残差平方和 | | | | | 0.044 | | | 0.109 | |

2.8 与传统模型比较

1976年, Kamlet 等人^[1]对芳香族炸药的特性落高和氧平衡指数 OB_{100} 之间的关系进行了研究, 发现 $\lg H_{50} \%$ 与 OB_{100} 之间成线性相关, 即所谓“感度-结构趋势”。

对于不含 α -H 的芳香族化合物, 其撞击感度的特征落高适用于下列关系式:

$$\lg(H_{50}) = 1.714 - 0.31OB_{100} \quad (3)$$

对于含 α -H 的芳香族化合物, 感度相对较高, 适合下面的关系式:

$$\lg(H_{50}) = 1.372 - 0.235OB_{100} \quad (4)$$

对于含 N-NO₂ 的化合物, 符合下式:

$$\lg(H_{50}) = 1.369 - 0.168OB_{100} \quad (5)$$

其中氧平衡指数 OB_{100} 的计算公式为:

$$OB_{100} = \frac{100(2n_O - n_H - 2n_C - 2n_{COO})}{M} \quad (6)$$

式中: n_O, n_H, n_C 分别为炸药分子中所含的氧、氢、碳的原子个数, n_{COO} 为分子中所含有的羧基个数, M 为炸药的摩尔质量。

表6也列出了传统公式的预测值。图2为测试组的预测值与真实值的关系图。

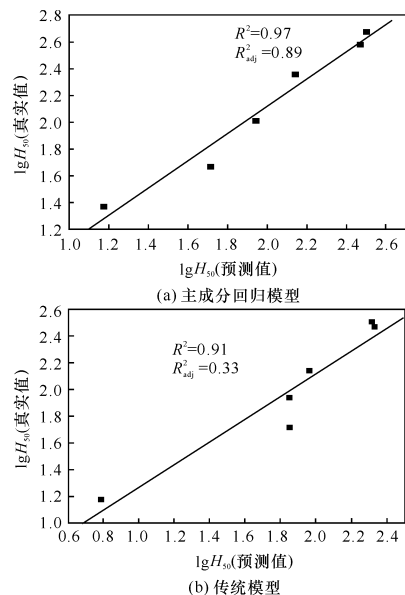


图2 测试组预测值与真实值的关系图
Fig. 2 Relationship between predicted values and real ones

3 结 论

(1) 采用 DFT6-311+G(d,p) 方法优化和计算了 36 种炸药分子的量子化学参数。在此基础上, 结合拓扑参数, 以 6 个对感度有重要影响的参数, 用主成分回归对前 30 种炸药分子进行训练, 得出预测方程为: $\lg H_{50} = 9.985 + 0.0305 \times \text{NICS}(1) - 4.359 \times \text{最长 C-NO}_2 \text{ 键键长} - 0.413 \times \text{硝基个数} + 0.229 \times \text{氨基个数} - 0.308 \times \alpha\text{-C-H}$ 。由所建立的模型可以看出增加硝基个数, 感度升高, 增加氨基个数, 感度降低, 这与硝基有致爆性及氨基致钝性是一致的。此外, 影响撞击感度的还有最长 C-NO₂ 键键长, 是否还有 $\alpha\text{-C-H}$ 以及芳香性判决核独立化学位移。

(2) 用所建模型对 6 种炸药分子的撞击感度进行了预测, 得到测试集的 R^2 (决定系数) 和 R_{adj}^2 (交叉验证系数) 为 0.97、0.89, 优于由氧平衡指数确定的传统模型的 0.91 和 0.33, 主成份回归分析法是撞击感度研究的一个非常有效的工具。

参 考 文 献:

[1] Kamlet M J, Adolph H G. The relationship of impact sensitivity with structure of organic high explosives. II. polynitroaromatic explosives [J]. Propellants, Explosives, Pyrotechnics, 1979, 4(1):30-34.

[2] Politzer P, Murray J S. Relationships between dissociation energies and electrostatic potentials of C-NO₂ bonds: Applications to impact sensitivities[J]. J Mol Struct, 1996, 376(1-3):419-424.

[3] Svatopluk Z, Miloslav K. New aspects of impact reactivity polynitro compounds, Part III. impact sensitivity as a function of the inter-molecular interactions [J]. Propellants, Explosives, Pyrotechnics, 2003, 28(6):301-307.

[4] 杜军良, 舒远杰, 周阳, 等. 硝基苯胺炸药的芳香性与其撞击感度的关系[J]. 火炸药学报, 2008, 31(4):6-9. DU Jun-liang, SHU Yuan-jie, ZHOU Yang, et al. Relationship between the impact sensitivities and aromaticity for nitroanilines explosives [J]. Chinese Journal of Explosives and Propellants, 2008, 31(4):6-9.

[5] 杜军良, 舒远杰, 周阳, 等. 一种表征芳香族炸药撞击感度的简单方法[J]. 含能材料, 2008, 16(6):182. DU Jun-liang, SHU Yuan-jie, ZHOU Yang, et al. A simple way to characterize impact sensitivity of

aromatic explosives[J]. Chinese Journal of Energetic materials, 2008, 16(6):182.

[6] 郭东星. 主成分回归方法研究及其在医学方面的应用[D]. 太原:山西医科大学, 2000.

[7] Frisch M J, Trucks G W, Schlegel H B, et al. Gaussian 98, Revision D. 3 [CP/CD]. Pittsburgh PA: Gaussian, Inc., 1998.

[8] Becke A D. Density-functional thermochemistry. III. The role of exact exchange[J]. J Chem Phys, 1993, 98(7):5648-5652.

[9] Lee C, Yang W, Parr R G. Development of the correlation-energy formula into a functional of the electron density [J]. Phys Rev B, 1988, 37(2):785-789.

[10] Miehlich B, Savin A, Stoll H, et al. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr [J]. Chem Phys Lett, 1989, 157(3):200-206.

[11] Stanger A. Nucleus-independent chemical shift (NICS): Distance dependence and revised criteria for aromaticity and antiaromaticity [J]. J Org Chem, 2006, 71(3):883-893.

[12] 封继康. 什么是“芳香性”? 芳香性的定义、判据和无机芳香性化合物[J]. 分子科学学报, 2005, 21(4):1-11. FENG Ji-kang. What is aromaticity? The definition and criteria of aromaticity and inorganic Aromatic compounds [J]. Journal of Molecular Science, 2005, 21(4):1-11.

[13] 王开明, 程新路, 杨向东. 硝基苯酚类炸药 C-NO₂ 键中点静电势最大值与其实验撞击感度关系的研究[J]. 四川大学学报(工程版), 2002, 34(2):104-107. WANG Kai-ming, CHENG Xin-lu, YANG Xiang-dong. Investigation of relationships between impact sensitivities and the maxima of electrostatic potentials at the midpoints of C-NO₂ bonds in nitrophenols [J]. Journal of Sichuan University (Engineering Science Edition), 2002, 34(2):104-107.

[14] 周霖. 爆炸化学基础[M]. 北京:北京理工大学出版社, 2005. ZHOU Lin. Base of Explosive Chemistry [M]. Beijing: Beijing Institute of Technology Press, 2005.

[15] Keshavarz M H, Jaafari M. Investigation of the various structure parameters for predicting impact sensitivity of energetic molecule via artificial neural network [J]. Propellants, Explosives, Pyrotechnics, 2006, 3:161-167.