

文章编号: 1006-9941(2004)04-0207-07

用人工神经网络法预估芳香族多硝基化合物生成焓

王芳¹, 刘剑洪¹, 田德余^{1,2}, 陈莉¹, 赵凤起³, 洪伟良¹, 李翠华¹, 张雪利¹

- (1. 深圳大学师范学院化生系, 广东 深圳 518060;
2. 国防科技大学航天与材料工程学院, 湖南 长沙 410073;
3. 西安近代化学研究所, 陕西 西安 710065)

摘要: 采用误差反向传播学习(BP)的一个双层神经网络,以分子结构中不同基团作为描述码,预估芳香族多硝基化合物的生成焓,研究了网络参数及分子结构描述码的影响,同时用分子子图法进行了多元线性回归,取得了较好的结果(其回归方程相关系数达到0.9967),计算结果的绝大多数相对误差在10%范围以内。

关键词: 物理化学; 生成焓; 芳香族多硝基化合物; 人工神经网络(ANN); 分子结构描述码(MSD)

中图分类号: O621.2

文献标识码: A

1 引言

芳香族多硝基化合物有着广泛的用途,可用作军用炸药,爆破工程,矿山开采,是某些染料中间体原料,也是特种推进剂,耐热和钝感炸药的主成分。研究它们的化学反应的平衡组成,计算推进剂的比冲、特征速度等能量特性参数^[2],炸药的爆压、爆速等爆轰参数,都需要生成焓数据。而这些数据却分散在各种手册^[1-3]或文献中不易查找,合成和实验研究又要花费大量经费,所以发展了许多评估方法,如分子子图法^[4,5],基团贡献法等。本实验收集了58种芳香族多硝基化合物的生成焓数据,用人工神经网络和分子结构描述码(或基团结构)进行模拟计算,选择其中部分(47种)作为训练和校正集,11种为预测集,总效果较好。其回归方程的相关系数达到了0.9967,此方法可以预估许多新的芳香族多硝基化合物的生成焓。

2 人工神经网络的基本原理^[6]

人工神经网络(ANN)是模拟人脑组织结构和智能行为的系统工程,具有非线性变换,平行处理,自学习,自适应能力,自组织及外推特性。误差反向传播

(BP)神经网络是常用的一种,它将一组样本作为训练集,提供一系列的输入输出模式令网络学习,依一定的学习规则调节各层节点间的连接权重,使网络的实际输出与期望值比较达到一定的精度要求:

$$e = \frac{1}{2} \sum_{j=1}^n (o_j - t_j)^2 \rightarrow \min$$

式中, e 为输出误差, n 为样本(训练集)数, t_j 是样本 j 的期望输出值, o_j 是网络实际计算的输出值。人工神经网络BP模型实质上是一个从输入到输出等维非线性映射,应选择合适的样本特征作为输入参数。分子结构信息数值化是最重要的参数,它能反映分子的结构特征,具有通用性、简便、灵活的特点。为此本文列出了芳香族多硝基化合物分子结构描述码(molecular structure describer, MSD),见表1、表2。

3 结果与讨论

3.1 网络参数的影响

芳香族多硝基化合物的结构相似、数据丰富,但受不同基团的影响,其生成焓明显不同,因此,本研究以基团为描述码对其生成焓进行预估。共收集了58种芳香族多硝基化合物的生成焓,其中47种作为学习训练和校验数据,11种为预测数据。本文以化合物的15个分子基团描述码作为输入参数,选用15-16-13-10-1的4层网络,输入和输出节点数由输入信息和输出结果数目决定,适当增加隐含层节点数可提高ANN性能预估的精度。经多次试算、比较,得到较好的网络参数:学习速率0.1,动量项为0.7。结果列于图1和表3中。

收稿日期: 2004-01-17; 修回日期: 2004-04-01

基金项目: 国防预研项目(12N1.9613-2)

作者简介: 王芳(1973-),女,讲师,从事高分子材料与计算化学研究。

联系人: 田德余 e-mail: TDY8181@sina.com.cn

表1 芳香族多硝基化合物分子子图

Table 1 Molecular subgraph of aromatic polynitro compounds






a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
	-NO ₂ (邻)	-NO ₂ (间)	-NO ₂ (对)	-N=N-			-CH ₃			-NH ₂	-CH ₃	-OH	-COOH	-COH

表2 各种芳香族多硝基化合物的分子结构描述码

Table 2 MSD of aromatic polynitro compounds

No.	compound name	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
1	tetranitrodibenzetraza-pentalene	2	4	0	0	2	0	0	0	0	0	0	0	0	0	0
2	bis(2,4,6-trinitrophenyl) diazene	2	4	0	2	1	0	0	0	0	0	0	0	0	0	0
3	2,4,6,2',4',6'-hexanitrodiphenylamine	2	4	0	2	0	0	0	0	1	0	0	0	0	0	0
4	3,3'-diamion-2,2',4,4',6,6'-hexanitrobiphenyl	2	4	0	2	0	0	0	0	0	0	2	0	0	0	0
5	3-nitrobiphenyl	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
11	(Z)-1,2-bis(2,4,6-trinitrophenyl) ethylene	2	4	0	2	0	0	1	0	0	0	0	0	0	0	0
12	(Z)-1,2-bis(4-nitrophenyl) ethylene	2	0	0	2	0	0	1	0	0	0	0	0	0	0	0
13	(E)-1,2-bis(4-nitrophenyl) ethylene	2	0	0	2	0	1	0	0	0	0	0	0	0	0	0
14	(Z)-1-nitro-4-(2-phenylethenyl) benzene	2	0	0	1	0	0	1	0	0	0	0	0	0	0	0
15	(E)-1-nitro-4-(2-phenylethenyl) benzene	2	0	0	1	0	1	0	0	0	0	0	0	0	0	0
21	1,3,5-trinitrobenzene	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0
22	1,3-dinitrobenzene	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0
23	1,4-dinitrobenzene	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0
24	2,3,4,6-tetranitro aniline	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0
25	2,4-dinitroaniline	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0
37	2,6-dinitroaniline	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0
38	2,4-dinitroaniline	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0
39	3-nitrophenol	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0
40	4-nitrophenol	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
41	2-nitrophenol	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0
42	3-nitrobenzoic acid	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0
43	4-nitrobenzoic acid	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0
44	2-nitrobenzoic acid	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
45	2-nitrobenzaldehyde	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1

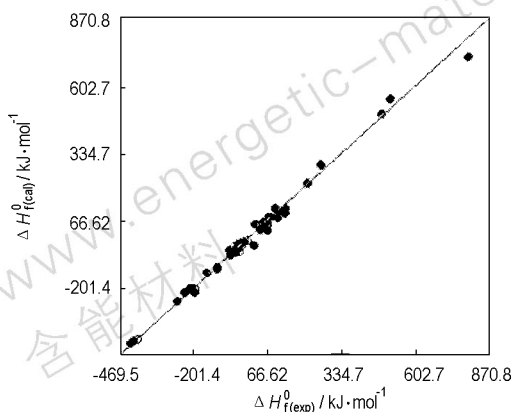


图1 芳香族多硝基化合物的生成焓实验与预测结果比较图

Fig. 1 Relation for exp. and cal. of the enthalpy of formation for aromatic polynitro compounds

3.2 回归方程的建立

本文利用分子基团描述码(当做分子子图)对47组用ANN求得的生成焓的数值进行多元线性回归(MLR),其回归方程为:

$$\begin{aligned}
 Y = & -134.567058 + 103.202327a_1 - 1.753095a_2 - \\
 & 0.370509a_3 - 18.101780a_4 + 251.500472a_5 + \\
 & 33.102832a_6 + 55.449498a_7 - 20.355580a_8 + \\
 & 13.195741a_9 + 74.065965a_{10} - 11.730829a_{11} - \\
 & 6.906941a_{12} - 170.616936a_{13} - 375.026808a_{14} - \\
 & 87.926808a_{15}
 \end{aligned}$$

$$N = 47 \quad R = 0.9967$$

回归方程的相关系数达0.9967,大部分的相对误差在10%范围以内,达到了较高的精度。

表3 芳香族多硝基化合物的生成焓
Table3 The enthalpy of formation for Aromatic polynitro compounds

No.	structure	compound name	formula	$\Delta H_f^0 / \text{kJ} \cdot \text{mol}^{-1}$		error%
				ref. [1-3]	ann.	
1		tetranitrodibenzotetraza-pentalene	$\text{C}_{12}\text{H}_4\text{N}_8\text{O}_8$	510.45	535.10	4.8
2		bis(2,4,6-trinitrophenyl)-diazene	$\text{C}_{12}\text{H}_4\text{N}_8\text{O}_{12}$	262.50	288.30	9.83
3		2,4,6,2',4',6'-hexanitro-diphenyl-amine	$\text{C}_{12}\text{H}_5\text{N}_7\text{O}_{12}$	40.10	42.00	4.70
4		3,3'-diamino-2,2',4,4',6,6'-hexanitro-biphenyl	$\text{C}_{12}\text{H}_6\text{N}_8\text{O}_{12}$	-14.90	-17.60	18.12
5		3-nitrobiphenyl	$\text{C}_{12}\text{H}_9\text{NO}_2$	65.30	70.10	7.35
6		4-nitrobiphenyl	$\text{C}_{12}\text{H}_9\text{NO}_2$	40.60	37.00	8.86
7		3,4-dinitrodiphenylamine	$\text{C}_{12}\text{H}_9\text{N}_3\text{O}_4$	64.31	59.00	8.26
8		4-nitrodiphenylamine	$\text{C}_{12}\text{H}_{10}\text{N}_2\text{O}_2$	70.90	64.00	9.73
9		4,4'-dinitrohydrazobenzene	$\text{C}_{12}\text{H}_{10}\text{N}_4\text{O}_4$	109.70	116.00	5.74
10		(E)-1,2-bis(2,4,6-trinitrophenyl)ethylene	$\text{C}_{14}\text{H}_6\text{N}_6\text{O}_{12}$	59.70	57.80	3.18
11		(Z)-1,2-bis(2,4,6-trinitrophenyl)ethylene	$\text{C}_{14}\text{H}_6\text{N}_6\text{O}_{12}$	78.24	78.80	0.72
12		(Z)-1,2-bis(4-nitrophenyl)ethylene	$\text{C}_{14}\text{H}_{10}\text{N}_2\text{O}_4$	74.10	73.80	0.40
13		(E)-1,2-bis(4-nitrophenyl)ethylene	$\text{C}_{14}\text{H}_{10}\text{N}_2\text{O}_4$	54.90	52.50	4.37
14		(Z)-1-nitro-4-(2-phenyl-ethenyl)benzene	$\text{C}_{14}\text{H}_{11}\text{NO}_2$	132.00	117.90	10.68
15		(E)-1-nitro-4-(2-phenyl-ethenyl)benzene	$\text{C}_{14}\text{H}_{11}\text{NO}_2$	102.70	92.58	9.80

Table 3 (Continued)

No.	structure	compound name	formula	$\Delta H_f^0 / \text{kJ} \cdot \text{mol}^{-1}$		error%
				ref. [1-3]	ann.	
16		2,2',2'',4,4',4'',6,6',6''-nona-nitroterphenyl	$\text{C}_{18}\text{H}_5\text{N}_9\text{O}_{18}$	132.20	129.70	1.90
17		2,2'',4,4',4'',6,6',6''-octa-nitro-terphenyl	$\text{C}_{18}\text{H}_6\text{N}_8\text{O}_{16}$	94.70	103.86	9.67
18		2,2',2'',2''',4,4',4'',4''',6,6',6'',6'''-dodecanitroquaterphenyl	$\text{C}_{24}\text{H}_6\text{N}_{12}\text{O}_{24}$	213.10	217.20	1.92
19		azo-bis(2,2',4,4',6,6'-hexa-nitro-biphenyl)	$\text{C}_{24}\text{H}_6\text{N}_{14}\text{O}_{24}$	480.00	479.70	0.06
20		2,2',2'',2''',4,4',4'',4''',6,6',6'',6'''-dodecanitro-3,3'-bis-(phenylazo)-biphenyl	$\text{C}_{24}\text{H}_6\text{N}_{16}\text{O}_{24}$	791.70	727.05	8.17
21		1,3,5-trinitrobenzene	$\text{C}_6\text{H}_3\text{N}_3\text{O}_6$	-37.00	-34.70	6.22
22		1,3-dinitrobenzene	$\text{C}_6\text{H}_4\text{N}_2\text{O}_4$	-27.40	-28.44	3.80
23		1,4-dinitrobenzene	$\text{C}_6\text{H}_4\text{N}_2\text{O}_4$	-38.70	-32.10	17.05
24		2,3,4,6-tetranitro aniline	$\text{C}_6\text{H}_3\text{N}_5\text{O}_8$	-48.90	-48.50	0.82
25		2,4-dinitroaniline	$\text{C}_6\text{H}_5\text{N}_3\text{O}_4$	-67.80	-61.30	9.58
26		2,5-dinitroaniline	$\text{C}_6\text{H}_5\text{N}_3\text{O}_4$	-44.30	-45.40	2.48

Table 3 (Continued)

No.	structure	compound name	formula	$\Delta H_f^0 / \text{kJ} \cdot \text{mol}^{-1}$		error%
				ref. [1-3]	ann.	
27		2,6-dinitroaniline	$\text{C}_6\text{H}_5\text{N}_3\text{O}_4$	-50.60	-44.17	12.71
28		3,5-dinitroaniline	$\text{C}_6\text{H}_5\text{N}_3\text{O}_4$	-38.90	-36.72	5.6
29		3-nitrobenzenamine	$\text{C}_6\text{H}_6\text{N}_2\text{O}_2$	-38.30	-38.68	0.01
30		2-nitrobenzenamine	$\text{C}_6\text{H}_6\text{N}_2\text{O}_2$	-26.10	-30.10	15.33
31		4-nitrobenzenamine	$\text{C}_6\text{H}_6\text{N}_2\text{O}_2$	-42.00	-38.20	9.05
32		2,4,6-trinitrotoluene	$\text{C}_7\text{H}_5\text{N}_3\text{O}_6$	-65.50	-63.00	3.82
33		3-methyl-2,4,6-trinitro-benzene	$\text{C}_7\text{H}_5\text{N}_3\text{O}_7$	-255.90	-259.30	1.33
34		1-methyl-2,4-dinitro-benzene	$\text{C}_7\text{H}_6\text{N}_2\text{O}_4$	-66.40	-62.10	6.48
35		1-methyl-2,6-dinitro-benzene	$\text{C}_7\text{H}_6\text{N}_2\text{O}_4$	-51.10	-51.73	1.23
36		1-methyl-4-nitrobenzene	$\text{C}_7\text{H}_7\text{NO}_2$	-48.10	-50.80	5.61
37		2,6-dinitrophenol	$\text{C}_6\text{H}_4\text{N}_2\text{O}_5$	-209.90	-208.20	0.81
38		2,4-dinitrophenol	$\text{C}_6\text{H}_4\text{N}_2\text{O}_5$	-232.70	-226.90	2.49
39		3-nitrophenol	$\text{C}_6\text{H}_5\text{NO}_3$	-210.50	-211.50	0.48
40		4-nitrophenol	$\text{C}_6\text{H}_5\text{NO}_3$	-194.30	-207.40	6.74
41		2-nitrophenol	$\text{C}_6\text{H}_5\text{NO}_3$	-194.14	-191.70	1.26
42		3-nitrobenzoic acid	$\text{C}_7\text{H}_5\text{NO}_4$	-414.00	-413.40	0.15
43		4-nitrobenzoic acid	$\text{C}_7\text{H}_5\text{NO}_4$	-426.90	-423.40	0.82
44		2-nitrobenzoic acid	$\text{C}_7\text{H}_5\text{NO}_4$	-398.50	-398.00	0.13

Table 3 (Continued)

No.	structure	compound name	formula	$\Delta H_f^0 / \text{kJ} \cdot \text{mol}^{-1}$		error%
				ref. [1-3]	ann.	
45		2-nitrobenzaldehyde	$\text{C}_7\text{H}_5\text{NO}_3$	-115.00	-113.80	1.04
46		4-nitrobenzaldehyde	$\text{C}_7\text{H}_5\text{NO}_3$	-151.00	-147.20	2.52
47		3-nitrobenzaldehyde	$\text{C}_7\text{H}_5\text{NO}_3$	-112.10	-114.20	1.87
* 48		2,4,6-trinitrophenol	$\text{C}_6\text{H}_3\text{N}_3\text{O}_7$	-214.30	-210.78	1.64
* 49		2,4-dinitroresorcinol	$\text{C}_6\text{H}_4\text{N}_2\text{O}_6$	-415.60	-398.97	4.00
* 50		2,4,6-trinitroaniline	$\text{C}_6\text{H}_4\text{N}_4\text{O}_6$	-74.50	-75.44	1.26
* 51		1,3-diamino-2,4,6-trinitrobenzene	$\text{C}_6\text{H}_5\text{N}_5\text{O}_6$	-97.90	-96.15	1.79
* 52		triaminotrinitrobenzene	$\text{C}_6\text{H}_6\text{N}_9\text{O}_6$	-139.75	-130.83	6.38
* 53		hexanitrobenzene	$\text{C}_6\text{N}_6\text{O}_{12}$	150.60	150.55	0.03
* 54		2,3,6-trinitrobenzoic acid	$\text{C}_7\text{H}_3\text{N}_3\text{O}_8$	-403.00	-422.56	4.85
* 55		3,5-dinitrobenzoic acid	$\text{C}_7\text{H}_4\text{N}_2\text{O}_6$	-432.60	-440.73	1.88
* 56		2,4,6-trinitro-m-cresol	$\text{C}_7\text{H}_5\text{N}_3\text{O}_7$	-255.90	-280.63	9.66
* 57		1-methyl-2,6-dinitrobenzene	$\text{C}_7\text{H}_6\text{N}_2\text{O}_4$	-51.10	-50.98	0.23
* 58		1-methyl-2-nitrobenzene	$\text{C}_7\text{H}_7\text{NO}_2$	-9.70	-9.10	6.19

注: 1) 列出的全部芳香族多硝基化合物常温下为固态; 2) 序号前有 * 者为预估值, 未参与训练; 3) ΔH_f^0 均为标准状态下的生成焓。

4 分子结构描述码的影响

分子结构描述码选择恰当与否对生成焓预估精度有着密切关系。若选择不当,将会引起很大的误差。为此本文所选的分子结构描述码必须能反映物质的结构特征,同时尽量选用最新文献测定的生成焓数据。

ANN方法是一种自然的非线性建模过程,把分子结构描述码当作分子子图对生成焓的数值进行回归,模拟计算结果精度都较高,绝大部分生成焓误差在10%的误差范围内,今后仍需进一步提高计算精度。

参考文献:

- [1] Pedley, J. B., et al. *Thermochemical Data of Organic Compounds* (second edition) [M]. Printed in Great Britain at the University Press, Cambridge, 1986.
- [2] 田德余,刘剑洪. 化学推进剂计算能量学[M]. 郑州:河南科技出版社,1999.
- TIAN De-yu, LIU Jian-hong. *Calculation Energetics of Propellants Chemistry*[M]. Zhengzhou: Henan Science and Technology Press,1999.
- [3] 田德余. 常用化合物性能数据手册[M]. 长沙:湖南科技出版社,1998.
- TIAN De-yu. *Common Compounds Properties Data Handbook*[M]. Changsha: Hunan Science and Technology Press,1988.
- [4] 刘剑洪,田德余,洪伟良,等. 用分子图法计算硝基咪唑化合物的生成热[J]. 化学物理学报,2002,15:351.
- LIU Jian-hong, TIAN De-yu, HONG Wei-liang, et al. Estimation and prediction on heats of formation for nitro furazan series compounds with novel molecular subgraph [J]. *Chin. J. Chem. Phys.*, 2002, 15: 351.
- [5] 刘剑洪,田德余,陈刚,等. 用分子图法计算烷烃热力学函数[J]. 国防科技大学学报,2000,22(2):37.
- LIU Jian-hong, TIAN De-yu, CHEN Gang, et al. Estimation and prediction on thermodynamical functions of alkanes with novel molecular subgraph [J]. *Journal of National University of Defense Technology*, 2000, 22 (2): 37.
- [6] 胡守仁. 神经网络导论[M]. 长沙:国防科技大学出版社,1993.
- HU Shou-ren, *The Guide of Artificial Neural Network* [M]. Changsha: National University of Defense Technology Press,1993.

Prediction of the Enthalpy of Formation for Aromatic Polynitro Compounds with Artificial Neural Network

WANG Fang¹, LIU Jian-hong¹, TIAN De-yu^{1,2}, CHEN Li¹, ZHAO Feng-qi³,
HONG Wei-liang¹, LI Cui-hua¹, ZHANG Xue-li¹

- (1. Normal College, Shenzhen University, Shenzhen 518060, China;
2. National University of Defence Technology, Changsha 410073, China;
3. Xi'an Modern Chemistry Research Institute, Xi'an 710065, China)

Abstract: With the molecular structure describers, the enthalpy of formation for aromatic polynitro compounds are predicted by using an artificial neural network. The influence of neural network parameters and molecular structure describers(MSD) are studied. The correlation coefficient is excellent. The relative errors are almost not more than 10%.

Key words: physical chemistry; enthalpy of formation; aromatic polynitro compound; artificial neural network(ANN); molecular structure describer(MSD)