

# 烯烃的分子距离边数矢量 ( $\mu$ ) 及沸点的估计与预测

## Estimation and Prediction for Normal Boiling Point of Alkenes Based on a Novel Molecular Distance-edge Vector ( $\mu$ )

尹情胜 刘树深 李志良\*

Yin Qingsheng Liu Shushen Li Zhiliang

(桂林工学院应用化学系 桂林市建干路 541004)

(Dept. of Applied Chem., Guilin Institute of Technology, Jianganlu, Guilin, Guangxi, 541004)

**摘要** 基于双键相对于单键键长的大小,提出了烯烃的分子距离边数(MDE)矢量( $\mu$ 矢量);进而通过多元线性回归方法建立 $\mu$ 矢量与烯烃沸点之间的关系模型并进行预测,计算沸点与实验值之间的均方根误差仅为3.78 K,相关系数为0.9993

**关键词** 分子距边矢量 烯烃 沸点 建模

中图法分类号 O 623.123

**Abstract** A novel molecular distance-edge vector (MDE) of alkenes, called  $\mu$  vector, was proposed based on the ratio of single bond's length to double bond's. Multiple linear regression was used to develop a linear model between  $\mu$  vector and the normal boiling point of alkenes with a high precision (mean squares error,  $MSE = 3.78$  K) and a good correlation ( $r = 0.9993$ ).

**Key words** molecular distance-edge vector, alkenes, normal boiling point, modelling

定量构效(性)关系(QSAR/QSPR)研究一直是许多化学家和相关科学家十分关注的课题<sup>[1~3]</sup>。其中分子结构的参数化是十分关键的一步。国内外已经提出了百余种描述分子结构的图论指数<sup>[4]</sup>。一个好的图论指数至少要满足两个基本条件:即对各种化合物特别是异构体具有良好分辨能力和与化合物物理性质的高度相关性。文献[5]基于分子结构的两个基本特征即原子间距和相互连接边数提出了分子距离边数矢量(MDE矢量,称之为 $\lambda$ 矢量),并将其用于烷烃沸点相关,结果表明该MDE矢量对150个烷烃具有良好的分辨能力和与沸点的良好相关性。本文首先以双键键长与单键键长的比值作为通过双键连接的两个碳原子之间的相对距离即边数对边进行着色,并结合两原子相互作用的大小与原子间距离平方成反比规律,提出了描述烯烃分子结构的新MDE矢量,我们称之为 $\mu$ 矢量;继而利用多元线性回归方法拟合 $\mu$ 矢量与烯烃沸点之间的线性关系(关系模型)。结果表明, $\mu$ 矢量与烯烃沸点具有良好的相关性。

## 1 原理

### 1.1 烯烃分子 $\mu$ 矢量计算方法

以4-甲基-1-己烯(其分子碳骨架结构如图1所示)为例来说明 $\mu$ 矢量计算过程

#### 1.1.1 分子中各碳原子的分类与编号

将各碳原子按与之相连的碳原子数分为四类,记为 $C_1, C_2, C_3, C_4$ ,分别表示与1, 2, 3和4个碳原子相连,并对骨架图中各碳原子进行编号(以上标表示),则第 $k$ 个属于 $i$ 类的碳原子表示为 $C_i^k$ 。4-甲基-1-己烯中各碳原子分类及编号如图1

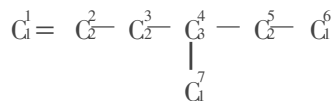


图1 4-甲基-1-己烯分子骨架示意图

Fig 1 4-methyl-1-hexence molecular bone structure

#### 1.1.2 计算每两类(包括同类)碳原子之间的所有距离

用 $d_{ij}^k$ 表示两个碳原子 $C_i^k$ 与 $C_j^l$ 之间的距离。设单键相对键长为1,由于双键键长是单键键长的0.864倍(单键键长为1.54 nm,双键键长为1.33 nm),故双键相对键长为0.864。考虑到一般情况下,相隔4个

1998-02-26收稿,1998-05-18修回。

\* 湖南大学化学化工学院,长沙,410082(College of Chemistry and Chemical Industry, Hunan University, Changsha, Hunan, 410082)

及 4 个以上碳原子的两个碳原子之间的相互作用可以忽略不计, 则上述烯烃的每两类 (包括同类) 碳原子  $C$  和  $C$  之间的所有距离  $D_{ij}$  如下:  $D_{11} = (3.864, 3), D_{12} = (0.864, 1.864, 3.864, 2, 2, 3, 1, 3, 4), D_{13} = (2.864, 1, 2), D_{22} = (1, 3, 2), D_{23} = (1, 2, 1)$  其余各项为 0

### 1.1.3 求两类原子间的相互作用

两类原子间的相互作用与各原子间的距离平方的倒数和成正比, 把上述  $D_{ij}$  各项距离平方后求倒数和, 即得到两类 (包括同类) 碳原子之间的相互作用的相对大小  $M_{ij}$ :

$$M_{ij} = \sum_{k \leq j}^{n_{ij}} \left(\frac{1}{d_{ij}^k}\right)^2 \quad (i = 1, 2, 3, 4; j \geq i) \quad (1)$$

式中,  $n_{ij}$  是  $D_{ij}$  中的距离项数, 称为作用项数。比如,  $D_{12}$  有 9 个距离,  $n_{12} = 9$ , 而  $D_{13}$  有 3 项,  $n_{13} = 3$   $M_{ij}$  中总计有 10 个元素:  $M_{11}, M_{12}, M_{13}, M_{14}, M_{22}, M_{23}, M_{24}, M_{33}, M_{34}, M_{44}$ , 将其组合成矢量即得表征烯烃结构的 MDE 矢量, 称为  $\mu$  矢量, 其元素对应为  $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6, \mu_7, \mu_8, \mu_9, \mu_{10}$  4-甲基-1-己烯的  $\mu$  矢量为:

表 1 本实验所研究的 62 个烯烃编号与名称对照表

Table 1 The number and relative name of alkenes in the study

No	化合物 Compound	No	化合物 Compound
1	丙烯 Propylene	32	3-甲基-1-戊烯 3-methyl-1-pentene
2	1,3-丁二烯 1,3-butadiene	33	乙烯 ethylene
3	1-丁烯 1-butene	34	1-壬烯 1-nonene
4	顺-2-丁烯 cis-2-butene	35	1-十一碳烯 1-undecene
5	异丁烯 isobutene	36	1-十三碳烯 1-tridecene
6	异丙烯 isoprene	37	1-十五碳烯 1-pentdecene
7	1-戊烯 1-pentene	38	1-十七碳烯 1-heptadecene
8	异戊烯 isopentene	39	1-十九碳烯 1-nonadecene
9	2,3-二甲基-1-丁烯 2,3-dimethyl-1-butene	40	1-二十碳烯 1-eicosene
10	2,3-二甲基-2-丁烯 2,3-dimethyl-2-butene	41	顺-2-戊烯 cis-2-pentene
11	2-乙基-1-丁烯 2-ethyl-1-butene	42	顺-3-己烯 cis-3-hexene
12	1-己烯 1-hexene	43	顺-2-庚烯 cis-2-heptene
13	顺-2-己烯 cis-2-hexene	44	顺-3-庚烯 cis-3-heptene
14	2-甲基-1-戊烯 2-methyl-1-pentene	45	顺-2-辛烯 cis-2-octene
15	4-甲基-1-戊烯 4-methyl-1-pentene	46	顺-3-辛烯 cis-3-octene
16	1-庚烯 1-heptene	47	顺-4-辛烯 cis-4-octene
17	1-辛烯 1-octene	48	3-甲基-顺-2-戊烯 3-methyl-cis-2-pentene
18	2,4,4-三甲基-1-戊烯 2,4,4-trimethyl-1-pentene	49	4-甲基-顺-2-戊烯 4-methyl-cis-2-pentene
19	2,4,4-三甲基-2-戊烯 2,4,4-trimethyl-2-pentene	50	2-甲基-1-己烯 2-methyl-1-hexene
20	1-癸烯 1-decene	51	3-甲基-1-己烯 3-methyl-1-hexene
21	1-十二碳烯 1-dodecene	52	4-甲基-1-己烯 4-methyl-1-hexene
22	1-十三碳烯 1-tridecene	53	2-甲基-1-庚烯 2-methyl-1-heptene
23	1-十六碳烯 1-hexadecene	54	2-乙基-1-戊烯 2-ethyl-1-pentene
24	1-十八碳烯 1-octadecene	55	3-乙基-1-戊烯 3-ethyl-1-pentene
25	2-甲基-1-丁烯 2-methyl-1-butene	56	2,3,3-三甲基-1-丁烯 2,3,3-trimethyl-1-butene
26	2-甲基-2-丁烯 2-methyl-2-butene	57	2,3-二甲基-1-己烯 2,3-dimethyl-1-hexene
27	3-甲基-1-丁烯 3-methyl-1-butene	58	2-乙基-1-己烯 2-ethyl-1-hexene
28	2,3-二甲基-1,3-丁二烯 2,3-dimethyl-1,3-butadiene	59	顺-1,3-戊二烯 cis-1,3-pentadiene
29	1,5-己二烯 1,5-hexadiene	60	1,4-戊二烯 1,4-pentadiene
30	2,3-二甲基-1-丁烯 2,3-dimethyl-1-butene	61	反-1,4-己二烯 trans-1,4-hexadiene
31	2-甲基-2-戊烯 2-methyl-2-pentene	62	顺-2,4-己二烯 cis-2,4-hexadiene

$$\mu = (0.178, 1, 3.479, 1, 1.371, 9, 0, 1.361, 1, 2.250, 0, 0, 0, 0)$$

## 1.2 多元线性回归建模

本文利用多元线性回归方法来拟合烯烃  $\mu$  矢量与其沸点 (BP) 之间的相关关系, 设回归模型为:

$$\log(T_{opt} - BP) = b_0 + \sum_{j=1}^{10} (b_j \times \mu_j) \quad (2)$$

式中  $T_{opt}$  为烯烃的最大或极限沸点温度, 由计算确定;  $b_0$  为回归直线在纵轴上的截距;  $b_j$  为结构描述子 (descriptor)  $\mu_j$  的回归系数, 表示该结构描述子对此烯烃沸点的贡献。用此方法得到的回归方程  $\log(T_{opt} - BP) \sim \mu$  来预测烯烃沸点。

## 2 结果

从文献 [6] 所列烃类化合物中选取其中 62 个烯烃 (见表 1) 作为研究对象, 由于  $\mu$  矢量对烯烃的顺反异构体无分辨能力, 故其余的 14 个反式异构体没有选取。但对于沸点来说, 顺反异构体相差很小, 所以用同一组  $\mu$  矢量表征其结构不影响结果。

### 3 讨论

在所研究的 62 个烯烃中,由于每个烯烃分子中都没有同时包含两个 C<sub>4</sub> 原子,故 M<sub>44</sub> 皆为 0,即  $\mu_{10} = 0$ ,则对各烯烃沸点的贡献也均等于 0,因此进行多元回归分析时只取  $\mu$  矢量的前 9 个元素作为烯烃结构的特征参数。

由于烯烃的沸点随着碳原子数目的增加其增大速度变慢,最终趋向于极限沸点温度 ( $T_{opt}$ )。为了反映这一变化规律,进行回归前先对沸点作对数变换,即采用方程 (2) 式为模型。烯烃的极限沸点温度  $T_{opt}$  通过试探法以均方根误差最小为目标经多次计算寻优而得,部分温度下计算的均方根误差 ( $MSE$ ) 见表 2 从表 2 可知,当温度达到 801 K 时  $MSE$  最小,  $MSE = 3.78$  K,故最优化极限温度  $T_{opt} = 801$  K

利用多元线性回归方法建立的 62 个烯烃  $\mu$  矢量与其沸点的回归方程如下:

$$\log(801 - BP) = 2.74472 + 0.040164 \mu_1 - 0.00107026 \mu_2 - 0.01065643 \mu_3 - 0.0205339 \mu_4 - 0.019986 \mu_5 - 0.023904 \mu_6 - 0.032019 \mu_7 - 0.025906 \mu_8 - 0.037472 \mu_9$$

由该方程计算的沸点与原实验值之间的相关系数  $r = 0.9993$ ,均方根误差  $MSE = 3.78$  K,解释方差  $EV = 0.9983$  显然,相关性良好,计算误差小。

表 3 列出本实验所涉及的 62 个烯烃沸点的实验值和由多元回归方程计算的沸点值。如果将计算沸点对实验沸点作图 (图 2),所有点都在直线上,表明计算值与实验值吻合良好。

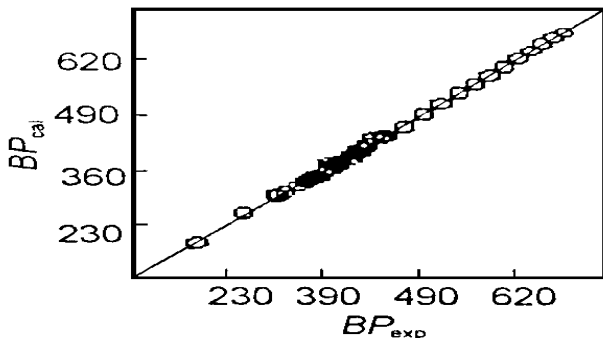


图 2 沸点的计算值对实验值相关图

Fig. 2 Plot of calculated and experimental boiling points

表 2 极限沸点温度与均方根误差的相关关系

Table 2 Relation of  $T_{opt}$  to mean square error

极限沸点温度 $T_{opt}$ (K)	均方根误差 Mean square error (K)	极限沸点温度 $T_{opt}$ (K)	均方根误差 Mean square error (K)
750	4.4085	801	3.7803
780	3.8663	802	3.7806
790	3.8019	803	3.7811
800	3.7804	810	3.8367

表 3 62 个烯烃沸点 ( $BP$ ) 的实验值和计算值<sup>[6]</sup>

Table 3 Experimental and calculated values of normal boiling points ( $BP$ ) of 62 alkenes<sup>[6]</sup>

No.	$BP_{exp}$	$BP_{cal}$	$e(\%)$	No.	$BP_{exp}$	$BP_{cal}$	$e(\%)$
1	225.4	233.7	3.68	32	327.3	327.6	0.09
2	268.7	268.1	-0.22	33	169.4	172.2	1.65
3	266.9	268.2	0.49	34	420.0	417.5	-0.60
4	276.9	276.2	-0.25	35	465.8	464.6	-0.26
5	266.2	273.1	2.59	36	505.9	505.9	0
6	307.2	303.1	-1.33	37	541.6	542.2	0.11
7	303.1	300.8	-0.76	38	573.5	574.0	0.09
8	301.1	298.9	-0.73	39	602.2	601.8	-0.07
9	328.3	324.2	-1.25	40	615.5	614.5	-0.16
10	346.3	344.5	-0.52	41	310.0	309.0	-0.32
11	337.8	332.5	-1.57	42	339.6	343.4	1.12
12	336.6	334.3	-0.68	43	371.6	371.3	-0.08
13	342.0	343.7	0.50	44	368.9	372.5	0.98
14	335.2	334.5	-0.21	45	398.8	398.6	-0.05
15	327.0	323.9	-0.95	46	396.0	399.5	0.88
16	366.8	363.9	-0.79	47	395.7	399.7	1.01
17	394.0	391.6	-0.61	48	340.8	336.9	-1.14
18	374.6	379.6	1.33	49	329.5	338.3	2.67
19	378.1	373.6	-1.19	50	365.0	367.9	0.79
20	443.8	441.9	-0.43	51	357.0	360.2	0.90
21	486.5	486.0	-0.10	52	359.9	360.0	0.03
22	524.2	524.7	0.10	53	392.4	394.7	0.59
23	558.0	558.6	0.11	54	367.1	364.6	-0.68
24	588.0	588.4	0.07	55	357.3	355.8	-0.42
25	304.3	302.1	-0.72	56	351.0	351.0	0
26	311.7	307.9	-1.22	57	383.6	393.3	2.53
27	293.2	292.9	-0.10	58	393.1	394.7	0.41
28	341.6	337.3	-1.26	59	317.2	309.3	-2.49
29	332.6	334.7	0.63	60	299.1	301.1	0.67
30	314.4	313.7	-0.22	61	338.1	342.6	1.33
31	340.4	339.8	-0.18	62	356.6	350.4	-1.74

### 参考文献

- 1 Randic M, Ring I D Numbers. J Chem Inf Comput Sci, 1988, 28: 142.
- 2 Needhan D E, Wei I, Seybold P G. Molecular modeling of the physical properties of the alkanes. J Am Chem Soc, 1988, 110: 4186.
- 3 Miyashita Y, Okuyama T, Ohsako H et al. Graph theoretical approach to carbon-13 chemical shift sum in alkanes. J Am Chem Soc, 1989, 110: 3469.
- 4 Mihhalec Z, N Trinajstic. A graph-theoretical approach to APR, J Chem Educ, 1992, 69: 701.
- 5 Liu S S, Cao C Z, Li Z L. Approach to estimation and prediction for normal boiling points of alkanes based on a molecular distance-edge vector (MDE), J Chem Inf Comput Sci, 1998. (in print)
- 6 Wessel M D, Jurs P C. Prediction for normal boiling points of hydrocarbons from molecular structure, J Chem Inf Comput Sci, 1995, 35: 68-76. (责任编辑:黎贞崇)