

新的价连接性指数与 烃标准熵和燃烧热的相关性研究

堵锡华*

(徐州教育学院化学系, 徐州 221006)

摘要: 定义原子特征值 $\beta_i = (n_i - 1)m_i \pm h_i$ 。由 β_i 建构新的价连接性指数 ${}^m X = \sum (\beta_i \cdot \beta_j \cdot \beta_k \cdots)^{0.5}$, 其中 0 阶指数 ${}^0 X = \sum (\beta_i)^{0.5}$, 1 阶指数 ${}^1 X = \sum (\beta_i \cdot \beta_j)^{0.5}$ 。并计算了 4 个系列 72 个烃分子的 ${}^0 X$ 、 ${}^1 X$ 值。发现 ${}^m X$ 与烃的标准熵、标准燃烧热有良好的相关性, 相关系数均在 0.99 以上。并采用 Jackknife 方法对模型稳健性进行了检验。

关键词: 原子特征值; 价连接性指数; 标准熵; 标准燃烧热; 相关性; 模型稳健性

中图分类号: O657.7, O62 文献标识码: A

1 前言

物质的定量结构—性质相关性(QSPR)一直是计算化学的研究热点^[1-3]。目前一种较为简便的方法就是利用拓扑指数与化合物的理化性质建立 QSPR 模型来对化合物的理化性质参数进行评估和预测, 国内学者辛厚文、江元生、许禄等作了卓有成效的研究^[4-7]。分子连接性指数是 Kier 和 Hall 等人改进了 Randic 提出的分子拓扑指数 χ 而发展起来的 ${}^m \chi_i$ 指数系列^[8], 它在有机分子的构效关系研究中得到了广泛应用。但目前一般的文献方法或是忽略氢原子影响, 或是忽略化学环境影响, 即使在点价中考虑了氢原子的贡献, 也只是采用单一不变模型, 事实上不同系列化合物的不同理化性质, 氢原子对其影响大小也有所不同。本文在邻接矩阵基础上, 把分子图形特征和化学环境有机结合起来, 构建新的价连接性指数 ${}^m X$, 使其与化合物理化性质参数的相关性得到进一步提高。利用修正的 Jackknife 法检验模型的稳健性, 从侧面验证了模型较为可靠。

2 价连接性指数 ${}^m X$ 的建构方法

成键原子的化学行为与其价电子数、主量子数及键连原子密切相关, 据此定义成键原子化学特征的值 β_i :

$$\beta_i = (n_i - 1)m_i \pm h_i \quad (1)$$

式中, n_i 为 i 原子的主量子数即电子层数; m_i 为 i 原子的价电子数; h_i 为与 i 原子连接的氢原子数。当 $\beta_i = (n_i - 1)m_i - h_i$, 即氢原子影响较小时, 相应伯、仲、叔、季碳原子的 β_i 依次为: 1、2、3、4, 它与 Randic 的碳原子支化度^[9]一致; 当 $\beta_i = (n_i - 1)m_i + h_i$, 即氢原子影响较大时, β_i 依次为: 7、6、5、4, 与 Randic 的碳原子支化度负相关($\beta_i = 8 - \delta_i$)。

* 通讯联系人, Email: wlxyl68@pub.xz.jsinfo.net

收稿日期: 2000-08-16; 修回日期: 2001-04-18。

以分子图的邻接矩阵为基础,在 Kier 指数^[8]基础上略加改进,由 β_i 建构新的价连接性指数(mX):

$${}^mX = \sum (\beta_i \cdot \beta_j \cdot \beta_k \cdots)^{0.5} \quad (2)$$

其中 0 阶项(0X)、1 阶项(1X)的计算公式为:

$${}^0X = \sum (\beta_i)^{0.5} \quad (3)$$

$${}^1X = \sum (\beta_i \cdot \beta_j)^{0.5} \quad (4)$$

式(3)表达分子的 0X 等于分子中每一原子的贡献之和;式(4)表达分子的 1X 等于分子中每一化学键的贡献之和。根据 $\beta_i = (n_i - 1)m_i - h_i$ 计算得到 0X_A 、 1X_A ;根据 $\beta_i = (n_i - 1)m_i + h_i$ 计算得到 0X_B 、 1X_B ;例如:2-甲基丙烷,其 0X_A 、 1X_A 的值为:

$${}^0X_A = (3)^{0.5} + (1)^{0.5} \times 3 = 4.732$$

$${}^1X_A = (1 \times 3)^{0.5} \times 3 = 5.196$$

其 0X_B 、 1X_B 的值为:

$${}^0X_B = (5)^{0.5} + (7)^{0.5} \times 3 = 10.173$$

$${}^1X_B = (5 \times 7)^{0.5} \times 3 = 17.748$$

mX 蕴含的物理意义:(1) mX 能反映氢原子在分子中对不同理化性质影响的大小。连接氢原子数越多的原子,对分子的贡献越大,则 β_i 越大,其 X_B 值越大;相反,连接氢原子数越多的原子,对分子的贡献越小,则 β_i 越小,其 X_A 值也越小。不同的理化性质,氢原子对其影响程度必然不同,使用 X_A 、 X_B 即可区别。(2) mX 反映分子的大小。碳原子数越多,成键越多,分子体积越大,相应的 mX 越大。(3) mX 反映分子的形状。同碳原子数的分子,分子形状不同,相应的 1X 也就不同。

3 价连接性指数 mX 与烃热力学性质的相关性

标准熵、标准燃烧热的值引自文献^[10]。把 72 个烃分子分成 4 个系列,用价连接性指数 mX 与其拟合,得线性回归方程。

3.1 mX 与烃标准熵(S_m°)的相关性

$$\text{烷烃类: } S_m^\circ = 16.5046 {}^0X_B + 124.6428 \quad (5)$$

$$n' = 25, S = 19.9594, r = 0.9933, F = 1699.19$$

$$\text{烯烃类: } S_m^\circ = -2.2458 {}^1X_A^{1.5} + 23.6351 {}^0X_B + 111.5839 \quad (6)$$

$$n' = 29, S = 7.1955, r = 0.9916, F = 764.07$$

$$\text{炔烃类: } S_m^\circ = 33.3201 {}^0X_B^{0.8} + 90.4267 \quad (7)$$

$$n' = 7, S = 2.5117, r = 0.9989, F = 2268.98$$

$$\text{芳香烃类: } S_m^\circ = -4619.9810 {}^0X_A + 713.6348 \quad (8)$$

$$n' = 10, S = 3.3505, r = 0.9966, F = 1170.47$$

3.2 mX 与烃标准燃烧热($-\Delta H_c^\circ$)的相关性

$$\text{烷烃类: } -\Delta H_c^\circ = 266.6824 {}^0X_B + 126.7255 \quad (9)$$

$$n' = 23, S = 37.5030, r = 0.9999, F = 104984.25$$

$$\text{烯烃类: } -\Delta H_c^\circ = 263.8012 {}^0X_B + 108.7557 \quad (10)$$

$$n' = 26, S = 39.8117, r = 0.9989, F = 10891.09$$

$$\text{炔烃类: } -\Delta H_c^\circ = 267.8345 {}^0X_B + 90.8336 \quad (11)$$

$$n' = 7, S = 12.8600, r = 0.9999, F = 2273.52$$

$$\text{芳香烃类: } -\Delta H_c^\circ = 266.6903 {}^0X_B - 313.8088 \quad (12)$$

$$n' = 9, S = 1.9889, r = 0.999995, F = 699993.00$$

式中, n' 为样本数; S 为估计标准误差; r 为相关系数; F 为 Fischer 检验值。

由上可见, 价连接性指数 mX 与4个系列烃分子的标准熵、标准燃烧热的相关系数均超过0.99, 特别是对燃烧热的一元相关即可达到高度相关, r 都在0.9999左右, 且均是 X_B 与其相关, 说明氢原子是影响燃烧热的重要因素。将文献值和按式(5)~(12)的计算值列于表1中(22~25号化合物是指环烷烃, 46~49、54号化合物是指环烯烃, 68~72号化合物的方括号是指甲基在苯环上的位置), 文献值与计算值基本吻合。

表1 烃的 mX 和标准熵(S_m°)、标准燃烧热($-\Delta H_c^\circ$)

Table 1 mX and S_m° , $-\Delta H_c^\circ$ of hydrocarbon

No.	Compound	0X_A	1X_A	0X_B	$S_m^\circ/\text{J} \cdot (\text{k} \cdot \text{mol})^{-1}$		$-\Delta H_c^\circ/\text{kJ} \cdot \text{mol}^{-1}$	
					Exp.	Cal.	Exp.	Cal.
Alkane								
1	CH ₄	0.000	0.000	2.828	186	171	889	881
2	CH ₃ CH ₃	2.000	1.000	5.292	230	212	1557	1538
3	CH ₃ CH ₂ CH ₃	3.414	2.828	7.741	270	252	2217	2191
4	CH ₃ (CH ₂) ₂ CH ₃	4.828	4.828	10.190	310	293	2874	2844
5	CH ₃ (CH ₂) ₃ CH ₃	6.243	6.828	12.640	349	333	3509	3498
6	CH ₃ (CH ₂) ₄ CH ₃	7.657	8.828	15.089	388	374	4158	4151
7	CH ₃ (CH ₂) ₅ CH ₃	9.071	10.828	17.540	428	414	4810	4804
8	CH ₃ (CH ₂) ₆ CH ₃	10.485	12.828	19.988	467	455	5464	5457
9	CH ₃ (CH ₂) ₇ CH ₃	11.899	14.828	22.438	506	495	6118	6111
10	CH ₃ (CH ₂) ₈ CH ₃	13.314	16.828	24.887	545	535	6770	6764
11	CH ₃ (CH ₂) ₉ CH ₃	14.728	18.828	27.337	584	576	7424	7417
12	CH ₃ (CH ₂) ₁₀ CH ₃	16.142	20.828	29.786	623	616	8072	8070
13	CH ₃ (CH ₂) ₁₄ CH ₃	21.799	28.828	39.584	778	778	10687	10683
14	CH ₃ (CH ₂) ₁₈ CH ₃	27.456	36.828	49.382	934	940	13300	13296
15	CH ₃ CH(CH ₃)CH ₃	4.732	5.196	10.173	295	293	2865	2840
16	CH ₃ CH(CH ₃)CH ₂ CH ₃	6.146	7.328	12.623	344	333	3498	3493
17	CH ₃ C(CH ₃) ₂ CH ₃	6.000	8.000	12.583	306	332	3512	3482
18	CH ₃ CH(CH ₃)(CH ₂) ₂ CH ₃	7.560	9.328	15.072	381	373	4153	4146
19	CH ₃ CH ₂ CH(CH ₃)CH ₂ CH ₃	7.560	9.459	15.072	380	373	4155	4146
20	CH ₃ C(CH ₃) ₂ CH ₂ CH ₃	7.414	10.243	15.032	358	372	4143	4135
21	CH ₃ CH(CH ₃)CH(CH ₃)CH ₃	7.464	9.928	15.055	366	373	4150	4142
22	CH ₂ CH ₂ CH ₂	4.243	6.000	7.348	237	246		2086

续表1。

No.	Compound	0X_A	1X_A	0X_B	$S_m^\ominus / J \cdot (k \cdot mol)^{-1}$		$-\Delta H_C^\ominus / kJ \cdot mol^{-1}$		
					Exp.	Cal.	Exp.	Cal.	
23	$\underline{CH_2(CH_2)_2}CH_2$	5.657	8.000	9.798	265	286		2740	
24	$\underline{CH_2(CH_2)_3}CH_2$	7.071	10.000	12.247	293	327	3287	3393	
25	$\underline{CH_2(CH_2)_4}CH_2$	8.485	12.000	14.697	298	367	3916	4046	
Alkene									
26	$CH_2 = CH_2$	2.828	2.000	4.899	220	221	1409	1401	
27	$CH_3CH = CH_2$	4.146	4.182	7.331	267	266	2056	2043	
28	$CH_3CH_2CH = CH_2$	5.560	6.313	9.781	306	307	2715	2689	
29	$CH_3CH = CHCH_3$ (cis)	5.464	6.464	9.764	301	305	2708	2685	
30	$CH_3CH = CHCH_3$ (trans)	5.464	6.464	9.764	297	305	2704	2685	
31	$CH_3(CH_2)_2CH = CH_2$	6.975	8.313	12.230	346	347	3346	3335	
32	$CH_3CH_2CH = CHCH_3$ (cis)	6.878	8.596	12.213	346	345	3338	3331	
33	$CH_3CH_2CH = CHCH_3$ (trans)	6.878	8.596	12.213	340	344	3334	3331	
34	$CH_3(CH_2)_3CH = CH_2$	8.389	10.313	14.680	385	384	3999	3981	
35	$CH_3(CH_2)_2CH = CHCH_3$ (cis)	8.293	10.596	14.663	387	381	3988	3977	
36	$CH_3(CH_2)_2CH = CHCH_3$ (trans)	8.293	10.596	14.663	381	381	3986	3977	
37	$CH_3CH_2CH = CHCH_2CH_3$ (cis)	8.293	10.727	14.663	380	379	3993	3977	
38	$CH_3CH_2CH = CHCH_2CH_3$ (trans)	8.293	10.727	14.663	375	379	3986	3977	
39	$CH_3(CH_2)_4CH = CH_2$	9.803	12.313	17.129	424	419	4652	4627	
40	$CH_3(CH_2)_5CH = CH_2$	11.217	14.313	19.579	463	453	5305	5274	
41	$CH_3C(CH_3) = CH_2$	5.414	6.828	9.741	294	302	2698	2678	
42	$CH_3CH_2C(CH_3) = CH_2$	6.828	9.071	12.190	340	338	3331	3324	
43	$CH_3CH(CH_3)CH = CH_2$	6.878	8.914	12.213	334	340	3364	3331	
44	$CH_3CH = C(CH_3)CH_3$	6.732	9.196	12.173	339	337	3323	3320	
45	$CH_3C(CH_3) = C(CH_3)CH_3$	8.000	12.000	14.583	365	363	3970	3956	
46	$\underline{CH_2CH = CH}$	4.878	7.899	6.922	244	225		1935	
47	$\underline{C H_2 C H_2 C H = C H}$	6.293	9.899	9.371	264	263		2581	
48	$\underline{C H_2 (C H_2)_2 C H = C H}$	7.707	11.899	11.821	290	299	3112	3227	
49	$\underline{C H_2 (C H_2)_3 C H = C H}$	9.121	13.899	14.270	311	332	3747	3873	
50	$CH_2 = C = CH_2$	4.828	5.657	6.899	244	244	1942	1929	
51	$CH_3CH = C = CH_2$	6.146	8.025	9.331	293	281	2591	2570	
52	$CH_2 = CHCH = CH_2$	6.293	7.899	9.371	279	283	2539	2581	
53	$CH_2 = CHC(CH_3) = CH_2$	7.560	10.742	11.781	316	311	3158	3217	
54	$\underline{C H_2 C H = C H C H = C H}$	8.342	13.899	11.394	268	265		3115	
Ethyne									
55	$CH \equiv CH$	3.464	3.000	4.472	201	201	1299	1289	
56	$CH_3C \equiv CH$	4.732	5.464	6.882	248	246	1936	1934	

续表1。

No.	Compound	0X_A	1X_A	0X_B	$S_m^\circ/\text{J} \cdot (\text{k} \cdot \text{mol})^{-1}$		$-\Delta H_c^\circ/\text{kJ} \cdot \text{mol}^{-1}$	
					Exp.	Cal.	Exp.	Cal.
57	$\text{CH}_3\text{CH}_2\text{C}=\text{CH}$	6.146	7.707	9.331	291	289	2594	2590
58	$\text{CH}_3\text{C}=\text{CCH}_3$	6.000	8.000	9.292	283	289	2549	2580
59	$\text{CH}_3(\text{CH}_2)_2\text{C}=\text{CH}$	7.560	9.707	11.781	330	330	3252(g)	3246
60	$\text{CH}_3\text{CH}_2\text{C}=\text{CCH}_3$	7.414	10.243	11.741	332	329	3236(g)	3235
61	$\text{CH}_3(\text{CH}_2)_3\text{C}=\text{CH}$	8.975	11.707	14.230	369	369	3910(g)	3902
Arene								
62	C_6H_6	10.392	18.000	13.416	269	269	3265	3264
63	$\text{C}_6\text{H}_5\text{CH}_3$	11.660	20.928	15.826	321	317	3907	3907
64	$\text{C}_6\text{H}_5\text{CH}_2\text{CH}_3$	13.074	23.171	18.276	361	360	4561	4560
65	$\text{C}_6\text{H}_5\text{CH}_2\text{CH}_2\text{CH}_3$	14.488	25.171	20.725	401	395	5214	5213
66	$\text{C}_6\text{H}_5\text{CH}(\text{CH}_3)\text{CH}_3$	14.392	25.856	20.708	389	393	5213	5209
67	$\text{C}_6\text{H}_5\text{C}(\text{CH}_3)_2\text{CH}_3$	15.660	28.928	23.118		419		5852
68	$\text{C}_6\text{H}_4(\text{CH}_3)_2[1,2]$	12.928	23.928	18.236	353	356	4549	4550
69	$\text{C}_6\text{H}_4(\text{CH}_3)_2[1,3]$	12.928	23.856	18.236	358	356	4547	4550
70	$\text{C}_6\text{H}_4(\text{CH}_3)_2[1,4]$	12.928	23.856	18.236	352	356	4549	4550
71	$\text{C}_6\text{H}_3(\text{CH}_3)_3[1,3,5]$	14.196	26.785	20.645	385	388	5189	5192
72	$\text{C}_6\text{H}_2(\text{CH}_3)_4[1,2,3,4]$	15.464	29.928	23.055	417	415		5835

4 结果讨论

4.1 mX 具有较高的结构区分能力

每种原子的电子层数(n_i)、价电子数(m_i)以及它在分子中所连接的氢原子数(h_i)通常不会完全相同,由式(1)计算的 β_i 也不会相同。即 β_i 对成键原子、特别是杂原子的结构差异具有较强的结构差异区分能力,成为表征原子结构特征的结构参数。有机分子中所含成键原子的种类和数目也不会相同,按同一数学规则—式(3)、(4)计算的 0X 、 1X 值通常不同,特别是 1X 对有机分子呈现出很好的结构选择性,本文的72个烃分子的 1X 不存在简并现象便是证明。

4.2 mX 揭示了影响有机物热力学性质的本质因素

考察有机分子的热力学性质与其结构的关系发现:(1)同系列有机物,其中碳原子数及杂原子数越多,分子体积越大,相应标准熵、标准燃烧热的值越大,其 0X 、 1X 值也越大。而且化学环境对分子的影响已蕴含在式(3)、(4)的“ Σ ”之中。(2)同碳数不同系列有机物,其标准熵、标准燃烧热的值与碳原子的 n_i 、 m_i 及所连接的氢原子数正相关,相应的 0X_B 、 1X_B 值也越大。例如烷烃的 $-\Delta H_c^\circ$ 大于相应烯烃的 $-\Delta H_c^\circ$ 大于相应炔烃的 $-\Delta H_c^\circ$ 值,相应的 0X_B 、 1X_B 值也逐渐变小,这与式(9)~(12)中 0X_B 前的系数为正是一致的,说明氢原子的影响较大。但对于标准熵(S_m°),它取决于质点运动的程度,故有时连接氢原子数越多的原子对分子的影响较大,有时则较小,这从 S_m° 值可以看出,在式(5)~(8)中也体现出这一点。 mX 与标准燃烧热

的相关程度更高,且均与 X_B 相关,说明燃烧热受氢原子影响较大,是重要影响因素。而 X_B 则正好符合这一点,它充分考虑了氢原子在分子中的影响。

4.3 mX 蕴含丰富的结构信息

mX 不但能反映有机分子的尺度大小,而且与分子支化度密切相关。碳原子连接氢原子数越少,其支化度越大, S_m° 、 $-\Delta H_c^\circ$ 越小,相应 0X_B 、 1X_B 值也越小。如:2-甲基丁烷(344、3498)、戊烷(349、3509), 0X_B 、 1X_B 也逐渐增大。

4.4 模型稳健性检验

以具有代表性的链烷烃为例,用 Jackknife 法^[11]对回归方程(5)的稳健性进行检验。从 25 种化合物中依次剔除 1 种化合物,对其余 24 种化合物用 mX 与 S_m° 进行回归,重复 25 次共得到 25 个 Jackknife 相对误差检验值,图为其频数分布,其中正负相对误差最大的化合物是环己烷(估算值 $S_m^\circ = 370$)和甲烷($S_m^\circ = 169$),平均相对误差为 5.03。Jackknifed 法 r 变化总趋势为:剔除结构简单且 S_m° 值较大的化合物,模型的 r 减小;剔除结构复杂且 S_m° 值较大的环烷烃化合物,模型的 r 增大;说明在所研究的化合物中,高 S_m° 化合物对模型稳健性影响较大,故大多数化合物呈现出对回归模型的系统负偏。从图可以看出,它们基本上呈正态分布,所建模型具有可接受的总体稳健性。对于其他回归方程,用同样方法检验,基本与此类似,说明笔者提出的价连接性指数用于估算烃类化合物的 S_m° 、 $-\Delta H_c^\circ$ 是合理的。用 Jackknife 法对其他方程的稳健性进行检验,基本类似。

本文的 mX 基本上满足了拓扑指数要有较低的简并度和良好的相关性这两条要求,说明本文建立的 mX 是合理的,可用于各系列有机物其他性质的相关性研究^[12]。与一般文献相比,本文还具有计算简单,所需参数少,应用方便等特点。

参 考 文 献

- [1] Zhang Hongguang (张宏光), Xin Houwen (辛厚文). *Chin. J. Chem. Phys.* (化学物理学报), 1989, 2: 413
- [2] Li Linfeng (李林峰), You Xiaozeng (游效曾). *Chin. Sci. Bull.* (科学通报), 1993, 38: 421
- [3] Yang Feng (杨峰), Luo Mingdao (罗明道), Qu Songsheng (屈松生). *Chin. J. Chem. Phys.* (化学物理学报), 1999, 12: 747
- [4] Xin Houwen (辛厚文). *Topology of Molecules* (分子拓扑学), University of Science and Technology of Chinese Press (中国科学技术大学出版社), Hefei (合肥), 1991: 1
- [5] Yang Jiaan (杨家安), Jiang Yuansheng (江元生). *Acta Chimica Sinica* (化学学报), 1983, 41(10): 15
- [6] Feng Changjun (冯长君). *Chin. J. Chem. Phys.* (化学物理学报), 2000, 13: 66
- [7] Yao Yuyuan (姚瑜元), Xu Lu (许禄), Yuan Xiushun (袁秀顺). *Acta Chimica Sinica* (化学学报), 1993, 51: 1041
- [8] Kier L B, Hall H. *Molecular Connectivity in Chemistry and Drug Research*, New York, Academic Press, 1976: 82

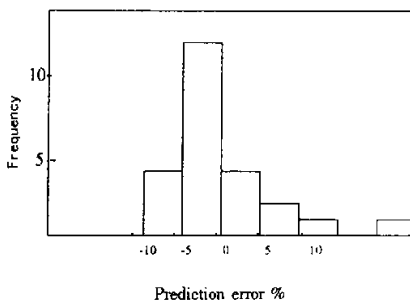


图1 Jackknife 检验相对误差分布
Fig. 1 Distribution of Jackknifed relative error with removing one chemical at one time

- [9] Randic M. *J. Amer. Chem. Soc.*, 1975, **97**: 6609
- [10] Aylward G H, Findlay T J V. Translated by Zhou Ninghui(周宁怀译). *SI Chemical Data(SI 化学数据表)*, Higher Education Press(高等教育出版社). Beijing(北京), 1985: 72
- [11] Dietrich W S, Dreyer N D, Hansch C. *J. Med. Chem.*, 1980, **23**: 1201
- [12] Du Xihua (堵锡华), Feng Changjun (冯长君). *Chin. J. Anal. Chem.* (分析化学), 2001, **29**: 348

Studies on the Relationship between Novel Connectivity Index ${}^m X$ and S_m^Φ or $-\Delta H_C^\Phi$ of Hydrocarbon

Du Xihua

(Department of Chemistry, Xuzhou Education College, Xuzhou 221006)

Abstract Atomic characteristic value (β_i) is defined as: $\beta_i = (n_i - 1) m_i \pm h_i$. Thenovel connectivity index (${}^m X$) of atomic characteristic are set up with the β_i , and ${}^m X$ being defined as ${}^0 X = \sum (\beta_i)^{0.5}$, ${}^1 X = \sum (\beta_i \cdot \beta_j)^{0.5}$. The ${}^0 X$ (${}^1 X$) values of 72 molecules for 4 series of hydrocarbon are calculated. It is found that X was highly correlated with the standard entropies and standard heat of combustion for these compounds. The results show that all the correlation coefficients are larger than 0.99. It has been demonstrate that the method is easy to compute, clear in physical meaning and convenient in applications. A modified jackknife test was performed to validate model robustness.

Key word Atomic characteristic value, Connectivity index, Standard entropies, Standard heat of combustion, Correlativity, Robustness