Approach to Estimation and Prediction for Normal Boiling Point (NBP) of Alkanes Based on a Novel Molecular Distance-Edge (MDE) Vector, λ

Shushen Liu,[†,‡] Chenzhong Cao,[†,§] and Zhiliang Li*[,†,⊥]

Department of Chemistry and Chemical Engineering, Institute of Chemometrics and Pharmacy (ICP), Hunan University, Changsha 410082 P. R. China, Department of Applied Chemistry, Guilin Institute of Technology, Guilin, 541004 P. R. China, Department of Chemistry, Xiangtan Teacher's College, Xiangtan, 411100 P. R. China, and Department of Modern Chemistry, University of Science and Technology of China (USTC), Heifei, 230026 P. R. China

Received December 18, 1996

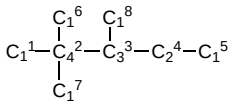
Models that estimate and predict the normal boiling point (NBP) of alkanes based on a molecular distance-edge (MDE) vector, λ, have been developed by using multiple linear regression (MLR) methods. The structures of the examined compounds are selectively described by an MDE vector structure descriptor, a novel molecular distance-edge vector recently developed in our laboratory. MLR was used to develop a linear model containing ten variables with a high precision root mean squares error (RMS = 4.985K) and a good correlation with the correlation coefficient (R = 0.9948). In addition, a predictive model has been developed by using 125 isomers in alkanes as the training set, and its performance was certified by employing 25 alkanes chosen randomly as the test set from a total of 150 alkane compounds; excellent predicted results were obtained with the RMS and R values found between the calculated value and observed NBP being RMS = 4.486K and R = 0.9945.

INTRODUCTION

The use of graph-theoretical invariants in quantitative structure−activity/property relationships (QSAR/QSPR) studies[1−6] has become a major interest to many chemists and related scientists in recent years; several excellent review papers on graph-theoretical indices have been published.[7−9] However, most of the existing graph-theoretical indices are related either to a vertex adjacency matrix, known simply as an adjacency matrix, or to a distance matrix in a molecular graph. As a suggestion proposed by Randić,[10] a new graph-theoretical index should at least have both good discrimination for diverse isomers and high correlation between the index and physical property/activity of the compound. This condition is frequently difficult to fulfil. In this paper, it was found that a novel molecular distance-edge (MDE) vector is based on the two most fundmental structural variables, one for distance between atoms in the molecular graph and another for edges of the adjacency in the graph.[11] The novel MDE vector has not only a good discrimination with no redundancy for 150 alkane isomers containing carbon atom numbers from one through ten but also a high correlation between the MDE vector and the logarithms of the normal boiling point (NBP) of these compounds with a correlation coefficient of R = 0.9948 and a root mean squares error of RMS = 4.985K, which shows quite good results.

The development of a model to predict boiling points as well as other physical properties from the structural features of a molecule is of high importance and utility.[12] Several references[13−15] were made to investigations regarding the relationship between normal boiling point (NBP) and molecular structure descriptors. The calculated boiling point value based on the QSPR model or relationship can also be used to estimate other properties such as critical temperature.[16] In these approaches, one of the most primary and difficult problems is selection of a few descriptors from a great variety of indices or parameters pools; this process is very tedious and time consuming. In the present work, in the MDE vector ten elements regarded as structural descriptors are utilized for describing the examined series of molecules and for indirectly creating an QSPR model without any variable selection. Thus, the developed method can be regarded as a simple but powerful technique for structural parameterization and QSPR modeling.

Chart 1. Graph Representing the Carbon Skeleton of 2,2,3-Trimethylpentane[a]



[a] The superscript, p, is a sequential number, and the subscript, q, is the number of C−C bonds connected with other carbon atoms in notation C_q^p.

THEORETICAL SECTION

Taking 2,2,3-trimethylpentane, shown in Chart 1, for example, the procedure of creating a molecular distance-edge vector in a molecular graph is illustrated as follows.

Based on connecting C−C bond numbers between two atoms, the carbon atoms in all alkanes can be classified as four basic types: type 1, 2, 3, and 4 for primary (CH_3−), secondary (CH_2<), ternary (-CH<), and quaternary (>C<) carbons, symbolized as C_1, C_2, C_3, and C_4 in Chart 1, which

[†] Hunan University.
[‡] Guilin Institute of Technology.
[§] Xiangtan Teachers' College.
[⊥] University of Science and Technology of China (USTC).

S0095-2338(97)00109-1 CCC: $15.00
Published on Web 03/07/1998

Table 1. The Value of MDE Vector for 150 Alkanes

no.	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{22}	λ_{23}	λ_{24}	λ_{33}	λ_{34}	λ_{44}
1	0	0	0	0	0	0	0	0	0	0
2	1.0000	0	0	0	0	0	0	0	0	0
3	0.5000	2.0000	0	0	0	0	0	0	0	0
4	0.3333	2.8284	0	0	1.0000	0	0	0	0	0
5	1.5000	0	3.0000	0	0	0	0	0	0	0
6	0.2500	3.3019	0	0	2.3811	0	0	0	0	0
7	1.1447	1.8899	2.3811	0	0	1.0000	0	0	0	0
8	3.0000	0	0	4.0000	0	0	0	0	0	0
9	0.2000	3.6144	0	0	3.9654	0	0	0	0	0
10	0.9449	2.9417	2.0801	0	1.0000	1.4142	0	0	0	0
11	0.9086	3.3019	1.8899	0	0.5000	2.0000	0	0	0	0
12	2.4495	2.3784	0	3.3636	0	0	1.0000	0	0	0
13	2.2894	0	5.6569	0	0	0	0	1.0000	0	0
14	0.1667	3.8385	0	0	5.6762	0	0	0	0	0
15	0.8143	3.6397	1.8899	0	2.3811	1.6510	0	0	0	0
16	0.7663	4.1123	1.6510	0	1.6510	2.3811	0	0	0	0
17	2.1213	3.7467	0	3.0393	1.0000	0	1.4142	0	0	0
18	1.9442	1.9420	4.9310	0	0	1.4142	0	1.0000	0	0
19	1.8899	2.0000	4.6188	0	0	2.0000	0	0.5000	0	0
20	2.0396	4.2983	0	2.8284	0.5000	0	2.0000	0	0	0
21	0.7500	4.3267	1.5000	0	1.5000	3.0000	0	0	0	0
22	3.9203	0	3.2988	3.7893	0	0	0	0	1.0000	0
23	0.1429	4.0083	0	0	7.4734	0	0	0	0	0
24	0.7211	4.1460	1.7544	0	3.9654	1.8072	0	0	0	0
25	0.6694	4.6538	1.5000	0	3.1473	2.5558	0	0	0	0
26	0.6552	4.8137	1.4422	0	2.9417	2.8284	0	0	0	0
27	1.8974	4.6696	0	2.8284	2.3811	0	1.6510	0	0	0
28	1.7203	3.1506	4.5217	0	1.0000	2.1491	0	1.0000	0	0
29	1.6398	3.4867	4.0859	0	0.5000	3.0393	0	0.5000	0	0
30	1.6287	3.2660	4.0000	0	1.0000	2.8284	0	0.3333	0	0
31	1.7855	5.4216	0	2.5558	1.6510	0	2.3811	0	0	0
32	1.6688	3.6144	4.2983	0	0.3333	2.8284	0	1.0000	0	0
33	0.6463	4.9473	1.3104	0	2.9417	3.3636	0	0	0	0
34	3.4532	2.2516	2.8717	3.4941	0	1.0000	0.5000	0	1.0000	0
35	3.2988	2.5000	2.5864	3.2220	0	1.0000	1.0000	0	0.5000	0
36	3.4178	2.4418	3.0418	3.2988	0	0.5000	1.0000	0	1.0000	0
37	3.2220	0	8.4810	0	0	0	0	2.3811	0	0
38	1.6837	3.5095	4.2983	0	0.5000	2.8284	0	1.0000	0	0
39	1.7321	5.8259	0	2.3784	1.5000	0	3.0000	0	0	0
40	5.8804	0	0	8.4853	0	0	0	0	0	1.0000
41	0.1250	4.1420	0	0	9.3333	0	0	0	0	0
42	0.6507	4.5342	1.6510	0	5.6762	1.9193	0	0	0	0
43	0.5984	5.0478	1.3925	0	4.8324	2.6481	0	0	0	0
44	0.5778	5.2865	1.3104	0	4.5088	3.0418	0	0	0	0
45	1.7321	5.3458	0	2.6750	3.9654	0	1.8072	0	0:	0
46	1.5598	4.0083	4.2420	0	2.3811	2.6207	0	1.0000	0	0
47	1.4708	4.4359	3.7771	0	1.6510	3.5328	0	0.5000	0	0
48	1.4325	4.4992	3.5671	0	1.6510	3.7798	0	0.3333	0	0
49	1.4422	4.1602	3.5777	0	2.3811	3.3019	0	0.2500	0	0
50	1.6079	6.1732	0	2.3784	3.1473	0	2.5558	0	0	0
51	1.4867	4.6538	3.9415	0	1.3104	3.5328	0	1.0000	0	0
52	1.4325	4.8137	3.6144	0	1.1906	4.1602	0	0.5000	0	0
53	1.5737	6.4465	0	2.3094	2.9417	0	2.8284	0	0	0
54	0.5724	5.3742	1.1906	0	4.6105	3.4941	0	0	0	0
55	0.5646	5.4742	1.1447	0	4.5566	3.7893	0	0	0	0
56	3.1380	3.7011	2.6481	3.2988	1.0000	1.4142	0.8165	0	1.0000	0
57	2.9626	4.1798	2.2516	3.0418	0.5000	2.0000	1.1547	0	0.5000	0
58	2.8854	4.0825	2.1764	2.8717	1.0000	1.4142	1.4142	0	0.3333	0
59	3.0813	3.9667	2.8717	3.0418	1.0000	0.8165	1.4142	0	1.0000	0
60	2.8750	2.0069	7.7323	0	0	1.6510	0	2.3811	0	0
61	2.7821	2.1257	7.2429	0	0	1.4142	0	1.0000	0	0
62	2.9091	4.4798	2.4418	2.8049	0.5000	1.1547	2.0000	0	0.5000	0
63	3.0257	4.4273	2.6481	13.0418	0.3333	1.4142	1.4142	0	1.0000	0
64	1.5059	4.5145	3.9415	0	1.6510	3.5328	0	1.0000	0	0
65	1.4510	4.6696	3.6144	0	1.5000	4.1602	0	0.5000	0	0
66	1.5326	6.7491	0	2.1491	2.9417	0	3.3636	0	0	0
67	1.4609	4.9473	3.7467	0	1.1447	4.2426	0	1.0000	0	0
68	5.2594	2.7495	0	7.7429	0	0	1.4142	0	0	1.0000
69	4.9654	0	6.4474	3.7063	0	0	0	1.0000	1.4142	0
70	4.9482	3.0000	0	6.9282	0	0	2.0000	0	0	0.5000
71	5.0219	0	6.6039	3.7798	0	0	0	0.5000	2.0000	0
72	3.0779	4.1524	2.5000	3.2220	0.5000	1.0000	2.0000	0	1.0000	0
73	3.0062	4.5032	2.8049	2.8717	0.5000	1.0000	2.0000	0	1.0000	0
74	2.8717	2.0762	7.6718	0	0	1.8899	0	2.3811	0	0
75	1.5000	7.0191	0	2.0000	3.0000	0	4.0000	0	0	0

Table 1 (Continued)

no.	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{22}	λ_{23}	λ_{24}	λ_{33}	λ_{34}	λ_{44}
76	0.1111	4.2504	0	0	11.2406	0	0	0	0	0
77	0.5953	4.8434	1.5683	0	7.4734	2.0041	0	0	0	0
78	0.5437	5.3503	1.3104	0	6.6320	2.7016	0	0	0	0
79	0.5200	5.6297	1.2164	0	6.2390	3.1473	0	0	0	0
80	0.5130	6.6731	1.1906	0	6.1205	3.3019	0	0	0	0
81	1.6036	5.8677	0	2.5558	5.6762	0	1.9193	0	0	0
82	1.4373	4.6598	4.0324	0	3.9654	2.9557	0	1.0000	0	0
83	1.3462	5.1191	3.5616	0	3.1473	3.8331	0	0.5000	0	0
84	1.2970	5.3066	3.3144	0	2.9417	4.2420	0	0.3333	0	0
85	1.2805	5.2253	3.2068	0	3.1473	4.1800	0	0.2500	0	0
86	1.3014	4.8342	3.2660	0	3.9654.	3.6144	0	0.2000	0	0
87	1.4747	6.7186	0	2.2494	4.8324	0	2.6481	0	0	0
88	1.3544	5.3811	3.6977	0	2.7016	3.9415	0	1.0000	0	0
89	1.2909	5.6608	3.3413	0	2.4069	4.6873	0	0.5000	0	0
90	1.2659	5.6193	3.1811	0	2.5251	4.7568	0	0.3333	0	0
91	1.4238	7.1354	0	2.1491	4.5088	0	3.0418	0	0	0
92	1.3308	5.6193	3.6144	0	2.4069	4.2983	0	1.0000	0	0
93	0.5165	5.6897	1.1052	0	6.4099	3.5328	0	0	0	0
94	0.5047	5.8442	1.0400	0	6.3325	3.9654	0	0	0	0
95	2.9054	4.7487	2.5000	3.1548	2.3811	1.6510	1.0400	0	1.0000	0
96	2.7253	5.2909	2.0762	2.9091	1.6510	2.3811	1.3104	0	0.5000	0
97	2.6233	5.4548	1.8946	2.7464	1.6510	2.3811	1.5000	0	0.3333	0
98	2.5864	5.2002	1.9037	2.6265	2.3811	1.6510	1.6510	0	0.2500	0
99	2.8413	5.0478	2.7464	2.8717	2.3811	1.0400	1.6510	0	1.0000	0
100	2.6340	3.3416	7.2741	0	1.0000	2.6207	0	2.3811	0	0
101	2.5190	3.6860	6.6843	0	0.5000	3.5328	0	1.6510	0	0
102	2.4735	3.5540	6.4429	0	1.0000	3.5328	0	1.3104	0	0
103	2.6481	5.7331	2.3352	2.5864	1.6510	1.3104	2.3811	0	0.5000	0
104	2.5028	3.8385	6.6320	0	0.3333	3.7798	0	1.6510	0	0
105	2.4418	3.7893	6.2929	0	0.5000	4.1602	0	1.1906	0	0
106	2.5654	5.7984	2.0814	2.5000	1.6510	1.5000	2.3811	0	0.3333	0
107	2.7608	5.6882	2.4418	2.8717	1.3104	1.8899	1.6510	0	1.0000	0
108	2.6233	6.0042	2.1257	2.6481	1.1906	2.0801	2.0801	0	0.5000	0
109	2.7430	5.7984	2.5000	2.8049	1.3104	1.6510	1.8899	0	1.0000	0
110	2.5759	3.8385	7.0497	0	0.2500	3.3019	0	2.3811	0	0
111	1.3747	5.2331	3.6977	0	3.1473	3.9415	0	1.0000	0	0
112	1.3156	5.4609	3.3413	0	2.9417	4.6873	0	0.5000	0	0
113	1.2849	5.4648	3.1811	0	2.9417	4.7568	0	0.3333	0	0
114	1.3896	7.5011	0	2.0000	4.6105	0	3.4941	0	0	0
115	1.3156	,5.8061	3.4358	0	2.4495	4.9310	0	1.0000	0	0
116	1.2762	5.8532	3.1974	0	2.3811	5.2987	0	0.5000	0	0
117	1.3104	5.8532	3.4358	0	2.3348	4.9310	0	1.0000	0	0
118	1.3654	7.5829	0	1.9420	4.5566	0	3.7893	0	0	0
119	0.5000	5.9078	1.0900	0	6.3771	4.2426	0	0	0	0
120	1.3561	5.4208	3.6144	0	2.9417	4.2983	0	1.0000	0	0
121	4.8406	4.5145	0	7.3084	1.0000	0	2.1491	0	0	1.0000
122	4.5342	2.2255	5.8834	3.5328	0	1.4142	0.3333	1.0000	1.4142	0
123	4.3705	2.4495	5.4216	3.3674	0	2.0000	0.5000	0.5000	1.1547	0
124	4.4831	5.1678	0	6.3844	0.5000	0	3.0393	0	0	0.5000
125	4.2875	2.6207	5.4216	3.1473	0	1.4142	1.0000	1.0000	0.8165	0
126	4.3281	4.8990	0	6.0000	1.0000	0	2.8284	0	0	0.3333
127	4.5662	2.3348	6.0856	3.5328	0	1.1547	0. 5000	0.5000	2.0000	0
128	4.3824	2.6207	5.6079	3.3019	0	1.4142	1.0000	0.3333	1.4142	0
129	4.4794	2.4980	6.0856	3.3019	0	0.8165	1.0000	1.0000	1.4142	0
130	4.2655	0	11.4880	0	0	0	0	3.9654	0	0
131	4.7243	5.2415	0	7.0656	0.3333	0	2.8284	0	0	1.0000
132	2.8150	5.4036	2.3053	3.0418	1.6510	2.3811	1.3104	0	1.0000	0
133	2.6922	5.5954	1.9601	2.8717	1.5000	3.0000	1.4422	0	0.5000	0
134	2.7318	5.8142	2.6481	2.6481	1.6510	1.3104	2.3811	0	1.0000	0
135	2.5958	3.7011	7.0497	0	0.5000	3.3019	0	2.3811	0	0
136	2.5791	3.8542	6.9946	0	0.3333	3.5328	0	2.3811	0	0
137	2.5958	6.1371	2.3053	2.4418	1.5000	1.4422	3.0000	0	0.5000	0
138	2.5119	3.8091	6.6035	0	0.5000	3.9654	0	1.6510	0	0
139	2.7142	6.0891	2.3053	2.8049	1.1447	2.3811	1.8899	0	1.0000	0
140	2.6825	6.2560	2.4418	2.6481	1.1447	1.8899	2.3811	0	1.0000	0
141	1.3416	7.7742	0	1.8072	4.7451	0	4.3528	0	0	0
142	1.2927	6.0787	3.2660	0	2.2894	5.6569	0	1.0000	0	0
143	2.6265	3.4532	7.1862	0	1.0000	2.9417	0	2.3811	0	0
144	7.1011	0	3.5860	8.4616	0	0	0	0	1.4142	1.0000
145	6.9481	0	3.8643	7.9185	0	0	0	0	2.0000	0.5000
146	4.7411	5.1396	0	7.0656	0.5000	0	2.8284	0	0	1.0000
147	4.5113	2.4019	5.8834	3.4641	0	1.4142	0.5000	1.0000	1.4142	0
148	4.5270	2.5698	6.1723	3.3674	0	1.0000	1.0000	0.5000	2.0000	0
149	2.6794	6.2287	2.5864	2.5000	1.5000	1.5000	3.0000	0	1.0000	0
150	4.3076	0	11.6518	0	0	0	0	4.2426	0	0

Table 2. Experimental Normal Boiling Point for Cl−C10 Alkanes and Predicted Value Based on Multiple Linear Regression for MDE Vector

no.	compound name	NBP^{e} [a]	NBP^{M1}	NBP^{M2}	NBP^{c1} [b]	NBP^{c2} [b]	NBP^{t1} [c]	NBP^{t2} [c]
1	methane	−164.00	−129.93		−89.28		−137.99	
2	ethane	−88.60	−90.78	−94.54	−91.68	−99.19	−88.90	−90.61
3	propane	−42.10	−39.51	−40.71	−39.01	−40.45	−36.86	−38.30
4	butane	−0.50	3.71	2.86	4.34	3.36	6.45	4.98
5	2-methylpropane	−11.70	−15.64	−12.32	−16.41	−12.63	−18.86[d]	−15.24[d]
6	pentane	36.10	41.72	41.06	42.49	41.76	44.24	43.00
7	2-methylbutane	27.80	23.72	25.37	23.03	24.76	22.67	24.08
8	2,2-dimethylpropane	9.50	−15.49		−23.33		−13.62	
9	hexane	69.00	74.98	74.50	75.71	75.20	77.16[d]	76.26[d]
10	2-methylpentane	60.30	60.81	61.76	60.46	61.46	60.65	61.25
11	3-methylpentane	63.30	60.25	60.92	59.62	60.32	60.31	60.68
12	2,2-dimethylbutane	49.70	34.43		32.64		35.49	
13	2,3-dimethylbutane	58.00	54.85	57.25	55.18	57.83	52.45	54.85
14	heptane	98.40	103.84	103.55	104.53	104.22	105.67[d]	105.13[d]
15	2-methylhexane	90.00	92.96	93.57	92.79	93.43	93.17	93.49
16	3-methylhexane	92.00	91.15	91.42	90.73	91.01	91.64	91.65
17	2,2-dimethylpentane	79.20	75.25	73.12	74.90	72.37	76.07[d]	72.70[d]
18	2,3-dimethylpentane	89.80	86.59	87.80	86.81	88.04	85.74	86.74
19	2,4-dimethylpentane	80.50	82.48	83.69	82.53	83.84	80.89[d]	82.11[d]
20	3,3-dimethylpentane	86.10	77.40	75.91	76.87	75.07	77.54	75.48
21	3-ethylpentane	93.50	94.09	94.51	93.68	94.12	94.52	94.79
22	2,2,3-trimethylbutane	80.90	74.15	73.11	73.45	71.85	73.99	71.70
23	octane	125.70	128.72	128.60	129.31	129.18	130.21	130.01
24	2-methylheptane	117.60	120.39	120.83	120.31	120.78	120.74	121.00
25	3-methylheptane	118.00	118.19	118.28	117.93	118.03	118.83	118.75
26	4-methylheptane	117.70	118.54	118.62	118.26	118.35	119.21	119.12
27	2,2-dimethylhexane	106.80	108.21	106.75	108.33	106.77	108.87	106.59
28	2,3-dimethylhexane	115.60	116.59	117.33	116.86	117.61	116.42[d]	116.92[d]
29	2,4-dimethylhexane	109.40	110.93	111.32	110.81	111.22	110.38	110.70
30	2,5-dimethylhexane	109.00	110.84	111.23	110.69	111.10	110.06	110.46
31	3,3-dimethylhexane	112.00	109.89	108.97	109.81	108.83	109.92	108.85
32	3,4-dimethylhexane	117.70	115.56	116.05	115.66	116.15	115.61	115.87
33	3-ethylhexane	118.50	121.57	121.84	121.34	121.62	122.13	122.34
34	2,2,3-trimethylpentane	110.00	107.60	107.18	107.28	106.77	107.68	106.46
35	2,2,4-trimethylpentane	99.20	99.88	98.08	99.74	97.84	99.50	97.17
36	2,3,3-trimethylpentane	114.70	109.96	110.05	109.67	109.71	109.46[d]	109.22[d]
37	2,3,4-trimethylpentane	113.40	114.16	115.04	115.73	116.69	113.10	113.92
38	2-methyl-3-ethylpentane	115.60	116.26	116.85	116.45	117.05	116.23	116.64
39	3-methyl-3-ethylpentane	118.20	114.35	114.33	114.22	114.17	113.58	114.03
40	2,2,3,3-tetramethylbutane	106.50	105.82	104.71	105.66	104.04	106.85[d]	104.02[d]
41	nonane	150.77	150.05	150.08	150.29	150.31	151.23	151.31
42	2-methyloctane	142.80	143.66	144.04	143.63	144.03	144.03[d]	144.33[d]
43	3-methyloctane	143.80	141.51	141.54	141.30	141.33	142.15	142.11
44	4-methyloctane	142.40	141.77	141.77	141.55	141.57	142.45	142.39
45	2,2-dimethylheptane	132.70	135.13	134.14	135.39	134.34	135.64	134.13
46	2,3-dimethylheptane	140.50	141.90	142.42	142.15	142.67	142.03	142.36
47	2,4-dimethylheptane	133.50	135.65	135.70	135.55	135.61	135.59	135.56
48	2,5-dimethylheptane	136.00	134.33	134.27	134.02	133.96	134.17	134.11
49	2,6-dimethylheptane	135.20	135.66	135.63	135.46	135.44	135.37	135.38
50	3,3-dimethylheptane	137.30	136.41	135.77	136.44	135.77	136.45	135.86
51	3,4-dimethylheptane	140.10	139.12	139.33	139.2	139.41	139.54	139.55
52	3,5-dimethylheptane	136.00	136.32	136.30	136.11	136.10	136.34	136.27
53	4,4-dimethylheptane	135.20	137.69	137.32	137.88	137.49	137.49	137.35
54	3-ethylheptane	143.00	144.62	144.78	144.47	144.64	145.20[d]	145.39[d]
55	4-ethylheptane	141.20	145.95	146.22	145.91	146.21	146.47	146.81
56	2,2,3-trimethylhexane	131.70	136.63	136.58	136.66	136.64	136.78	136.23
57	2,2,4-trimethylhexane	126.50	128.76	127.30	128.59	127.09	128.81	126.96
58	2,2,5-trimethylhexane	124.00	127.25	125.27	127.22	125.19	127.17	125.00
59	2,3,3-trimethylhexane	137.70	139.46	139.86	139.40	139.81	139.07	139.42
60	2,3,4-trimethylhexane	139.00	138.58	138.86	139.54	139.80	138.47	138.58
61	2,3,5-trimethylhexane	131.30	126.13	124.31	125.00	121.88	125.49[d]	123.62[d]
62	2,4,4-trimethylhexane	126.50	131.40	130.78	131.41	130.78	130.59	130.21
63	3,3,4-trimethylhexane	140.50	138.18	138.30	137.90	138.02	138.02	138.08
64	2-methyl-3-ethylhexane	138.00	141.16	141.50	141.37	141.71	141.47[d]	141.67[d]
65	2-methyl-4-ethylhexane	133.80	138.00	138.13	137.93	138.07	137.91	138.05
66	3-methyl-3-ethylhexane	140.60	141.69	141.97	141.87	142.17	140.90	141.88
67	3-methyl-4-ethylhexane	140.40	142.00	142.25	142.13	142.39	142.41	142.56
68	2,2,3,3-tetramethylpentane	140.27	138.79	139.26	138.48	139.02	138.86	138.77
69	2,2,3,4-tetramethylpentane	133.00	134.87	133.79	135.55	134.38	135.07[d]	133.51[d]
70	2,2,4,4-tetramethylpentane	122.70	122.29	118.39	122.33	117.70	122.99	118.46
71	2,3,3,4-tetramethylpentane	141.50	140.68	141.26	140.80	141.44	140.28	140.73
72	2,2-dimethyl-3-ethylpentane	133.83	135.20	137.08	135.17	137.27	133.91	136.05
73	2,3-dimethyl-3-ethylpentane	141.60	140.99	141.99	140.85	141.88	140.05	141.46
74	2,4-dimethyl-3-ethylpentane	136.73	139.88	140.19	141.07	141.38	139.75	139.94
75	3,3-diethylpentane	146.20	146.00	147.21	146.15	147.52	144.45	146.85
76	decane	174.12	168.26	168.40	166.99	167.16	169.18	169.48
77	2-methynonane	167.00	163.35	163.70	163.05	163.42	163.68[d]	164.06[d]

Table 2 (Continued)

no.	compound name	NBP^{e} [a]	NBP^{M1}	NBP^{M2}	NBP^{c1} [b]	NBP^{c2} [b]	NBP^{t1} [c]	NBP^{t2} [c]
78	3-methylnonane	167.80	161.40	161.44	161.02	161.06	161.98	162.04
79	4-methylnane	165.70	161.50	161.49	161.22	161.21	162.13	162.15
80	5-methylnonane	165.10	171.53	170.11	171.95	170.79	172.83[d]	171.40[d]
81	2,2-dimethyloctane	155.00	157.28	156.64	157.63	156.94	157.66	156.72
82	2,3-dimethyloctane	164.31	162.93	163.34	163.11	163.52	163.18	163.48
83	2,4-dimethyloctane	153.00	157.11	157.03	157.12	157.05	157.26	157.15
84	2,5-dimethyloctane	156.80	155.67	155.45	155.44	155.21	155.79	155.59
85	2,6-dimethyloctane	158.54	' 155.13	154.85	154.79	154.50	155.21	155.00
86	2,7-dimethyloctane	159.87	156.80	156.59	156.50	156.27	156.76	156.65
87	3,3-dimethyloctane	161.20	158.30	157.85	158.25	157.77	158.33	158.06
88	3,4-dimethyloctane	166.00	159.63	159.72	159.54	159.63	160.17	160.13
89	3,5-dimethyloctane	160.00	157.11	156.93	156.84	156.66	157.39	157.19
90	3,6-dimethyloctane	160.00	155.81	155.54	155.41	155.12	156.00	155.76
91	4,4-dimethyloctane	157.50	159.81	159.66	160.07	159.90	159.59	159.81
92	4,5-dimethyloctane	162.10	159.94	160.00	159.94	159.99	160.54	160.46
93	3-ethyloctane	166.00	164.07	164.18	163.86	163.98	164.62	164.84
94	4-ethyloctane	163.64	165.81	166.06	165.85	166.12	166.28[d]	166.68[d]
95	2,2,3-trimethylheptane	158.00	160.12	160.30	160.10	160.29	160.27	160.16
96	2,2,4-trimethylheptane	147.70	152.40	151.29	152.34	151.22	152.57	151.22
97	2,2,5-trimethylheptane	148.00	150.29	148.71	150.18	148.55	150.45[d]	148.75[d]
98	2,2,6-trimethylheptane	148.20	150.71	148.85	150.71	148.80	150.81	148.96
99	2,3,3-trimethylheptane	160.10	162.83	163.33	162.86	163.39	162.53	163.14
100	2,3,4-trimethylheptane	163.00	161.54	161.64	162.19	162.29	161.80	161.75
101	2,3,5-trimethylheptane	157.00	155.23	154.88	155.52	155.15	155.16[d]	154.78[d]
102	2,3,6-trimethylheptane	155.70	154.60	154.15	154.78	154.30	154.27	153.91
103	2,4,4-trimethylheptane	153.00	155.87	155.57	155.86	155.57	155.20	155.29
104	2,4,5-trimethylheptane	157.00	155.70	155.36	155.99	155.62	155.64	155.29
105	2,4,6-trimethylheptane	144.80	152.73	152.36	153.25	152.89	152.27	152.02
106	2,5,5-trimethylheptane	152.80	152.87	151.98	152.77	151.85	152.31	151.84
107	3,3,4-trimethylheptane	164.00	160.15	160.40	159.79	160.04	160.10	160.42
108	3,3,5-trirnethylheptane	155.68	155.21	154.44	155.02	154.23	154.95	154.41
109	3,4,4-trimethylheptane	164.00	161.01	161.42	160.71	161.13	160.78	161.40
110	3,4,5-trimethylheptane	164.00	160.36	160.28	160.76	160.65	160.80	160.55
111	2-methyl-3-ethylheptane	166.00	162.19	162.40	162.22	162.43	162.62	162.77
112	2-methyl-4-methylheptane	160.00	160.00	160.02	159.87	159.89	160.13	160.21
113	2-methyl-5-ethylheptane	159.70	158.11	157.99	157.83	157.71	158.18	158.16
114	3-methyl-3-ethylheptane	163.80	164.67	164.75	164.88	164.97	i64.15	164.92
115	t3-methyl-4-ethylheptane	167.00	163.00	163.16	162.84	162.99	163.53	163.66
116	3-methyl-5-ethylheptane	158.30	159.87	159.83	'159.75	159.71	160.08	160.10
117	4-methyl-3-ethylheptane.	167.00	162.45	162.57	162.25	162.36	163.02	163.08
118	4-methyl-4-ethylheptane	167.00	165.09	165.70	165.07	165.74	164.20[d]	165.68[d]
119	4-propylheptane	160.00	167.37	167.75	167.90	168.37	167.77	168.32
120	4-isopropylheptane	160.00	162.78	163.01	163.00	163.23	163.22	163.40
121	2,2,3,3-tetramethylhexane	158.00	164.81	165.87	166.16	167.46	164.58	165.52
122	2,2,3,4-tetramethylhexane	154.90	158.21	157.57	158.71	158.03	158.62	157.66
123	2,2,3,5-tetramethylhexane	148.40	152.93	151.59	153.18	151.81	152.88[d]	151.40[d]
124	2,2,4,4-tetramethylhexane	153.30	151.37	149.04	151.22	148.54	151.50	149.21
125	2,2,4,5-tetramethylhexane	148.20	151.50	149.72	152.03	150.19	151.56[d]	149.77[d]
126	2,2,5,5-tetramethylhexane	137.46	146.01	141.87	147.45	143.08	146.69	142.50
127	2,3,3,4-tetramethylhexane	164.59	163.53	164.20	163.46	164.21	163.38	164.10
128	2,3,3,5-tetramethylhexane	153.00	156.62	156.32	156.84	156.55	156.01	155.95
129	2,3,4,4-tetramethylhexane	162.20	160.63	160.57	160.88	160.80	160.46	160.48
130	2,3,4,5-tetramethylhexane	161.00	160.25	159.90	162.10	161.58	160.22	159.84
131	3,3,4,3-tetramethylhexane	170.50	166.09	167.53	164.96	166.71	165.46	167.16
132	2,2-dimethyl-3-ethylhexane	159.00	161.06	161.40	i60.96	161.33	161.09	161.35
133	2,2-dimethyl-4-ethylhexane	147.00	155.52	154.80	155.62	154.91	155.48	154.67
134	2,3-dimethyl-3-ethylhexane	169.00	164.34	165.39	164.02	165.10	163.56[d]	165.15[d]
135	2,3-dimethyl-4-ethylhexane	164.00	161.29	161.30	161.80	161.79	161.65	161.55
136	2,4-dimethyl-3-ethylhexane	164.00	161.67	161.67	162.17	162.15	162.06	161.94
137	2,4-dimethyl-4-ethylhexane	158.00	159.28	159.79	159.29	159.85	157.98[d]	159.30[d]
138	2,5-dimethyl-3-ethylhexane	157.00	157.42	157.22	157.78	157.56	157.28	157.11
139	3,3-dimethyl-4-ethylhexane	165.00	163.40	163.89	163.13	163.65	163.16	163.92
140	3,4-dimethyl-3-ethylhexane	170.00	164.61	165.45	164.15	165.02	164.00	165.38
141	3,3-diethythexane	166.30	168.89	170.19	169.63	171.21	167.43[d]	169.96[d]
142	3,4-diethylhexane	162.00	165.41	165.68	165.79	166.10	165.89	166.20
143	2-methyl-3-isopropylhexane	163.00	163.09	163.22	163.84	163.96	163.34	163.36
144	2,2,3,3,4-pentamethylpentane	166.05	165.94	166.57	165.88	166.70	166.04	166.42
145	2,2,3,4,4-pentamethylpentane	159.29	158.88	157.10	158.84	156.63	159.88	157.67
146	2,2,3-trimethyl-3-ethylpentane	168.00	166.61	168.09	166.26	168.09	165.93	167.71
147	2,2,4-trimethyl-3-ethylpentane	155.30	160.06	159.47	160.63	160.01	160.36	159.59
148	2,3,4-trimethyl-3-ethylpentane	169.44	166.08	167.26	165.76	16,7.06	165.46	167.02
149	2-trimethyl-3,3-diethylpentane	174.00	167.57	169.29	166.80	168.63	166.23[d]	168.88[d]
150	2,4-dimethyl-3-isopropylpentane	157.04	163.00	162.83	164.66	164.44	163.08	162.86

[a] NBP^{e} experimental normal boiling point. [b] NBP^{c1} and NBP^{c2} calculated by cross validation. [c] NBP^{t1} and NBP^{t2} calculated by internal training set and external prediction set. [d] NBP of external rediction set predicted by internal training set. [e] NBP^{M1} calculated by model with using all 150 alkanes and NBP^{M2} by one with leaving out three alkanes.

represents the carbon atoms connecting one, two, three, and four C−C bonds, respectively. Let d_{ij} express a geometric average of graph theoretical distance between all atoms having i, j valences, which is defined as follows

$$d_{ij} = \prod_{i \le j} (ik, jl)^{1/(2 \times n_{ij})} \quad (i = 1, 2, 3, 4; j \ge i) \qquad (1)$$

where (ik,jl) refers to the distance, which is actually the C−C bond number between C_i^k and C_j^l, where the kth carbon atom belongs to type C_i while the lth carbon atom belongs to type C_j. n_{ij} stands for the pathway number from the kth atom(k) through the lth atom(l), i.e., item number of multiplication on the right side. As for the sample molecule, 2,2,3-trimethylpentane, ten values of various distance d_{ij} are obtained by the following calculations. It may help a reader to insert, before listing d_{11}, d_{12}, ..., d_{44}, a small table of distances as follows:

d_{11}	d_{12}	d_{13}	d_{14}	d_{22}
(1,5) = 4	(1,4) = 3	(1,3) = 2	(1,2) =1	(4,4) = 0
(1,6) = 2	(4,5) = 1	(3,5) = 2	(2,5) = 3	
(1,7) = 2	(4,6) = 3	(3,6) = 2	(2,6) = 1	
(1,8) = 3	(4,7) = 3	(3,7) = 2	(2,7) = 1	
(5,6) = 3	(4,8) = 2	(3,8) = 1	(2,8) = 2	
(5,7) = 4				
(5,8) = 3				
(6,8) = 3				
(7,8) = 3				

d_{23}	d_{24}	d_{33}	d_{34}	d_{44}
(3,4) = 1	(2,4) = 2	(3,3) = 0	(3,2) = 1	(2,2) = 0

Then, we can get

$d_{11} = (2 \times 2 \times 3 \times 4 \times 2 \times 3 \times 4 \times 3 \times 4 \times 3)^{1/(2\times10)} = 1.7017$; $n_{11} = 10$
$d_{12} = (3 \times 3 \times 3 \times 2 \times 1)^{1/(2\times5)} = 1.4902$; $n_{12} = 5$
$d_{13} = (2 \times 2 \times 2 \times 1 \times 2)^{1/(2\times5)} = 1.3195$; $n_{13} = 5$
$d_{14} = (1 \times 1 \times 1 \times 2 \times 3)^{1/(2\times5)} = 1.1962$; $n_{14} = 5$
$d_{22} = 0$; $n_{22} = 0$
$d_{23} = (1)^{1/(2\times1)} = 1$; $n_{23} = 1$
$d_{24} = (2)^{1/(2\times1)} = 1.4142$; $n_{24} = 1$
$d_{33} = 0$; $n_{33} = 0$
$d_{34} = (1)^{1/(2\times1)} = 1$; $n_{34} = 1$
$d_{44} = 0$; $n_{44} = 0$

In general, every interaction degree (energy) of two arbitrary types of carbon atoms can be regarded as a function of $1/d_{ij}^2$. For alkane molecules, a set of vectors, called λ vector or molecular distance-edge (MDE) vector in our present paper, can be used to approach the interaction energy of all atoms of the molecule. Then λ vector can be defined as follows:

$$\lambda_{ij} = (n_{ij}/d_{ij}^2) \quad (i = 1, 2, 3, 4; j \ge i) \qquad (2)$$

In the λ vector, there are only ten elements, which represent ten varied interaction models in alkanes. Some values of the λ vector are listed in Table 1; see ref 11 for more detailed data.

A multiple linear regression (MLR) method was used to develop a model of relationship between the logarithm of NBP of 150 alkanes and ten descriptor elements of the λ or MDE vector. The model was given in the following form

$$\ln(V_{\text{opt}} - \text{BP}) = b_0 + \sum_{j=1, j\ge i}^{10} b_{ij} \times \lambda_{ij} \qquad (3)$$

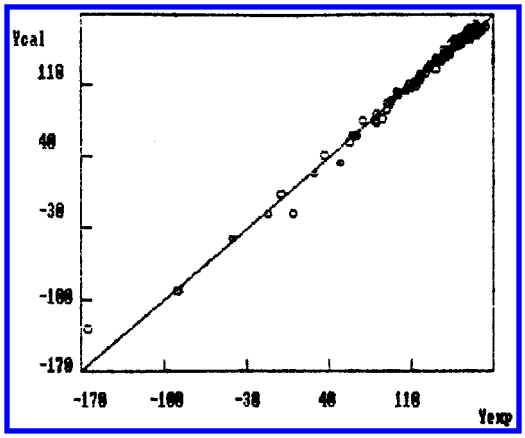


Figure 1. Plot of calculated (Ycal) vs observed boiling boints (Yexp) for working data set containing 150 compounds.

where V_{opt} is the limit, not the maximum, value of NBP for examined compounds, which was taken as $V_{\text{opi}} = 266.7$ °C by our screening optimization technique; BP is the normal boiling point for the compound; b_0 is the intercept term; and b_{ij} is the regression coefficient for descriptor λ_{ij}, which is the contribution value element j and can be regarded as the λ_{ij} in λ vector for the compound. And then $\ln(V_{\text{opi}} - \text{BP}) \sim \lambda$ model is utilized to predict NBP of external compounds in the test set.

RESULTS AND DISCUSSION

Data Set. The experimentally measured normal boiling points of all 150 alkane compounds are taken from ref 17. The range of experimental NBP was from −164 to 174.12 °C; the number of carbons per alkanes spanned from one through ten carbons. These 150 compounds composed a working data set. Of these 150 saturated hydrocarbons, 25 compounds were chosen randomly as an external prediction set, and the remaining 125 compounds comprised the training set. The external prediction set was never used until after a model had been developed using the training set. The experimental error in determining NBP root mean square (RMS) falls in general within the approximate range of 5.4−11.4K according to ref 18.

Descriptors. Molecular distance-edge λ vector with ten descriptor elements has a good discriminating ability for isomers of alkanes. For the examined 150 alkanes, according to the defined eq 1, the values of elements in the λ vector for compounds are different from each other. On the other hand, the MDE or λ vector has a good correlation with the physicochemical properties of the compounds as well. The results in our previous papert[11] showed that MDE vectors can be correlated with many physical properties and/or themodynamical functions such as Gibbs free-energy, enthalpy, heat capacity, molar volume, molar refraction, and so on.

Logarithm Transformation. Through many examinations, it was revealed that the normal boiling point of organic compounds, together with the other related physicochemical properties such as critical temperature, critical pressure, and density lean toward a limit value, a maximum value, or a minimum value with an increment of carbon bonds; the plot of NBP vs carbon atom numbers can be regarded as a logarithm function. Therefore, the dependent variables represented in the logarithm transformation mode of NBP

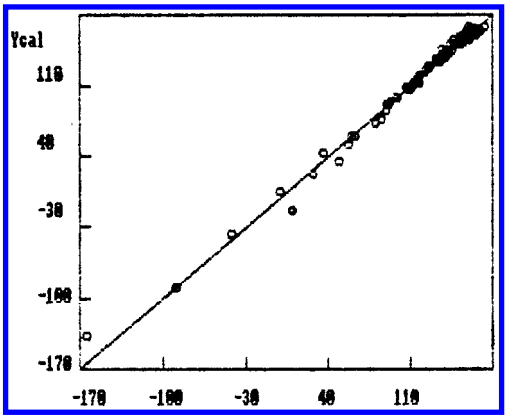


Figure 2. Plot of calculated (Ycal) vs observed boiling points (Yexp) for training set containing 125 compounds.

is well correlated with MDE, instead of the mode of NBP with MDE vector.

Regression Analysis. Multiple linear regression was used to develop linear models that linked the boiling point to λ vector of the compunds. This regression equation is given by eq 2 as stated above. Applying a MLR, the best model containing 10 variables is developed with the RMS error of 4.985K and an explained variance of EV = 99.44%, which is quite satisfactory and better than those reported in various literature.[12−18] This model can then be represented as follows

$$\ln(266.7 - \mathrm{BP}) = 5.9830 - 0.1039\lambda_{11} - 0.1034\lambda_{12} - 0.06133\lambda_{13} - 0.007161\lambda_{14} - 0.08385\lambda_{22} - 0.02965\lambda_{23} - 0.01056\lambda_{24} - 0.04224\lambda_{33} - 0.08573\lambda_{34} - 0.2304\lambda_{44}$$

$$R = 0.9948; \quad \mathrm{EV} = 99.44\%; \quad \mathrm{RMS} = 4.985\mathrm{K}; \quad n = 150; \quad F = 1321.24 \quad (4)$$

The observed normal boiling point (NBP^e) and calculated boiling point, (NBP^{M1}) from eq 4, are listed in Table 2. In the same way, the NBP^{M1} data are plotted against NBP^e, given in Figure 1. From Table 2 and Figure 1, it is shown that there are no observable patterns or deviations from normal behavior except the three compounds methane, 2,2-dimethylpropane, and 2,2-dimethylbutane. These results indicate that the model (4) possesses high internal stability.

In order to further validate stability of a model, a new model obtained by using 125 compounds in the training set from all 150 compounds was developed to predict the boiling points for the remaining 25 compounds in the external prediction set. The RMS error was 4.820K for the training set and 4.486K for the external prediction set. The new prediction model was given as follows:

$$\ln(266.7 - \mathrm{BP}) = 6.0003 - 0.1293\lambda_{11} - 0.1114\lambda_{12} - 0.05156\lambda_{13} + 0.005196\lambda_{14} - 0.08318\lambda_{22} - 0.02442\lambda_{23} + 0.004590\lambda_{24} - 0.04822\lambda_{33} - 0.08455\lambda_{34} - 0.2125\lambda_{44}$$

$$R = 0.9954; \quad \mathrm{EV} = 99.50\%; \quad \mathrm{RMS} = 4.820\mathrm{K}; \quad n = 125; \quad F = 1223.48 \quad (5)$$

The NBP^{t1}, normal boiling points calculated by eq 5, are also listed in Table 2 and are given in Figures 2 and 3.

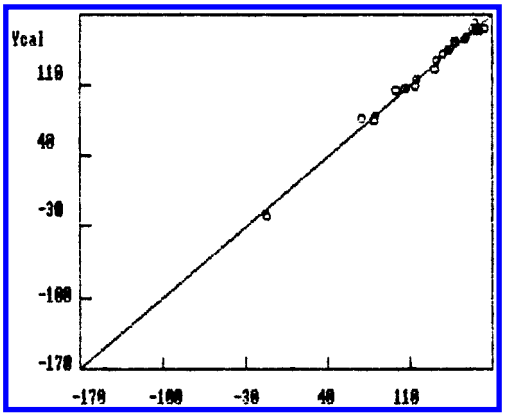


Figure 3. Plot of calculated (Ycal) vs observed boiling points (Yexp) for prediction set containing 25 compounds.

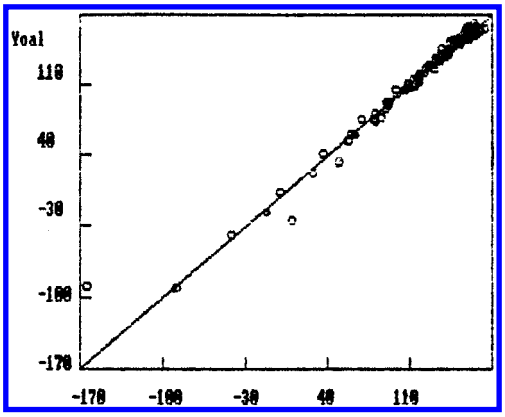


Figure 4. Plot of calculated by cross validation vs observed boiling points for working data set containing 150 compounds.

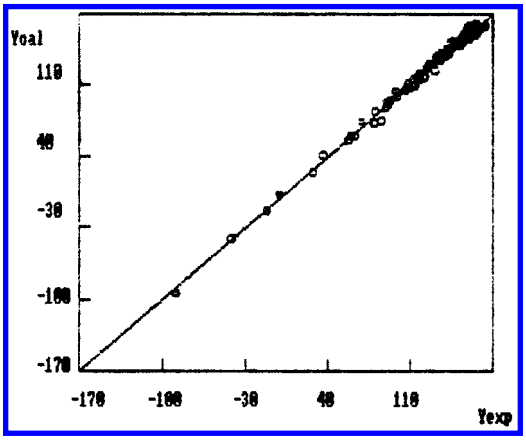


Figure 5. Plot of calculated (Ycal) vs observed boiling points (Yexp) for working data set containing 147 compounds.

Moreover, a cross-validation technique was employed to predict NBP of each compound in a data set of a total of 150 compounds; the PRESS (prediction residual sum of squares) statistics, obtained by leaving out each observation and predicting its value from the remaining $n-1$ observations, is of the leave-one-out precision with $S = 7.749$K. The NBP of each alkane was also listed in Table 2, and the plot of predicted NBP^{c1} by cross-validation vs observed NBP^e was given in Figure 4. It was found that there was no observable pattern exclusive of methane.

In the same way, the normal boiling points calculated by discarding three compounds, methane, 2,2-dimethylpropane, and 2,2-dimethylbutane, which behave quite differently from the other compounds, are also listed in Table 2 and are given

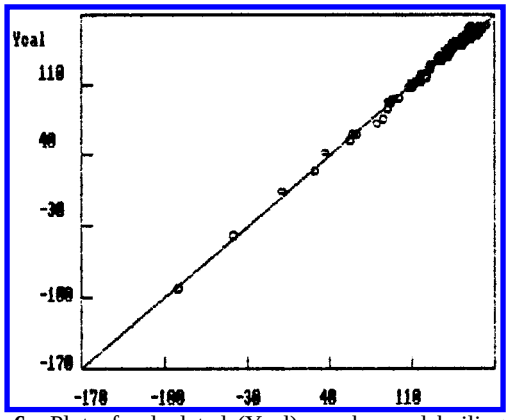


Figure 6. Plot of calculated (Ycal) vs observed boiling points (Yexp) for trading set containing 122 compounds.

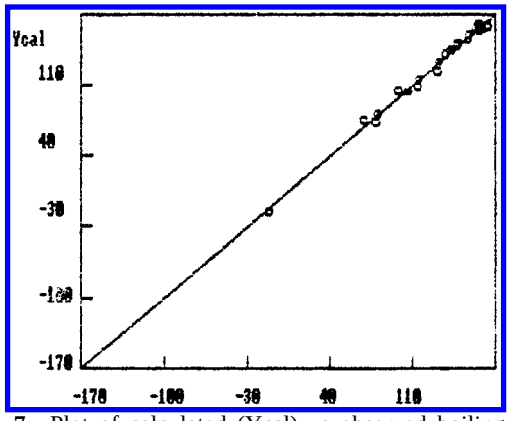


Figure 7. Plot of calculated (Ycal) vs observed boiling points (Yexp) for prediction set containing 25 compounds.

in Figures 5−8; two models are obtained by the MLR method, one for 147 alkanes as modeling; data and another one for the 122 alkanes remaining after leaving out 25 compounds randomly. As for 2,2-dimethylbutane, why it is an outlier is not yet clear; this needs further examination. These equations are given as follows:

$$\ln(266.7 - \mathrm{BP}) = 5.9184 - 0.02883\lambda_{11} - 0.08797\lambda_{12} - 0.08129\lambda_{13} + 0.04263\lambda_{14} - 0.08483\lambda_{22} - 0.03956\lambda_{23} + 0.03381\lambda_{24} - 0.04807\lambda_{33} - 0.1099\lambda_{34} - 0.2995\lambda_{44}$$

$$R = 0.9967; \quad \mathrm{EV} = 99.64\%; \quad \mathrm{RMS} = 3.342\mathrm{K}; \quad n = 147; \quad F = 2030.98 \quad (6)$$

$$\ln(266.7 - \mathrm{BP}) = 5.9432 - 0.06456\lambda_{11} - 0.09529\lambda_{12} - 0.06822\lambda_{13} + 0.02065\lambda_{14} - 0.08487\lambda_{22} - 0.03529\lambda_{23} + 0.02383\lambda_{24} - 0.05361\lambda_{33} - 0.1138\lambda_{34} - 0.2965\lambda_{44}$$

$$R = 0.9966; \quad \mathrm{EV} = 99.63\%; \quad \mathrm{BMS} = 3.350\mathrm{K}; \quad n = 122; \quad F = 1632.84 \quad (7)$$

CONCLUSION

Four models have been developed that estimate and predict the normal boiling point of alkanes with a high degree of accuracy. These models provided more accurate calculated NBP than various previous methodologies in literature, primarily because the MDE vector in this paper can more effectively describe the structure and interaction between atoms in a molecular graph. It is foreseeable that the MDE vector would be utilized widely in QSPR/QSAR and/or the other related studies. This research is in progress.

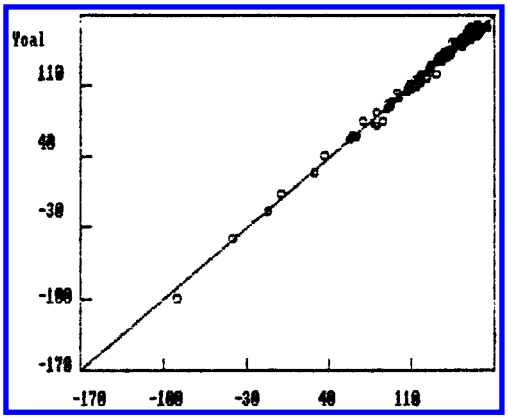


Figure 8. Plot of calculated by cross validation vs observed boiling points for working data set containing 147 compounds.

ACKNOWLEDGMENT

The project was supported by National Natural Science Foundation of China (NSFC), State Educational Comission Foundation (SECF), Ministry of Mechanical Industrial Fund (MMIF) and awarded by funding Hok Yinh-Tong Educational Foundation (HYEF). Professor Shin-ichi Sasaki and late Professor Yoshikatsu Miyashita were thanked for their invaluable suggestions and Professor Menglong Li and Professor Leming Shi for their technical assistance. Professor Milan Randić was acknowledged for his helpful recomments.

REFERENCES AND NOTES

(1) Randić, M. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 142.
(2) Needhan, D. E.; Wei, I.; Seybold, P. G. *J. Am. Chem. Soc.* **1988**, *110*, 4186.
(3) Miyashita, Y.; Okuyama, T.; Ohsako, H.; Sasaki, S. *J. Am. Chem. Soc.* **1989**, *110*, 3469.
(4) Miyashita, Y.; Li, Z.; Sasaki, S. *Trend Anal. Chem.* **1993**, *12*(2), 50.
(5) Basak, S. C. *Med. Sci. Res.* **1987**, *15*, 605.
(6) Li, Z.; Muramatsu, Y.; Hu, F. *Acta Pharm. Sin.* **1996**, *31*(1), 38.
(7) Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 334.
(8) Hansen, P. J.; Jurs, P. C. *J. Chem. Educ.* **1988**, *65*, 574.
(9) Mihalić, Z.; Trinajstic, N. *J. Chem. Educ.* **1992**, *69*, 701.
(10) Randić, M. *J. Math. Chem.* **1991**, *7*, 155.
(11) Cao, C.; Liu, S.; Li, Z. To be submitted for publication.
(12) Wessel, M. D.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 68.
(13) Egolf, L. M.; Wessel, M. D.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 947.
(14) Egolf, L. M.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 616.
(15) Hall, L. H.; Kier, L. B. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1039.
(16) Fisher, C. H. *Chem. Eng.* **1989**, *96*, 157.
(17) Mihalić, Z.; Nikolić, S.; Trinajstic, N. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28.
(18) Wessel, M. D.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 841.

CI970109Z