Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2010, 2(4): 200-205
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X

Development & comparison of single and multi-parameter QSPR models

S. T. Shukla[1] and V. V. Sawant[2]

[1]*Department of Chemistry, S. P. Mahavidyalaya, Lonand*
[2]*Department of Chemistry, Yashwantrao Chavan Institute of Science, Satara*

ABSTRACT

*A fundamental concept of chemistry is that the structural characteristics of a molecule are responsible for its properties. This concept is utilized to develop QSPR (Quantitative Structure Property Relationship) models. With respect to environmental considerations, the application of QSPR models may reduce the use of undesirable materials and help in reduction of pollution . These models help in the prediction of properties even before they are synthesized. The correlation between three topological indices viz, Wiener index (**W**) ,Randic index (**χ**) and Harary number(**H**) and boiling point (bp) has been studied for 40 alkanes up to 8 carbon atoms. Three single parameter QSPR models have been developed using these indices separately. These models have been used to predict the boiling points of nonanes. A fourth QSPR model has been developed using the three indices together to study their combined effect. The values predicted by this multi-parameter model have been compared with the values obtained by single parameter QSPR models.*

Key words: QSPR, Wiener number, Randic index, Harary number.

INTRODUCTION

Graph theory is a branch of discrete mathematics related to topology and combinatory. It deals with the way objects are connected and with all the consequences of the connectivity. The connectivity in a system is , thus a fundamental quality of graph theory.[1]

Graph theory offers the means to numerically characterize structures of chemicals. Chemical graph theory is concerned with graphs that represent chemical systems[2-6]

A special class of chemical graphs are molecular graphs which represent the constitution of molecules. In these graphs , vertices correspond to individual atoms and edges correspond to

bonds between them. On the basis of these molecular graphs a number of topological indices have been proposed.[3-10]

A topological index or a graph theoretical index is a numeric quantity derived from the chemical graph which is characterized by degree of vertex i.e the number of edges ending on the vertex and the path i.e the sequence of edges connecting two vertices.[1]

Literature reveals use of various topological indices for encoding the molecular structural information . Attempts are made to add many more indices to the existing list. Depending on the fact that structural characteristics of a molecule are responsible for its properties. Various QSPR studies have been done and successful QSPR models have been developed to predict the properties of molecules.[11-13] Correlations of octane number of heptane and octane isomers with various topological indices have been studied recently[14].

It is now realized that each different structural aspect of the behaviour of molecule can be characterized by a separate index .According to the literature, a QSPR model involving a single topological index does not serve as an ideal model[15].

Therefore in the present investigation a combined effect of three indices viz, Randic index (χ), Wiener index (W) and Harary number (H) has been undertaken.

Methodology

A set of 40 alkanes has been chosen as the training set on the basis of the values given in the literature[16]. It has been subjected to rigorous statistical analysis[17] and QSPR models have been developed using HCL Pentium 4 .Values of indices and experimental boiling points have been listed in Table 1. Plots of boiling point Vs each topological index have been shown in Figures 1-3.

Following structure –property relationship models have been developed for each index considered.

Model 1

$$\mathbf{b.p = a\,\chi + b\,\chi^2 + c\,\chi^3 + d} \quad \text{----------(1)}$$

where a = 80.977 , b = 6.704, c = -2.184, d = -165.714

r = 0.9902 s = 4.19

Model 2

$$\mathbf{b.p = a\,W + b\,W^2 + c\,W^3 + d} \quad \text{----------(2)}$$

where a = 9.588, b= -0.158 c = 0.00093 d = -106.231

r = 0.8913 s = 13.33

Model 3

$$\mathbf{b.p = a\,H^{1/2} + bH + cH^{3/2} + d} \quad \text{----------(3)}$$

where a = 51.683, b = 23.441 c = -3.956 d = -162.81

r = 0.9847 s = 9.98

(r = correlation coefficient , s = standard deviation }

The combined effect of these three indices on b.p has been studied by developing a model with all the three indices together as follows :

Model 4

$$\mathbf{b.p = a\,\chi + b\,H^{1/2} + c\,W + d} \text{ ----------(4)}$$

where a = 58.765, b = 30.3105, c = -0.412, d = -168.348

s = 3.302

Table I. Values of Harary number(H), Wiener number(W), Randic index (χ), and boiling points (in ^{0}C) of training set of alkanes (up to 8 carbon atoms)

Alkane	H	W	χ	bp
methane	0	0	0	-164
ethane	1.000	1	1	-88.6
propane	2.250	4	1.4142	-42.4
2-mpropane	3.750	9	1.7321	-11.7
butane	3.611	10	1.9142	-0.5
2,2-dmpropane	5.500	16	2.0000	9.5
2methyl butane	5.222	18	2.2701	27.8
pentane	5.035	20	2.4142	36.1
2,2-dmbutane	7.083	28	2.5607	49.7
2,3-dmbutane	6.944	29	2.6427	58
2-methylpentane	6.708	32	2.7701	60.3
3-methylpentane	6.757	31	2.8081	63.3
hexane	6.498	35	2.9142	69
2,2,3-tmbutane	8.917	42	2.9434	80.9
2,2-dmpentane	8.632	46	3.0607	79.2
3,3-dmpentane	8.729	44	3.1213	86.1
2,3-dmpentane	8.542	46	3.1807	89.8
2,4-dmpentane	8.444	48	3.1259	80.5
2-methylhexane	7.355	52	3.2701	90
3-methylhexane	8.283	50	3.3081	92
3-ethylpentane	8.354	48	3.3461	93.5
heptane	7.989	56	3.4142	98.4
2,2,3,3-tmbutane	11.000	58	3.2500	106.5
2,2,3-tmpentane	10.576	63	3.4814	110
2,3,3-tmpentane	10.625	62	3.5040	114.7
2,2,4-tmpentane	10.431	66	3.4166	99.2
2,2-dmhexane	10.176	71	3.5607	106.8
3,3-dmhexane	10.318	67	3.6213	112
3-e-3-mpentane	10.438	64	3.6820	118.2
2,3,4-tmpentane	10.389	65	3.5534	113.4
2,3-dmhexane	10.108	70	3.6807	115.6
3-e-2-mpentane	10.201	67	3.7187	115.6
3,4-dmhexane	10.179	68	3.7187	117.7
2,4-dmhexane	10.059	71	3.6639	109.4

2,5-dmhexane	9.966	74	3.6259	109
2-methylheptane	9.731	79	3.7701	117.6
3-metylheptane	9.814	76	3.8081	118
4-methylheptane	9.837	75	3.8081	117.7
3-ethylhexane	9.920	72	3.8461	118.5
octane	9.502	84	3.9142	125.7

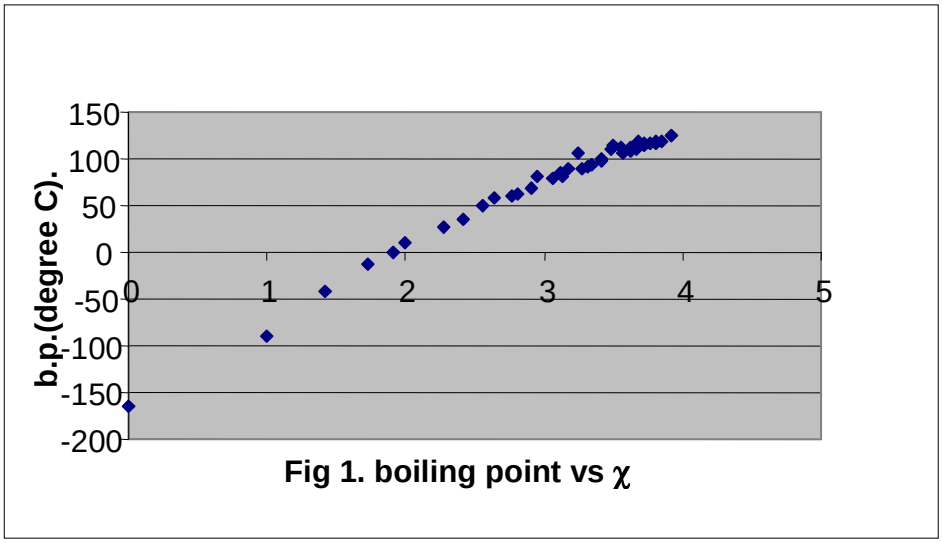

Fig 1. boiling point vs χ

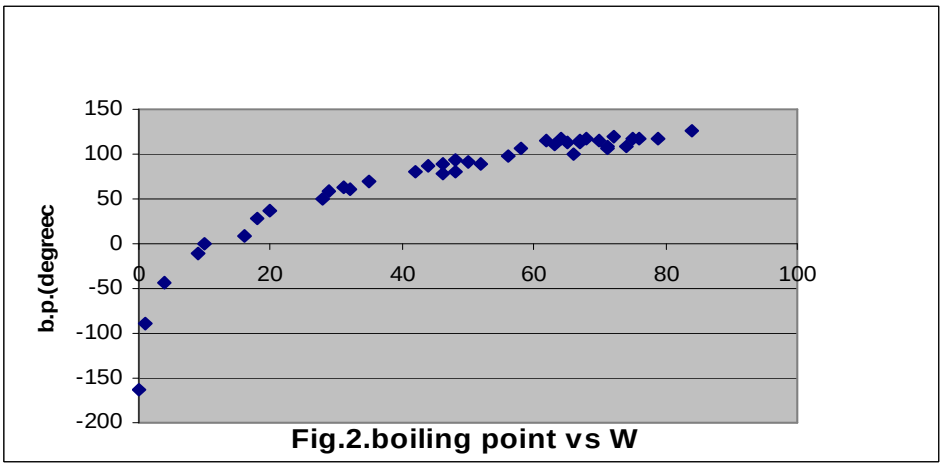

Fig.2.boiling point vs W

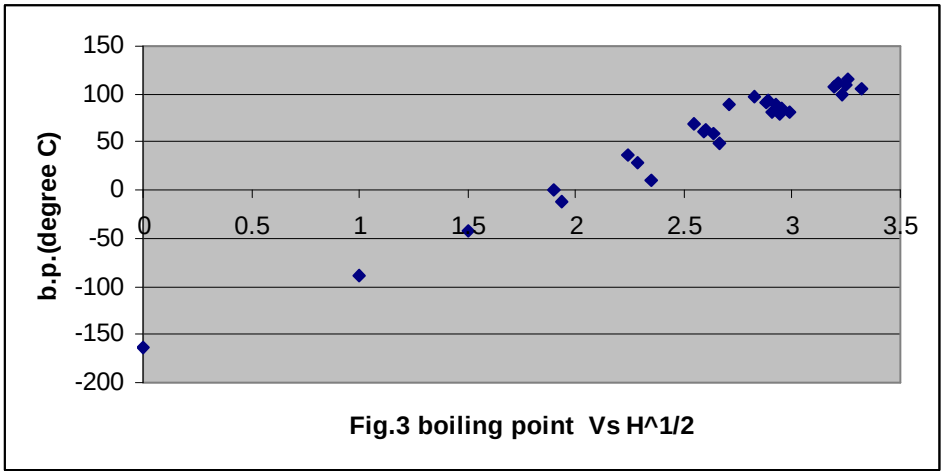

Fig.3 boiling point Vs H^1/2

TableII. Comparison between Experimental (exp) & predicted (pred) boiling points (^{O}C) of higher alkanes

Alkane	bp (exp)	pred bp (model 1)	bp(exp)-bp(pred)	pred bp(model 2)	bp(exp)-bp(pred)	pred bp (model 3)	bp(exp)-bp(pred)	pred bp (model 4)	bp(exp)-bp(pred)
2,2,3,3-tetramethylpentane	140.27	119.38	20.88	131.02	9.24	140.70	-0.43	130.15	10.11
2,2,3,4-tetramethylpentane	133	120.95	12.04	141.97	-8.97	137.96	-4.96	129.84	3.15
2,2,3- trimethylhexane	131.7	125.15	6.54	163.41	-31.715	134.93	-3.23	133.54	-1.844
2,2-dimethyl-3-ethlpentane	133.83	126.28	7.54	148.40	-14.57	136.10	-2.27	137.92	-4.097
3,3,4-trimethylhexane	140.5	126.92	13.5	148.40	-7.90	136.36	4.13	139.36	1.13
2,3,dimethyl3-ethylpentane	141.6	127.55	14.04	141.97	-0.37	137.07	4.52	141.83	-0.23
2,2,4,4-tetramethylpentane	122.7	115.36	7.33	148.40	-25.71	137.89	-15.19	120.35	2.34
2,3,3-trimethylhexane	137.7	125.83	11.86	155.54	-17.84	135.65	2.04	136.00	1.69
2,2,4-trimethylhexane	126.5	124.31	2.18	172.07	-45.57	133.94	-7.44	130.71	-4.21
2,4,4-trimethylhexane	126.5	125.02	1.47	163.41	-36.91	134.67	-8.17	133.18	-6.68
2,2,5-trimethylhexane	124	123.08	0.91	191.94	-67.94	132.750	-8.75	126.33	-2.33
2,2-dimethylheptane	132.7	127.44	5.25	228.76	-96.06	130.15	2.54	131.24	1.45
3,3-dimethylheptane	137.3	129.03	8.26	191.94	-54.64	131.91	5.38	138.01	-0.71
3,3-diethylpentane	146.2	131.76	14.43	148.40	-2.21	135.19	11.00	150.65	-4.45
2,3-dimethylheptane	140.5	130.44	10.05	215.49	-74.99	129.55	10.94	138.87	1.62
3-methyl octane	143.3	132.97	10.32	275.07	-131.77	126.13	17.16	141.67	1.62
4-methyloctane	142.4	132.97	9.42	258.50	-116.11	126.52	15.87	142.65	-0.25
3-ethylheptane	143	133.59	9.40	228.76	-85.76	127.52	15.47	146.94	-3.93
4-ethylheptane	141.2	133.59	7.60	215.49	-74.29	127.90	13.29	147.91	-6.71
nonane	150.77	134.55	16.21	376.62	-225.85	122.46	28.30	142.39	8.45
2,3,3,5-tetramethylhexane	153	133.80	19.19	376.62	-223.62	152.58	0.41	152.41	0.58
3,3diethyl2-methylpentane	174	136.18	37.81	311.81	-137.81	153.35	20.64	170.87	3.12
2,2,4,4-tetramethylhexane	153.3	132.25	21.04	364.97	-211.67	153.48	-0.18	147.86	5.43
2,2,5-trimethyheptane	148	135.02	12.97	579.32	-431.32	148.40	-0.40	150.15	-2.15
5-methylnonane	165.1	135.90	29.19	890.59	-725.49	142.43	22.66	161.97	3.12

RESULTS AND DISCUSSION

Predictions are made for the values of the boiling points of structurally similar compounds that are not part of the training set via the QSPR models developed during this investigation.(Table2).Thus the above models can be used for prediction of boiling points of higher alkanes also . Satisfactory results are obtained for alkanes containing up to 10 carbon atoms . This can be observed by studying table 2.

Comparison of the values predicted by the 4 models reveals that lowest deviation is observed when the 3 indices are taken together. (model 4). This may be because each of these parameters takes care of a certain structural detail of a large alkane.

CONCLUSION

On the basis of the observations found in the present investigation it can be concluded that the most accurate QSPR model for alkane boiling points is based on H , W and χ taken together (model no. 4) The accuracy of the model is decided by the low standard deviation value($< 4^0$C).

Acknowledgement
We are thankful to Dr. A.S. Bhoite, Mr. A.S. Mahadar, Mr. V.A. Shinde and Y.C. Institute of Science for encouragement.

REFERENCES

[1] Douglas .B. West Introduction to graph theory Prentice Hall Delhi (2001)
[2] I.S.Dmitriev, Molecules without chemical bonds Mir publishers Moscow (1981)
[3] Diudea M.V. and Gutman *J Croatica Chimica Acta* **71(1)** 21-51 (1998)
[4] Wiener. H, *J.Am Chem. Soc*. **69**, 17-20 (1947)
[5] Wiener. H, *J.Am Chem. Soc*. **69**, 2636-2638 (1947)
[6] Schultz H.P.*J.Chem.Inf.Comput.Sci*. **29**, 227-228 (1989)
[7] Wiener .H. *J.Phys.Chem.* **52**, 425-430 (1948)
[8] Wiener .H. *J.Phys.Chem*. **52**, 1082-1089 (1948)
[9] Randic.M.*J.Amer.Chem.Soc*.**97**, 6609-6615 (1975)
[10] Hollas B., Gutman I. and Trinajstic N. *Croatica Chimica Acta* **78(4)** 489-492(2005)
[11] Nicolic S., Trinajstic N., and Mihalic Z., *Croatica Chimica Acta* **68(1)** 105-129(1995)
[12] Natarajan R, Nirdosh L and Muthuswami S.V. *Current Science* **77(9)** 10 Nov 1170-1174 (1999)
[13] Agarwal V.K. , Bano S. , Mathur K.C. ,Khadikar P.V. Mandloi M and Sapre N. *Proc. Indian.Acad.Sci(Chem Sci)***vol III** N0.6 ,6697-708 (1999)
[14] Hosoya H. *Croatica Chimica Acta* **75(2)**, 433-445(2002)
[15] Perdih .A and Perdih.M *Acta Chim Slov* **47**, 231-259(2000)
[16] Mihalic'Z and Trinajstic' *J.Chem.Edu* **69(9)** ,701-712(1992)
[17] Ray.M and Sharma S., Mathematical Statistics, Prasad & Sons (1985)