Topological estimation of aromatic stabilities of polyacenes and helicenes: modeling of resonance energy and benzene character

Padmakar V. Khadikar*

*Research Division, Laxmi Fumigation and Pest Control, Pvt. Ltd., 3 Khatipura,
Indore 452 007, India*
E-mail: pvkhadikar@rediffmail.com

Jyoti Singh

Department of Chemistry, A.P.S. University, Rewa 486 003, India

Maya Ingle

Computer Center, Devi Ahilya Vishwavidyalaya, Khadwa Road, Indore 452010, India
E-mail: maya_ingle@rediffmail.com

The distance-based topological indices viz Wiener (W)-, Szeged (Sz)-, Padmakar-Ivan (PI)- and Sadhana (Sd)-indices have been used for estimating aromatic stabilities as well as % benzene character of polyacenes and helicenes. Excellent models are obtained in combined set and there was no need to split the data set into polyacenes and helicenes. The newly introduced Sd index was found the best index among all the four indices used. In bi-parametric modeling with the combination of the Sd index with the PI index yielded excellent results. The results are discussed critically on the basis of variety of statistical parameters.

KEY WORDS: aromatic stability, polyacenes, helicenes, Padmakar-Ivan index, Sadhana index, regression analysis, topological index, resonance energy, % benzene character

1. Introduction

Of late Trinajstic and Randic have predicted aromatic stabilities of polyacenes and helicenes using conjugate circuits model [1]. These benzenoid hydrocarbons have a common molecular formula $C_N H_{(N+6)/2}$ but differ structurally in the mode of annelation. Polyacenes are linearly para-annelated benzenoids while

*Corresponding author.

433

helicenes are angularly ortho-annelated benzenoids having helical symmetries. Our interest in these classes of benzenoids is due to their considerable difference in physical, chemical, and biological properties [2,3]. Furthermore their chemistry is still very much of interest to synthetic chemists environmental chemists, cancer research chemists, theoretical chemists, analytical chemists, structural chemists, etc [4–27]. It is because of these properties that we have undertaken the present investigation of estimating aromatic stabilities as well as % benzene character of polyacenes and helicenes (table 1). The main objective of our work is that aromatic stabilities of these compounds are responsible for the exhibition of their physicochemical properties as well as biological activities. Furthermore, other theoretical methods of the estimation of aromatic stabilities are cumbersome and time consuming. Here, we present the method for the estimation of aromatic stabilities of polyacenes and helicenes using four distance-based topological indices viz Wiener (W) [28],- Szeged (Sz) [29–32],- Padmakar-Ivan (PI)[33-35] - and Sadhana (Sd)-indices [36]. The choice of these indices being: W index is the first and widely used index, while other three are the modifications of the W index for cyclic compounds. It is worthy to mention that Sz index is introduced by Gutman and coworkers [29–32] as a modification of W index for cyclic graph. For acyclic graphs (trees) Sz index coincides with W index. In view of this, Khadikar [33,34] and Jaiswal and Khadikar [35] proposed yet another W-like index viz PI index. Unlike Sz index, PI index is different for acyclic and cyclic compounds. In these sense PI index can be considered as the modification of Sz index. The Sd [36] index is useful only for cyclic compounds and was discovered by Khadikar while computing PI index.

The work in this paper is presented under two different headings:

 (i) Topological estimation of normalized resonance energy (NRE), and

(ii) Topological estimation of % benzene character (% BC).

Consequently, we have used the earlier reported NRE and % BC [1]. In estimating these parameters we have used maximum-R^2 method and adopted stepwise regression analysis [37–40].
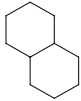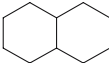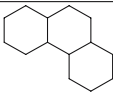
2.   Result and discussion
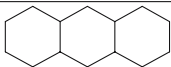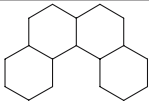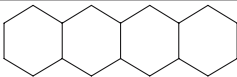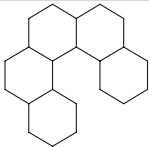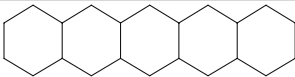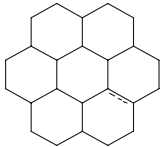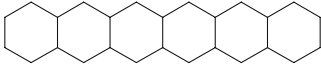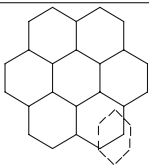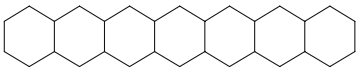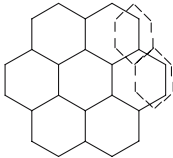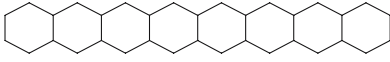
2.1. Topological estimation of NRE

The set of polyacenes and helicenes is the same as used before [1] (table 1). The adopted resonance energy (RE) and NRE are presented in table 2. This table 1 also records the calculated values of W, Sz, PI, and Sd indices.

It is worthy to mention that the RE is given by the following expression:

$$RE = 1/K \Sigma_{n} R_n R_n, \tag{1}$$

Table 1. Structural details of polyacenes and helicenes.

	1.
2.	2.
3.	4.
5.	6.
7.	8.
9.	10.
11.	12.
13.	14.

where, R_n represents the number of R_n circuits for a given value of n. However, in order avoid the size effect, the RE is normalized:

$$NRE = RE/N, \tag{2}$$

Table 2

Values molecular resonance energy (RE) normalized molecular resonance energy (NRE), % benzene character (% BC) of polyacenes and helicenes along with topological indices (W, Sz, PI, Sd).

Compd no.	h	RE(ev)	NRE(ev)	% BC	W	Sz	PI	Sd
1.	1	0.809	0.145	100	27	54	24	12
2.	2	1.323	0.133	91	109	243	96	44
3.	3	1.975	0.141	97	271	632	218	96
4.	3	1.600	0.114	79	279	656	216	96
5.	4	2.540	0.141	80	529	1253	384	168
6.	4	1.780	0.099	68	569	1381	384	168
7.	5	3.126	0.142	96	899	2138	612	260
8.	5	1.904	0.087	60	1011	2506	600	260
9.	7	3.728	0.143	99	2039	4828	1206	504
10.	6	1.991	0.076	53	1637	4119	864	372
11.	8	4.321	0.144	90	2841	6697	1578	656
12.	7	2.056	0.069	47	2419	6308	1176	504
13.	9	4.911	0.144	99	3819	8918	2000	828
14.	8	2.106	0.062	43	3569	9161	1536	656

where NRE stands for RE per π-electron, while N is the number of π electrons in the conjugated molecules.

A perusal of table 2 shows that RE values for polyacenes increases with the size of the molecule. However, gradually this increase in RE slows down and reaches to the asymptotic value ≈ 2.512. The NRE values of polyacenes decrease steadily from benzene to octacenes. The situation is quite different with helicenes, which are predicted to be of high aromatic stability. The aromatic stability is so high in helicenes that the considerable steric interaction pushing the overlapping terminal rings apart cannot harm the molecule. A given helicene in comparison to the energies of isomeric polyacenes is always a more stable system. Generally, a benzenoid system with ortho- annelated rings are always more stable than the corresponding system with para-annelated rings [1].

A perusal of table 2 shows that NRE values of helicenes remain rather higher aromatic with the increase in the size.

A preliminary regression analysis indicated that no mono-parametric regression is possible for modeling NRE. Thus, one should try multiple regression analysis for obtaining statistically significant model for modeling NRE. Before a multivariate analysis is undertaken it is convenient to normalize the data in certain ways in order to make the detection of significant correlations easier. Normally, it is sufficient to pre-process the data by means of auto-scaling and mean-centering the variables. Auto scaling gives each variable unit variance and hence the same chance to contribute to a calculated model, while mean scaling facilitates interpretation. This can be achieved by obtaining correlation matrix. Such a correlation matrix obtained in the present case is shown in table 3.

Table 3
Correlation matrix for the inter-correlation of structural descriptors and their correlation with the activity.

	h 1.000	RE(ev)	NRE(ev)	%BC	W	Sz	PI	Sd
h								
RE(ev)	0.778	1.000						
NRE(ev)	−0.266	0.375	1.000					
%BC	−0.289	0.333	0.974	1.000				
W	0.956	0.695	−0.273	−0.277	1.000			
Sz	0.950	0.651	−0.328	−0.332	0.998	1.000		
PI	0.976	0.795	−0.169	−0.179	0.987	0.976	1.000	
Sd	0.978	0.784	−0.189	−0.199	0.989	0.979	0.999	1.000

The correlation matrix (table 3) indices that the topological indices used are linearly correlated (i.e. are auto-correlated) and that none of the four topological indices correlate significantly with NRE indicating once again that mono-para-metric models are not adequate for modeling the NRE. In view of this we have used the maximum-R^2 improvement method to derive predictive models. This method finds the "best" one variable model, the "best" two variable models and so forth for the prediction of NRE. Several models (combination of topological

Table 4
Statistical parameters and qualities of correlation for estimating NRE.

Model	Topological indices	Se	R	R^2A	F
(i) Mono-parametric modeling					
1.	W	0.0319	−0.2721	–	0.959
2.	Sz	0.0313	−0.3273	–	1.448
3.	PI	0.0327	−0.1681	–	0.349
4.	Sd	0.0326	−0.1882	–	0.441
(ii) bi-parametric modeling					
5.	W, Sz	0.0179	0.8561	0.6844	15.096
6.	W, PI	0.0216	0.6721	0.3521	4.532
7.	W, Sd	0.0276	0.6851	0.2509	3.177
8.	Sz, PI	0.0226	0.7569	0.4952	7.376
9.	PI, Sd	0.0138	0.9171	0.8122	29.189
(iii) tri-parametric modeling					
10.	W, Sz, PI	0.0186	0.8588	0.6587	9.363
11.	W, PI, Sd	0.0143	0.9198	0.7999	18.320
12.	Sz, PI, Sd	0.0143	0.9191	0.7982	18.137

W = Wiener index; Sz = Szeged index; PI = Padmakr-Ivan index; Sd = Sadhana index; Se = Stan-dered error of estimation; R = Multiple correlation coefficient; R^2A = adjusted R^2; F = Fisher sta-tistics.

Table 5
Comparison of estimated biological activity NRE(ev) with their observed values.

S.N.	Actual	NRE(ev)					
		Pred.	Res.	Pred.	Res.	Pred.	Res.
1.	0.145	0.149	−0.004	0.146	−0.001	0.146	−0.001
2.	0.133	0.138	−0.005	0.136	−0.003	0.136	−0.003
3.	0.141	0.131	0.010	0.130	0.011	0.130	0.011
4.	0.114	0.126	−0.012	0.126	−0.012	0.126	−0.012
5.	0.141	0.114	0.027	0.116	0.025	0.116	0.025
6.	0.099	0.114	−0.015	0.115	−0.016	0.115	−0.016
7.	0.142	0.125	0.017	0.127	0.015	0.127	0.015
8.	0.087	0.101	−0.014	0.103	−0.016	0.102	−0.015
9.	0.143	0.133	0.010	0.135	0.008	0.135	0.008
10.	0.076	0.086	−0.010	0.088	−0.012	0.088	−0.012
11.	0.144	0.142	0.002	0.143	0.001	0.142	0.002
12.	0.069	0.071	−0.002	0.072	−0.003	0.071	−0.002
13.	0.144	0.154	−0.010	0.152	−0.008	0.153	−0.009
14.	0.062	0.056	0.006	0.051	0.011	0.052	0.010

indices) were examined to identity combinations of the used topological indices with good predicted capacity. The results are given in table 4.

The data presented in table 4 indicate that among the bi-parametric modeling attempted, the bi-parametric model containing PI and Sd indices as the correlating parameters yielded excellent model:

$$NRE = 0.1596 + 0.0020(\pm 2.7370 \times 10^{-4})PI - 0.0049(\pm 6.5914 \times 10^{-4})Sd,$$
$$n = 14, \quad Se = 0.0138, \quad R = 0.9171, \quad R^2A = 0.8122, \quad F = 29.019. \quad (3)$$

The positive coefficient of PI index in this equation (3) indicates that NRE increases with increase in the magnitude of PI index. While the negative coefficient of Sd is probably due to its high collinearity with PI index. This problem is discussed in details in following section.

The perusal of table 4 also shows that there is only slight improvement in the statistics as we go for tri-parametric modeling. In view of this we have confined to the above model (equation (3)) only. Using this model we have calculated NRE values for the entire set of 14 compounds and compared them with the observed NRE. Such a comparison is shown in table 5, which shows that the estimated NRE are very close to their observed values. A careful observation of this table 5 shows that the residue (i.e. difference between observed and estimated NRE) for the compound 5 is high (0.027) as compared to other 13 compounds. Upon deletion of this compound 5 from the regression procedure there was a tremendous improvement is the statistics to the extent that now R increases from 0.9171 to 0.9480 and the slandered error of estimation (Se)

Table 6
Statistical parameters and quality of correlation for estimating % benzene character.

Model	Topological indices	Se	R	R^2A	F
(i) Mono-parametric modeling					
13.	W	20.7531	−0.2767	-	0.995
14.	Sz	20.3740	−0.3317	-	1.483
15.	PI	21.2512	−0.1781	-	0.395
16.	Sd	21.1655	−0.1988	-	0.494
(ii) bi-parametric modeling					
17.	W, Sz	11.7272	0.8542	0.6806	14.848
18.	W, PI	17.2634	0.6436	0.3078	3.890
19.	W, Sd	18.5078	0.5716	0.2044	2.670
20.	Sz, PI	15.2849	0.7354	0.4573	6.478
21.	PI, Sd	7.4370	0.9441	0.8713	45.097
22.	Sz, Sd	16.1543	0.6979	0.3939	5.224
(iii) tri-parametric modeling					
23.	W, Sz, PI	7.7991	0.9441	0.8587	27.338
24.	Sz, PI, Sd	7.7966	0.9441	0.8588	27.358
25.	W, PI, Sd	7.7991	0.9441	0.8587	27.338

For the parameters used see the footnote of table 4.

decreases from 0.0138 to 0.0113. This improvement model is found as below:

$$NRE = 0.1563 + 0.0021(\pm 2.2442 \times 10^{-4})PI - 0.0050(\pm 5.4031 \times 10^{-4})\ Sd,$$
$$n = 13, \quad Se = 0.0113, \quad R = 0.9480, \quad R^2A = 0.8784, \quad F = 44.337. \quad (4)$$

Comparison of equations (3) and (4) indicates that the coefficient of both PI and Sd terms remains almost constant. This means that the models are quite stable and the chosen correlating parameters (PI, Sd) are the most appropriate parameters.

In order to confirm our results we have estimated NRE using (equation (4)) and compared them with the observed NRE. The correlation of observed and estimated NRE gave predictive correlation coefficient (R^2_{Pred}) of the order of 0.8981 ($R = 0.9477$) indicating that the proposed model has excellent predictive power too (figure 1).

2.2. *Topological estimation of % benzene character*

We now discuss the topological estimation of % benzene character using the same set of topological indices and follow simple as well as multiple regression analysis. The results of such modelings are are given in table 6.

Before discussing the results it is interesting to comment on % benzene character of polyacenes and helicenes used. It is estimated using the following
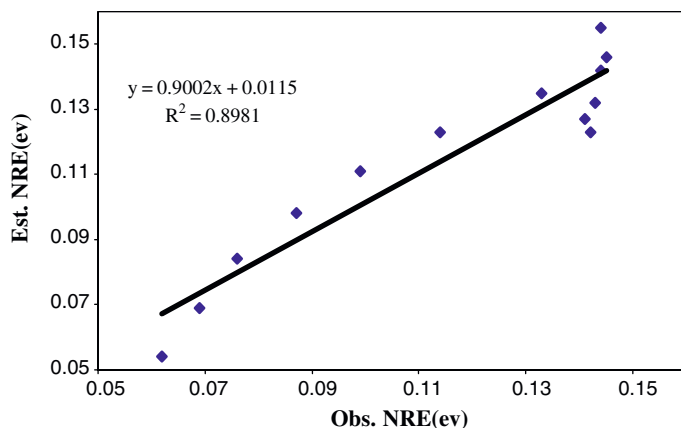
Figure 1. Correlation of observed and estimated NRE using model (1), i.e. bi-parametric modeling containing PI and Sd indices are correlating parameters.

expression:

$$\% \text{ benzene character (BC)} = \frac{\text{NRE (molecule)}}{\text{NRE (benzene)}} \times 100. \tag{5}$$

This parameter is used for the reason that some people believe that the only truly aromatic compound is benzene. The above equation (5) is a measure of the benzene character of a given polyacene or helicene. It may also serve as a finer classification within an isomeric class of aromatic molecules. A perusal of table 2 shows that % BC for polyacene rapidly decrease and thus follows the decrease in the stability of polyacenes with the size.

From table 6, we observed that the results obtained are more or less similar to those obtained in the estimation of NRE. Here also, no mono-parametric correlations are possible to estimate both NRE and % BC and that tremendous improved in the statistic occured as we go to bi-parametric regression, while there is no further improvement in statistics while attempting tri-parametric regression analysis. Also, that here too the best model is the bi-parametric model containing PI and Sd as the correlating parameters. This model is found as below:

$$\% \text{ BC} = 107.3098 + 1.3688(\pm 0.1474)\text{PI} - 3.3115(\pm 0.3551)\text{Sd},$$
$$n = 14, \quad Se = 7.4370, \quad R = 0.9441, \quad R^2 A = 0.8715, \quad F = 45.097. \tag{6}$$

Thus, similar model is resulted for modeling % BC also. We have estimated % BC using this model (equation (6)) and compared with the observed % BC. This comparison is shown in table 7. The residue i.e. difference between observed and estimated % BC shows that the estimated % BC are very close to their

Table 7

Comparison of estimated biological activity %BC with their observed values.

S.N.	Actual	%BC					
		Pred.	Res.	Pred.	Res.	Pred.	Res.
1.	100	100.424	−0.424	100.254	−0.254	100.791	−0.791
2.	91	93.012	−2.012	92.915	−1.915	93.212	−2.212
3.	97	87.811	9.189	87.790	9.210	87.854	9.146
4.	79	85.073	−6.073	85.057	−6.057	85.096	−6.096
5.	80	76.609	3.391	76.686	3.314	76.427	3.573
6.	68	76.609	−8.609	76.661	−8.661	76.485	−8.485
7.	96	84.044	11.956	84.142	11.858	83.862	12.138
8.	60	67.618	−7.618	67.707	−7.707	67.417	−7.417
9.	99	89.123	9.877	89.232	9.768	88.930	10.070
10.	53	58.102	−5.102	58.174	−5.174	57.932	−4.932
11.	90	94.980	−4.980	95.028	−5.028	94.898	−4.898
12.	47	48.059	−1.059	48.081	−1.081	48.070	−1.070
13.	99	103.048	−4.048	102.979	−3.979	103.156	−4.156
14.	43	37.490	5.510	37.295	5.705	37.871	5.129

observed values. However, compared to other compounds, the compound 7 has slightly large residue. Deleting this compound 7 from the regression procedure yielded much-improved model:

$$\% \, BC = 105.7682 + 1.3531(\pm 0.1337) \, PI - 3.2723(\pm 0.3220)Sd,$$
$$n = 13, \quad Se = 6.7297, \quad R = 0.9561, \quad R^2A = 0.9141, \quad F = 53.238. \quad (7)$$

In obtaining this model, we observed that R is increased from 0.9441 to 0.9561 and Se decreased from 7.4370 to 6.7297.

Using this improved model (equation (7)) we have estimated % BC and compared with their observed values. Then, we have correlated observed and estimated % BC for obtaining using predictive correlation coefficient R^2_{Pred}, which found as 0.9141 ($R = 0.9561$). This indicates that the proposed model has excellent predictive power also (figure 2).

3. Comments on co-linearity defect

The correlation matrix (table 3) has indicated that all the four topological indices are highly linearly correlated. This means that a model containing any combination of these indices will suffer from the defect due to co linearity. Thus, the proposed models for estimating NRE and % BC appears to suffer from the defect due to co-linearity. The change in sign of PI/Sd further supports the occurrence of collinearity defect. Statistically such models are not considered
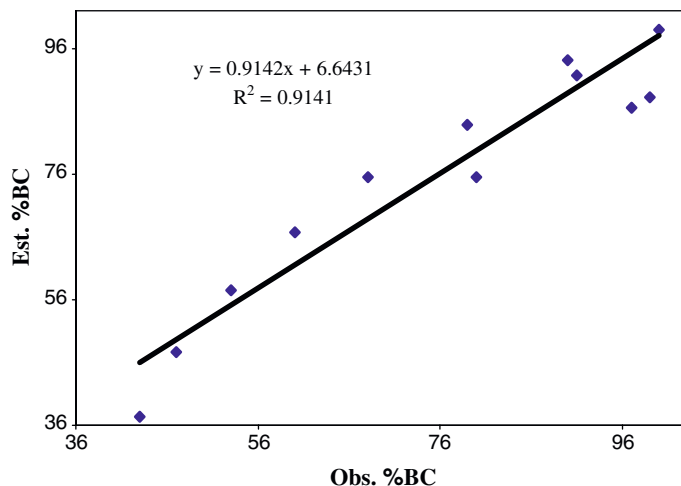
Figure 2. Correlation of observed and estimated % BC using model (1), i.e. bi-parametric modeling containing PI and Sd indices are correlating parameters.

good. However, Randic (MR) has thoroughly investigated such a problem and proposed that one can use model containing highly linearly topological indices safely for modeling the activity. According to Randic [41,42] under certain situations even highly correlated descriptors could be retained in the model. We will, therefore, used Randic [41,42] recommendations in the present case also. Randic stated that if a descriptor strongly correlates with another descriptor already used in a regression, such a descriptor in most studies should be discarded. For example $^1\chi$ and $^2\chi$, $^1\chi$ often strongly correlate and in many structure–property–activity studies $^2\chi$ has been discarded. This is not theoretically justified and despite the widespread practice should be stopped. Although two highly correlated descriptors overall depict the same features of molecular structure, it is important to recognize that even highly interrelated descriptors differ in some other structural traits. The difference between them may be relatively small but nevertheless very important for structure–property regression.

The criteria for inclusion or exclusion of descriptors should not be based on parallelism between descriptors even if overwhelming, but should be based on whether the part in which two descriptors disagree is or is not relevant for the characterization of the property considered. If the part in which the second descriptor differ from the first, regardless of how small it is, is relevant for the property under consideration, then the descriptor should be included. Randic further stated that the selection of descriptors to be used in structure–property–activity studies should not be delegated solely to computers [41,42] although statistical criteria will continue to be useful for preliminary screening of descriptors taken from a large pool. Often in an automated selection of descriptors, a descriptor will be discarded because it is highly correlated with another

descriptor already selected. But what is important is not whether two descriptors parallel one another, i.e. duplicates much of the same structural information, but whether they are complementary in those parts that are important for structure–property–activity correlations. Hence, the residual of the correlation between two descriptors should be examined and kept or discarded depending on how well it can improve the correlation based on already selected descriptors. In view of Randic [41,42], recommendations and the fact that PI and Sz indices have different information contents, these highly correlated descriptors can be retained in the proposed models.

At this stage, it is worth mentioning that problems caused by co-linearity; and how to deal with them, continue to be of prime concern to theoretical statistician. From a decision maker's viewpoint, one should be aware of that co-linearity can (and usually does) exist and recognize the basic problems it can cause. Some of the most obvious problems and indications of severe multi-colinearity are:

(1) Incorrect signs on the coefficient.

(2) A change in the values of the previous coefficient when a new variable is added to the model.

(3) Change to insignificant of a previously significant variable when a new variable is added to the model.

(4) An increase in the standard error of the estimate when a variable is added to the model.

In the present case none of the correlating variables have their coefficients smaller than the respective standard error.

4. Experimental

4.1. NRE and % BC

We have adopted NRE as well as % BC from our earlier work [1].

4.2. Topological indices

All the four topological (W, Sz, PI, and Sd) indices were calculated from the completely hydrogen-suppressed molecular graphs of polyacenes and helicenes. The calculations of these indices are well documented in the literature and thus, they are not given here. However, below we give the expressions used for their calculations:

(a) *Wiener index (W):*

$$\text{(i) Polyacenes: } 1/3[16h^3 + 36h^2 + 26h + 3], \tag{8}$$

$$\text{(ii) Helicenes: } 1/3[8h^3 + 27h^2 - 26h + 27]. \tag{9}$$

(b) *Szeged index (Sz):*

$$\text{(i) Polyacenes: } 1/3[44h^3 + 72h^2 + 43h + 3], \tag{10}$$

$$\text{(ii) Helicenes: } 1/3[16h^3 + 204h^2 - 157h + 99]. \tag{11}$$

(c) *Padmakar-Ivan index (PI):*

$$\text{(i) Polyacenes: } 24h^2. \tag{12}$$

$$\text{(ii) Helicenes: } 25h^3 - 3h + 2. \tag{13}$$

(d) *Sadhana index (Sd):*

$$\text{(i) Polyacenes: } 2h(5h + 1). \tag{14}$$

$$\text{(ii) Helicenes: } 2h(5h + 1). \tag{15}$$

These indices were calculated either by using the software of Lukovits or that of Sisodia

4.3. *Regression analysis*

Regression analysis was made using maximum-R^2 method following stepwise regression analysis. This is done using Regress-1program of Lukovits.

5.   Conclusions

The results and discussion made above indicated that we can use distance-based topological indices such as W, Sz, PI, and Sd for modeling, monitoring and estimated NRE and % BC and thus the aromatic stabilities of polyacenes and helicenes. Even though these classes of compounds differ in their properties there is no need to deal them as a separate class of compounds. This is probably due to the fact that out of the four topological indices used only Sd index has a common expression for the calculation of Sd of polyacenes and helicenes. The results also indicate that though in both the cases excellent models are obtained, the methodology used by us gives slightly better results for modeling % BC.

We will like to comment that although the number of data point available was probably too small to make meaningful regression analysis, the correlation attempted by us gave excellent results. Consequently, we conclude that our approach should be viewed as a viable additional new technique in the growing

armory of data-reduction techniques applicable to multi cyclic systems. To conclude, it we stressed that the correlating ability of the technique we have introduced is confined to benzenoid system, such as polyacenes and helicenes. The long-term goal to obtaining correlation, which should place polyacenes and helicenes on a single curve, is to some extent, fulfilled.

Acknowledgments

References

[1] M. Randic, M. Benjamin and N. Trinajstic, On the aromatic stabilities of polyacenes and helicenes croatia, Chem. Acta 59(2) (1986) 345.
[2] E. Clar, *Polycyclic Hydrocarbons* (Academic Press, London, 1964).
[3] R.H. Martin, Angew. Chem. Int. Edit. English 13 (1974) 649.
[4] E. Clar, J.M. Robertson, R. Schlogel and W. Schmidt, J. Amer. Chem. Soc. 103 (1981) 1320.
[5] M.L. Lee, M.V. Novotny and K.D. Bartle, *Analytical Chemistry of Polycyclic Aromatic Hydrocarbons* (Academic, New York 1981).
[6] A.T. Balaban, Pure Appl. Chem. 54 (1982) 1075.
[7] H. Angliker, E. Rommel and J. Wirz, Chem. Phys. Lett. 87 (1982) 208.
[8] J. Pataki and R.G. Harvey, J. Org. Chem. 47 (1982) 20.
[9] J.K. Ray and R.G. Harvey, J. Org. Chem. (1982) 47 3335.
[10] H. Meier, T. Molz, U. Merkle, E. Echter and M. Lorch, Liebigs Ann. Chem. (1982) 314.
[11] B.N. Cyvin, P. Klaeboe, J.C. Whitner and S.J. Cyvin, Z. Naturforsch. 37a, (1982) 251.
[12] J.V. Knop, K. Szymanski, Z. Jericevic and N. Trinajstic, J. Comput. Chem. 4 (1983) 23.
[13] D.N. Nicolaides and K.E. Litinas, J. Chem. Res. (S) (1983) 57.
[14] S. Mondal, T.K. Bandyopadhyay and A.J. Bhattacharaya, Ind. J. Chem. 22B (1983) 448.
[15] K. Yamamoto, T. Harada, M. Nakazaki, T. Naka, Y. Kai, S. Harada and N. Kasai, J. Amer. Chem. Ber. (1983) 105, 7171.
[16] H. Staab and F. Diederich, Chem. Ber. 116 (1983) 3487.
[17] L. Klasinc, B. Kovac and H. Gusten, Pure Appl. Chem. 55 (1983) 289.
[18] A.T. Balaban and I. Tomescu, Math. Chem. (Mulheim/Ruhr) 14 (1983) 155.
[19] S. El-Basil, Croat. Chem. Acta 57 (1984) 1.
[20] M. Randic and N. Trinajstic, J. Amer. Chem. Soc. 106 (1984) 4428.
[21] J.R. Dias, J. Chem. Inf. Comput. Sci. 22 (1982) 15; ibid. 22 (1982) 139; ibid. 24 (1984) 124.
[22] J.V. Knop, W.R. Muller, K. Szymanski and N. Trinajstic, Computer Generation of Some Classes of Molecules (SKTH, Zagreb, 1985).
[23] D.R. Boyd and D.M. Jerina, in: Small Ring Heterocycles, *Part III* Ed. A. Hassner (Wiley, New York, 1985), p. 197.
[24] R.A. Aleksejczyk, G.A. Berchtold and A.G. Braun, J. Amer. Chem. Soc. 107, (1985) 2554.
[25] G.W. Gribble, C.S. Lettoullier, M.P. Sibi and R.W. Allen, J. Org. Chem. 50 (1985) 1611.
[26] L.A. Levy and V.P. Sashikumar J. Org. Chem. 50 (1985) 1760.
[27] J.R. Dias, Theoret. Chim. Acta in Press.
[28] H. Wiener, J. Am. Chem. Soc. 69 (1947) 17.
[29] P.V. Khadikar, N.V. Deshpande, P.P. Kale, A. Dabrynin, I. Gutman and G. Domotor, The Szeged Index and an Analogy with the Wiener J. Index, Chem. Inf. Comput. Sci. 35 (1995) 545.

[30] I. Gutman, P.V. Khadikar, P.V. Rajput and S. Karmarkar, The szeged index of polyacenes, J. Serb. Chem. Soc. 60 (1995) 759.

[31] P.V. Khadikar, S. Karmarkar, S. Joshi and I. Gutman, Estimation of protonation constants of salicylhydroxamic acids by means of the wiener topological index, J. Serb. Chem. Soc. 61 (1996) 89.

[32] P.V. Khadikar, S. Karmarkar, V.K. Agrawal, J. Singh, A. Shrivastava, I. Lukovits and M.V. Diudea, SzegedIndex – applications for drug modeling, Lett. Drug Design Discovery 2 (2005) 606.

[33] P.V. Khadikar, V. Sharma, S. Karmarkar and C.T. Supuran, QSAR studies on Benzene sulfon-amide carbonic anhydrase inhibitors: need of hydrophobic parameter for topological modeling of binding constants of sulfonamides to human CA-II, Bioorg. Med. Chem. Lett. 15 (2005) 15 923.

[34] P.V. Khadikar, V. Sharma, S. Karmarkar and C.T. Supuran, Novel use of chemical shift in NMR as molecular descriptor: a first report on modeling carbonic anhydrase inhibitory activity and related parameters, Bioorg. Med. Chem. Lett. 15 (2005) 931.

[35] M. Jaiswal and P.V. Khadikar, Use of distance-based topological indices for the estimation of ^{13}C NMR shifts : a case of Benzene derivatives, J. Indian Chem. Soc. 82 (2005) 247.

[36] P.V. Khadikar, S. Singh, M. Jaiswal and D. Mandoli, Topological estimation of electronic absorption bands of arene absorption spectra as a tool for modeling their toxcity environmental pollution, Bioorg. Med. Chem. Lett. 14 (2004) 4795.

[37] S. Chaterjee, A.S. Hadi and B. Price, Regression Analysis by Examples, 3rd edn. (Wiley, New York, 2000).

[38] C. Hansch, in: Correlation Analysis in Chemistry, Vol. 397 Eds. N.B. Chapman and J. Shorter) (Plenum Press, NY, 1978). N. Draper and H. Smith, *Applied Regression Analysis* (Wiley, NY, 1966).

[39] J. Devillers and A.T. Balaban, Topological Indices and Related Descriptors in QSAR and QSPR, (Gorden & Breach, Williston, VT, 2000).

[40] L. Pogliani, Amino acids, J. Phys. Chem. 98 (1994) 1494.

[41] M. Randic, Acta Chem. Slov. 45 (1998) 239.

[42] M. Randic, J. Chem. Inf. Comput. Sci. 37 (1997) 672.