Iranian J. Math. Chem. 13 (1) March (2022) 73 – 84

IRANIAN JOURNAL OF
Mathematical Chemistry
Journal homepage: ijmc.kashanu.ac.ir

Original Scientific Paper

QSPR Analysis of Some Novel Extension and Generalization of Sombor Index

ANKUR BHARALI[1], JIBONJYOTI BURAGOHAIN[2,•] AND AMITAV DOLEY[3]

[1]Department of Mathematics, Dibrugarh university, India-786004
[2]Department of Mathematics, Pragjyotish College, India-781009
[3]Departmrnt of Mathematics, DSHK College, India-786001

ARTICLE INFO

Article History:

Received: 28 February 2022
Accepted : 4 March 2022
Published online: 30 March 2022
Academic Editor: Ivan Gutman

Keywords:

Topological index
Distance
Status of a vertex
QSPR analysis
Sombor index

ABSTRACT

Although many indices are strongly correlated with various chemical as well as physical properties of a molecular compound but some of them lack proper geometrical interpretations. A newly introduced index called Sombor index is able to catch the attention of the researchers because of its connection with the notion of "2-norm". In this paper, we propose the status version of the Sombor index. Further, we discuss a generalization of our proposed index and carry out QSPR analysis. Some mathematical properties of the generalization are also discussed.

1. INTRODUCTION

A graph is any collection of vertices and the pairwise relationships between them known as edges. For a graph G, the vertex set is denoted as $V(G)$ and the edge set is denoted as $E(G)$. Vertices in a graph may represent any objects or any kind of physical situations [3, 5]. The degree of a vertex u in a graph G is a whole number that is assigned to the vertex by counting the number of vertices in G that are directly connected to the vertex u. We use

•Corresponding author (Email: j.bgohain75@yahoo.com).
DOI: 10.22052/IJMC.2022.246111.1613

$d_G(u)$ or $d(u)$ to denote the degree of the vertex u in G. An edge between two vertices u and v is denoted as uv. Other graph theoretic notions used but not defined are taken from [19].

Chemical graph theory is the branch of graph theory where molecular graphs of chemical compounds are studied. In the field of chemical graph theory a mathematical quantity called topological index is proven to be very useful in various QSPR/QSAR studies. A topological index can be termed as a numerical value that can be computed from a graphical representation of a molecular constitution. The general form of a topological index for a graph G is given as $TI(G) = \sum_{edges} F(x, y)$, where x, y may be degree or distance or eccentricity etc. of vertices of the graph G.

It is interesting to note that the number of topological indices defined till today is growing enormously. Sometimes the indices are defined without proper justification to their mathematical formulation. Again, a lot of them do not have any notable chemical correlation as well as any geometrical interpretation [15]. Some of the important and widely applied topological indices to different fields of chemistry can be found in [2, 7–9, 12]. In 2021, Gutman, being motivated by the notion of Euclidean metrics, defined a novel degree based topological index named Sombor index [6]. The Sombor index is defined as $SO(G) = \sum_{uv \in E(G)} \sqrt{d(u)^2 + d(v)^2}$. Some of the mathematical properties and other study related to Sombor index can be found in [16, 17, 20].

Although various physico-chemical properties of molecules can be predicted by many classical topological indices, but there are still some important properties of molecules that can not be captured by the existing topological indices. For example, the inverse sum indeg (ISI) index is a strong predictor of total surface area of octane isomers [18], but this index may not be a good predictor of Acentric Factor of octane isomers. In 2020, Doley et al. [4] introduced a new status based topological index, called inverse sum indeg status index, which is an extension of ISI index and they have shown that this index has a strong correlation with Acentric Factor of octane isomers. Now a days, extension and generalization of the classical topological indices become as an interesting area of research. Some extension and generalization of the classical topological indices may be found in [1, 12–14].

A molecule is often represented by a graph. A Molecular graph or chemical graph is a simple finite connected graph whose nodes and edges correspond the atoms and chemical bonds between the atoms, respectively, of a molecule. Hydrogen atoms are often dropped in the representation.

Let G be a graph. The status of a vertex $v \in V(G)$ is defined by $\sigma_G(v) (\text{or } \sigma(v)) = \sum_{u \in V(G)} d(u, v)$, where $d(u, v)$ is the distance between u and v in G. The *diameter*, $diam(G)$ of G is the maximum eccentricity among the vertices of G, where eccentricity $\epsilon(v)$ of a vertex v is the maximum distance from v to other vertices in G. The *complement graph*, $\overline{G}$ of G is the graph with $V(G) = V(\overline{G})$ and $uv \in E(\overline{G})$ if and only if $uv \notin E(G)$. A

graph G is called a *k-distanced balanced* graph if $\sigma_G(u) = k$ for all $u \in V(G)$. Let G_1 and G_2 be two vertex disjoint graphs. Then, *join* of G_1 and G_2, denoted by $G_1 \vee G_2$ is the graph union $G_1 \cup G_2$ together with all the edges joining $V(G_1)$ and $V(G_2)$. Note that $diam(G_1 \vee G_2) = 2$.

In this papar, we communicate the status version of the newly proposed Sombor index. The chemical correlation and isomer discrimination power of the status Sombor index is established for octane isomers and benzenoid hydrocarbons. The status Sombor index is further generalized to check if the generalization can better the correlation or not. The worth of the generalization is then confirmed with the help of a linear regression analysis. Some of the basic mathematical properties of our proposed generalized index is also derived for some graph classes.

2. SOMBOR STATUS INDEX

Motivated by the works of [1, 12–14], we analogously define the status version of the Sombor index and called it Sombor status index (SOS) whose definition goes as follows.

$$SOS(G) = \sum_{uv \in E(G)} \sqrt{\sigma(u)^2 + \sigma(v)^2}.$$

This section aims to discuss the chemical correlations of the SOS index with various physico-chemical properties and physical properties [Acentric factor (AcentFac), Entropy (S), Total surface area (TSA), Standard enthalpy of vaporization (DHVAP), enthalpy of vaporization (HVAP)] of octane isomers and also to investigate its correlation with the boiling points (bp) of the benzenoid hydrocarbons.

To carry out the regression analysis we use the following linear models:

$$\text{AcentFac} = 0.0019 \times [SOS(G)] + 0.0304,$$
$$\text{TSA} = 0.5772 \times [SOS(G)] + 295.54,$$
$$\text{DHVAP} = 0.0177 \times [SOS(G)] + 6.3191,$$
$$\text{HVAP} = 0.084 \times [SOS(G)] + 55.805,$$
$$\text{S} = 0.2224 \times [SOS(G)] + 70.059.$$

The Table 1 enlists the experimental vales of AcentFac, TSA, DHVAP, HVAP of all the 18 molecules of octane isomers and the calculated values of SOS index of the molecules. The Table 1 contains the experimental values of BP of 21 benzenoid hydrocarbons and their calculated values of SOS index.

Table 1: Acentric factor, Entropy, Total surface area, Standard enthalpy of vaporization, enthalpy of vaporization and SOS index (*correct upto fourth decimal place*) *of octane* (C_8H_{18}) *isomers.*

S. No.	Octane isomers	AcentFac	Entropy	TSA	DHVAP	HVAP	SOS index
1	n-octane	0.3979	111.67	415.3	9.915	73.19	198.8625
2	2-methylheptane	0.3779	109.84	407.85	9.484	70.3	184.9815
3	3-methylheptane	0.3710	111.26	397.34	9.521	71.3	176.6908
4	4-methylheptane	0.3715	109.32	396.04	9.483	70.91	173.9367
5	3-ethylhexane	0.3625	109.43	379.04	9.476	71.7	165.6485
6	2,2-dimethylhexane	0.3394	103.42	405.11	8.915	67.7	162.9075
7	2,3-dimethylhexane	0.3482	108.02	384.93	9.272	70.2	160.1412
8	2,4-dimethylhexane	0.3442	106.98	388.11	9.029	68.5	162.8715
9	2,5-dimethylhexane	0.3568	105.72	395.08	9.051	68.6	171.1313
10	3,3-dimethylhexane	0.3226	104.74	389.79	8.973	68.5	151.936
11	3,4-dimethylhexane	0.3403	106.59	376.91	9.316	70.2	154.6431
12	3-ethyl, 2-methylhexane	0.3324	106.06	368.1	9.209	69.7	151.8889
13	3-ethyl, 3-metylpentane	0.3069	101.48	366.99	9.081	69.3	143.7212
14	2,2,3-trimethylpentane	0.3008	101.31	371.75	8.826	67.3	140.9582
15	2,2,4-trimethylpentane	0.3054	104.09	392.19	8.402	64.87	149.1358
16	2,3,3-trimethylpentane	0.2932	102.06	377.4	8.897	68.1	138.2408
17	2,3,4-trimethylpentane	0.3174	102.39	368.93	9.014	68.37	146.3941
18	2,2,3,3-tetramethylbutane	0.2553	93.06	390.47	8.41	66.2	127.3499

Table 2: Boiling point(BP) and SOS index of benzenoid hydrocarbons.

Sl. no.	BP(^{0}C)	SOS index	Sl. no.	BP(^{0}C)	SOS index
1	218	333.4249791	12	542	2760.221354
2	338	856.9944559	13	535	2957.290328
3	340	884.6924979	14	536	3182.320878
4	431	1759.608486	15	531	3126.510917
5	425	1787.535181	16	519	3154.316864
6	429	1647.305353	17	590	3461.12216
7	440	1843.287308	18	592	3601.892365
8	496	2255.263219	19	596	3814.186703
9	493	2155.615974	20	594	3814.173975
10	497	2164.047945	21	595	3546.099715
11	547	2847.175304			

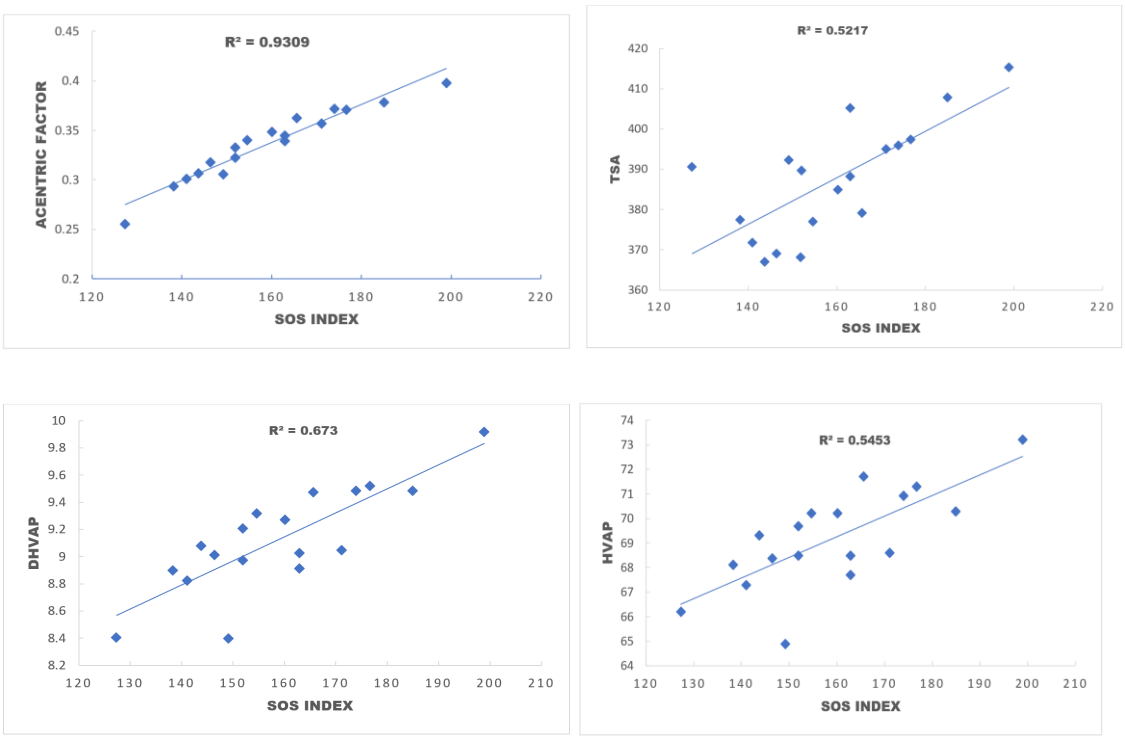

Figure 1: Correlation of SOS index with AcentFac, TSA, DHVAP, HVAP, S of octane isomers.

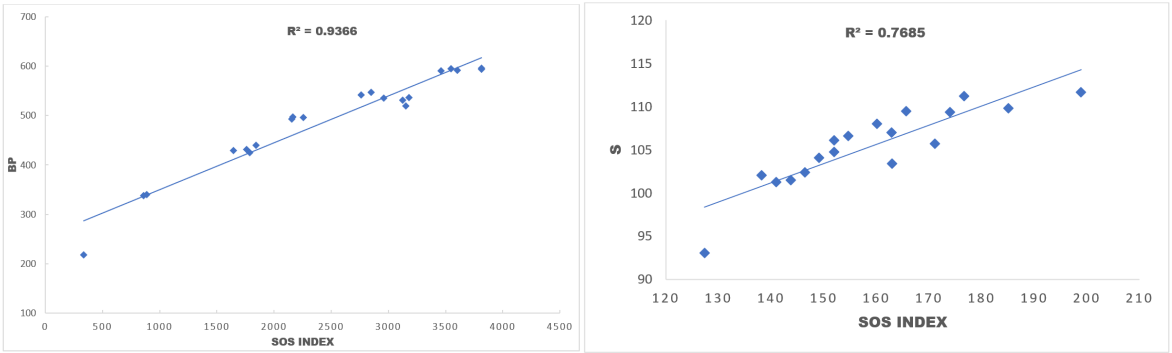

Figure 2: Correlation of SOS index with BP of benzenoid hydrocarbons.

The correlations of the proposed SOS index with various properties of octane isomers are depicted in the Figure 1. It is found that the correlation coefficient between Acentric Factor of octane isomers and the SOS index is 0.96 (*correct upto two decimal places*). See Table 3. The values of R^2, SEE and F-value are also given in Table 3.

Table 3: Correlaion coefficient(R), Coefficient of determination(R^2), Standard error estimates(SEE) and F-value for SOS index.

Properties	R	R^2	SEE	F-value
AcentFac	0.9648	0.9309	0.0096	3.02931E-42
TSA	0.7223	0.5217	10.1439	0.364213984
DHVAP	0.8204	0.673	0.2259	1.13567E-24
HVAP	0.7385	0.5453	1.4083	1.84854E-12
S	0.8766	0.7685	2.2407	7.24867E-07

The degeneracy of a topological index is a good parameter to determine the isomer discrimination power of the index. We use sensitivity, a measure of degeneracy introduced by Konstantinova [10], given by the following formula:

$$S_{TI} = \frac{N - N_{TI}}{N},$$

where N =no. of isomers considered, N_{TI} =no. of isomers that have the same TI value.

The isomer discriminating power of TIs increases as the value of S_{TI} gets closer to 1. In general, the isomer discriminating power of degree based topological indices are more than the other classes of TIs [11]. But we found that the sensitivity value of the SOS index for both the octane and benzenoid hydrocarbons are 1, which indicates that the SOS index has a good isomer discriminating power.

3. GENERALIZATION OF SOMBOR STATUS INDEX AND REGRESSION ANALYSIS

In this section we first propose a generalization of the SOS index and then we present a linear regression analysis to study its correlations with properties of octane and benzenoid hydrocarbons and hence establish its applicability as a molecular property predictor. We adopt the symbol "SOS_p" to denote the generalized SOS index and for a graph G it is defined by the following formula:

$SOS_p(G) = \sum_{uv \in E(G)} \sqrt{\sigma(u)^p + \sigma(v)^p}$, where p is a positive real number.

Although this generalization is proposed keeping p norm in mind but we restrict ourselves to square root rather than p^{th} root to reduce computational complexities as well

as standard errors. We now calculate the values of SOS_p index for the octane isomers and benzenoid hydrocarbons and then investigate correlations.

The regression line between various properties of octane isomers, benzenoid hydrocarbons and SOS_p index for different values of p are depicted in Figure 3, 4 and 5. We observe that the correlation between the SOS_p index and some properties (more precisely BP, HVAP, DHVAP and TSA) of ocatne isomers gets stronger with the increase in the value of p while for some properties (S, AcentFac), the correlation gets weaker with the increase in the value of p. In the case of BP of benzenoid hydrocarbons, it is seen that the correlation gets weaker with the increase in the value of p.

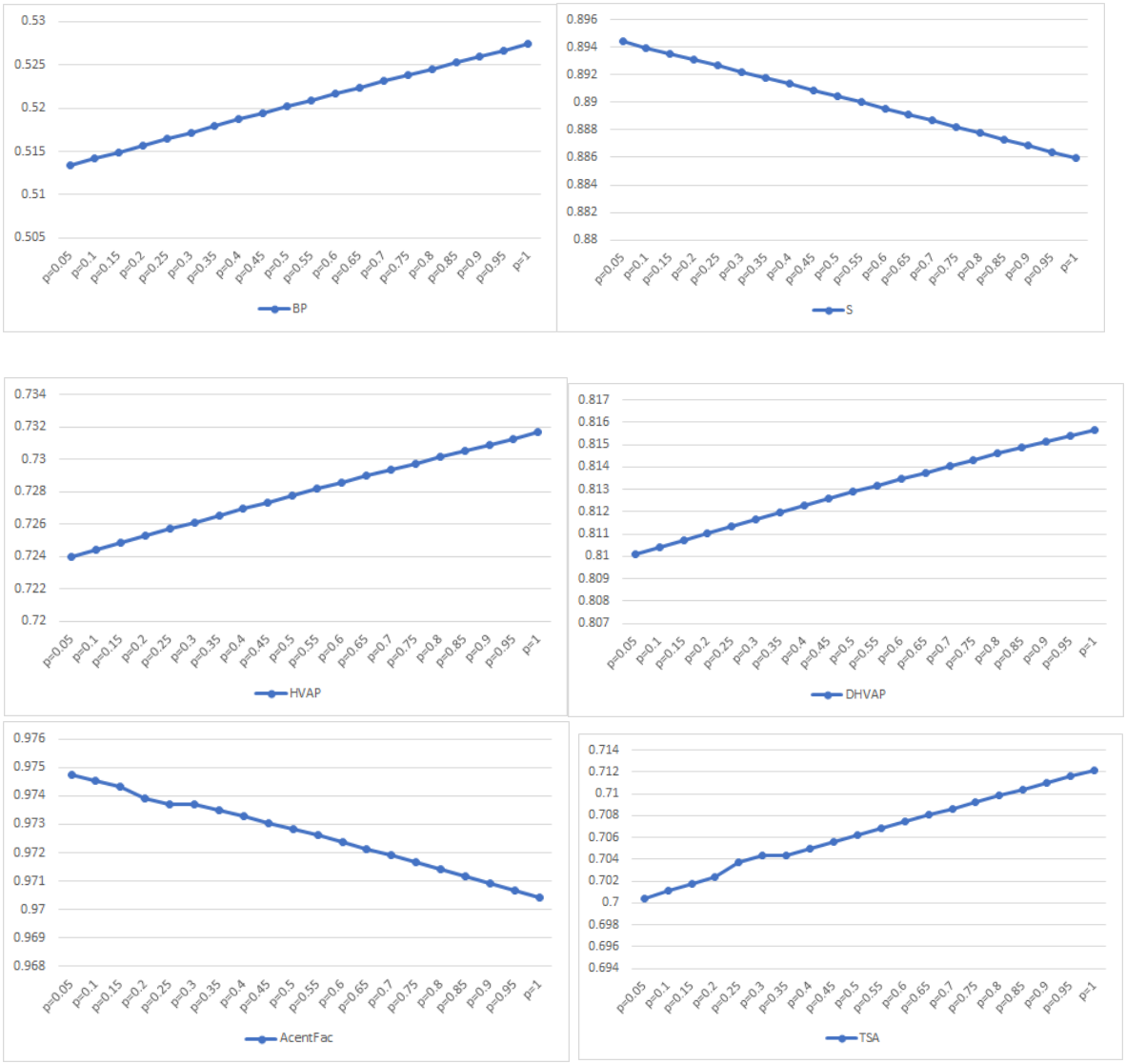

Figure 3: Regression line between BP, S, HVAP, DHVAP, AccentFac, TSA of octane isomers and SOS_p index; $0 < p \leq 1$.

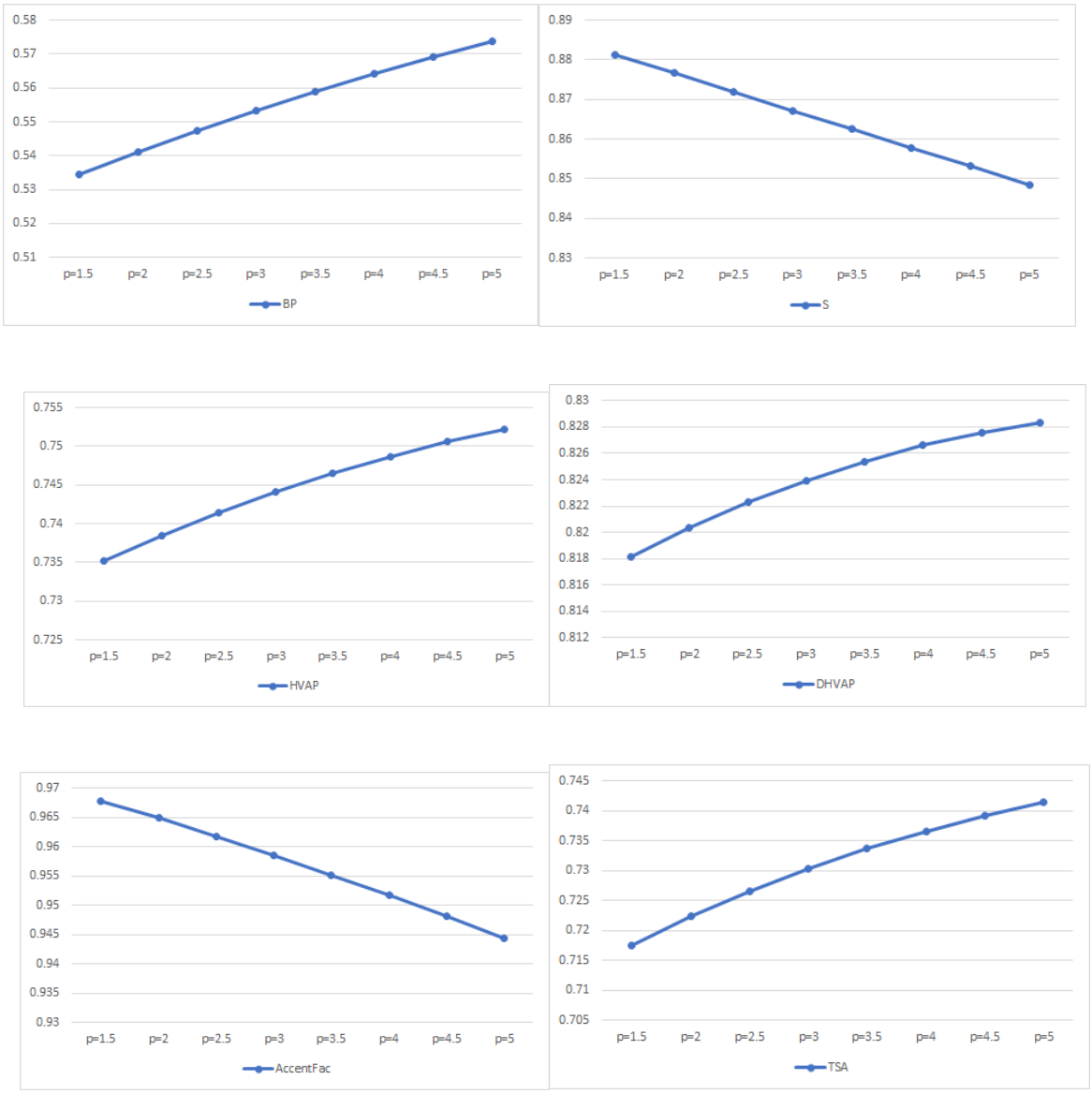

Figure 4: Regression line between BP, S, HVAP, DHVAP, AccentFac, TSA of octane isomers and SOS_p index; $1 < p \leq 5$.

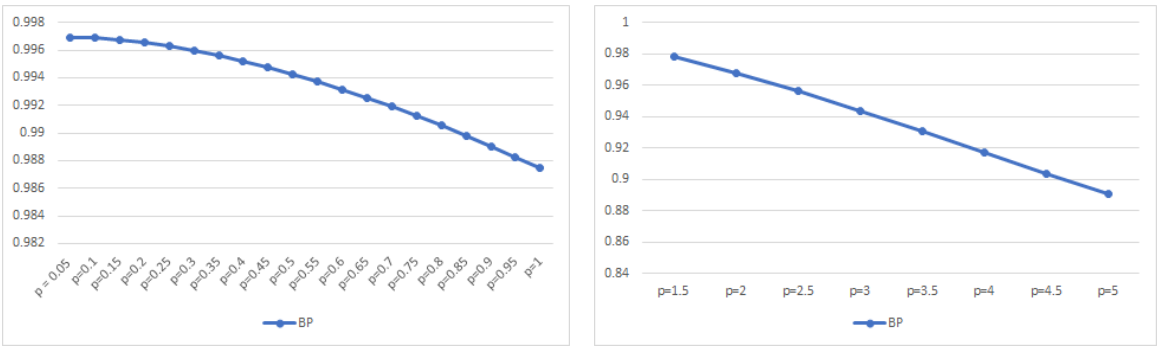

Figure 5: Regression line between boiling point (BP) of benzenoid hydrocarbons and SOS_p index.

4. SOS_p Index of Graphs with Given Diameter

Here, we obtain some basic properties of SOS_p index in connection with some graphs with given diameter.

Theorem 1. *Let G be a connected graph with n nodes and $diam(G) = D$. Then,*

$SOP_p(G) \leq \sum_{uv \in E(G)} \sqrt{[D(n-1) - (D-1)d(u)]^p + [D(n-1) - (D-1)d(v)]^p}$, where $p \in (0, \infty)$, with equality if and only if $D \leq 2$.

Proof. For any node u of G, the no. of nodes at distance 1 from u is $d(u)$. Again, the no. of nodes at distance at most $diam(G) = D$ from u is $(n-1) - d(u)$. Therefore,

$$\sigma(u) \leq d(u) + D(n-1-d(u)) = D(n-1) - (D-1)d(u)$$
$$\Rightarrow (\sigma(u))^p \leq [D(n-1) - (D-1)d(u)]^p, p \in (0, \infty).$$

So we have, $\sigma(u)^p + \sigma(v)^p \leq [D(n-1) - (D-1)d(u)]^p + [D(n-1) - (D-1)d(v)]^p$ which follows that

$SOS_p(G) \leq \sum_{uv \in E(G)} \sqrt{[D(n-1) - (D-1)d(u)]^p + [D(n-1) - (D-1)d(v)]^p}$.

Equality holds if and only if $D = 1$ or $D = 2$. ■

Theorem 2. *If graph G is isomorphic to $G_1 \vee G_2$, where G_1 and G_2 are any two vertex disjoint graph with n_1 and n_2 nodes respectively. Then,*

$$SOS_p(G) = \sum_{uv \in E(G_1)} \sqrt{[2\lambda - (d_{G_1}(u) + n_2)]^p + [2\lambda - (d_{G_1}(v) + n_2)]^p}$$
$$+ \sum_{uv \in E'(G)} \sqrt{[2\lambda - (d_{G_1}(u) + n_2)]^p + [2\lambda - (d_{G_2}(v) + n_1)]^p}$$
$$+ \sum_{uv \in E(G_2)} \sqrt{[2\lambda - (d_{G_1}(u) + n_1)]^p + [2\lambda - (d_{G_2}(v) + n_1)]^p},$$

where $E'(G) = \{uv | u \in V(G_1), v \in V(G_2)\}$ and $\lambda = n_1 + n_2 - 1$.

Proof. We have $diam(G) = 2$. The set $\{E_1, E_2, E_3\}$ is a partition of the edge set $E(G)$ where $E_1 = \{uv | u, v \in V(G_1)\}$, $E_2 = \{uv | u \in V(G_1), v \in V(G_2)\}$ and $E_3 = \{uv | u, v \in V(G_2)\}$. Again, if $u \in V(G)$, then

$$d_G(u) = \begin{cases} d_{G_1}(u) + n_2, & \text{if } u \in V(G_1), \\ d_{G_2}(u) + n_1, & \text{if } u \in V(G_2). \end{cases}$$

Now, from Theorem 1 we have that

$$SOS_p(G) = \sum_{uv \in E(G)} \sqrt{[2(n-1) - d(u)]^p + [2(n-1) - d(v)]^p}.$$

So,

$$SOS_p(G) = \sum_{uv \in E_1} \sqrt{[2(n_1 + n_2 - 1) - (d_{G_1}(u) + n_2)]^p + [2(n_1 + n_2 - 1) - (d_{G_1}(v) + n_2)]^p}$$

$$+ \sum_{uv \in E_2} \sqrt{[2(n_1 + n_2 - 1) - (d_{G_1}(u) + n_2)]^p + [2(n_1 + n_2 - 1) - (d_{G_2}(v) + n_1)]^p}$$

$$+ \sum_{uv \in E_3} \sqrt{[2(n_1 + n_2 - 1) - (d_{G_2}(u) + n_1)]^p + [2(n_1 + n_2 - 1) - (d_{G_2}(v) + n_1)]^p}$$

from which the result follows. ■

Theorem 3. *Let G be a connected graph with n nodes and $diam(G) = D$. Then,*

$$SOS_p(G) \geq \sum_{uv \in E(G)} \sqrt{(2n - 2 - d(u))^p + (2n - 2 - d(v))^p}$$

where $p \in (0, \infty)$, with equality if and only if $D \leq 2$.

Proof. There are $n - 1 - d(u)$ number of nodes that are at distance at least 2 from the node u of G. So, for any node u, $\sigma(u) \geq d(u) + 2(n - 1 - d(u)) = 2n - 2 - d(u)$. Therefore,

$$\sqrt{\sigma(u)^p + \sigma(v)^p} \geq \sqrt{(2n - 2 - d(u))^p + (2n - 2 - d(v))^p}.$$

Hence,

$$SOS_p(G) \geq \sum_{uv \in E(G)} \sqrt{(2n - 2 - d(u))^p + (2n - 2 - d(v))^p}.$$

If $D = 1$, then $n - 1 - d(u) = 0$, where $n - 1 - d(u)$ is the no. of nodes at distance at least 2 from the nodes u, and then $\sigma(u) = d(u)$ and so the equality holds good. If $D = 2$, then $\sigma(u) = d(u) + 2(n - 1 - d(u))$. Hence, for $D = 2$ we have

$$SOS_p(G) = \sum_{uv \in E(G)} \sqrt{(2n - 2 - d(u))^p + (2n - 2 - d(v))^p},$$

which completes the proof. ■

Theorem 4. *Let $\overline{G}$ be a connected graph and let $|V(G)| = n$ and $diam(\overline{G}) = D$. Then,*

1. $SOS_p(\overline{G}) \geq \sum_{uv \in E(\overline{G})} \sqrt{[(n-1) - d_G(u)]^p + [(n-1) - d_G(v)]^p}$,
2. $SOS_p(\overline{G}) \leq \sum_{uv \in E(\overline{G})} \sqrt{[(n-1) - (D-1)d_G(u)]^p + [(n-1) - (D-1)d_G(v)]^p}$.

Proof. Let $u \in \overline{G}$. Then the no. of nodes at distance 1 from u in $\overline{G}$ is $n-1-d_G(u)$. Again, no. of nodes at distance at least 2 and at most D from u in $\overline{G}$ is $d_G(u)$. So,

$$\sigma_{\overline{G}}(u) \geq n-1-d_G(u),\ \sigma_{\overline{G}}(u) \leq (n-1)+(D-1)d_G(u).$$

$$\sqrt{\sigma_{\overline{G}}^p(u)+\sigma_{\overline{G}}^p(v)} \geq \sqrt{(n-1-d_G(u))^p+(n-1-d_G(v))^p}$$

and

$$\sqrt{\sigma_{\overline{G}}^p(u)+\sigma_{\overline{G}}^p(v)} \leq \sqrt{(n-1+(D-1)d_G(u))^p+(n-1+(D-1)d_G(v))^p}$$

and so the result follows. ■

5. CONCLUSION

In this presentation, we propose the status version of the newly introduced Sombor index and called it as SOS index and analyze its chemical correlation with some of the important properties like BP, AcentFac, DHVAP, etc., of octane isomers and benzenoid hydrocarbons. It is found that SOS index is highly correlated with the AcentFac of octane isomers with correlation coefficient $R = 0.96$. The correlations of SOS index with BP, DHVAP, HVAP, TSA and S of octane are also impressive. In addition to the correlations, we confirm the excellent isomer discrimination power of the index with the help of a sensitivity parameter. Further, we discuss a generalization of SOS index which we introduce as SOS_p index and carry out a linear regression analysis for some values of p between 0.05 to 5. It is observed that R increases for BP, HVAP, DHVAP, TSA while R decreases for AcentFac and S of octane isomers as the p value increases. But, in case of benzenoid hydrocarbons, R decreases for BP as the p value increases. We also establish some mathematical properties of the SOS_p index for some classes of graphs.

ACKNOWLEDGEMENT. We are very thankful to the reviewers for their comments and suggestions to improve the presentation of the manuscript.

REFERENCES

1. J. Buragohain, B. Deka and A. Bharali, A generalized ISI index of some chemical structures, *J. Molecular Struct.* **1028** (2020) 127843.
2. K. C. Das, Atom-bond connectivity index of graphs, *Discrete Appl. Math.* **158** (2010) 1181–1188.
3. N. Deo, *Graph Theory with Application to Engineering and Computer Science*, PHI Learning Pvt. Ltd., New Delhi, 2016.

4. A. Doley, J. Buragohain and A. Bharali, Inverse sum indeg status index of graphs and its applications to octane isomers and benzenoid hydrocarbons, *Chemometrics Intell. Lab. Sys.* **203** (2020) 104059.
5. F. Harary, *Graph Theory*, Addison-Wesley: Reading, MA, 1969.
6. I. Gutman, Geometric Approach to Degree-Based Topological Indices: Sombor Indices, *MATCH Commun. Math. Comput. Chem*. **86** (2021) 11–16.
7. I. Gutman, Degree-based topological indices, *Croat. Chem. Acta.* **86** (2013) 351–361.
8. I. Gutman, B. Furtula and V. Katanić, Randić index and information, *AKCE Int. J. Graphs Comb.* **15** (2018) 307–312.
9. M. Hasani, Study of inverse sum indeg index, *J. Math. Nanosci*. **7** (2) (2017) 103–109.
10. E. V. Konstantinova, The discrimination ability of some topological and information distance indices for graphs of unbranched hexagonal system, *J. Chem. Inf. Comput. Sci.* **36** (1996) 54–57.
11. S. Mondal, A. Dey, N. De and A. Pal, QSPR analysis of some neighbourhood degree-based topological descriptors, *Complex Intel. Syst*. 7 (2021) 977–996.
12. H. S. Ramane, A. S. Yalnaik and R. Sharafdini, Status connectivity indices and coindices of graphs and its computation to some distance-balanced graphs *AKCE Int. J. Graphs Comb.***17** (1) (2020) 98–108.
13. H. S. Ramane and A. S. Yalnaik, Status connectivity indices of graphs and its applications to the boiling point of benzenoid hydrocarbons, *J. Appl. Math. Comput.* **55** (2017) 609–627.
14. H. S. Ramane, B. Basavanagoud and A. S. Yalnaik, Harmonic Status Index of Graphs, *Bull. Math. Sci. Appl.* **17** (2016) 24–32.
15. M. Randić and J. Zupan, On interpretation of well-known topological indeces, *J. Chem. Inf. Comput. Sci.* **41** (2001) 550–560.
16. I. Redžepović, Chemical applicability of Sombor indices, *J. Serb. Chem. Soc*. **86** (5) (2021) 445–457.
17. H. Liu, H. Chen, Q. Xiao, X. Fang and Z. Tang, More on Sombor indices of chemical graphs and their applications to the boiling point of benzenoid hydrocarbons, *Int. J. Quantum Chem.* **121** (17) (2021) e26689.
18. D. Vukičević and M. Gašperov, Bond additive modelling 1. Adriatic indices, *Croat. Chem. Acta* **83** (2010) 243–260.
19. D. B. West, *Introduction to Graph Theory*, 2nd Ed., Prentice Hall, Inc., Upper Saddle River, NJ, USA, 2001.
20. W. Zhang, L. You, H. Liu and Y. Huang, The expected values and variances for Sombor indices in a general random chain, *Appl. Math. Comput*. **411** (2021) 126521.