EasyChair Preprint

№ 4265

Structural Modelling of Hydrocarbons for the Prediction of Octane Number and Designing of Sustainable Synthetic Fuel

Sanjay Kumar, Mamta Thakur, Naman Shah and Sarthak Jain

September 25, 2020

Structural modelling of Hydrocarbons for the prediction of Octane Number and Designing of Sustainable Synthetic Fuel

Sanjay Kumar
Symbiosis University of Applied Sciences, Indore, M.P, INDIA
vice-chancellor@suas.ac.in

Mamta Thakur*, Naman Shah, Sarthak Jain
Symbiosis University of Applied Sciences, School of Mechatronics Engineering, Indore, M.P., INDIA
mamta.thakur@suas.ac.in, namans504@gmail.com, sjsarthakjain@gmail.com

Abstract: **The objective of the present study is to develop the regression based mathematical model to demonstrate the structural impact on the Octane number of the hydrocarbons. To achieve this goal, a set of 66 (training set of 41 hydrocarbons and test set of 25 hydrocarbon) hydrocarbons has been consider and encoded into their structural descriptors viz., Wiener Index (W), Zero order, First order and Second order Connectivity Index(χ0,χ1,χ2), Shultz Molecular Topological Index(SMTI), Balaban branching Index(J) and Indicator parameter(Ic). The multiple linear regression analysis (MLR) has been performed to obtain the structure property relationship in the form of mathematical model. Moreover, the same model has also been used to predict the Octane numbers of all the hydrocarbons. The model demonstrating pivotal role of Schutz Molecular Topological Index (SMTI), Balaban branching Index (J), Connectivity Indices and indicator parameter in regulating the Octane number of the hydrocarbons. In addition to the statistical parameters, the predictive ability and robustness of the model is further cross validated by the external validation method by applying the model on the test set of 25 hydrocarbons. The proposed QSPR model provides a valuable insight to design the novel Ssynthetic fuel that can be a sustainable energy solution for the future.**

Keyword: **Fuel, Hydrocarbons, Octane Number, Modelling.**

I. INTRODUCTION

As it is well known fact that the knocking property of the fuel and hence the octane rating of the fuel is a function of the molecular structure of hydrocarbons.[1] Octane number (ON) is a figure of merit representing the resistance of gasoline to premature detonation when exposed to heat and pressure in the combustion chamber of an internal-combustion engine.[2]

It is measured against a prescribed binary mixture of isooctane (2,2,4-trimethylpentane, ON = 100) and n-heptane (ON = 0) under standard conditions. The relationship between the molecular structure and Octane number already attended certain basic set of rule like; ON increases with the number of tertiary and quaternary carbon atoms[3-7], it increases with the number of methyl groups[8-10], it decreases with the total number of carbon atom along the chain, ON increases with the shift of branch towards the centre of the longest carbon chain.

In the present study certain additional structural features were identified to add value in the structural studies of the hydrocarbon. It is a well-known fact that the physical and chemical properties of a compound are a function of its molecular structure. Quantitative structure property relationship (QSPR) is empirically defined the relationship between molecular structure and observed properties of the compounds and represented as -

Property = f (structural descriptors)

The structural descriptors tested in the present study are - Wiener Index (W)[11], Zero order Connectivity Index(χ0)[12], First Order Connectivity Index (χ1)[13], Second Order Connectivity Index (χ2)[14], Shultz Molecular Connectivity Index (SMTI)[15], Balaban J Index (J)[16], with an Indicator parameter Ic (possess value 1 if cyclic structure present in the molecule, otherwise possesses 0). The topological descriptors were considered in the present study due to the direct physical significance of molecular size and shape in the Octane number of the hydrocarbon. With these descriptors and experimental Octane number, the efforts have been made to develop a quantitative structure property relationship (QSPR) model expedient- to explore structural impact, to predict Octane number and finally to design a molecule(s) exhibiting desirable Octane number.

The multiple linear regression method has been employed to establish predictive Quantitative Structure Property Relationship (QSPR) model, which reflects the effect of structural features on which Octane number of the hydrocarbon relies. Also, this model is useful to

predict the Octane number of the hydrocarbons that do not belong to a given set of compounds. The work has a potential to provide a significant input in the area of structural modelling of the hydrocarbons with the desirable Octane number.

II. Computational Methodology

Experimental data of Octane number for the data set of 66 compounds were taken from literature[17]. The dataset of 66 compounds has been classified as training set of 41 compounds and test set of 25 compounds. The molecular structures were drawn and 3D optimized using ACD Chemsketch software and the structural descriptors for each molecule have been calculated using E-dragon (java-based program). The calculated descriptors viz., Wiener Index (W), Zero order connectivity Index ($\chi 0$), First Order connectivity Index ($\chi 1$), Second Order Connectivity Index ($\chi 2$), Schultz Molecular Topological Index (SMTI), Balaban J Index (J) and Indicator parameter(I_c) to indicate presence and absence of cyclic group in a compound were listed in Table 1 for training as well as for the test set of Hydrocarbons.

The selection of structural descriptor has been done by multiple linear regression method using SPSS software. Octane number has been classified as a dependent parameter and rest all descriptors were classified as independent parameters in variable selection step of regression analysis in SPSS. Only training set was used to develop QSPR model, whereas test set is used for the external validation of the QSPR model.

The correlations between all the descriptors and Octane number is shown in Table 2. Step up Multiple linear regression (MLR) method has been adopted using SPSS software for the selection of the descriptors which regulates the Octane number of the hydrocarbons. MLR analysis subsequently leads to the mathematical model which shows quantitative relationship between the selected descriptors and Octane number. In the final step the Octane numbers of the hydrocarbons in a training and a test set has been predicted using QSPR model obtained in the MLR analysis.

Validation of the model: The validation of QSPR model (Eq (1)) has been performed on the basis of two strategies: (i) Internally validated by the statistical parameters shown below Eq (1) & (2) Property prediction of test set compounds: In general, R^2 of the test set greater than 0.6 represents good prognostic ability of the model.[18]

III Results

The mathematical model obtained from MLR analysis using SPSS has been represented below as Eq (1).

$$ON = 0.125\,SMTI\ (\pm 0.125) + 49.9\,j\ (\pm 11) + 70.7\,\chi^0(\pm 21.1) - 143.8\,\chi^1\ (\pm 27.99) - 31.89\,\chi^2\ (\pm 12.5) + 109.533\,I_c\ (\pm 15.6) + 35.02 \quad (1)$$

N = 41, R = 0.946, R^2 = 0.894, Adjusted R^2 = 0.875, Standard Error of estimate = 9.4, Predicted residual sum of squares (PRESS) = 2893.177 & Sum of square of Y (SSY) = 27361.010, PRESS/SSY = 0.106

The internal validation of the mathematical model [Eq (1)] has been done by investigating the statistical parameters shown below the equation viz., regression coefficient (R), R^2, Standard error of estimation, F-Ratio. The developed models were further validated by the calculation of the following statistical parameters: predicted residual sum of squares (PRESS), total sum of squares deviation (SSY) and cross-validated correlation coefficient (r^2 adj).

PRESS[19,20] is an important cross-validation parameter as it is a good approximation of the real predictive error in the model. Its value being less than SSY points out that the model predicts better than chance and can be considered statically significant. The smaller PRESS value means the better of the model predictability. Also, for reasonable QSPR model, the PREES/SSY ratio should be lower than 0.4[21]. The data presented below the Eq (1) indicate that for the developed model this ratio is 0.106.

The predicted Octane number for the training set of 40 compounds is given in Table – 3 along with the Experimental Octane number, unstandardized residues and standardized residues. For the graphical visualization of outlier, the Williams plot has been shown as Figure 1.

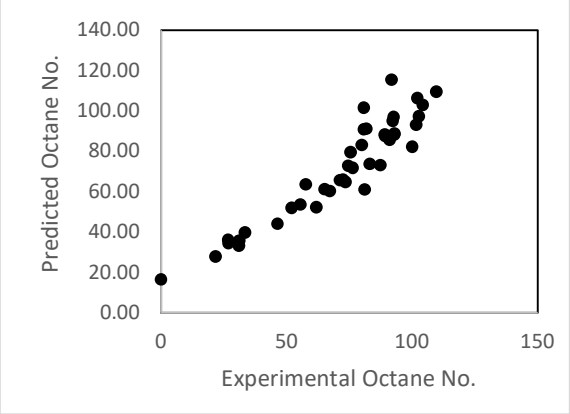

Figure 1. Correlation between experimental and predicted Octane no of training set.

The difference between the experimental ON and predicted ON is depicted as unstandardized residuals in Table 3 whereas the residual divided by an estimate of its standard deviation is depicted as standardized residuals [22] in Table 3. Standardized residuals quantify how large the residuals are in standard deviation units, and therefore can be easily used to identify outliers. It is an assumption in the statistics that the data showing standardized residue within the ±3 values will not be considered as outlier or a misfit data in the set.[23]

It is worthy to show the standardized residue in the present study, since some of the predicted Octane number showing deviation from experimental Octane number by approximately 10 units or more. The reason for high difference is the usage of real data (data without normalization) of Octane number in the QSPR analysis, this subsequently leads to the high values of regression coefficient of independent parameters e.g.,49.9 for J, 70.7 for χ0, 143.8 for χ1, 31.89 for χ2 & 109.533 for I_c this will bring an abrupt change in the predicted Octane number on the unit increment in the independent parameter. In order to cope up with this misleading information due to unstandardized residue, the fitness of the data has been ensured on the basis of Standardized residual values.

Discussion & Interpretation of QSPR Model: Quantitative Structure Property Relationship provides useful insight about the dependence of properties of molecules on its structure. The structural aspects that largely affects Octane number were selected by MLR method, but the interpretation part of these descriptors actually reveal their importance. In the present study SMTI plays a positive role in regulating Octane number of hydrocarbons. Let Γ be a molecular graph on N vertices. The "molecular topological index" (MTI) of the graph Γ introduced by Schultz1 in 1989[15] is defined in the following way:

$$MTI = MTI(\Gamma) = \sum_{i=1}^{N}[v(A+D)]_I \qquad (2)$$

Eq (2) comprises of summation of adjacency matrix(A) and distance matrix(D) of the compound, which in turn is a multiple of valencies(v) of the vertices in the graph. The positive coefficient of SMTI indicates that the higher SMTI is required in the compound to increase its Octane number, but the magnitude of coefficient is very low, which leads to the inference that out of the three component participating in the SMTI descriptor (i.e., Adjacency, distance vector and valencies of vertices), optimization of either one or two components is required.

Another descriptor included in Eq (1) is Balaban J Index (J) effectively discriminate cyclic and acyclic structures and branching in the structure. Positive coefficient of J with high magnitude indicate branched cyclic structure exhibits highest Octane no then unbranched cyclic structure which in turn possess higher octane no. than their acyclic or linear hydrocarbon analogues. The descending order of Octane number of hydrocarbons with reference to J index is expected as –

Branched cyclic analogue > cyclic Analogue > Branched acyclic > n-hydrocarbon analogue.

By the virtue of J this has been observed that adjacency and valencies of the atoms in hydrocarbons is required to be higher as compare to distances between the atoms. Inclusion of J with SMTI optimized three components of SMTI and also justified lower coefficient of SMTI.

Zero order Randic` connectivity index, the values of χ0 index increase with the increase in length and branches of hydrocarbon chains. This descriptor is vertex-based descriptor and applicable to molecule all along its structure. Represented as Eq (3)

$$\chi^0 = \sum_{i=1}^{i=n} \frac{1}{\sqrt{\delta_i}} \qquad (3)$$

Where,

χ0 = zero order connectivity index,

n = total no of vertices in the molecular graph

δ = Valance of Carbon atom in hydrogen supressed molecular graph.

χ0 is a vertex weighted connectivity index

Zero order connectivity index gives an information about the chain length and branching in 2-dimension, positive coefficient of χ0 shows higher value of this descriptor increases the Octane number.

First order Randic` connectivity index χ1 is an edge weighted descriptor expressed as

$$\chi^1 = \sum_{i=1}^{no.\ of\ edges\ or\ bonds} \frac{1}{\sqrt{\delta_i \times \delta_j}} \qquad (4)$$

χ1 contain information about the molecular volume and molecular surface area, higher value of χ1 indicate less molar volume and low surface area of the molecule and subsequently reduces Octane number. χ1 in present study is significantly reflects all specifications of branching present in a hydrocarbon.

The length of the branch should not extend beyond certain limit. For this reason, most of the fuel were methyl substituents and not ethyl, propyl, butyl and so on. Below are the two examples Figure 2 and Figure 3 demonstrating inverse relationship of χ1 with molar volume and Octane No.

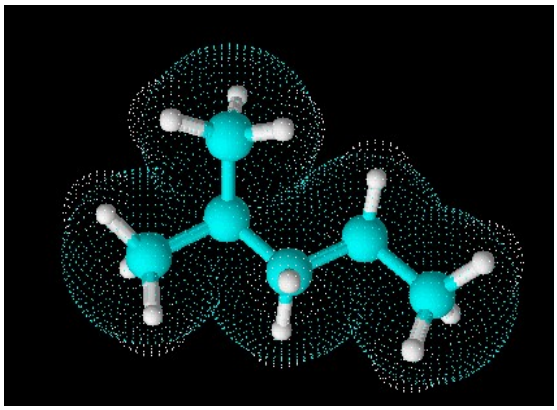

Figure 2. 2 - methyl pentane (χ1 = 2.77, Molar Volume = 127.9 cm3, O.No = 46.4)

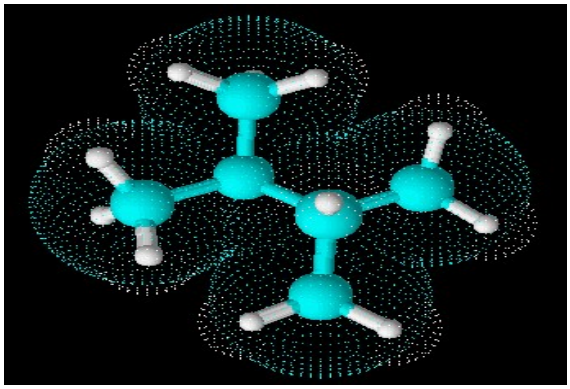

Figure 3. 2,3, di methyl butane (χ1 = 2.643, Molar volume = 128.3, Octane no = 104.3)

χ1 descriptor differentiate chain isomers in terms of Octane number, a chain isomer possessing lower χ1 value possess high volume and surface area which subsequently leads to a high-Octane rating.

Second order connectivity index χ2 has been represented as –

$$\chi^1 = \sum_{i=1}^{no.\ of\ edges\ or\ bonds-1} \frac{1}{\sqrt{\delta_i \times \delta_j \times \delta_k}} \quad (5)$$

χ2 is derived from fragment of two bond length and hence provides information about position and type of branching and indicate structural flexibility of the molecule. It is more suitable to compare positional isomerism in hydrocarbon. As per the QSPR model obtain and shown in Eq (1) χ2 is with negative coefficient, signifying inverse relationship between χ2 and Octane rating, molecule with higher χ2 will shows lower octane rating, e.g., 2- methyl pentane having χ2 = 2.183 and octane rating 73.4 and 3-methyl pentane with χ2 = 1.992 showing octane rating = 74.5. This support the hypothesis that shifting of branch to the centre of the molecule favours Octane rating and presence of branch on right or left extremes of molecule reduces Octane rating.

Indicator parameter I_c is used in present study to highlight the significance of cyclic structure in the molecule, the positive coefficient of I_c supports the presence of cyclic structure in a molecule with reference to octane number.

The predicted Octane number of training set of 40 compounds using Eq (1) has been shown in Table 3.

Validation Set: External cross validation of the model has also been done by applying the QSPR model obtained as Eq (1) on the test set of 25 compounds. All the independent variables of 25 compounds viz., SMTI, J, χ0, χ1, χ2 and I_c is shown in Table 1 & their corresponding Predicted and experimental Octane number is shown in Table 4. The linear correlation between Experimental and predicted Octane numbers is graphically represented in Figure 4 with all statistical parameters.

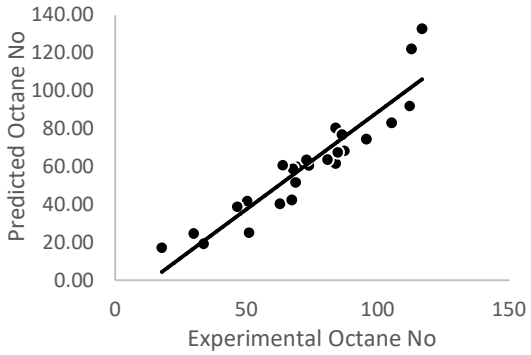

Figure 4. The predictive ability of the model [Eq(1)]

Eq (1) has also been validated by the test set which is characterised by the statistical summary of the Predicted and Experimental Octane number with R = 0.931, R^2 = 0.866, Adjusted R^2 = 0.86, Standard error of estimate = 0.96

The statistical parameter obtained for the test set of 25 compound further validating the predictive ability of the QSPR model.

IV CONCLUSIONS

QSPR model obtained in the present study reveals the importance of structural features on the Octane number of the hydrocarbons. The presence of SMTI in a model, with a positive coefficient leads to the inference, that the hydrocarbon with optimum chain length possessing Carbon in it's at-most valency will exhibit high Octane Number.

Also, the positive coefficient of Balaban J Index reflects increase in Octane number with the presence of branched cyclic structures. Positive coefficient of χ0 indicate that with increase in chain length there will be an increase in Octane number. Also, it supports the effect of branching on the Octane number of hydrocarbons. Negative coefficient of χ1 shows with increase in χ1 there will be a decrease in the Octane number, since the χ1 is a function of molar volume and molecular surface area, and it has also been found that lower the surface area or molecular volume greater will be the χ1. Therefore, negative coefficient of χ1 indicate lower values of χ1 increases Octane number, χ1 descriptor differentiate chain isomers in terms of Octane number, a chain isomer possessing lower χ1 value possess high volume and surface area which subsequently leads to a high-Octane rating. As per the QSPR model obtain as Eq (1) χ2 possess negative coefficient, signifying inverse relationship between χ2 and Octane rating, it supports in differentiating positional isomers of hydrocarbon with reference to Octane number and it has been found that isomer with higher positioning of branch shows higher Octane number.

Table 1. Training set & Test set of hydrocarbons with their structural descriptors & Indicator parameter.

Training set of Hydrocarbons								
S.No	**Compond Name**	**SMTI**	**W**	**J**	**χ0**	**χ1**	**χ2**	**Indicator**
1	2,2,3-Trimethylpentane	230	63	3.623	6.784	3.481	3.675	0
2	2,3-Dimethylbutane	108	29	2.993	5.155	2.643	2.488	0
3	2,3,4-Trimethylpentane	236	65	3.464	6.732	3.553	3.347	0
4	Methylpropane	36	9	2.324	3.577	1.732	1.732	0
5	Cyclopentane	80	15	2.083	3.536	2.5	1.768	1
6	2,2,4 TrimethylPentane	342	94	3.467	7.492	3.955	4.278	0
7	Methylbutane	68	18	2.54	4.284	2.27	1.802	0
8	2,2-Dimethylpentane	170	46	3.154	5.914	3.061	3.311	0
9	1,1,2,4-Tetramethylcyclopentane	334	79	2.583	7.069	4.022	4.435	1
10	1,1-Dimethylcyclopentane	180	39	2.4	5.328	3.207	3.371	1
11	2,2-Dimethylbutane	106	28	3.168	5.207	2.561	2.914	0
12	2,3-Dimethylpentane	168	46	3.144	5.862	3.181	2.63	0
13	Methylcyclopentane	126	26	2.184	4.406	2.894	2.39	1
14	1,2,3-Trimethylcyclopentane	252	58	2.436	6.146	3.715	3.391	1
15	3-Ethyl,2-Methylpentane	242	67	3.355	6.569	3.719	2.821	0
16	2,4-Dimethylpentane	176	48	2.953	5.862	3.126	3.023	0
17	1,1,3-Trimethylcyclopentane	254	58	2.435	6.199	3.601	4.012	1
18	isopropylcyclopentane	270	62	2.242	5.983	3.805	3.289	1
19	3,3-Dimethylpentane	162	44	3.36	5.914	3.121	2.871	0
20	3-Ethyl,3-Methylpentane	232	64	3.583	6.621	3.682	2.871	0
21	1,3-Dimethylcyclopentane	186	41	2.257	5.276	3.288	3.023	1
22	3,4-Dimethylhexane	246	68	3.292	6.569	3.719	2.771	0
23	3,3-Dimethylhexane	244	67	3.373	6.621	3.621	3.268	0
24	3-Methylpentane	114	31	2.754	4.992	2.808	1.922	0
25	2-Methylpentane	118	32	2.627	4.992	2.77	2.183	0
26	2,2-Dimethylhexane	260	71	3.112	6.621	3.561	3.664	0
27	2,3-Dimethylhexane	254	70	3.171	6.569	3.681	3.01	0
28	ethylcyclopentane	194	43	2.141	5.113	3.432	2.559	1
29	2,4-Dimethylhexane	258	71	3.099	6.569	3.664	3.143	0
30	n-pentane	74	20	2.191	4.121	2.414	1.354	0
31	2-ethyl,1-methylcyclopentane	264	61	2.303	5.983	3.843	3.077	1
32	2,5-Dimethylhexane	270	64	2.928	6.569	3.626	3.365	0
33	3-Methylhexane	182	50	2.832	5.699	3.308	2.302	0
34	2-Methylhexane	190	52	2.678	5.699	3.27	2.536	0
35	Isobutylcyclopentane	394	93	2.113	6.69	4.288	3.78	1
36	n-proplycyclopentane	290	67	2.058	5.81	3.932	2.939	1
37	n-Hexane	128	35	2.339	4.828	2.914	1.707	0
38	3-Methylheptane	276	76	2.862	6.406	3.808	2.656	0
39	4-Methylheptane	272	75	2.92	6.406	3.808	2.683	0
40	2-Methylheptane	288	79	2.716	6.406	3.77	2.89	0
41	n-Heptane	204	56	2.447	5.536	3.414	2.061	0
Test Set of Hydrocarbons								
42	2,2-Dimethylheptane	380	104	3.073	7.328	4.061	4.018	0
43	Diethylpentane	316	92	3.825	7.328	4.243	2.914	0
44	2,2-Dimethyl-3-ethylpentane	318	88	3.793	7.492	4.019	3.879	0
45	2,4-Dimethyl-3-ethylpentane	324	90	3.678	7.439	4.091	3.56	0
46	2,2,3,3-Tetramethylpentane	298	82	4.145	7.707	3.811	4.487	0
47	3,3,4-Trimethylheptane	444	123	3.778	8.199	4.542	4.032	0
48	2,2,3,3-Tetramethylhexane	416	115	4.282	8.414	4.371	4.475	0
49	Cyclohexane	132	27	2	4.243	3	2.121	1
50	Methylcyclohexane	193	42	2.123	5.113	3.394	2.743	1
51	Ethylcyclohexane	280	64	2.125	5.82	3.932	2.912	1
52	1,1-Dimethylcyclohexane	262	59	2.328	6.036	3.707	3.725	1
53	1,2-Dimethylcyclohexane	264	60	2.279	5.983	3.805	3.239	1
54	1,3-Dimethylcyclohexane	268	61	2.231	5.983	3.788	3.377	1
55	1,4-Dimethylcyclohexane	272	62	2.192	5.983	3.788	3.365	1
56	n-Propylcyclohexane	399	94	2.078	6.527	4.432	3.293	1
57	Isopropylcyclohexane	375	88	2.228	6.69	4.305	3.642	1

58	1-Methyl-1-ethylcyclohexane	359	84	2.366	6.743	4.268	3.725	1
59	1,1,2-Trimethylcyclohexane	343	80	2.491	6.906	4.128	4.135	1
60	1,2,3-Trimethylcyclohexane	349	82	2.413	6.853	4.215	3.745	1
61	1,2,4-Trimethylcyclohexane	357	84	2.346	6.853	4.198	3.873	1
62	1,3,5-Trimethylcyclohexane	357	84	2.341	6.853	4.182	4.023	1
63	Isobutylcyclohexane	526	126	2.131	7.397	4.788	4.134	1
64	sec-Butylcyclohexane	504	121	2.24	7.397	4.843	3.784	1
65	1-Isopropyl-4-methylcyclohexane	498	120	2.26	7.56	4.698	4.264	1
66	1-Methyl-2-n-propylcyclohexane	502	121	2.251	7.397	4.843	3.81	1

Table 2: Correlation matrix, representing intercorrelations between the descriptors.

		ON	SMTI	J	x0	x1	x2	Indicator
Pearson Correlation	ON	1.000	-.391	.275	-.175	-.402	.163	.090
	SMTI	-.391	1.000	.123	.875	.972	.764	.269
	J	.275	.123	1.000	.567	.155	.270	-.696
	x0	-.175	.872	.567	1.000	.872	.789	-.101
	x1	-.402	.972	.155	.872	1.000	.709	.273
	x2	.163	.764	.270	.789	.709	1.000	.336
	Indicator	.090	.269	-.696	-.101	.273	.336	1.000

Table 3: Experimental, Predicted Octane number, unstandardized residues & Standardized residue

S.No	Compound Name	Experimental Octane No.	Predicted Octane No	Unstandardized residues	Standardized residue
1	2,2,3-Trimethylpentane	109.6	109.50	0.10	0.01
2	2,3-Dimethylbutane	104.3	103.02	1.28	0.13
3	2,3,4-Trimethylpentane	102.7	97.41	5.29	0.54
4	Methylpropane	102.1	106.42	-4.32	-0.44
5	Cyclopentane	101.6	93.19	8.41	0.85
6	2,2,4 trimethylPentane	100	82.38	17.62	1.78
7	Methylbutane	93	88.91	4.09	0.41
8	2,2-Dimethylpentane	92.8	88.32	4.48	0.45
9	1,1,2,4-Tetramethylcyclopentane	92.6	97.02	-4.42	-0.45
10	1,1-Dimethylcyclopentane	92.3	95.16	-2.86	-0.29
11	2,2-Dimethylbutane	91.8	115.56	-23.76	-2.40
12	2,3-Dimethylpentane	91.1	85.69	5.41	0.55
13	Methylcyclopentane	89.3	87.69	1.61	0.16
14	1,2,3-Trimethylcyclopentane	89.2	88.42	0.78	0.08
15	3-Ethyl,2-Methylpentane	87.3	73.20	14.10	1.43
16	2,4-Dimethylpentane	83.1	73.80	9.30	0.94
17	1,1,3-Trimethylcyclopentane	81.7	91.30	-9.60	-0.97
18	isopropylcyclopentane	81.1	61.21	19.89	2.01
19	3,3-Dimethylpentane	80.8	101.61	-20.81	-2.11
20	3-Ethyl,3-Methylpentane	80.8	90.94	-10.14	-1.03
21	1,3-Dimethylcyclopentane	79.9	83.21	-3.31	-0.34
22	3,4-Dimethylhexane	76.3	71.92	4.38	0.44
23	3,3-Dimethylhexane	75.5	79.58	-4.08	-0.41
24	3-Methylpentane	74.5	72.86	1.64	0.17
25	2-Methylpentane	73.4	64.90	8.50	0.86
26	2,2-Dimethylhexane	72.5	66.09	6.41	0.65
27	2,3-Dimethylhexane	71.3	65.72	5.58	0.56
28	ethylcyclopentane	67.2	60.49	6.71	0.68
29	2,4-Dimethylhexane	65.2	61.32	3.88	0.39
30	n-Pentane	61.8	52.41	9.39	0.95
31	2-Ethyl,1-methylcyclopentane	57.6	63.74	-6.14	-0.62
32	2,5-Dimethylhexane	55.5	53.65	1.85	0.19
33	3-Methylhexane	52	52.01	-0.01	0.00
34	2-Methylhexane	46.4	44.13	2.27	0.23
34	Isobutylcyclopentane	33.4	39.80	-6.40	-0.65
36	n-Proplycyclopentane	31.2	35.87	-4.67	-0.47
37	n-Hexane	31	33.31	-2.31	-0.23

38	3-Methylheptane	26.8	34.58	-7.78	-0.79
39	4-Methylheptane	26.7	36.23	-9.53	-0.96
40	2-Methylheptane	21.7	27.96	-6.26	-0.63
41	n-Heptane	0	16.56	-16.56	-1.68

Table 4: Experimental and Predicted Octane number of the Test set using Eq(1).

S.NO	Compounds	Octane No	Predicted O.No
41	2,2-Dimethylheptane	50.30	41.85
42	Diethylpentane	84.00	80.41
43	2,2-Dimethyl-3-ethylpentane	112.10	92.09
44	2,4-Dimethyl-3-ethylpentane	105.30	83.18
45	2,2,3,3-Tetramethylpentane	116.80	132.88
46	3,3,4-Trimethylheptane	86.40	76.99
47	2,2,3,3-Tetramethylhexane	112.80	122.30
48	Cyclohexane	84.00	61.79
49	Methylcyclohexane	73.80	60.57
50	Ethylcyclohexane	46.50	38.78
51	1,1-Dimethylcyclohexane	87.30	68.36
52	1,2-Dimethylcyclohexane	80.90	63.82
53	1,3-Dimethylcyclohexane	69.30	59.97
54	1,4-Dimethylcyclohexane	67.70	58.91
55	n-Propylcyclohexane	17.80	17.24
56	Isopropylcyclohexane	62.80	40.39
57	1-Methyl-1-ethylcyclohexane	68.70	51.69
58	1,1,2-Trimethylcyclohexane	95.70	74.51
59	1,2,3-Trimethylcyclohexane	84.80	67.55
60	1,2,4-Trimethylcyclohexane	72.90	63.57
61	1,3,5-Trimethylcyclohexane	63.80	60.84
62	Isobutylcyclohexane	33.70	19.26
63	sec-Butylcyclohexane	51.00	25.20
64	1-Isopropyl-4-methylcyclohexane	67.30	42.52
65	1-Methyl-2-n-propylcyclohexane	29.90	24.67

CONFLICT OF INTEREST

"The authors declare no conflict of interest."

AUTHOR CONTRIBUTIONS

Dr. Sanjay Kumar provided the motivation and guidance and acted as a reviewer; Dr. Mamta Thakur conceptualized the work developed the methodology, and investigated the results; Naman Shah & Sarthak Jain did the calculation, chemical structure modelling and all the software operations. All authors had approved the final version of the paper

REFERENCES

1. W. Leppard the chemical origin of fuel octane sensitivity *SAE Transactions* Journal of fuel and lubricants Vol. 99, Section 4, 1990, 862-876
2. Vikram Mittal and John B. Heywood The Shift in Relevance of Fuel RON and MON to Knock Onset in Modern SI Engines SAE International Journal of Engines Vol. 2, No. 2 ,2010, 1-10
3. https://www.cpp.edu › ~psbeauchamp › pdf › 314_supp_6_isom_form
4. Tal'roze, V. L.; Lyubimova, A. L. *Dokl. Akad. Nauk SSSR* **1952**, *86*, 909; *Chem Abstr.* **1953**, *47*, 2590.
5. Wexler, S.; Jesse, N. J. Am. Chem. Soc. 1962, 84, 3425.
6. East, A. L. L.; Liu, Z. F.; McCague, C.; Cheng, K.; Tse, J. S. J. Phys. Chem. A 1998, 102, 10903.
7. Hunter, K. C.; East A. L. L. J. Phys. Chem. A 2002, 106, 1346.
8. J. Warnatz, Combustion: physical and chemical fundamentals, modeling and simulation, experiments, pollutants formation, J. Warnatz, U. Maas, R. W. Dibble, 2nd Ed., Springer, Berlin etc 1999.
9. W. Lovell, Ind. Eng. Chem. 1948, 40, 2388–2438.
10. C. Morley, Combust. Sci. Technol. 1987, 55, 115–123.
11. Wiener, H. Journal of the American Chemical Society, 1947, 1 (69): 17–20.
12. Randić M.On characterization of molecular branching J. Am. Chem. Soc., 97 (1975), pp. 6609-6615.
13. Randić M.On history of the Randić index and emerging hostility toward chemical graph theory MATCH Commun. Math. Comput. Chem., 59 (2008), pp. 5-124.
14. Kier L.B., Hall L.H.Molecular Connectivity in Chemistry and Drug Research Academic Press, New York (1976).
15. H.P. Schultz, J.Chem.Inf.Comput.Sci. 1989, 29, 227-228.
16. A. T. Balaban, Pure Appl. Chem. 1983, 55, 199–206.
17. Jabir H. Al-Fahemi, Nahla A. Albis, and Elshafie A.M. Gad, Journal of Theoretical Chemistry Volume

2014, Article ID 520652, 6 pages http://dx.doi.org/10.1155/2014/520652.
18. TSAR reference guide.
19. Cronim MTD, Gregory BW & Shutz JW Chem. Res. Toxicol., 11, 1998, 902-12.
20. Box GEB, Hunter WG & Hunter JS Statistics for Experiments, Wiley, New York, 2000.
21. V. A. McNally, M. Rajabi, A. Gbaj, I. J.Stratford, P. N. Edwards, K. T. Douglas, R. A. Bryce, M. Jaffar, S. Freeman, J. Pharm. Pharmacol. 59, 2007, 537-538.
22. www.ibm.com
23. Richard J Rossi, Applied Biostatistics for the Health Sciences, Wiley, 2010.

Dr. Sanjay Kumar, Vice Chancellor Symbiosis University of Applied Sciences, Indore also serve as Vice chancellor of ITM University Raipur. He completed his doctorate degree from IIT Delhi in Microwave Communication & RADAR Engg. In addition he is also awarded by Ph.D degree in Human Resource, MS Counselling Psychotherapy, Satellite Communication and Application Course from ISRO, Space Application Centre, Ahmedabad, Intelligence gathering and handling Course from Intelligence School Pune, A number of courses on Russian Radar and Indian radars, Electronic warfare, Avionics etc, Intelligence Equipment state of art Communication Set up like, Raven, Schimitter etc. Dr. Kumar is having 34 year of Academic, Research and administrative experience. He published 07 Books in the areas like Microwave Engg. Wave and Antenna Engg. He is a recipient of many awards like Presidential Award (Vishisht Seva Medal) for world class research in Reverse Engineering (Electronics) and utmost dedication to the nation, Bihar Gaurav Samman by Hon. Governor of Bihar and many more. His research papers in the area of Electronics Engg. are published in reputed international journal.

Dr. Mamta Thakur completed her masters in Chemistry with specialization in Organic Chemistry in 1998, she was also awarded Ph.D degree in 2003 on “Modelling of Structure Activity: Study on Bio-Active Organic Systems”. She is also a recipient of most cited paper award 2003- 06 by Elsevier publication. Currently she is working as a Professor at Symbiosis University of Applied Sciences. In her entire career she has taught wide spectrum of subjects viz., Organic Chemistry, Inorganic Chemistry, Pharmaceutical Chemistry, Material Science & Engineering, Reliability Engineering, Retail Information Technology, Business Computing, Retail Market research, Excel Spreadsheet Environmental Studies. Apart from Science and technology, she also pursuing post-graduation in Management with specialization in Marketing and Operation & Production. Dr Mamta Thakurs Area of research is Computer aided Molecular Modelling, Cheminformatics and Drug design. Completed funded project in this area.

Naman Shah is a graduate student within the Engineering program at the Symbiosis University of Applied Sciences, Indore. He will graduate with a Mechatronics Engineering Degree in 2021.

Sarthak Jain is a graduate student within the Engineering program at the Symbiosis University of Applied Sciences, Indore. He will graduate with a Mechatronics Engineering Degree in 2021.