Graph Energy and Entropy: Insights into Molecular Complexity and Properties

S. Sripriya^{1,2,a}, A. Anuradha^{2,b}, R. Sundara Rajan^{1,c}, Sandi Klavžar^{3,4,5,d}

¹ Department of Mathematics, Hindustan Institute of Technology and Science, Chennai - 603 103, India

 a sripriyas@hindustanuniv.ac.in ^c vprsundar@gmail.com

² Department of Mathematics, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai - 603 203, India ^b anuradha@srmist.edu.in

³ Faculty of Mathematics and Physics, University of Ljubljana, Slovenia ⁴ Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia

⁵ Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

d sandi.klavzar@fmf.uni-lj.si

Abstract

Graph entropy is an emerging concept in graph theory that has not yet been extensively explored. Most research focuses on figuring out this parameter for different graph families. In this paper, the eigenvalue-based entropies, the first Zagreb entropy, the second Zagreb entropy, and the sum-connectivity entropy are introduced. The importance of graph energy and entropy in comprehending molecular activity in the context of alkanes is emphasized. By employing linear regression, the correlation between the adjacency energy, first and second Zagreb energy, sum-connectivity energy, and molecular properties of 72 alkane isomers starting from butane to nonane are explored. Multicollinearity is addressed using double regression, which shows that energy and entropy descriptors act as independent variables with variance inflation factors below 2.5, ensuring minimal redundancy. Our findings indicate a strong correlation between energy-entropy pairs and the physicochemical properties, reinforcing their predictive significance. Additionally, a 10-fold cross-validation approach is implemented to validate model robustness and predictive reliability. Entropy, when included as an independent variable in double regression, exhibits a moderate positive correlation with melting point and a strong positive correlation with respect to critical pressure and vapor pressure, which is not observed in linear regression. Further, sum-connectivity spectral radius and spectral spread exhibit high correlations with density and refractive index with correlation values of 0.92456 and 0.89765, respectively. The dataset, spectral parameters, and regression models presented in this study offer a structured

framework for property prediction in hydrocarbons, with potential applications in chemical informatics.

1 Introduction

The primary focus of organic chemistry is the study of molecules that contain carbon, including their compositions, structures, properties, reactions, and synthesis. In organic chemistry, alkanes [?], are a fundamental class of hydrocarbons composed entirely of carbon and hydrogen atoms. They are known as saturated hydrocarbons because each carbon atom in them is only connected by a single bond, allowing for the maximum number of hydrogen atoms to be bound to each carbon. The formula for alkanes is C_nH_{2n+2} , where n is the number of carbon atoms. Single bonds between the hydrogen and carbon atoms make alkanes nonpolar, unreactive, and comparatively stable molecules.

Alkanes exist in multiple structural configurations called isomers, in addition to their stability. This is pertinent when a carbon chain has four or more atoms. Despite having the same molecular formula, isomers differ in their structural configurations, giving them unique physical and chemical characteristics. Alkanes are important for understanding functionalized organic molecules and complicated hydrocarbons. As alkanes have well-defined structures, it is beneficial to learn about the fundamentals of molecular structure, bonding, and reactivity in organic chemistry. Chemical graph theory [?] denotes molecules as graphs with bonds as edges and atoms as vertices. This representation helps to predict molecular properties based on structure. Graph spectral parameters are one of the important tools in understanding the properties of molecular graphs. Graph entropies [?], based on the eigenvalues of these graphs, help to measure the complexity and stability of molecular structures.

In this paper, QSPR analysis on 72 isomers from butane to nonane is conducted to determine the correlation between molecular structure and various physiochemical properties. The main objective is to develop predictive models that could accurately estimate properties such as boiling point, melting point, critical temperature, critical pressure, molar volume, molar refraction, the heat of vaporization, density, refractive index, surface tension, vapor pressure, logP, complexity, flash point, and polarizability based on the molecular structure of the alkanes. Double regression analysis enables us to assess the impact of various structural factors while simultaneously mitigating the effects of multicollinearity, a common issue in regression models. This statistical method helps us to establish a relationship between the independent variables and a dependent variable that allows for prediction and an evaluation of how variations in input factors influence the outcome. The coefficient of determination R^2 helps us to assess how well the independent variables account for the variability in the dependent variable. The F-test is used to determine whether the overall regression model is significant or not. It helps to evaluate whether the independent variables explain a significant amount of variance in the dependent variable. A high F-value with a low p-value < 0.05 expresses that the model is statistically significant. Multicollinearity occurs when independent variables in a regression model exhibit strong correlations, making it difficult to interpret their individual effects. To quantify the extent of this

issue, the Variance Inflation Factor (VIF) is utilized. VIF measures how much the variance of a regression coefficient is increased due to collinearity among predictors. A VIF > 10 indicates a serious multicollinearity problem, whereas a VIF between 1 and 5 suggests minimal impact on the model's reliability and is considered acceptable, and a VIF equal to 1 indicates no multicollinearity.

2 Preliminaries

Given a chemical compound, the molecular graph is constructed in such a way that the carbon atoms are considered as vertices, bonds between the atoms are considered as edges, and hydrogen atoms are ignored. The hydrogen atoms are ignored due to the fact that they are univalent, and their main role is in saturating the carbon atoms.

Let A(G) be the adjacency matrix of a graph G of order r, and let ξ_i , $i \in [r] = \{1, 2, \ldots, r\}$ be the eigenvalues of A(G) such that $\xi_1 \geq \xi_2 \geq \cdots \geq \xi_r$. The arrangement of distinct eigenvalues, along with their algebraic multiplicities, is called the spectrum of G. The highest among all the eigenvalues is known as spectral radius [?]. The spectral gap is the difference between the largest and the second-largest eigenvalue of A(G) and is used as an indicator of graph connectivity. The spectral spread is the difference between the maximum and the minimum eigenvalue. It gives insight into the overall range of the eigenvalues of the graph's adjacency matrix is termed as energy [?] and is written as $\mathscr{E}[A(G)] = \sum_{i=1}^{r} |\xi_i|$. This value has been proved to be an appropriate indicator for the π -electron energy in organic compounds.

The first Zagreb matrix [?] of a graph G is a square matrix FZ(G) such that each entry in the matrix will be $\delta_i + \delta_j$, if the vertices ν_i and ν_j are adjacent to each other or zero otherwise. Here δ_i and δ_j are the degrees of the vertices ν_i and ν_j respectively. If $\alpha_i, i \in [r]$, are the eigenvalues of FZ(G), then the first Zagreb energy is given by $\mathscr{E}[FZ(G)] = \sum_{i=1}^r |\alpha_i|$.

The second Zagreb matrix [?] $SZ(G) = [d_{ij}]$ of a graph G is a square matrix of order r defined by

$$d_{ij} = \begin{cases} \delta_i \cdot \delta_j; & \nu_i \sim \nu_j, \\ 0; & \text{otherwise} \end{cases}$$

If $\mu_1, \mu_2, \ldots, \mu_r$ are the eigenvalues of SZ(G), then its corresponding energy is $\mathscr{E}[SZ(G)] = \sum_{i=1}^r |\mu_i|$.

The sum connectivity matrix [?] $SC(G) = [s_{ij}]$ of a graph G is defined as

$$s_{ij} = \begin{cases} \frac{1}{\sqrt{\delta_i + \delta_j}}; & (\nu_i, \nu_j) \in E(G), \\ 0; & \text{otherwise.} \end{cases}$$

The summation of the absolute values of the eigenvalues of SC(G) is the sumconnectivity energy $\mathscr{E}[SC(G)]$ of G. The eigenvalue-based modulus adjacency entropy of G [?] is defined as

$$\mathcal{I}[A(G)] = -\sum_{i=1}^{r} \frac{|\xi_i|}{\mathscr{E}[A(G)]} \log \frac{|\xi_i|}{\mathscr{E}[A(G)]}$$

In order to carry out the QSPR analysis on the physical properties of alkanes, in parallel to the adjacency entropy we now introduce the first Zagreb entropy, the second Zagreb entropy, and the sum-connectivity entropy as follows.

• The first Zagreb modulus entropy is

$$\mathcal{I}[FZ(G)] = -\sum_{i=1}^{r} \frac{|\alpha_i|}{\mathscr{E}[FZ(G)]} \log \frac{|\alpha_i|}{\mathscr{E}[FZ(G)]},$$

where $\alpha_i, i \in [r]$, are the eigenvalues of FZ(G).

• The second Zagreb modulus entropy is

$$\mathcal{I}[SZ(G)] = -\sum_{i=1}^{r} \frac{|\mu_i|}{\mathscr{E}[SZ(G)]} \log \frac{|\mu_i|}{\mathscr{E}[SZ(G)]}$$

where μ_i are the eigenvalues of SZ(G).

• The sum-connectivity modulus entropy is

$$\mathcal{I}[SC(G)] = -\sum_{i=1}^{r} \frac{|\gamma_i|}{\mathscr{E}[SC(G)]} \log \frac{|\gamma_i|}{\mathscr{E}[SC(G)]},$$

where γ_i are the eigenvalues of SC(G).

3 Literature Survey

The notion of graph energy was introduced by Gutman and Trinajstić [?], and a number of findings pertaining to variations in energy, spectral radius, and its bounds can be found in the literature. D H Rouvray and C Crafford [?] applied topological invariants derived from molecular graphs to study and predict physicochemical properties of organic molecules. They examined the relationship between boiling point, refractive index, surface tension and density with the topological indices and showed that these structural descriptors correlate strongly with experimental property values, supporting the idea that molecular topology influences macroscopic behavior. Verbruggen and Tolls [?] carried out a literature study on the physicochemical properties of aliphatic hydrocarbons containing ten or more carbon atoms. Raja and Anuradha [?] performed regression analysis with Sombor indices and its variants along with their associated entropies in analysing the physical attributes of family of alkanes with carbon range of four to nine. Mondal et al. [?] developed an algorithm model to make the computation of the energy and its Estrada index along with its integral representation of energy. Sripriya and Anuradha [?] compared first and second Zagreb energy for

certain families of graphs and provided a criterion leading to an infinite set of nonisomorphic Zagreb equi-energetic graphs for all r > 1 within-partite graphs. Prakasha et al. [?] have obtained interesting results on sum-connectivity energy for some families of graphs. Sripriya and Anuradha [?] determined sum-connectivity energy on graph operations and given spectra in terms of base graphs themselves. A survey describing the techniques involved in measuring graph entropy and demonstrating its broad applicability was conducted by Dehmer and Mowshowitz [?]. Dehmer et al. [? have created a broad structure for determining the graph entropy based on a local information graph and related functions derived from the graph topology. The entropy based on eigenvalues of the directed bipartite network has been defined by Sun and Zhao [?]. Altassan et al. [?] discussed the QSPR analysis and carried out statistical modeling (linear, logarithmic, and quadratic) of the physiochemical properties of anticancer drugs with the ISI-index (energy). Bhadre et al. [?] relate energies with the QSPR analysis of 67 alkanes. Andrew and Anuradha [?] found that the spectral gap has a great correlation with the density of nonane isomers. They also established a correlation between the spectral gap and refractive Index of the compound decane and all its isomers [?]. Kumar et al. [?] computed QSPR analysis of alkanes on certain degree-based indices. Sharon et al. [?] investigated the transmission properties of H-naphtalenic nanosheets using graph-theoretical approaches. Raj et al. explore the application of graph-theoretical concepts to analyze the molecular structures of key anti-HIV drugs [?]. Dong et al. [?] derived Wiener-entropy and the eccentricityentropy and by deriving its extremal behaviour concluded that the Wiener-entropy of graphs of a given order is more spread than the eccentricity-entropy. In addition, numerous researchers address a wide range of QSPR analysis results on different drugs, nanotubes, various operations, molecules, etc., employing various topological indices [? ????].

4 Results and Discussion

The molecular graph for the 72 alkane isomers from butane to nonane is constructed. The matrices, such as the adjacency matrix, the first Zagreb matrix, the second Zagreb matrix, and the sum-connectivity matrix, are determined and the eigenvalues and energies are computed using MATLAB. The eigenvalue-based entropies are calculated using the program.

Input n (number of eigenvalues) Initialize sum = 0, result = 0 For i = 1 to n: Input eigenvalue γ_i Add $|\gamma_i|$ to sum If sum = 0, print error and stop For i = 1 to n: If $\gamma_i = 0$, skip Compute $p_i = |\gamma_i|/sum$ Compute entropy = $p_i * log(p_i)$ Add to result

SI.No	Name	$\mathscr{E}[\mathbf{A}(\mathbf{G})]$	$\mathscr{E}[\mathbf{FZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SC}(\mathbf{G})]$	$\mathcal{I}[\mathbf{A}(\mathbf{G})]$	$\mathcal{I}[\mathbf{FZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SC}(\mathbf{G})]$
2 1	n-butane isobutane(2	$4.472 \\ 3.464$	$14.424 \\ 13.856$	$11.312 \\ 10.392$	$2.516 \\ 1.732$	$\frac{1.28266}{0.693147}$	$1.22343 \\ 0.693147$	$\frac{1.10955}{0.693147}$	$\frac{1.30509}{0.693147}$
4 3	metnyl propane) n-pentane isopentane (2	5.464 5.226	$18.806 \\ 20$	$\begin{array}{c} 16\\ 17.318\end{array}$	$2.98 \\ 2.72$	$1.34996 \\ 1.29776$	$1.3193 \\ 1.21603$	$\frac{1.25548}{1.07983}$	$\frac{1.36065}{1.32638}$
5	metnyl butane) neopentane (dimethyl-	4	20	16	1.788	0.693147	0.693147	0.693147	0.693147
© ~ ∞ o	propane) n-hexane 2-methyl pentane 3-methyl pentane 2 3	$\begin{array}{c} 6.988\\ 6.156\\ 6.9 \end{array}$	$\begin{array}{c} 24.56 \\ 24.182 \\ 25.9 \\ 25.6 \\ 612 \end{array}$	21.452 21.808 23.696 24.74	3.764 3.15 3.652 9.044	$\begin{array}{c} 1.66482 \\ 1.35821 \\ 1.66149 \\ 1 & 32066 \end{array}$	$\begin{array}{c} 1.6186\\ 1.32369\\ 1.54711\\ 1.57919\end{array}$	$\begin{array}{c} 1.53402 \\ 1.25839 \\ 1.34971 \\ 1.00118 \end{array}$	$\begin{array}{c} 1.6813 \\ 1.36889 \\ 1.69779 \\ 1.34737 \end{array}$
10	butane (2,2 reohexane (2,2	5.818	26.226	23.976	2.822	1.29263	1.17273	1.03569	1.33383
$11 \\ 13 \\ 13 \\ 14 \\ 112 \\ 11$	dimethyl butane) n-heptane 2 methyl hexane 3 methyl hexane 2 2 dimethyl hexane	8.054 7.728 7.88 6.73	29.072 30.096 30.214 30.318	26.222 27.364 28.388 38.408	$\begin{array}{c} 4.282 \\ 3.96 \\ 4.108 \\ 2.238 \end{array}$	$\begin{array}{c} 1.7336\\ 1.67698\\ 1.70914\\ 1.35456\end{array}$	$\begin{array}{c} 1.7044 \\ 1.62558 \\ 1.63202 \\ 1.90141 \end{array}$	$\begin{array}{c} 1.65178 \\ 1.54131 \\ 1.50102 \\ 1.2020 \end{array}$	$\begin{array}{c} 1.74434 \\ 1.69494 \\ 1.73312 \\ 1.73512 \end{array}$
15 15	2,2 dimethyl pen- tane 3,3 dimethyl pen-	7.596	32.338	31.324	3.834	1.67269	1.50661	1.28458	1.72167
$16 \\ 17$	tane 2,3 dimethyl pen- tane 2.4 dimethyl pen-	7.664 6.828	31.476 29.424	30.858 27.46	3.872 3.31	1.6749 1.37149	1.57352 1.35938	1.35885 1.31147	1.70803 1.37565
$\frac{18}{19}$	tane 3 ethyl pentane 2,2,3 trimethyl	$\frac{8}{6.602}$	30.33 31.874	$29.166 \\ 32.79$	$4.24 \\ 3.056$	1.73287 1.33492	1.6386 1.26728	1.47075 1.06987	$\frac{1.75966}{1.35438}$
$\begin{array}{c} 20\\ 21 \end{array}$	butane <i>n</i> -octane 2 methyl heptane	$9.516 \\ 8.764$	$34.712\ 34.48$	$31.62\ 32.056$	$5.028 \\ 4.462$	$\begin{array}{c} 1.94373 \\ 1.74188 \end{array}$	$\frac{1.90781}{1.71134}$	$\frac{1.84414}{1.66169}$	$1.95661 \\ 1.75229$

 ${\bf Table \ 1} \ \ {\rm The \ four \ energy \ variants \ of \ alkanes \ with \ respective \ entropies$

SI.No	Name	$\mathscr{E}[\mathbf{A}(\mathbf{G})]$	$\mathscr{E}[\mathbf{F}\mathbf{Z}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SC}(\mathbf{G})]$	$\mathcal{I}[\mathbf{A}(\mathbf{G})]$	$\mathcal{I}[FZ(G)]$	$\mathcal{I}[\mathbf{SZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SC}(\mathbf{G})]$
$22 \\ 24 \\ 22 \\ 24 \\ 22 \\ 22 \\ 22 \\ 22 \\$	3 methyl heptane 4 methyl heptane 2.3 dimethyl hex-	$\begin{array}{c} 9.408 \\ 8.828 \\ 8.646 \end{array}$	36.018 34.438 35.808	$33.824 \\ 32.966 \\ 35.594$	$\begin{array}{c} 4.896 \\ 4.55 \\ 4.33 \end{array}$	$\begin{array}{c} 1.93903 \\ 1.75289 \\ 1.72864 \end{array}$	$\begin{array}{c} 1.85733\\ 1.70672\\ 1.66467\end{array}$	$\begin{array}{c} 1.72032 \\ 1.62522 \\ 1.5112 \end{array}$	$\begin{array}{c} 1.966 \\ 1.76577 \\ 1.74659 \end{array}$
25	ane 2,4 dimethyl hex-	8.564	35.532	34.114	4.278	1.71371	1.65049	1.54617	1.73679
26	ane 2,5 dimethyl hex-	8.472	35.668	33.316	4.156	1.69592	1.66057	1.58631	1.71043
27	ane 3,4 dimethyl hex-	9.332	37.358	37.164	4.8	1.93662	1.81033	1.57346	1.97649
28	ane 2,2 dimethyl hex-	8.312	36.314	33.986	4.06	1.67209	1.58614	1.49877	1.69961
29	ane 3,3 dimethyl hex-	8.52	36.484	35.912	4.262	1.71285	1.59384	1.43847	1.74581
$30 \\ 31$	ane 3 ethyl hexane 2,2,3 trimethyl	$9.446 \\ 8.262$	$35.892\ 37.722$	$34.516\ 38.796$	$4.97 \\ 3.982$	$\frac{1.94165}{1.67056}$	$\frac{1.84711}{1.54984}$	$\frac{1.68504}{1.32713}$	$\frac{1.96828}{1.71067}$
32	$\begin{array}{c} \text{pentane} \\ 2,2,4 \\ \text{trimethyl} \end{array}$	7.384	35.516	33.962	3.396	1.37276	1.35316	1.30861	1.37803
33	pentane 2,3,3 trimethyl	8.376	37.96	39.74	4.064	1.69358	1.56409	1.30871	1.73154
34	pentane 2,3,4 trimehyl	8.424	37.024	37.774	4.088	1.69475	1.61826	1.40897	1.72002
35	pentane 3 ethyl 2 methyl	8.72	35.738	36.08	4.44	1.74136	1.65883	1.48616	1.76524
36	pentane 3 ethyl 3 methyl	9.292	38.23	38.2	4.796	1.93705	1.75066	1.47515	1.98675
37	pentane 2,2,3,3 tetra	7.212	38.156	42.332	3.176	1.34733	1.2956	1.06411	1.36124
38	пеция рисале nonane	10.628	39.316	36.428	5.574	2.00696	1.98125	1.93752	2.01637

 ${\bf Table \ 1} \ \ {\rm The \ four \ energy \ variants \ of \ alkanes \ with \ respective \ entropies \ (continued)$

Sl.No	Name	$\mathscr{E}[\mathbf{A}(\mathbf{G})]$	$\mathscr{E}[\mathbf{FZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SC}(\mathbf{G})]$	$\mathcal{I}[\mathbf{A}(\mathbf{G})]$	$\mathcal{I}[\mathbf{FZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SC}(\mathbf{G})]$
$ \begin{array}{c} 39 \\ 41 \\ 41 \end{array} $	2 methyl octane 3 methyl octane 4 methyl octane	$10.252 \\ 10.472 \\ 10.384 \\ 0.384$	$\begin{array}{c} 40.234 \\ 40.494 \\ 40.324 \\ 20.324 \end{array}$	37.526 38.598 38.482 38.482	5.22 5.416 5.354	$\begin{array}{c} 1.95333\\ 1.99205\\ 1.97619\\ \end{array}$	$\begin{array}{c} 1.91611\\ 1.92986\\ 1.92031\\ 1.92031\\ \end{array}$	$\begin{array}{c} 1.85535\\ 1.82868\\ 1.82668\\ 1.82614\\ \end{array}$	$\begin{array}{c} 1.96663 \\ 2.01281 \\ 1.99355 \\ \end{array}$
$\begin{array}{c} 42\\ 43\end{array}$	2,2 dimethyl hep- tane 2,3 dimethyl hep-	9.336 10.176	40.646 41.602	38.658 41.016	4.556 5.118	1.73813 1.95078	1.67866 1.87645	1.62566 1.71777	1.75513 1.97462
44	tane 2,4 dimethyl hep-	9.508	39.722	38.66	4.716	1.76014	1.72705	1.6646	1.77029
45	tane 3,4 dimethyl hep-	10.312	41.682	41.836	5.262	1.97412	1.88043	1.697	2.00264
46	tane 2,5 dimethyl hep-	10.152	41.574	39.76	5.056	1.94911	1.87788	1.7597	1.97436
47	2,6 dimethyl hep-	9.464	39.84	37.848	4.64	1.75349	1.73495	1.69376	1.76059
48	3,3 dimethyl hep-	10.096	42.434	41.42	5.07	1.94783	1.82405	1.66108	1.98358
49	4,4 dimethyl hep-	9.43	40.602	40.432	4.682	1.75323	1.67285	1.56879	1.77211
50	3,5 dimethyl hep-	10.29	41.598	40.682	5.24	1.97286	1.87895	1.73313	2.00498
$52 \\ 53 \\ 53 \\ 53 \\ 53 \\ 53 \\ 53 \\ 53 \\ $	tante 3 ethyl heptane 4 ethyl heptane 2,2,3 trimethyl	$\begin{array}{c} 10.564 \\ 10.492 \\ 9.246 \end{array}$	$\begin{array}{c} 40.55 \\ 40.356 \\ 42.058 \end{array}$	$39.362 \\ 39.282 \\ 43.534$	$5.522 \\ 5.47 \\ 4.442$	$\begin{array}{c} 2.00574 \\ 1.99308 \\ 1.72903 \end{array}$	$\begin{array}{c} 1.93199 \\ 1.92114 \\ 1.64772 \end{array}$	$\begin{array}{c} 1.80609 \\ 1.79909 \\ 1.47833 \end{array}$	2.02754 2.01254 1.75095
54	hexane 2,2,4 trimethyl	9.13	41.652	40.636	4.366	1.70869	1.6297	1.53871	1.73747
55	$\begin{array}{ccc} nexane \\ 2,2,5 \\ trimethyl \\ relation \\ relati$	9.06	41.904	39.964	4.254	1.69628	1.6449	1.57313	1.71548
56	2,3,4 trimethyl	10.096	42.912	44.166	5.02	1.94783	1.83899	1.60965	1.98374
57	2,3,5 trimethyl	9.336	41.14	41.366	4.498	1.73813	1.69724	1.57664	1.75161
58	3,3,4 tri methyl hexane	10.036	43.816	45.948	4.994	1.94642	1.7896	1.52156	1.99294

 ${\bf Table \ 1} \ \ {\rm The \ four \ energy \ variants \ of \ alkanes \ with \ respective \ entropies \ (continued)$

Sl.No	Name		$\mathscr{E}[\mathbf{A}(\mathbf{G})]$	$\mathscr{E}[\mathbf{FZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SZ}(\mathbf{G})]$	&[SC(G)]	$\mathcal{I}[\mathbf{A}(\mathbf{G})]$	$\mathcal{I}[\mathbf{FZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SZ}(\mathbf{G})]$	$\mathcal{I}[\mathbf{SC}(\mathbf{G})]$
59	2,4,4	trimethyl	9.19	41.716	41.546	4.42	1.71946	1.632	1.51052	1.74798
60	4 ethyl	2 methyl	10.176	41.38	40.346	5.156	1.95078	1.86668	1.72928	1.97719
61	a ethyl	4 methyl	10.424	41.716	42.554	5.39	1.99149	1.8817	1.68215	2.02353
62	a ethyl	2 methyl	10.198	41.426	41.51	5.184	1.95227	1.86584	1.69281	1.97631
63	pentant $2,3,3,4$	tetra	9.152	43.562	47.78	4.298	1.71881	1.62575	1.36223	1.7435
64	2,2,3,4	pentane tetra	9.02	43.252	45.548	4.2	1.69548	1.61236	1.40549	1.72356
65	2,2,3,3	pentane tetra	8.98	44.234	49.07	4.182	1.6942	1.569	1.29186	1.73385
66	2,2,4,4	pentane tetra	7.936	41.568	40.388	3.48	1.37821	1.37234	1.3362	1.38109
67	a et	pentane hyl 2,2	9.3	41.918	43.9	4.542	1.73773	1.63855	1.45999	1.76657
68	annerny 3 et	hyl 2,4	9.43	41.07	42.778	4.636	1.75323	1.6904	1.52681	1.77133
69	3,3 die	yl pentane sthyl pen-	10.472	42.738	44.25	5.46	2.00507	1.83652	1.57551	2.0465
70	ane 3 et	hyl $2,3$	10.062	43.814	46.318	5.026	1.94808	1.78809	1.49394	1.99222
71	anneun 3 ethyl	yr pentane 3 methyl	10.262	42.5	42.938	5.248	1.97331	1.82523	1.6111	2.01146
72	1,0,2,3,3 2,3,3 hexane	trimethyl	9.3	42.114	44.384	4.494	1.73773	1.64927	1.45473	1.75904

 ${\bf Table \ 1} \ \ {\rm The \ four \ energy \ variants \ of \ alkanes \ with \ respective \ entropies \ (continued)$

Set entropy = -result Print entropy.

Physiochemical properties such as boiling point, melting point, critical temperature, critical pressure, molar volume, molar refraction, the heat of vaporization, density, refractive index, surface tension, vapor pressure, logP, complexity, flash point, and polarizability are taken to find the relationship among the molecular descriptors. All the physical properties of alkanes are taken from online resources: Chemspider[?], National Institute of Standards and Technology[?] and Pubchem [?]. Table 1 gives the energy and entropy for the 72 alkane isomers. The regression analysis for the energies found in Table 1 is shown in Tables 2-5. The linear regression is obtained using the equation Y = a + b(MD), where Y is the physiochemical property considered, MD represents the energy that is being examined, a and b are constant and regression coefficients, respectively. Further, r, F, p, and S.E.E. represent the correlation coefficient, F-value, and standard error of estimate respectively.

End

Figure 1 gives the correlation coefficient between energy and physical properties of alkanes. Some interesting results observed in the process are listed below.



Fig. 1 Correlation Coefficient between the properties of alkanes and energies.

Observation

- 1. The following isomer pairs (referred using serial numbers from Table 1) (2,3), (5,9,17,18,23,26,40,47), (7,38), (8,12), (10,24,50), (14,35), (15,34,55,61), (16,52),(25,51),(28,45), (42,57), (43,60), (66,67,69,72), (31,71), (49,68), (48,56) have the same adjacency spectral radius and same adjacency spectral spread. The alkane isomers (2,5), (3,8,9,15,18,36,69), (11,12,17,23,24,34,49,52,58,61,62,63,65,67,68,72) have the same adjacency second largest eigenvalue. Further, (1,9,18,37), (17,23), (67,72), (25,51) (26,40), (31,71), (49,68), (43,60), (42,57), and (53,62) have the same adjacency spectral gap.
- The alkane isomers (2,5), (3,8,15,18,36,69), (11,23,27,49,52), (17,34,63,68), (21,43), (25,40), (45,67) have the same first Zagreb eigenvalue.

Physiochemical Property	Ν	a	q	r	R^2	F	p-value	S.E.E.	Indicator
Boiling point Melting point Critical temperature Critical pressure Molar volume Molar refraction Heat of vaporization Density Index of Refraction Surface tension Vapour pressure logP Complexity Flash point Polarizability	$\begin{array}{c} & & & & & & \\ & & & & & & & \\ & & & & $	$\begin{array}{c} -66.970\\ -108.936\\ 86.470\\ 86.470\\ 59.882\\ 10.539\\ 10.539\\ 10.547\\ 1.330\\ 13.985\\ 1.330\\ 1.3752\\ -15.752\\ -15.752\\ -15.752\\ -15.752\\ 4.140\end{array}$	$\begin{array}{c} 20.566\\ 0.464\\ 0.464\\ 122.565\\ -2.044\\ 11.672\\ 3.287\\ 3.287\\ 3.242\\ 0.018\\ 0.018\\ 0.008\\ 0.018\\ 0.008\\ 0.812\\ -177.745\\ 0.364\\ 7.063\\ 1.306\end{array}$	$\begin{array}{c} 0.940\\ -0.021\\ 0.919\\ 0.917\\ 0.901\\ 0.914\\ 0.895\\ 0.914\\ 0.865\\ 0.916\\ 0.916\\ 0.940\\ 0.900\\ 0.902\\ 0.902\\ 0.902\\ \end{array}$	$\begin{array}{c} 0.884\\ 0.0005\\ 0.8046\\ 0.812\\ 0.812\\ 0.812\\ 0.835\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.832\\ 0.810\\ 0.813\\ 0.810\\ 0.813\\ 0.811\end{array}$	$\begin{array}{c} 535.555\\ 0.031\\ 380.106\\ 158.077\\ 158.077\\ 158.077\\ 287.276\\ 387.276\\ 387.365\\ 301.857\\ 357.910\\ 733.365\\ 357.910\\ 71.496\\ 51.410\\ 71.496\\ 51.410\\ 299.807\\ 299.807\end{array}$	$\begin{array}{c} 0.000\\ 0.861\\ 0.000\\ 0.$	$\begin{array}{c} 12.4278\\ 36.8421\\ 16.1856\\ 2.2740\\ 9.6300\\ 2.6454\\ 2.4110\\ 0.0175\\ 0.0065\\ 0.0065\\ 0.0065\\ 0.0065\\ 0.0065\\ 13.7757\\ 13.7757\\ 13.0819\\ 1.0459\end{array}$	significant significant significant significant significant significant significant significant significant significant significant significant significant significant significant
Table 3 QSPR analysis of alkan	le isom	ters using fire	st Zagreb ei	nergy					
Physiochemical Property	N	a	q	r	R^2	F	p-value	S.E.E.	Indicator
Boiling point Melting point Critical temperature Critical pressure Molar volume Molar refraction Heat of vaporization Density Index of Refraction Surface tension Vapour pressure log P Complexity Flash point Polarizability	222222222222222222222222222222222222	-59.3964 -136.882 90.8786 42.8168 57.5612 10.0145 11.7682 0.5442 1.7682 0.5442 11.7682 0.5442 11.7682 0.5442 12.0617 610.3673 1.7516 -33.9896 -118.684 3.9523	$\begin{array}{c} 4.6912\\ 0.8849\\ 5.2553\\ -0.4775\\ 0.777\\ 0.7977\\ 0.7977\\ 0.7977\\ 0.7263\\ 0.0044\\ 0.0019\\ 0.1638\\ -40.818\\ 0.0791\\ 0.1638\\ 0.0791$	$\begin{array}{c} 0.946\\ -0.180\\ 0.944\\ 0.944\\ 0.9645\\ 0.9645\\ 0.903\\ 0.923\\ 0.923\\ 0.923\\ 0.923\\ 0.923\\ 0.923\\ 0.923\\ 0.923\\ 0.966\\ 0.887\\ 0.887\\ 0.966\end{array}$	$\begin{array}{c} 0.895\\ 0.032\\ 0.032\\ 0.891\\ 0.736\\ 0.931\\ 0.852\\ 0.852\\ 0.852\\ 0.852\\ 0.852\\ 0.815\\ 0.852\\ 0.812\\ 0.523\\ 0.687\\ 0.523\\ 0.531\\ 0.531\\ 0.791\\ 0.791\\ 0.791\\ 0.787\\ 0.533\\ 0.931\\ 0.931\\ 0.931\\ 0.952\\ 0.931\\ 0.932\\ 0.931\\ 0.$	$\begin{array}{c} 598.258\\ 2.2659\\ 194.727\\ 194.727\\ 928.653\\ 935.834\\ 534.434\\ 534.434\\ 534.434\\ 534.434\\ 138.137\\ 75.75$	$\begin{array}{c} 0.000\\ 0.137\\ 0.000\\ 0.$	$\begin{array}{c} 11.830\\ 36.252\\ 13.5434\\ 2.1107\\ 5.7600\\ 1.6085\\ 2.5498\\ 0.0134\\ 0.0134\\ 0.0134\\ 0.0134\\ 0.0134\\ 0.0134\\ 0.0134\\ 0.02545\\ 0.8545\\ 0.8545\\ 0.2822\\ 8.2978\\ 13.8545\\ 0.6254\end{array}$	significant significant significant significant significant significant significant significant significant significant significant significant significant significant

 ${\bf Table}~{\bf 2}~~{\rm QSPR}$ analysis of alkane isomers using adjacency energy

Physiochemical Property	Z	a	q	r	R^2	F	<i>p</i> -value	S.E.E.	Indicator
Boiling point Melting point Critical temperature Critical pressure Molar volume Molar refraction Heat of vaporization Density Density Surface tension Vapour pressure log Complexity Flash point Polarizability	1222222222222222222222222222222222222	$\begin{array}{c} -25.5568\\ -137.230\\ 126.2234\\ 38.8709\\ 78.0203\\ 15.7715\\ 15.7715\\ 15.7715\\ 15.7715\\ 15.7715\\ 15.7715\\ 15.7715\\ 15.7438\\ 0.5707\\ 1.3426\\ 16.2405\\ 2.4059\\ 2.4059\\ 2.4059\\ 2.2465\\ 10.2246\\ 2.2405\\ 10.2246\\ 2.2405\\ 10.2246\\ 2.2405\\ 10.2246\\ 10.2246\\ 2.2265\\ 10.2246\\ 2.230\\ 6.230\end{array}$	$\begin{array}{c} 3.8270\\ 0.9119\\ 0.9119\\ 4.3595\\ -0.3754\\ 2.3233\\ 0.6507\\ 0.6507\\ 0.6587\\ 0.6507\\ 0.6528\\ 0.0037\\ 0.0037\\ 0.0016\\ 0.1253\\ -33.423\\ 0.0622\\ 1.9157\\ 1.9157\\ 1.9157\\ 0.0528\\ \end{array}$	$\begin{array}{c} 0.908\\ -0.218\\ 0.921\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.926\\ 0.916\\ 0.734\\ 0.916\\ 0.848\\ 0.916\\ 0.916\\ 0.916\end{array}$	$\begin{array}{c} 0.824\\ 0.047\\ 0.047\\ 0.857\\ 0.856\\ 0.856\\ 0.847\\ 0.847\\ 0.846\\ 0.839\\ 0.485\\ 0.692\\ 0.838\\ 0.838\\ 0.719\\ 0.860\\ 0.860\\ \end{array}$	$\begin{array}{c} 327.834\\ 3.3801\\ 3.3801\\ 391.054\\ 118.548\\ 420.757\\ 416.848\\ 1192.677\\ 387.707\\ 387.707\\ 387.707\\ 365.0413\\ 157.281\\ 157.281\\ 157.281\\ 157.281\\ 157.281\\ 157.281\\ 157.281\\ 179.372\\ 423.132\\ \end{array}$	$\begin{array}{c} 0.000\\ 0.070\\ 0.070\\ 0.000\\ 0.$	$\begin{array}{c} 15.333\\ 35.967\\ 15.9923\\ 2.5010\\ 8.2166\\ 2.3120\\ 3.0616\\ 0.0136\\ 0.0136\\ 0.0136\\ 0.0136\\ 0.0059\\ 1.0049\\ 300.545\\ 0.3595\\ 1.2969\\ 115.8978\\ 0.3595\\ 0.3595\\ 0.3595\\ 0.3595\\ 115.8978\\ 0.9054 \end{array}$	significant significant significant significant significant significant significant significant significant significant significant significant significant significant significant
Table 5 QSPR analysis of alkan	ie isom	ners using sun	a-connectivi	ity energy					
Physiochemical Property	N	a	q	r	R^2	F	<i>p</i> -value	S.E.E.	Indicator
Boiling point Melting point Critical temperature Critical pressure Molar volume Molar refraction Heat of vaporization Density Index of Refraction Surface tension Vapour pressure logP Complexity Flash point Polarizability	122222222222222222222222222222222222222	$\begin{array}{c} -45.7457\\ -94.916\\ 110.909\\ 110.909\\ 75.3583\\ 14.8278\\ 13.3318\\ 0.5682\\ 1.33318\\ 0.5682\\ 1.3393\\ 14.4718\\ 1.3933\\ 14.4718\\ 1.3933\\ 1.7923\\ 0.9056\\ -111.6216\\ 5.8467\end{array}$	36.1766 -2.341 39.4251 -3.5126 19.7336 5.5728 5.5728 0.0310 0.01310 0.01310 0.01310 1.5096 1.5096 1.5096 1.5096 1.5096 1.5096 1.5096 1.5096 2.2137 2.2137	$\begin{array}{c} 0.871 \\ -0.056 \\ 0.845 \\ -0.753 \\ 0.798 \\ 0.798 \\ 0.850 \\ 0.851 \\ 0.851 \\ 0.895 \\ 0.895 \\ 0.888 \\ 0.888 \\ 0.895 \\ 0.891 \\ 0.891 \\ 0.891 \\ 0.891 \\ 0.804 \\ 0.804 \end{array}$	$\begin{array}{c} 0.758\\ 0.714\\ 0.567\\ 0.567\\ 0.637\\ 0.647\\ 0.647\\ 0.657\\ 0.614\\ 0.614\\ 0.614\\ 0.614\\ 0.616\\ 0.789\\ 0.789\\ 0.789\\ 0.789\\ 0.705\\ 0.705\\ 0.705\end{array}$	$\begin{array}{c} 219.460\\ 0.2164\\ 174.941\\ 91.6055\\ 122.740\\ 122.740\\ 131.976\\ 111.112\\ 134.272\\ 134.272\\ 134.272\\ 134.272\\ 2281.325\\ 22.956\\ 167.435\\ 125.911\\ 125.911\end{array}$	$\begin{array}{c} 0.000\\ 0.6433\\ 0.000\\ 0$	$\begin{array}{c} 17.9754\\ 36.7920\\ 21.9409\\ 2.7014\\ 13.1112\\ 3.1259\\ 0.217\\ 0.0086\\ 0.6600\\ 314.747\\ 0.2977\\ 15.7436\\ 16.2925\\ 16.2925\\ 1.4387\end{array}$	significant significant significant significant significant significant significant significant significant significant significant significant significant significant

 ${\bf Table \ 4} \ {\rm QSPR} \ {\rm analysis} \ {\rm of} \ {\rm alkane} \ {\rm isomers} \ {\rm using} \ {\rm second} \ {\rm Zagreb} \ {\rm energy}$

- 3. Similarly, (2,5), (3,8,15,18,36,69), (11,23,49,52), (12,24,62,72), (17,34,68), (54,66) have the same second Zagreb eigenvalue, and (1,3) have the same second Zagreb spectral gap.
- 4. The pairs of isomers (3,4), (12,14,17), (13,19), (18,23,25,66), (21,26,28,32), (22,39,42,47,55), (24,37,40), (27,33,43,57), (30,34), (31,41,44), (35,53), (36,51,60), (45,64,72), (48,59), (49,50), (63,68), (65,67,71) have the same sum-connectivity spectral radius and spectral spread. Also (2,5), (3,8,15,18,36,69), (11,23,49,52), (24,61), (62,72), (34,63,69), (43,50), (22,60) have the same sum-connectivity second largest eigenvalue, and (16,35), (27,29), and (63,68) have the same sum-connectivity spectral gap.
- 5. A comparative analysis of Tables 2-5 shows that the first Zagreb energy is highly correlated for the properties like boiling point, critical temperature, molar volume, molar refraction, density, index of refraction, and polarizability with values 0.946, 0.944, 0.964, 0.964, 0.923, 0.945, and 0.966 respectively, whereas adjacency energy performs relatively higher with respect to the descriptors heat of vaporization, surface tension, logP, flash point with correlation values 0.914, 0.916, and 0.9 respectively. The second Zagreb energy exhibits a high correlation in complexity with a correlation value of 0.916.
- 6. The low correlation values of -0.021, -0.18, -0.218, and -0.056 observed between the graph-theoretical descriptors of adjacency energy, first and second Zagreb energy, sum-connectivity energy, and the melting points of isomers stem from the fact that these descriptors fail to identify the precise factors that influence melting points. Melting points are influenced by a solid's molecular interactions, such as how effectively molecules pack or adhere to one another. These descriptors primarily describe the molecular structures but do not indicate the behavior of the molecules during the transition from solid to liquid state.
- 7. The strong negative correlation of -0.832, -0.858, -0.793, and -0.753 suggests that these graph energies serve as a good predictor for critical pressure, and there is a clear relationship indicating an increase in these energy values corresponds to a decrease in the critical pressure.
- 8. Also, these energy descriptors show an inverse relationship with respect to vapor pressure (from the table), which suggests that when the structural complexity of the molecule increases, the vapor pressure tends to decrease. This enables us to understand that molecules with higher structural descriptors are less likely to evaporate easily, and hence, they have lower vapor pressure.
- 9. The low *p*-values suggest that the correlations are not only statistically significant but also robust and consistent. This provides a strong validation to describe the relation between the molecular descriptors and physical properties. This enables us to understand that these energy descriptors are likely accurate and may be trusted for predictive purposes.
- 10. The low standard error of estimate suggests that the difference between the observed and predicted values of the physical properties is small and hence makes accurate predictions.

5 Double Regression Analysis on Isomers

From the previous section, it can be seen that although statistical regression techniques yielded interesting results for many of the physiochemical properties, they do not give satisfactory results in some cases, such as melting point. The linear regression yielded a negative correlation for some of the properties. Hence, further detailed analysis is warranted, and also a study to understand the role of graph entropy in such analysis is found necessary at this stage. With this in view, our study is extended using double regression analysis. The term double regression refers to a statistical modeling technique where two descriptors are concurrently used to forecast the value of a single response. Here, the four energy variants and their corresponding entropies are considered independent variables, and each of the physical properties of 72 isomers of alkanes are taken as a dependent variable, and statistical analysis is carried out. It is observed that there is a very high correlation for all the properties of alkanes using double regression. It is also noticed that there is a moderate positive correlation in melting point while using energy and entropy as an independent variable than linear regression. This is an interesting factor to be reckoned with since a survey of the existing literature shows only a negative correlation for this particular property. Critical pressure and vapor pressure also find a very high positive correlation when entropy is introduced as one of the independent variables. This is a crucial point, as earlier studies only yield a negative correlation for this property. This high positivity occurred only due to the presence of the indicator graph entropy. Hence, entropy serves as a very good descriptor for understanding the relationship among the compounds. The inclusion of eigenvalue-based entropy in the analysis suggests that the complexity and distribution of energy levels within the molecular graph structure are also being considered and hence, influence the relationship between molecular descriptors and physical properties differently in multiple regression models compared to simple linear regression. Table 6 exhibits the results obtained from the double regression using various combinations of energy and their entropy pairs. Table 7 gives the variance inflation factor, RMSE, and MAE of the predictor variables using a 10-fold cross-validation process, which validates our findings.

Observation

- 1. Physical properties like boiling point, critical temperature, heat of vaporization, density, surface tension, logo, and flash point are highly correlated with second Zagreb energy and its corresponding entropy with values 0.974, 0.965, 0.942, 0.933, 0.942, 0.956, and 0.924 respectively. A comparison of Tables 4 and 6 indicates the significant role of entropy with improved correlation, validating the usage of the double regression technique.
- 2. Properties like molar volume, molar refraction, index of refraction, vapor pressure, complexity, and polarizability exhibit a strong correlation with first Zagreb energy and its corresponding entropy with correlation values of 0.966, 0.966, 0.950, 0.768, 0.953, and 0.968 respectively. As in previous observations, the energy-entropy pair has resulted in improved correlation (Tables 3 and 6).

Properties	Energy and Entropy pair	r	R^2	a	q	c	<i>p</i> - value	<i>F</i> - value	S.E.E
Boiling point	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SZ(G)] \\ \end{split} $	$\begin{array}{c} 0.953 \\ 0.969 \\ 0.974^{*} \\ 0.874 \end{array}$	$\begin{array}{c} 0.908 \\ 0.940 \\ 0.948 \\ 0.763 \end{array}$	-45.126 -89.869 -82.665 -31.200	$\begin{array}{c} 29.520\\ 3.550\\ 2.867\\ 47.308\end{array}$	-57.938 43.987 61.201 -36.101	0.0000 0.0000 0.0000 0.0000	339.235 536.812 635.594 111.364	11.186 9.047 8.354 17.905
Melting point	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SZ(G)] \\ \end{split} $	$\begin{array}{c} 0.335 \\ 0.368^{*} \\ 0.350 \\ 0.182 \end{array}$	$\begin{array}{c} 0.112 \\ 0.136 \\ 0.122 \\ 0.033 \end{array}$	-59.783 -90.068 -92.985 -59.827	20.244 2.646 1.660 24.145	-128.76 -67.960 -47.640 -86.263	$\begin{array}{c} 0.018 \\ 0.008 \\ 0.012 \\ 0.324 \end{array}$	$\begin{array}{c} 4.239 \\ 5.265 \\ 4.679 \\ 1.145 \end{array}$	34.977 34.512 34.775 36.506
Critical tem- perature	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SZ(G)] \\ \end{split} $	$\begin{array}{c} 0.926 \\ 0.962 \\ 0.965 \\ 0.845 \end{array}$	$\begin{array}{c} 0.857 \\ 0.926 \\ 0.932 \\ 0.714 \end{array}$	$\begin{array}{c} 104.522 \\ 60.599 \\ 73.628 \\ 108.738 \end{array}$	$\begin{array}{c} 29.965\\ 4.121\\ 3.475\\ 37.763\end{array}$	-47.879 43.708 56.364 5.389	0.0000 0.0000 0.0000 0.0000	$\begin{array}{c} 206.917\\ 430.803\\ 472.203\\ 86.261 \end{array}$	15.627 11.256 10.787 22.096
Critical pres- sure	$ \mathscr{E}[A(G)], \mathcal{I}[A(G)] $ $ \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] $ $ \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] $ $ \mathscr{E}[SC(G)], \mathcal{I}[SC(G)] $	$\begin{array}{c} 0.893 \\ 0.863 \\ 0.861 \\ 0.861 \\ 0.799 \end{array}$	$\begin{array}{c} 0.798\\ 0.745\\ 0.742\\ 0.638\end{array}$	37.972 44.392 44.988 34.7136	-4.174 -0.418 -0.272 -8.091	$\begin{array}{c} 13.778 \\ -2.273 \\ -6.555 \\ 14.849 \end{array}$	0.0000 0.0000 0.0000 0.0000	$135.882 \\ 100.780 \\ 99.211 \\ 60.824$	1.860 2.088 2.1 2.487
Molar vol- ume	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SZ(G)] \\ \end{split} $	$\begin{array}{c} 0.947 \\ 0.966^{*} \\ 0.957 \\ 0.815 \end{array}$	$\begin{array}{c} 0.897\\ 0.933\\ 0.917\\ 0.664\end{array}$	85.860 52.982 54.571 94.814	$\begin{array}{c} 22.320\\ 2.674\\ 1.929\\ 34.622\end{array}$	-68.904 6.610 25.129 -48.286	0.0000 0.0000 0.0000 0.0000	300.574 478.386 379.174 68.064	7.031 5.683 6.328 12.709

 Table 6
 QSPR analysis using double regression

Properties	Energy and Entropy pair	r	R^{2}	a	<i>q</i>	c	<i>p</i> - value	F- value	S.E.E
Molar refrac- tion	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SC(G)] \\ \end{split} $	$\begin{array}{c} 0.950\\ 0.966*\\ 0.959\\ 0.821\end{array}$	$\begin{array}{c} 0.902 \\ 0.934 \\ 0.92 \\ 0.674 \end{array}$	$17.694 \\ 8.513 \\ 8.952 \\ 20.319$	6.220 0.742 0.536 9.775	-18.979 2.168 7.308 -13.628	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	315.943 490.444 396.739 71.261	$\begin{array}{c} 1.927 \\ 1.574 \\ 1.737 \\ 3.508 \end{array}$
Heat of vapourisa- tion	$ \begin{split} & \& [A(G)], \mathcal{I}[A(G)] \\ & \& [FZ(G)], \mathcal{I}[FZ(G)] \\ & \& [SZ(G)], \mathcal{I}[SZ(G)] \\ & \& [SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.927 \\ 0.933 \\ 0.942^{*} \\ 0.857 \end{array}$	$\begin{array}{c} 0.860 \\ 0.870 \\ 0.887 \\ 0.734 \end{array}$	$13.79 \\ 6.253 \\ 6.938 \\ 16.908$	$\begin{array}{c} 4.755\\ 0.520\\ 0.413\\ 8.465\end{array}$	-9.789 7.961 11.044 -8.876	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	$\begin{array}{c} 211.905\\ 231.622\\ 272.136\\ 95.387\end{array}$	2.235 2.151 2.004 3.079
Density	$ \begin{split} & \& [A(G)], \mathcal{I}[A(G)] \\ & \& [FZ(G)], \mathcal{I}[FZ(G)] \\ & \& [SZ(G)], \mathcal{I}[SZ(G)] \\ & \& [SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.876 \\ 0.928 \\ 0.933* \\ 0.784 \end{array}$	$\begin{array}{c} 0.767 \\ 0.861 \\ 0.871 \\ 0.614 \end{array}$	$\begin{array}{c} 0.566\\ 0.532\\ 0.547\\ 0.564\end{array}$	$\begin{array}{c} 0.026 \\ 0.004 \\ 0.003 \\ 0.0276 \end{array}$	-0.0508 0.0182 0.0256 0.011	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	$\begin{array}{c} 113.459\\ 213.322\\ 32.873\\ 54.888\end{array}$	$\begin{array}{c} 0.0169 \\ 0.0131 \\ 0.0229 \\ 0.0218 \end{array}$
Index of Refraction	$ \begin{split} & \& [A(G)], \mathcal{I}[A(G)] \\ & \& [FZ(G)], \mathcal{I}[FZ(G)] \\ & \& [SZ(G)], \mathcal{I}[SZ(G)] \\ & \& [SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.927 \\ 0.950* \\ 0.946 \\ 0.820 \end{array}$	$\begin{array}{c} 0.859 \\ 0.902 \\ 0.895 \\ 0.672 \end{array}$	$\begin{array}{c} 1.3441 \\ 1.3243 \\ 1.3273 \\ 1.3490 \end{array}$	$\begin{array}{c} 0.0135\\ 0.0017\\ 0.0013\\ 0.0209\end{array}$	-0.0364 0.0077 0.0163 -0.0240	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	209.713 315.844 293.327 70.648	$\begin{array}{c} 0.0056 \\ 0.0046 \\ 0.0048 \\ 0.0085 \end{array}$
Surface ten- sion	$ \begin{split} & \& [A(G)], \mathcal{I}[A(G)] \\ & \& [FZ(G)], \mathcal{I}[FZ(G)] \\ & \& [SZ(G)], \mathcal{I}[SZ(G)] \\ & \& [SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.916\\ 0.917\\ 0.942*\\ 0.898\end{array}$	$\begin{array}{c} 0.838 \\ 0.840 \\ 0.888 \\ 0.807 \end{array}$	$\begin{array}{c} 14.0311\\ 12.6075\\ 12.6777\\ 15.0853\end{array}$	$\begin{array}{c} 0.8303 \\ 0.0735 \\ 0.0609 \\ 1.9724 \end{array}$	-0.1195 3.5037 4.1236 -1.5055	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	176.441 178.718 270.701 142.133	$\begin{array}{c} 0.5995 \\ 0.5962 \\ 0.4982 \\ 0.6552 \end{array}$
Vapour pres- sure	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.723\\ 0.768*\\ 0.764\\ 0.714\end{array}$	$\begin{array}{c} 0.523\\ 0.589\\ 0.583\\ 0.510\\ 0.510\end{array}$	1861.49 2034.83 1897.13 2112.60	-97.823 -24.945 -23.733 166.413	-518.64 -612.96 -618.33 - 1558.08	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	37.276 48.764 47.588 35.393	$\begin{array}{c} 291.434\\ 270.454\\ 272.396\\ 295.364\end{array}$

 ${\bf Table \ 6} \ {\rm QSPR \ analysis \ using \ double \ regression \ (continued)}$

Table 6 QSP	R analysis using doubl	e regressic	on (contin	(pən				
Properties	Energy and Entropy pair	٤	R^2	a	p	ల	<i>p</i> - value	<i>F</i> - value
logP	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.954 \\ 0.942 \\ 0.956* \\ 0.904 \end{array}$	$\begin{array}{c} 0.911 \\ 0.887 \\ 0.914 \\ 0.817 \end{array}$	$\begin{array}{c} 1.8876\\ 1.0396\\ 1.0555\\ 2.383\end{array}$	$\begin{array}{c} 0.5346\\ 0.0524\\ 0.0394\\ 1.1059\end{array}$	-1.1019 1.0277 1.4471 -1.4659	0.0000 0.0000 0.0000 0.0000	351.738 271.571 364.908 153.616
complexity	$\& [A(G)], \mathcal{I}[A(G)] \\ \& [FZ(G)], \mathcal{I}[FZ(G)] \\ \& [SZ(G)], \mathcal{I}[SZ(G)] \\ \& [SC(G)], \mathcal{I}[SC(G)] \end{pmatrix}$	$\begin{array}{c} 0.788\\ 0.953*\\ 0.941\\ 0.502 \end{array}$	$\begin{array}{c} 0.621 \\ 0.908 \\ 0.886 \\ 0.252 \end{array}$	$15.811 \\ -9.3761 \\ -5.2185 \\ 8.1946$	20.001 3.1102 2.2118 15.826	-83.717 -35.529 -18.869 -18.090	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	$\begin{array}{c} 156.478\\ 342.092\\ 288.271\\ 21.645\end{array}$
Flash point	$ \begin{split} & \& [A(G)], \mathcal{I}[A(G)] \\ & \& [FZ(G)], \mathcal{I}[FZ(G)] \\ & \& [SZ(G)], \mathcal{I}[SZ(G)] \\ & \& [SC(G)], \mathcal{I}[SC(G)] \end{split} $	$\begin{array}{c} 0.907 \\ 0.920 \\ 0.924^{*} \\ 0.841 \end{array}$	$\begin{array}{c} 0.822 \\ 0.846 \\ 0.854 \\ 0.708 \end{array}$	-114.13 147.53 -140.97 -103.55	$\begin{array}{c} 21.5735\\ 2.5298\\ 2.1165\\ 34.8368\end{array}$	-35.052 41.644 52.173 -20.095	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	$\begin{array}{c} 159.948\\ 189.263\\ 201.040\\ 83.4983\end{array}$
Polarizability	$ \begin{split} & \mathscr{E}[A(G)], \mathcal{I}[A(G)] \\ & \mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)] \\ & \mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)] \\ & \mathscr{E}[SC(G)], \mathcal{I}[SZ(G)] \end{split} $	$\begin{array}{c} 0.949 \\ 0.968^{*} \\ 0.961 \\ 0.820 \end{array}$	$\begin{array}{c} 0.901 \\ 0.937 \\ 0.923 \\ 0.672 \end{array}$	6.9878 3.3414 3.5457 7.9866	$\begin{array}{c} 2.4511 \\ 0.2940 \\ 0.2136 \\ 3.8281 \end{array}$	-7.4389 0.8721 2.8651 -5.251	$\begin{array}{c} 0.0000\\ 0.0000\\ 0.0000\\ 0.0000\end{array}$	310.742 506.848 407.416 69.6166

 $\begin{array}{c} 11.253\\ 5.5337\\ 6.1627\\ 15.800\end{array}$

S.E.E

 $\begin{array}{c} 0.1950\\ 0.2190\\ 0.1918\\ 0.2794\end{array}$

 $\begin{array}{c} 12.7304 \\ 11.8673 \\ 111.5668 \\ 83.4983 \end{array}$

 $\begin{array}{c} 0.7649 \\ 0.6107 \\ 0.676 \\ 1.3953 \end{array}$

17

Note: Here * indicates the highest value among the various double regression pairs considered.

- 3. The correlation values are stronger, indicating a better fit and a more meaningful relationship between the descriptors (energy and entropy) for the property: melting point. Further, it can interpreted from Table 6 that a combination of the first Zagreb energy and its corresponding entropy is more effective in predicting the melting point for these alkanes compared to using individual descriptors in linear regression (refer to Tables 2-5).
- 4. The correlation values are not only strong but also positive, indicating that the combination of energy and entropy pairs is highly predictive for the critical pressure and for vapor pressure, which is not so in the linear regression. Hence the double regression model suggests that when energy and entropy are combined, they provide a more accurate and positively correlated prediction.
- 5. The high R^2 values and highly significant *F*-test results revealed remarkable predictive performance. A *p*-value < 0.05 indicates that the results are statistically significant, meaning there is strong evidence against the null hypothesis, which usually states that there is no relationship between the variables. Except for the melting point, the p-values for physical properties considered here are much smaller than 0.05, indicating that the relationship observed between the descriptors (energy and its corresponding eigenvalue-based entropy) and the properties are highly statistically significant and highly predictive in QSPR analysis. Thus the high *F*-value with significantly low *p*-values R^2 indicates that the overall regression model is statistically significant.

6 Validation of Model Performance Using 10-Fold Cross-Validation

To ensure the robustness and reliability of our regression models, we employed 10-fold cross-validation to compute Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), and to assess the potential issue of multicollinearity, the Variance Inflation Factor (VIF) was calculated. The results from Table 7 indicate that the first Zagreb energy and second Zagreb energy descriptors, along with their entropy parameter, provide statistically reliable models with minimal multicollinearity, as evidenced by VIF values less than 2.5, whereas adjacency energy and sum-connectivity, along with their entropy parameters, exhibited a higher VIF, suggesting potential multicollinearity concerns that may require further refinement through dimensionality reduction or alternative modeling techniques. The low RMSE and MAE values (Table 8) for a majority of the attributes indicate that the models perform extremely well in terms of prediction. The resilience and predictive ability of the models in estimating physiochemical attributes are thereby clearly demonstrated by the consistently low RMSE and MAE values observed over 10-fold cross-validation with significantly low *p*-values. Figure 2-3 illustrates the optimal regression line within the double regression model, highlighting the model's goodness of fit.





Table 7 Variance inflation factor for independent variables

$\mathscr{E}[\mathbf{A}(\mathbf{G})], \mathcal{I}[\mathbf{A}(\mathbf{G})]$	$\mathscr{E}[\mathbf{FZ}(\mathbf{G})], \mathcal{I}[\mathbf{FZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SZ}(\mathbf{G})]{,}\mathcal{I}[\mathbf{SZ}(\mathbf{G})]$	$\mathscr{E}[\mathbf{SC}(\mathbf{G})]{,}\mathcal{I}[\mathbf{SC}(\mathbf{G})]$
8.199	2.194	1.417	14.522

Table 8 VIF, RMSE, MAE values of the predictor variables

Properties	Energy and Entropy pair	VIF	RMSE	MAE	p-value
Boiling point	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.417	6.065	4.286	3.57e-45
Melting point	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.233	28.274	21.986	0.00752
Critical temperature	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	1.417	9.944	7.893	5.51e-41
Critical pressure	$\mathscr{E}[A(G)], \mathcal{I}[A(G)]$	8.199	1.718	1.384	1.18e-24
Molar volume	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.194	4.658	3.404	3.62e-41
Molar refraction	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.194	1.286	0.952	1.63e-41
Heat of vapourisation	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.417	1.614	1.078	1.85e-33
Density	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.417	0.012	0.009	2.09e-31
Index of Refraction	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.194	0.004	0.004	1.86e-35
Surface tension	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.408	0.485	0.396	4.16e-33
logP	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.417	0.177	0.125	4.16e-33
complexity	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.194	4.532	5.243	1.59e-36
Flash point	$\mathscr{E}[SZ(G)], \mathcal{I}[SZ(G)]$	1.417	10.579	8.518	1.65e-29
Polarizability	$\mathscr{E}[FZ(G)], \mathcal{I}[FZ(G)]$	2.170	0.484	0.376	1.4e-41

7 Comparative analysis of other spectral parameters for density and refractive index

In this section, the spectral properties (excluding energy and entropy) are used as molecular descriptors to build predictive models that relate the structure of a molecule to its physiochemical properties. Statistical analysis with properties such as density and refractive index of 72 isomers with spectral radii, second largest eigenvalue, spectral gap, and spectral spread for the adjacency, first and second Zagreb, sumconnectivity matrices are established in Tables 9 and 10, and Figure 4 gives the line of best fit for the spectral parameters of SC(G).

Observation

- 1. It is observed from Tables 9 and 10 that the spectral radius of sum-connectivity matrix SC(G) is highly correlated for both density and refractive index with correlation values 0.92456 and 0.89765, respectively.
- 2. A notable positive linear relationship between density and the second-largest eigenvalue SC(G) is seen with a correlation of 0.75069. Similarly, a positive and high correlation between the refractive index and the second-largest eigenvalue of the

 ${\bf Table \ 9} \ {\rm QSPR} \ {\rm analysis} \ {\rm of} \ {\rm density} \ {\rm with \ spectral} \ {\rm properties}$

Jensity	$\substack{\text{Spectral}\\ \text{radius}\\ A(G)$	A(G)	$_{\substack{\text{gap}\\A(G)}}^{\text{Spectral}}$	$\begin{array}{c} \text{Spectral} \\ \text{spread} \\ A(G) \end{array}$	$\begin{array}{c} \text{Spectral} \\ \text{radius} \\ FZ(G) \end{array}$	eigen- value $FA(G)$	$\substack{\text{gap}\\FZ(G)}$	$\begin{array}{c} \operatorname{Spectral} \\ \operatorname{spread} \\ FZ(G) \end{array}$	$\begin{array}{c} \operatorname{Spectral} \\ \operatorname{radius} \\ SZ(G) \end{array}$	eigen- value $SZ(G)$	$\substack{\text{gap}\\SZ(G)}$	Spectral spread $SZ(G)$	Spectral radius $SC(G)$	largest eigen- value SC(G)	$\substack{\text{gap}\\SC(G)}$	$\substack{\text{Spectral}\\\text{spread}\\SC(G)}$
	0.72598	0.72629	-0.44347	0.72598	0.60456	0.69005	-0.02537	0.60457	0.67251	0.61453	0.29803	0.66853	0.92456	0.71592	-0.63315	0.92456
	-0.0118	-3.5029	3.4915	-0.0236	-11.7127	17.7837	6.0709	-23.425	-32.2108	-17.7578	-14.1323	-63.777	0.4914	-1.5304	2.0218	0.9828
	2.9729	6.9131	-3.9406	5.9458	31.1578	32.678	-1.5199	62.3149	62.587	31.4104	30.7004	124.215	0.6746	3.1748	-2.5002	1.3491
3.E.E	0.0981	0.2279	0.2774	0.1962	11.4299	1.1938	2.086	2.8597	2.399	1.4045	3.4249	4.8128	0.0097	0.1078	0.1065	0.0194
	78.0056	78.1454	17.1368	78.0056	40.3227	63.6321	0.0451	40.3236	57.8011	42.4765	6.8235	56.5677	412.1122	73.6011	46.8379	412.1122
3 ²	0.5279	0.5275	0.1967	0.527	0.3655	0.4762	0.0006	0.3655	0.4523	0.3776	0.0888	0.4469	0.8548	0.5125	0.4009	0.8548
-value	0.000	0.000	0.0001	0.000	0.000	0.000	0.8325	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000

Table to K	אוזשווא זר מ			nde mit w v	oni ai biol	con rod										
Refractive index	Spectral radius $A(G)$	Second largest eigen-value $A(G)$	$\substack{ \text{Spectral} \\ gap \\ A(G) \\ \end{array}$	$\begin{array}{c} \text{Spectral} \\ \text{spread} \\ A(G) \end{array}$	$\operatorname{Spectral}_{radius} FZ(G)$	Second largest eigen-value $FZ(G)$	$\operatorname{Spectral}_{\operatorname{gap}} FZ(G)$	$\operatorname{Spectral}_{\operatorname{spread}}$	$\operatorname{Spectral}_{radius} SZ(G)$	Second largest eigen-value $SZ(G)$	$\begin{array}{c} \operatorname{Spectral} \\ \operatorname{gap} \\ SZ(G) \end{array}$	Spectral spread SZ(G)	$\operatorname{Spectral}_{radius} SC(G)$	Second largest eigen-value $SC(G)$	Spectral gap SC(G)	Spectral spread $SC(G)$
ı	0.68009	0.77167	-0.51324	0.68009	0.54888	0.74702	-0.11831	0.54889	0.58569	0.71274	0.17319	0.5841	0.89765	0.75069	-0.67717	0.89765
8	1.2524	1.3515	1.4151	1.2524	1.352	1.3633	1.4016	1.352	1.3665	1.3723	1.3922	1.3665	0.8977	1.3478	1.4168	0.8977
р	0.0699	0.0341	-0.0243	0.035	0.0045	0.0066	-0.0008	0.0022	0.0027	0.0059	0.0007	0.0013	0.5181	0.0713	-0.0722	0.259
S.E.E	0.0108	0.0093	0.0126	0.0108	0.0123	0.0098	0.0146	0.0123	0.0119	0.0103	0.0144	0.0119	0.0065	0.0097	0.0108	0.0065
Ĺ	60.239	103.0433	25.0329	60.239	30.1821	88.3827	0.9937	30.1833	36.5499	72.2775	2.1644	36.2485	290.3958	90.3806	59.2859	290.3958
\mathbf{R}^2	0.4625	0.5955	0.2634	0.4625	0.3013	0.558	0.3223	0.3013	0.343	0.508	0.03	0.3412	0.8058	0.5635	0.4586	0.8058
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.1457	.000	0.000	0.000	0.000	0.000



Fig. 3 Best-fit line obtained through double regression(Continued)

sum-connectivity matrix r = 0.71592 is noted. However, as compared to other spectra, the second largest eigenvalue of the adjacency matrix is comparatively high for both density and refractive index with correlations of 0.77167 and 0.72629, respectively.

3. There is an inverse relation between the spectral gap SC(G) and density r = -0.63315 as well as for refractive index with a correlation value -0.67717. This shows that as the spectral gap increases, the density (refractive index) tends to decrease, and vice versa.



Fig. 4 Line of best fit for the spectral parameters of SC(G)

Table 11 Comparison of r values with other graph-based indices

Index	BP	\mathbf{CT}	CP	\mathbf{MV}	\mathbf{MR}	\mathbf{HV}	\mathbf{ST}
ZM_1	0.848	0.896	-0.645	0.778	0.833	0.740	0.802
ZM_2	0.875	0.909	-0.698	0.819	0.862	0.803	0.830
ABC	0.748	0.815	-0.451	0.585	0.656	0.638	0.756
R	0.486	0.518	-0.474	0.528	0.546	0.301	0.336
SCI	0.625	0.657	-0.579	0.649	0.673	0.450	0.471
RR	0.879	0.917	-0.705	0.832	0.879	0.783	0.820
GO_1	0.867	0.908	-0.675	0.804	0.854	0.777	0.822
GO_2	0.806	0.844	-0.639	0.759	0.800	0.727	0.746
GA	0.751	0.789	-0.680	0.771	0.799	0.612	0.620
SO	0.728	0.906	-0.133	0.829	0.983	0.975	0.845
SO_{red}	0.260	0.629	0.647	-0.62	-0.537	0.356	0.62
M_{ij}	0.743	0.817	-0.734	0.794	0.816	0.689	0.670
m_{ij}	0.879	0.910	-0.828	0.822	0.824	0.912	0.869
A(G)	0.940	0.919	-0.832	0.897	0.901	0.914	0.916
FZ(G)	0.946	0.944	-0.858	0.964	0.964	0.903	0.817
SZ(G)	0.908	0.921	-0.793	0.926	0.925	0.856	0.734
SC(G)	0.871	0.845	-0.753	0.798	0.804	0.850	0.895
$A(\hat{G}), I[A(G)]$	0.953	0.926	0.893	0.947	0.950	0.927	0.916
FZ(G), $I[FZ(G)]$	0.969	0.962	0.863	0.966	0.966	0.933	0.917
SZ(G), I[SZ(G)]	0.974	0.965	0.861	0.957	0.959	0.942	0.942
SC(G), I[SC(G)]	0.874	0.845	0.799	0.815	0.821	0.857	0.898

- 4. The strong positive correlation with r = 0.92456 and 0.89765 suggests that density and refractive index is closely related to the spectral spread of SC(G) with density (refractive index) increases as spectral spread increases.
- 5. A *p*-values of 0.000 suggest that there is strong statistical evidence to reject the null hypothesis, claiming there is no relationship between the spectral properties such as spectral radius, second-largest eigenvalue, spectral gap, and spectral spread. This implies that these spectral properties are highly relevant and serve as significant predictors for the property of interest.
- 6. The smaller the standard error, the better the model's fit to the data. This indicates that the spectral properties are effectively capturing the underlying relationship with the properties.

8 Comparison with other correlation values for various descriptors

The correlation analysis presented in Table 11 reveals a comprehensive comparison between various topological indices and key physicochemical properties of molecular structures, highlighting the effectiveness of different descriptors in QSPR modeling. Among all the indices, such as the first Zagreb ZM_1 , second Zagreb ZM_2 index, ABCindex, Randić index R, Reciprocal Randić RR, sum-connectivity index SCI, Gourava indices GO_1 and GO_2 , GA index, Sombor index SO and its variants, maximum degree energy index M_{ij} , minimum degree energy index m_{ij} , the composite descriptors that integrate with entropy namely (SZ(G), I[SZ(G)]), (FZ(G), I[FZ(G)]), and

(A(G), I[A(G)]), exhibit the highest overall correlation values across most properties. Specifically, the energy-entropy descriptor (SZ(G), I[SZ(G))] demonstrates the strongest correlations with boiling point BP (0.974), critical temperature CT (0.965), molar volume MV (0.957), molar refraction MR (0.959), enthalpy of vaporization HV (0.942), and surface tension ST (0.942), underscoring its robustness as a predictive tool. Similarly, (FZ(G), I[FZ(G)]) show excellent performance, particularly in predicting properties like molar volume, molar refraction, boiling point and (A(G), I[A(G)]) is high in critical pressure, suggesting that these indices capture structural and energetic information. In contrast, individual descriptors such as R, SCI, and SO_{red} display generally weak or inconsistent correlations, with SO_{red} even showing negative values for molar volume and molar Refraction, indicating their limited standalone utility. Interestingly, while the SO index without entropy enhancement performs poorly for critical pressure (-0.133), it yields exceptionally high correlations with molar refractivity (0.983) and enthalpy of vaporization (0.975), suggesting a selective but potent predictive ability. Adding entropy consistently enhances the predictive strength of most indices, emphasizing the synergistic role of entropy-based measures in capturing molecular complexity and variability. Overall, the findings strongly support using composite energy-entropy descriptors, especially those based on the second Zagreb and first Zagreb matrices, as superior tools in the QSPR modeling of a wide range of molecular properties.

9 Conclusion

A comprehensive study of 72 alkane isomers starting from butane to nonane provided critical insights through statistical analysis by using linear and double regression models. Linear regression demonstrates that the graph energies effectively predict specific physicochemical properties of alkane isomers. Double regression models incorporating graph energy and its corresponding entropy provided a more nuanced understanding of the relationship between molecular descriptors and the physical properties. To ensure the robustness and predictive reliability of these models, a 10-fold crossvalidation was performed. The consistently low RMSE and MAE values through the cross-validation folds indicate strong model generalizability. Moreover, the high R^2 and F-Value, significantly small p-values confirmed a strong linear relationship between the molecular descriptors and the physical properties. An essential aspect of the study was the evaluation of multicollinearity among the descriptors. Variance inflation factor analysis reveals minimal multicollinearity for the first and second Zagreb energies, along with their corresponding entropies, except for the adjacency energy and sumconnectivity energy, which exhibited a higher VIF. This suggests that Zagreb energies and their corresponding entropy are statistically independent variables, whereas adjacency and sum-connectivity energies may contribute redundant information in some models. Furthermore, spectral properties such as spectral radius, second-largest eigenvalue, spectral gap, and spectral spread were identified as highly significant predictors for density and refractive index in alkane isomers. These findings highlight strong statistical evidence, high model accuracy, and valuable insights from both linear and double regression analysis. The integration of 10-fold cross-validation further supports the reliability of the models, while the multicollinearity assessment provides guidance on descriptor selection for future studies. This work suggests new avenues for further research with practical applications in chemical compound analysis, drug design, and materials science.

Funding. Not applicable.

Conflict of interest. The authors declare that they have no competing interests.

Ethics approval. Not applicable.

Consent to participate. Not applicable.

Consent for publication. The authors approve to publish in the journal.

References

- A. Arora, Hydrocarbons (Alkanes, Alkenes and Alkynes). Discovery Publishing House (2006)
- [2] A.T. Balaban, Applications of graph theory in chemistry. J. Chem. Inf. Comp. Sci. 25(3), 334–343 (1985)
- [3] M. Dehmer, Information processing in complex networks: Graph entropy and information functionals. Appl. Math. Comput. 201(1-2), 82-94 (2008)
- [4] R. Balakrishnan, K. Ranganathan, A textbook of Graph Theory. Springer Science & Business Media (2012)
- [5] I. Gutman, N. Trinajstić, Graph theory and molecular orbitals. Total φ -electron energy of alternant hydrocarbons. Chem. Phys. Lett. **17**(4), 535–538 (1972)
- [6] N.J. Rad, A. Jahanbani, I. Gutman, Zagreb energy and Zagreb Estrada index of graphs. MATCH Commun. Math. Comput. Chem. 79, 371–386 (2018)
- [7] B. Lučić, S. Nikolić, N. Trinajstić, B. Zhou, S. Ivaniš Turk, Sum-connectivity index. In: Gutman I, Furtula B (Eds.), Novel Molecular Structure Descriptors: Theory and Applications I, University of Kragujevac, 101–136 (2010)
- [8] D.H. Rouvray, C. Crafford, The dependence of physico-chemical properties on topological factors. South Afr. J. Sci. 72, 47–51 (1976)
- [9] E. Verbrugen, J. Tolls, Physicochemical properties of higher nonaromatic hydrocarbons: A literature study. J. Phys. Chem. Ref. Data 29(6), 1435–1446 (2000)
- [10] N.J.M.M. Raja, A Anuradha, Bonding alkane attributes with topological indices: a statistical intervention. J. Math. Chem. 62, 2889–2911 (2024)

- [11] S. Mondal, S. Barik, D. Nilanjan, A. Pal, A note on neighborhood first Zagreb energy and its significance as a molecular descriptor. Chemometr. Intell. Lab. Syst. 222, 104494 (2022)
- [12] S. Sripriya, A. Anuradha, On Zagreb energy of certain classes of graphs. Math. Stat. 12(4), 331–338 (2024)
- [13] K.N. Prakasha, P.S.K. Reddy, I.N. Cangul, Sum connectivity energy of graphs. Adv. Math. Sci. Appl. 28(1), 85–98 (2019)
- [14] S. Sripriya, A. Anuradha, Sum-connectivity energy on graph operations. Int. J. Appl. Math. 37(3), 293–304 (2024)
- [15] M. Dehmer, A. Mowshowitz, A history of graph entropy measures. Inf. Sci. 181(1), 57–78 (2011)
- [16] M. Dehmer, F. Emmert-Streib, Z. Chen, X. Li, Y. Shi, Mathematical Foundations and Applications of Graph Entropy. John Wiley & Sons (2017)
- [17] Y. Sun, H. Zhao, Eigenvalue-based entropy and spectrum of bipartite digraph. Complex Intell. Syst. 8(4), 3451–3462 (2022)
- [18] A. Altassan, B.A. Rather, M. Imran M, Inverse sum indeg index (energy) with applications to anticancer drugs. Mathematics **10**(24), 4749 (2022)
- [19] B.S. Bhadre, J.B. Veeragoudar, S.V.Patil, QSPR analysis on graph energies of chemical graphs. Eurasian Chem. Commun. 5, 246–263 (2023)
- [20] B.I. Andrew, A. Anuradha, Graph spectral analysis of nonane isomers. NanoSyst-Phys. Chem. Math. 15(1), 16–22 (2024)
- [21] B.I. Andrew, A. Anuradha, Analysing spectral parameters of decane—A graph theoretical perspective. Results Control Optim. 14, 100392 (2024)
- [22] K.A. Kumar, N.S. Basavarajappa, M.C. Shanmukha, QSPR analysis of alkanes with certain degree based topological indices. Malaya J. Matematik 8, 314–330 (2020)
- [23] J.O. Sharon, T.M. Rajalaxmi, S. Klavžar, R.S. Rajan, I. Rajasingh, Transmission in *H*-naphtalenic nanosheet. Appl. Math. Comput. 406, 126285 (2021)
- [24] R.N. Raj, R.S. Rajan, H. Ahmad, A comprehensive Analysis of partition dimensions in Efavirenz Abacavir Lamivudine Doravirine of anti-HIV Drug structures, Iraqi J. Comp. Sci. Math. 5(4), 100–107 (2024)
- [25] F. Bashir Farooq, N. Ul Hassan Awan, S. Parveen, N. Idrees, S. Kanwal, T.A. Abdelhaleem. Topological indices of novel drugs used in cardiovascular disease treatment and its QSPR modeling, J. Chem. **2022**, 9749575 (2022)

- [26] M. Randić, On characterisation of molecular branching. J. Amer. Chem. Soc. 97, 6609–6615 (1975)
- [27] M. Randić, Comparative structure property studies: Regression using a single descriptor. Croat. Chem. Acta 66, 289–312 (1993)
- [28] P. Preetha, M. Suresh, F.T. Tolasa, E. Bonyah, QSPR/QSAR study of antiviral drugs modeled as multigraphs by using TIs and MLR method to treat COVID 19 disease. Sci. Reports 14, 13150 (2024)
- [29] M.C. Shanmukha, N.S. Basavarajappa, K.C. Shilpa, A. Usha, Degree-based topological indices on anticancer drugs with QSPR analysis. Heliyon 6(6), e04235 (2020)
- [30] CSID:217, https://www.chemspider.com/
- [31] NIST Chemistry WebBook, https://webbook.nist.gov/chemistry/
- [32] Pubchem-Online Resource, https://pubchem.ncbi.nlm.nih.gov/
- [33] Y.Dong and C. Stijn, On the main distance-based entropies: the eccentricity-and Wiener-entropy. arXiv preprint arXiv:2208.12209 (2022)