

Predictive Modeling of Physicochemical Properties of *Cissus quadrangularis* Compounds Using Topological Indices

Ayush Agarwal ¹, Anuradha D S ², Jaganathan B ^{2,*}

¹ Department of Computer Science, Vellore Institute of Technology, Chennai, India; ayush.shyamsunder2022@vitstudent.ac.in;

² Department of Mathematics, Vellore Institute of Technology, Chennai, India; anu.radha2020@vitstudent.ac.in;

³ Mathematics Division, School of Advanced Sciences, Vellore Institute of Technology, Chennai, India; jaganathan.b@vit.ac.in;

* Correspondence: jaganathan.b@vit.ac.in;

Received: 6.01.2025; Accepted: 17.07.2025; Published: 30.09.2025

Abstract: *Cissus quadrangularis* (CQ) is a medicinal plant with extensive pharmaceutical applications. The intricate phytochemical composition of CQ presents challenges in isolating active compounds, determining their physicochemical properties, and calculating structure-activity relationships, bioavailability, and extract standardization. The accurate prediction of properties is essential to address these challenges, necessitating sophisticated modeling techniques. This study concentrates on predictive modeling of the physicochemical properties of this species. The six characteristics examined were boiling point (BP), melting point (MP), molar volume, polarizability, flash point, and IC₅₀, which are crucial for evaluating compound stability and bioactivity. Molecular descriptors and topological indices were used for the predictions. Models such as linear regression (LR), multiple linear regression (MLR), random forest regression (RFR), and AdaBoost have been developed. These findings indicate that the properties correlate with specific indices, thereby facilitating precise predictions. This modeling approach offers an efficient means of anticipating properties without experimental methods.

Keywords: QSAR; *Cissus quadrangularis*; regression model; degree-based topological indices.

© 2025 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The authors retain copyright of their work, and no permission is required from the authors or the publisher to reuse or distribute this article, as long as proper attribution is given to the original source.

1. Introduction

Cissus quadrangularis, commonly known as the "Veldt Grape," is a medicinal plant widely used in traditional medicine for its anti-inflammatory, bone-healing, and antioxidant properties [1,2]. Furthermore, while *Cissus* extracts show promise in various applications, such as bone regeneration and weight management, more studies are needed to establish their clinical efficacy and safety profile [3].

Cissus quadrangularis (CQ) presents several challenges in drug design despite its promising therapeutic potential. One major issue is the lack of standardization in extract preparation, which leads to variability in bioactive compound concentrations across studies. This inconsistency makes it difficult to establish optimal dosages and compare results across different research efforts. Another challenge is the limited understanding of the specific physiological effects of individual chemical constituents isolated from *Cissus* extracts.

Although numerous compounds have been identified, including steroids, flavonoids, and triterpenes, their precise mechanisms of action remain unclear. The bioavailability and stability of *Cissus*-derived compounds pose challenges in drug formulation. For instance, clumping of particles observed at higher concentrations of *Cissus* extracts can reduce the active surface area and potentially decrease the efficacy [4].

Given the cost and resource intensity associated with experimental methods, computational approaches provide a viable alternative for efficiently predicting these properties. The combination of topological indices and QSAR modeling can be a powerful tool for addressing drug design challenges. These methods can help researchers predict the properties and activities of compounds, guide the selection of promising candidates for further development, and ultimately streamline the drug discovery process [5]. By leveraging these computational techniques, researchers can identify novel therapeutic agents more efficiently and cost-effectively [6].

QSAR models, are widely applied in drug design, as they link biological activity to structural and physicochemical descriptors, including TIs, which play a crucial role in drug discovery research [7,8]. By applying QSAR techniques to compounds, researchers can predict their potential biological activities and physicochemical properties without the need for extensive experimental testing. This approach can significantly accelerate the drug design process and help to identify promising lead compounds. TIs can help researchers understand physical properties and chemical compounds using the topology of their chemical bonding between atoms [7,9–11]. TIs, derived from molecular structures, offer powerful molecular descriptors in cheminformatics as they capture essential aspects of compound structures that relate to physical and chemical behaviors [12].

Topological indices have been applied beyond the traditional pharmaceutical chemistry. They have been used to study the growth of microorganisms in biological networks, potentially aiding vaccine design and discovery. Degree-based topological indices have been used in the analysis of heart attack drugs and cancer drug research [13–17]. Pylonephritis drugs [18], anti-hepatitis drugs [19], and anti-psychotic drugs [20] have applications in treating lung [21] and eye disorders [22].

These indices have been applied to analyze potential antiviral drugs for COVID-19, demonstrating strong correlations with the physicochemical properties and biological activities of these compounds. The quantitative structure-property relationship (QSPR) analysis explored the connection between edge vertex, vertex edge degree based TIs and the observed physicochemical properties of phytochemicals tested against SARS-CoV-2 3CLpro. The findings indicated that the edge vertex, vertex edge degree based Zagreb indices, beta Zagreb index $M1\beta$ ve are noteworthy TIs that can successfully forecast the topological polar surface area and molecular weight of these phytochemicals [23].

The existing literature does not offer an extensive list of phytochemicals analyzed using topological indices in recent research articles, highlighting the growing importance of topological indices in various areas of pharmaceutical and chemical research, including the analysis of phytochemicals against SARS-CoV-2 [24]. A recent study presented a novel approach for analyzing *Azadirachta indica* (neem) compounds using topological indices and regression models. By employing TIs, this study offers potential acceleration of drug discovery processes, enhanced virtual screening efficiency, and a means to bridge geographical gaps in neem research. This methodology combines regression analysis with multi-criteria decision-making, allowing for a comprehensive evaluation of neem compounds considering multiple

properties simultaneously. While focusing on neem, this approach could potentially be applied to other phytochemicals, broadening its impact on natural-product-based drug discovery [25].

As the field evolves, mathematical descriptors are likely to play an increasingly important role in understanding and predicting the behavior of phytochemicals and other biologically active compounds [26]. Despite their potential, few studies have systematically applied these indices to property prediction in plant-based bioactive compounds, creating an opportunity to assess their predictive power for compounds derived from CQ [27].

In this study, several degree-based indices, such as the atom-bond connectivity index, forgotten index, and Zagreb indices, are used for prediction modelling. These indices reflect structural features, such as molecular branching, connectivity, and shape, and hence are used for predictive modelling of CQ bioactive compounds [28]. Physicochemical properties, such as Boiling Point (BP), Melting Point (MP), Molar Volume, Polarizability, Flash Point, and IC50, are crucial for determining the stability, reactivity, and biological activity of a compound. The key compound properties were estimated through a computationally efficient approach by establishing correlations and constructing accurate predictive models using regression analysis [29–31]. This predictive modelling can advance the understanding of CQ compounds and enhance early-stage assessments in drug discovery and development.

2. Materials and Methods

In this study, various regression techniques were used to analyze the relationship between topological indices and key physical properties of compounds extracted from *Cissus quadrangularis* (CQ). Four distinct regression methodologies were employed: Linear regression (LR), multiple linear regression (MLR), random forest regression (RFR), and AdaBoost, each with its specific strengths in capturing the relationships between molecular descriptors and physicochemical properties. The regression models, TIs, and CQ extracts used in this study are described in the following sections.

2.1. Topological indices.

The skeletal structure of a molecule represents a simplified, two-dimensional depiction of its molecular framework, displaying only the arrangement of atoms and the bonds between them, without including hydrogen atoms bonded to carbon[32–35]. This structural representation is particularly useful for calculating topological indices because it highlights the connectivity and branching patterns of atoms in a molecule, which are essential for deriving molecular descriptors that correlate with physical and chemical properties[36].

A graph $G(V, E)$ with vertex set $V(G)$ and edge set $E(G)$ is considered connected if there is a path linking every pair of vertices in G [37–40]. The distance between two vertices u and v , written as $d(u, v) = d_G(u, v)$ is the length of the shortest path between the vertices in the graph G [12]. The degree of a vertex v in G , denoted $\delta_v(G)$ or simply d_v , when unambiguous, is the count of vertices in G that are adjacent to v . This concept of degree is analogous to the notion of valence in chemistry [30,41]. The indices used were defined as follows:

ABC index, proposed by Estrada as[42]:

$$ABC(G) = \sum_{e=uv \in E(G)} \sqrt{\frac{\delta_u + \delta_v - 2}{\delta_u \delta_v}} \quad (1)$$

The Randić index, proposed by Milan Randić as[43]:

$$R(G) = \sum_{e=uv \in E(G)} \sqrt{\frac{1}{\delta_u \delta_v}} \quad (2)$$

The sum connectivity index[44], proposed by Zhou and Trinajstić as:

$$S(G) = \sum_{e=uv \in E(G)} \sqrt{\frac{1}{\delta_u + \delta_v}} \quad (3)$$

The GA index, proposed by Vukicević et al., as [37,45,46]:

$$GA(G) = \sum_{e=uv \in E(G)} \frac{2\sqrt{\delta_u \delta_v}}{\delta_u + \delta_v} \quad (4)$$

The first and second Zagreb indices, proposed by Gutman and Trinajstić as [45,47]:

$$M_1(G) = \sum_{e=uv \in E(G)} (\delta_u + \delta_v) \quad (5)$$

$$M_2(G) = \sum_{e=uv \in E(G)} (\delta_u \delta_v) \quad (6)$$

The Harmonic index, proposed by Fajtlowicz et al., as[48]:

$$H(G) = \sum_{e=uv \in E(G)} \frac{2}{\delta_u + \delta_v} \quad (7)$$

The Hyper Zagreb index[47], proposed by Shirdel et al. as:

$$HM(G) = \sum_{e=uv \in E(G)} (\delta_u + \delta_v)^2 \quad (8)$$

The third Zagreb index[46], proposed by Fath-Tabar et al. as:

$$ZG_3(G) = \sum_{e=uv \in E(G)} |\delta_u - \delta_v| \quad (9)$$

The forgotten index[38], proposed by Furtula et al. as:

$$F(G) = \sum_{e=uv \in E(G)} [(\delta_u)^2 + (\delta_v)^2] \quad (10)$$

The symmetric division index[49], proposed as:

$$SSD(G) = \sum_{e=uv \in E(G)} \left[\frac{P}{Q} + \frac{Q}{P} \right] \quad (11)$$

where, $P = \min(\delta_u, \delta_v)$ and $Q = \max(\delta_u, \delta_v)$.

2.2. *Cissus quadrangularis* extracts.

CQ contains phytoconstituents, such as carbohydrates, phytosterols, flavonoids, triterpenoids, glycosides, tannins, saponins, proteins, Vitamin C, alkaloids, and calcium. These compounds have various pharmacological properties, including antimicrobial, antiulcer, anti-inflammatory, antitumor, antihemorrhagic, antiallergic, antidiabetic, and antioxidant effects [50]. Notably, CQ plays a significant role in bone healing, fracture repair, and regeneration [51].

Different CQ extracts exhibited varying levels of bioactivity. The ethyl acetate fraction of CQ stem extracts showed the highest antioxidant activity, which was attributed to the presence of sterols, vitamin C, and tannins. The ethanolic extracts of CQ contain campesterol, β -sitosterol, stigmasterol, pterostilbene, resveratrol, and various flavonoids, which may contribute to its therapeutic effects [52]. CQ's diverse bioactive compounds make it a versatile medicinal plant with applications in treating various conditions, particularly in bone health and regeneration.

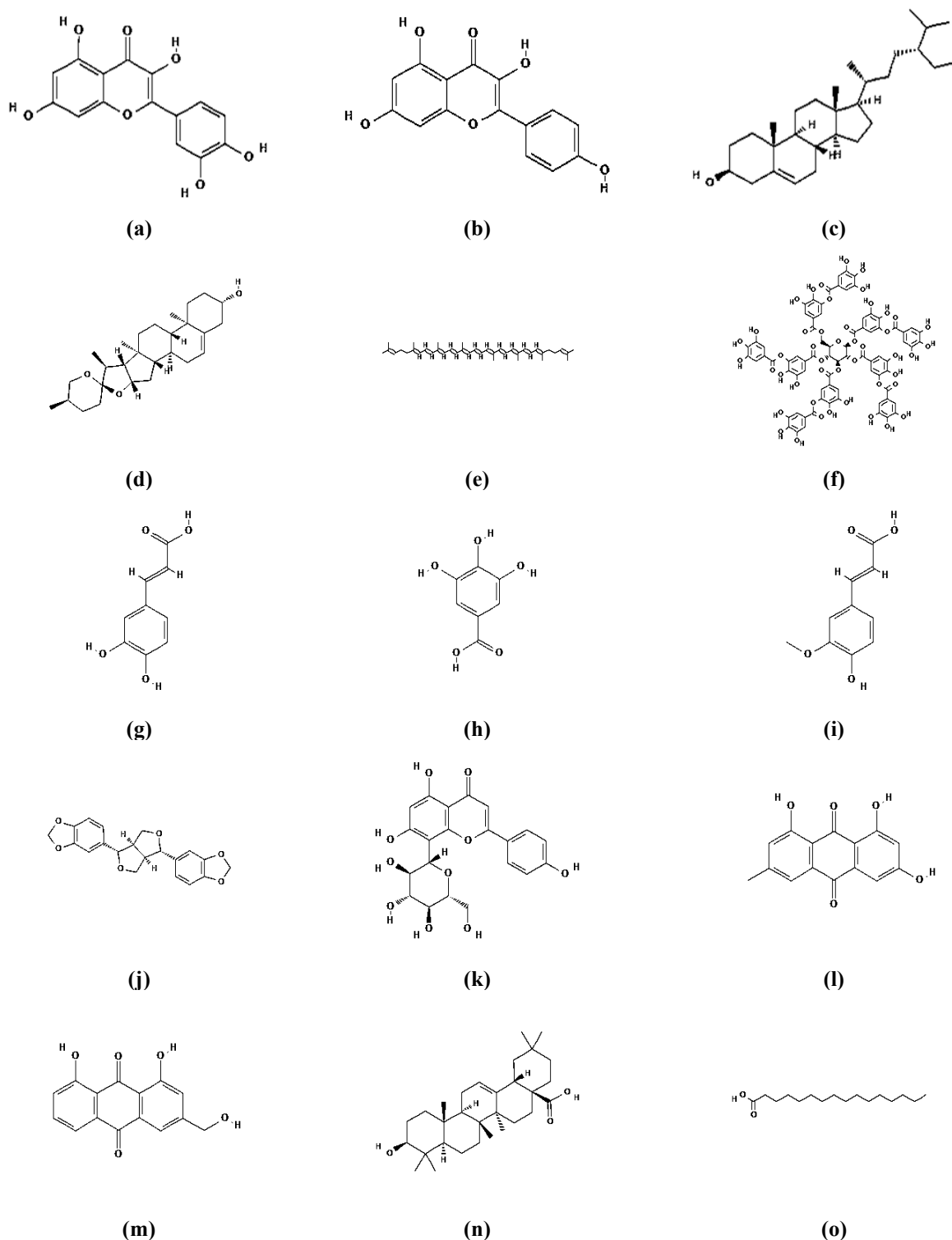


Figure 1. Molecular Structure of CQ extracts of (a) Quercetin; (b) Kaempferol; (c) Beta-sitosterol; (d) Diosgenin; (e) Lycopene; (f) Gallotannins; (g) Caffeic acid; (h) Gallic acid; (i) Ferulic acid; (j) Sesamin; (k) Vitexin; (l) Emodin; (m) Aloe-emodin; (n) Oleanolic acid; (o) Palmitic acid;

Their antioxidant, antimicrobial, and anti-inflammatory properties further enhance their therapeutic potential. Recent research has also explored the use of CQ in green synthesis of nanoparticles and bone scaffolds, opening up new avenues for biomedical applications [53]. The most effective extracts are considered for the study. The corresponding molecular structures obtained from PubChem are shown in Figure 1.

The physicochemical properties of CQ extracts taken from the globally recognized databases ChemSpider and PubChem are presented in Table 1.

Table 1. Physico-chemical properties of CQ extracts.

Phytochemical	BP (°C)	MP (°C)	Flash point (°C)	Molar volume (cm ³)	Polarizability (cm ³)	IC ₅₀ (nM)
Quercetin	642.4	316	248.1	168	29.1	10
Kaempferol	582.1	275	226.1	169.5	28.3	28
Beta-sitosterol	501.9	140	220.4	424.3	51.2	2500
Diosgenin	527.1	203	272.6	366.9	47.3	6960
Lycopene	660.9	175	350.7	604.2	74.1	5800
Gallo tannins	1103.6	200	363	320.4	55.4	1000
Caffeic acid	416.8	225	220	121.9	18.8	2
Gallic acid	501.00	258	271	97.3	15.4	10000
Ferulic acid	372.3	168	150.5	147.5	20.7	6440
Sesamin	504.4	123	212.3	255.8	35.7	7100
Vitexin	767.7	256	273.1	256.3	41.1	72
Emodin	586.9	256	322.8	170.6	27.4	240
Aloe-emodin	568.8	222	311.9	169.7	27.3	1000
Oleanolic acid	553.5	310	302.6	414.9	53	90
Palmitic acid	351.5	61.8	192	287.3	30.8	2600

The physicochemical properties of the predicted CQ extracts calculated using well-established degree-based indices are presented in Table 2.

Table 2. TIs of CQ extracts.

CQ extracts	ABC	R	S	GA	H	ZG ₃	SSD	M ₁	M ₂	F	HM
Quercetin	17.34	10.38	10.83	22.99	9.83	22	57.67	120	145	324	614
Kaempferol	16.57	9.97	10.42	22.13	9.50	20	54.33	114	137	304	578
Beta-sitosterol	24.55	14.70	15.30	32.43	13.91	37	173.00	173	212	491	915
Diosgenin	24.92	14.27	15.38	33.64	13.63	38	83.25	186	239	544	1022
Lycopene	28.67	18.99	18.76	37.38	18.10	34	93.67	170	174	402	750
Gallo tannins	95.79	57.35	59.61	126.06	54.13	132	321.00	658	787	1776	3350
Caffeic acid	9.59	6.09	6.09	12.34	5.73	14	32.33	60	65	152	282
Gallic acid	8.91	5.52	5.51	11.25	5.10	14	31.33	58	66	156	288
Ferulic acid	10.19	6.63	6.62	13.40	6.30	14	33.67	64	70	160	300
Sesamin	21.64	12.87	14.01	30.68	12.73	16	64.67	154	191	398	780
Vitexin	24.36	14.72	15.38	32.74	14.03	28	80.00	170	208	458	874
Emodin	15.89	9.40	9.84	21.03	8.87	20	53.33	112	138.	308	584
Aloe-emodin	15.07	9.00	9.45	20.21	8.57	16	49.67	106	131	288	550
Oleanolic acid	26.98	15.14	16.01	34.61	14.00	54	98.75	204	264	644	1172
Palmitic acid	12.24	8.77	8.52	16.65	8.57	6	37.33	68	66	142	274

2.3. Regression models.

Regression models play a crucial role in QSAR and drug design by enabling the prediction of the biological activities and properties of molecules based on their structural features. These models help to establish quantitative relationships between molecular descriptors and target properties, facilitating drug discovery. In QSAR modeling, regression techniques, such as multiple linear regression and machine learning approaches, are commonly

employed to develop predictive models. These models can prioritize candidate molecules for laboratory experiments, thereby saving time and resources in the drug discovery pipeline [54].

For instance, ensemble machine learning techniques, such as Random Forest and AdaBoost, have shown exceptional predictive accuracy across diverse datasets and target properties in QSAR modeling [5]. Interestingly, the integration of topological descriptors and chemical graph theory in regression models has proven significant in QSAR studies, allowing researchers to correlate chemical structures with physical properties and biological activity [55].

The use of cloud computing resources has enhanced computational capabilities for handling large amounts of chemical data, leading to more accurate and generalizable QSAR models. Regression models in QSAR and drug design provide valuable insights into structure-activity relationships, aid in the prediction of bioactivities of new compounds, and help elucidate potential molecular mechanisms of receptor-ligand interactions. These models streamline the drug development process and contribute to the advancement of telemedicine applications by enabling rapid and accurate prediction of compound properties [56].

A good regression model is typically evaluated using several statistical metrics, each of which provides unique insights into its performance. The coefficient of determination (R^2) is widely considered a standard metric for measuring the proportion of variance in the dependent variable explained by the model. The mean squared error (MSE) and its root variant, root mean squared error (RMSE), quantify the average squared difference between predicted and actual values [57,58]. These metrics provide valuable information that can be used to interpret specific contexts and goals.

Hence, the performance of the models was evaluated using statistical metrics. R -squared (R^2) measures the proportion of variance in the response variable explained by the model, providing insight into its overall fit; mean squared error (MSE) captures the average of squared differences between predicted and observed values; and root mean squared error (RMSE) provides a measure in the same units as the response variable, making interpretation more intuitive. Using these metrics, we assessed the model accuracy and determined the topological indices that contributed most effectively to predicting the physical properties of CQ compounds. For each property, the performance of all four regression techniques was used to identify the most suitable modeling approach. This facilitates the identification of indices having the strongest correlation with properties such as melting point (MP), boiling point (BP), flash point (FP), polarizability (P), molar volume (MV), and IC_{50} .

Linear Regression (LR) and Multiple Linear Regression (MLR) models were used to analyze the relationship between the selected topological indices and key physical properties of compounds extracted from CQ. LR and MLR are statistical techniques commonly used for predictive modeling, where LR establishes a linear relationship between a single predictor variable and a response variable, whereas MLR extends this to multiple predictors, capturing more complex associations [59,60].

Linear Regression (LR) was used to assess the individual predictive power of each topological index for a given property. MLR was then applied to incorporate multiple indices simultaneously, aiming to improve the prediction accuracy by accounting for the combined influence of indices on each property. Using both LR and MLR enables the determination of individual and collective contributions of topological indices, enhancing the ability to accurately predict physical properties from the molecular structure.

Eleven well-established degree-based topological indices have been utilized to predict nine physical properties: boiling point (BP), melting point (MP), flash point (F), molar volume (MV), polarizability (P), and IC50(nM) for 15 medicinal compounds extracted from CQ, as shown in Figure 1, which are used to improve bone health.

A linear regression model is of the form.

$$P = a + \{b \times TI\} \quad (12)$$

where P represents the physical properties of the drug, a is a constant, b is the regression coefficient, and TI denotes the topological index. The constant ‘ a ’ and the regression coefficient ‘ b ’ were computed using Python Libraries for seven physical properties and eleven degree-based topological indices of the molecular structures of the 14 drugs. Based on equation (1), the following are the linear regression models derived for the specified degree-based topological indices:

Multiple linear regression (MLR) was then applied to incorporate multiple indices simultaneously, aiming to improve the prediction accuracy by accounting for the combined influence of indices on each property.

A multilinear regression model is of the form.

$$P = [a_i] + \sum_{i=1}^n [b_i \times TI_i] \quad (13)$$

where P represents the physical properties of the drug, a_i is a constant, b_i is the regression coefficient, and TI_i denotes the respective topological index.

Random Forest Regression (RFR) was implemented as an ensemble learning method that constructs multiple decision trees during training and outputs the average prediction of individual trees for regression tasks. RFR handles nonlinear relationships effectively, capturing complex interactions between topological indices and properties that linear models might miss. The algorithm is less prone to overfitting than individual decision trees, resulting in a better generalization to unseen compounds. It provides built-in feature importance metrics, allowing the identification of the most influential topological indices for each property prediction. RFR manages high-dimensional data efficiently, accommodating the multiple topological indices used in our study. Our implementation of RFR involved optimizing key hyperparameters, including the number of trees in the forest, maximum tree depth, minimum number of samples required to split a node, and minimum number of samples required at a leaf node. Cross-validation was employed to determine the optimal hyperparameter settings for each of the predicted physicochemical properties [61].

For the CQ compound analysis, Random Forest Regression with optimized hyperparameters was implemented for each property prediction. The number of trees varied between 100 and 500, depending on the specific property being predicted, with the maximum depth ranging from 10 to 30 to prevent overfitting while maintaining high predictive accuracy. Feature importance analysis from the Random Forest models provided valuable insights into which topological indices were the most influential in predicting each physicochemical property[62].

Random Forest Regression is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of individual trees for regression tasks. The model can be expressed as:

$$P = \frac{1}{N} \sum_{i=1}^N (t_i) \times (TI_1, TI_2, TI_3, \dots, TI_m) \quad (14)$$

where P is the predicted property value, N is the number of trees in the forest, t_i represents the prediction of the i -th tree, and $(TI_1, TI_2, TI_3, \dots, TI_m)$ are the m topological indices used as input features [42].

Adaptive Boosting (AdaBoost) regression was employed as a meta-learning algorithm that combines multiple weak learners to create a strong predictive model. For our implementation, we used decision trees as base learners. Sequential training of weak learners, with each subsequent model focusing on instances where previous models performed poorly. Weighting training instances to emphasize difficult-to-predict compounds. Combining predictions through a weighted sum of outputs from individual weak learners. AdaBoost has proven particularly effective for datasets with complex relationships between molecular descriptors and properties because it can capture intricate patterns that simpler models might miss. For our CQ compound analysis, AdaBoost was configured with optimized parameters, including the maximum number of estimators, learning rate, and loss function [63].

AdaBoost Regression combines multiple weak learners (typically decision trees with a limited depth) to form a strong predictor. In our implementation, each weak learner was trained sequentially, with higher weights assigned to instances that previous learners predicted poorly. The final prediction is the weighted sum of all the weak learners' predictions.

$$P = \sum_{i=1}^M \alpha_i h_i(TI_1, TI_2, TI_3 \dots TI_m) \quad (15)$$

where P is the predicted property value, M is the number of weak learners, α_i is the weight assigned to the i -th learner (determined during training based on its accuracy), h_i is the prediction of the i -th weak learner, and $(TI_1, TI_2, TI_3 \dots TI_m)$ are the topological indices are used as features. The AdaBoost implementation utilizes decision trees with a maximum depth of three as base learners. The number of estimators was set between 50 and 200 based on the cross-validation performance for each property, with a learning rate of 0.01-0.1 to control the contribution of each weak learner to the final ensemble model [64].

3. Results and Discussion

Statistical studies have covered areas such as antioxidant activity, antimicrobial properties, bone healing, weight management, and various phytochemical constituents [65]. As per the literature, studies that provide insight into physicochemical property predictions have not yet been explored. Hence, the aforementioned regression analysis could potentially be used to explore the relationships between extract concentrations, phytochemical compositions, and biological effects. The statistical metrics obtained from this analysis are discussed in this section.

Correlation values play a crucial role in regression analysis, providing insights into the relationships between variables and guiding model interpretation. These correlation measures are essential for understanding the underlying structure of the data and can inform variable selection in regression models. Pearson's correlation coefficient is commonly used to analyze continuous data. The coefficients range from -1 to +1, with 0 indicating no association and values closer to ± 1 suggesting stronger relationships. The correlation between the properties and indices is shown in Table 3.

Table 3. Correlation between properties and indices.

Property	ABC	R	S	GA	M ₁	M ₂	H	HM	ZG ₃	F	SSD
BP	0.87	0.87	0.87	0.87	0.87	0.84	0.79	0.87	0.87	0.86	0.86
MP	-0.03	-0.06	-0.05	-0.04	-0.07	0.10	-0.09	0.00	0.03	0.04	0.04
F	0.57	0.57	0.57	0.56	0.56	0.59	0.49	0.57	0.56	0.57	0.57
MV	0.37	0.38	0.37	0.36	0.38	0.38	0.45	0.35	0.34	0.35	0.34
P	0.57	0.58	0.57	0.57	0.58	0.58	0.61	0.56	0.55	0.56	0.55
IC50	-0.17	-0.16	-0.16	-0.17	-0.16	-0.20	-0.18	-0.18	-0.19	-0.20	-0.20

The correlations ≥ 0.6 exhibit good predictive ability. These data are presented pictorially in Figure 2 for better understanding.

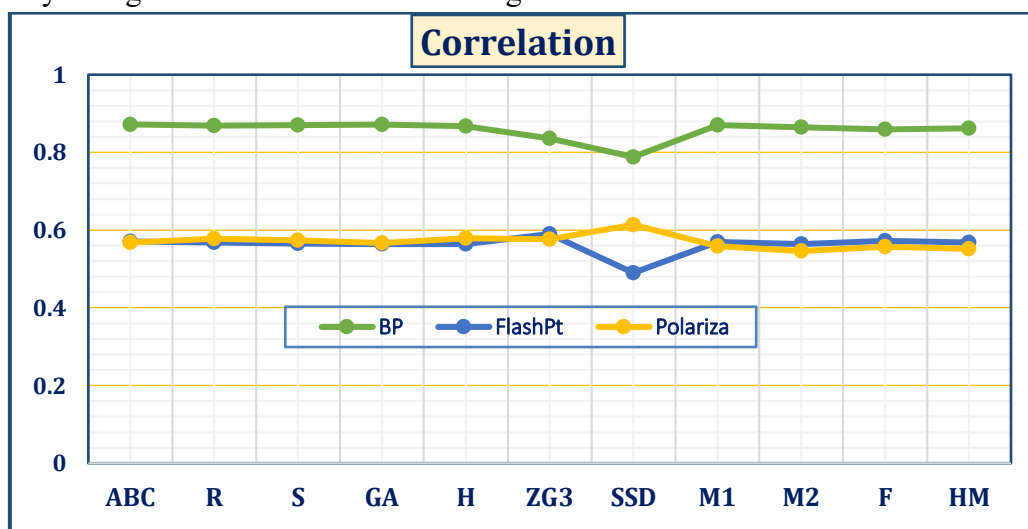


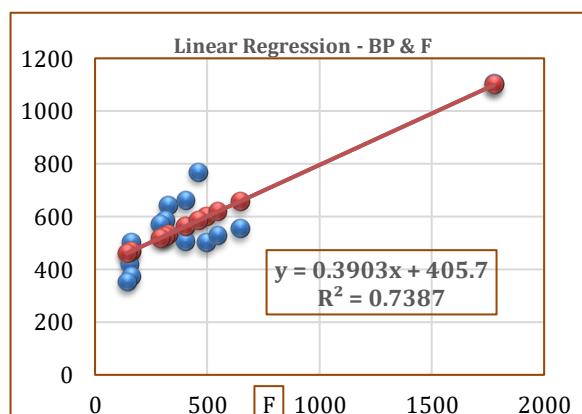
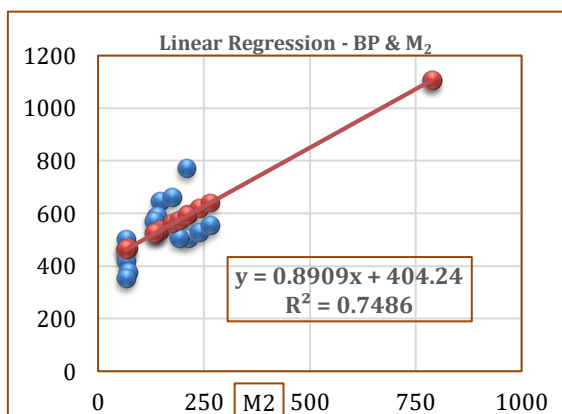
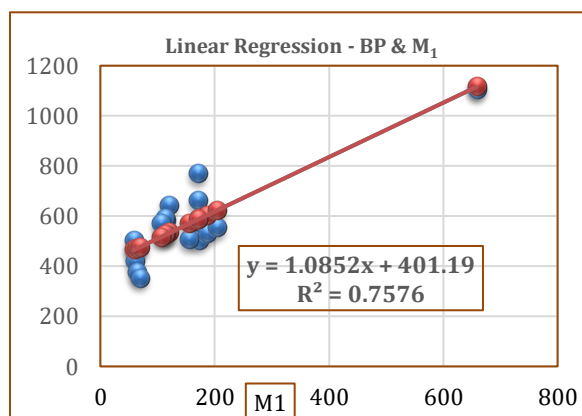
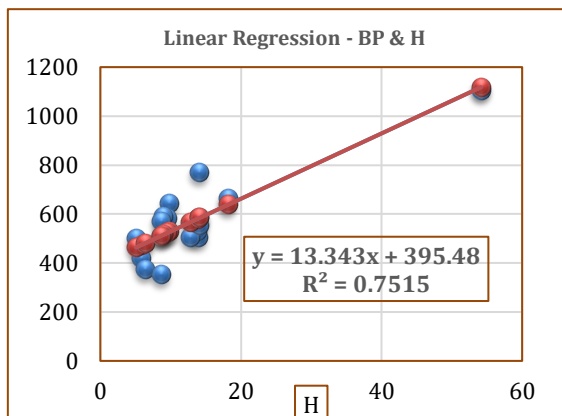
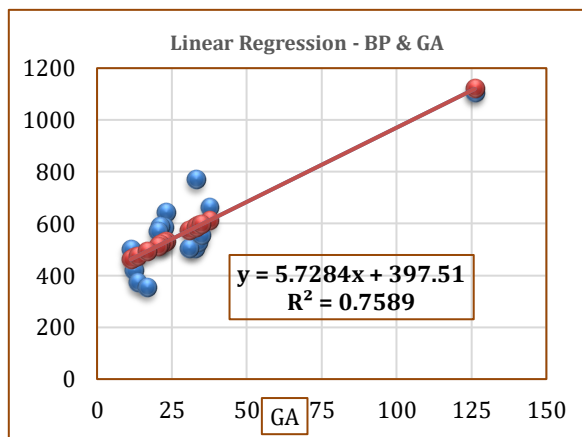
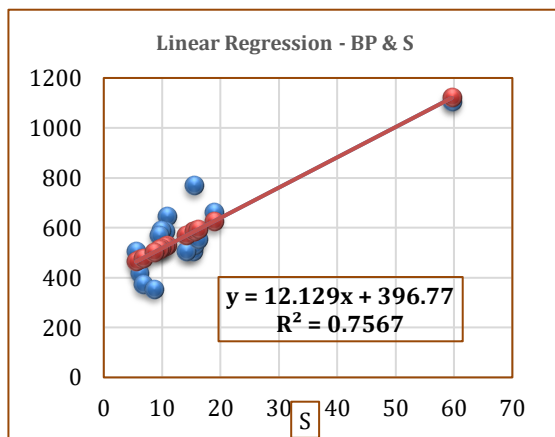
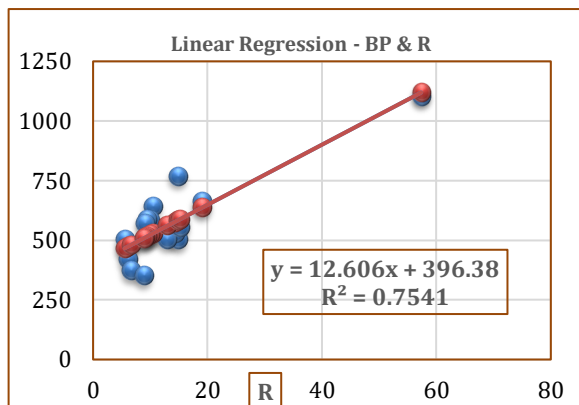
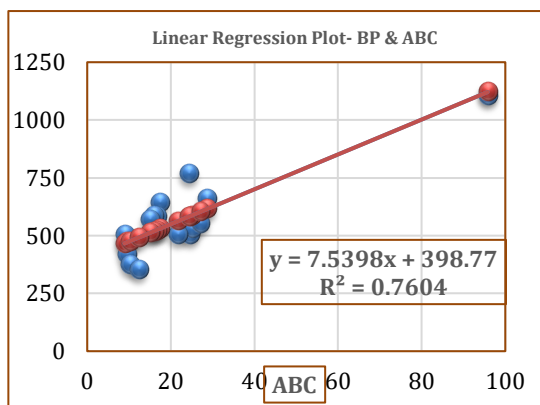
Figure 2. Correlation graphs between properties and indices of CQ extracts.

To evaluate the performance of the models, we used the R-squared (R^2), mean squared error (MSE), and root mean squared error (RMSE) as evaluation metrics. The R-squared measures the proportion of variance in the response variable explained by the model, providing insight into its overall fit. MSE and RMSE quantify the model’s prediction errors, with MSE capturing the average of squared differences between predicted and observed values, and RMSE providing a measure in the same units as the response variable, making the interpretation more intuitive. Using these metrics, we could assess the model accuracy and determine which contributed most effectively to predicting physical properties. The statistical parameters of the linear regression models are listed in Table 4.

Table 4. Statistical parameters of CQ - linear regression.

BP	a	b	R ²	MSE	RMSE
ABC(G)	7.54	398.77	0.76	7399.14	86.02
R(G)	12.61	396.38	0.75	7595.07	87.15
S(G)	12.13	396.75	0.76	7513.25	86.68
GA(G)	5.73	397.50	0.76	7447.74	86.30
H(G)	13.34	395.49	0.75	7675.54	87.61
SSD(G)	0.39	405.70	0.74	8070.75	89.84
ZG3(G)	1.91	414.81	0.62	11691.12	108.13
M1(G)	426.04	4.02	0.76	7512.14	86.67
M2(G)	1.09	401.19	0.76	7488.32	86.54
F(G)	0.89	404.24	0.75	7764.53	88.12
HM(G)	0.21	404.94	0.74	7917.30	88.98

The linear equation derived from linear regression is represented in Figure 3.



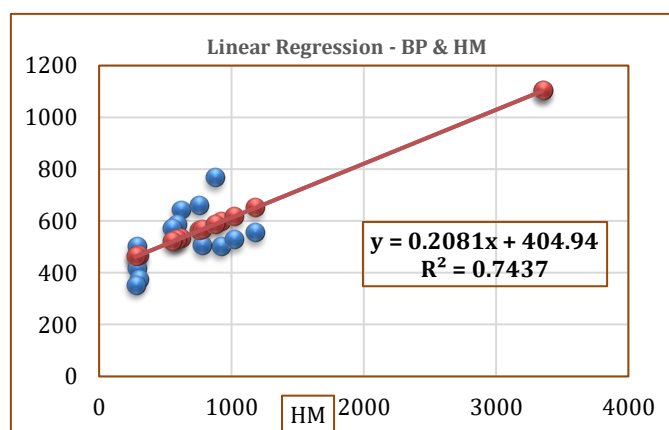


Figure 3. Linear regression of physicochemical properties of CQ.

In the regression analysis of molecular structures, the actual and predicted values play a crucial role in assessing the performance and accuracy of the model. Actual values represent the experimentally observed or measured properties of molecules, while the predicted values are those estimated by the regression model based on molecular descriptors. The difference between actual and predicted values helps evaluate the model's accuracy and predictive power, and different statistical methods can yield varying levels of accuracy. The actual and predicted values for each property are presented in the following sections.

The data of the actual physicochemical properties and the corresponding predictions from the linear regression analysis are listed in Tables 5 to 7.

Table 5. Comparison between actual BP and predicted BP from indices.

Phyto chemical	Actual Bp	ABC	R	S	Ga	H	SSD	ZG ₃	M1	M2	F	Hm
Quercetin	642.40	529.54	527.23	528.11	529.22	526.69	525.16	531.32	531.42	533.42	532.16	532.73
Kaempferol	582.10	523.69	522.05	523.16	524.26	522.24	518.78	521.37	524.91	526.29	524.36	525.24
Beta-sitosterol	501.90	583.85	581.72	582.33	583.27	581.07	745.86	605.89	588.94	593.11	597.35	595.37
Diosgenin	527.10	586.70	576.24	583.28	590.20	577.32	574.11	610.86	603.05	617.16	618.03	617.64
Lycopene	660.90	614.94	635.76	624.32	611.62	636.98	594.05	590.97	585.68	559.25	562.61	561.03
Gallotannins	1103.60	1121.02	1119.36	1119.78	1119.64	1117.76	1029.06	1078.18	1115.28	1105.36	1098.90	1102.15
Caffeic acid	416.80	471.07	473.18	470.64	468.21	471.98	476.68	491.54	466.31	462.15	465.03	463.63
Gallic acid	501.00	465.95	465.96	463.63	461.94	463.53	474.77	491.54	464.14	463.04	466.59	464.88
Ferulic acid	372.30	475.58	479.96	477.00	474.26	479.54	479.23	491.54	470.65	466.60	468.15	467.38
Sesamin	504.40	561.91	558.56	566.72	573.23	565.38	538.55	501.49	568.32	574.40	561.05	567.28
Vitexin	767.70	582.46	581.97	583.33	585.07	582.72	567.90	561.15	585.68	589.54	584.46	586.84
Emodin	586.90	518.57	514.84	516.15	518.00	513.79	516.87	521.37	522.74	527.18	525.92	526.49
Aloe-emodin	568.80	512.41	509.87	511.37	513.27	509.79	509.85	501.49	516.23	520.94	518.11	519.41
Oleanolic acid	553.50	602.16	587.25	590.94	595.79	582.31	603.77	690.41	622.58	639.43	657.06	648.86
Palmitic acid	351.50	491.05	506.94	500.15	492.91	509.79	486.25	451.77	474.99	463.04	461.13	461.97

Table 6. Comparison between actual FP and predicted FP from indices.

Phyto Chemical	Actual FP	ABC	R	S	GA	H	SSD	ZG ₃	M1	M2	F	HM
Quercetin	248.1	252.37	251.88	252.13	252.42	251.83	251.98	252.00	252.79	253.25	252.78	253.00
Kaempferol	226.1	251.10	250.76	251.06	251.36	250.87	250.66	249.67	251.38	251.71	251.06	251.36
Beta-sitosterol	220.4	264.17	263.70	263.83	264.02	263.55	297.48	269.46	265.27	266.16	267.17	266.70
Diosgenin	272.6	264.78	262.51	264.03	265.51	262.75	266.18	270.62	268.32	271.37	271.74	271.57
Lycopene	350.7	270.92	275.42	272.88	270.11	275.61	266.18	265.97	264.56	258.84	259.50	259.19
Gallotannins	363	380.86	380.34	379.76	379.15	379.26	355.87	380.04	379.39	376.99	377.94	377.54
Caffeic acid	220	239.66	240.15	239.73	239.32	240.04	241.98	242.68	238.68	237.83	237.95	237.88
Gallic acid	271	238.55	238.59	238.22	237.98	238.21	241.59	242.68	238.20	238.02	238.30	238.16
Ferulic acid	150.5	240.64	241.62	241.11	240.62	241.67	242.51	242.68	239.62	238.79	238.64	238.70
Sesamin	212.3	259.40	258.68	260.46	261.87	260.17	254.74	245.01	260.79	262.11	259.16	260.55

Phyto Chemical	Actual FP	ABC	R	S	GA	H	SSD	ZG3	M1	M2	F	HM
Vitexin	273.1	263.86	263.76	264.04	264.41	263.91	260.79	258.98	264.56	265.39	264.33	264.83
Emodin	322.8	249.98	249.19	249.55	250.01	249.05	250.27	249.67	250.91	251.90	251.40	251.63
Aloe-emodin	311.9	248.65	248.11	248.52	249.00	248.19	248.82	245.01	249.50	250.55	249.68	250.08
Oleanolic acid	302.6	268.14	264.90	265.68	266.71	263.82	268.19	289.25	272.56	276.18	280.36	278.40
Palmitic acid	192	244.01	247.48	246.10	244.63	248.19	243.96	233.37	240.56	238.02	237.09	237.52

Table 7. Comparison between actual P and predicted P from indices.

Phyto chemical	Actual P	ABC	R	S	GA	H	SSD	ZG ₃	M1	M2	F	HM
Quercetin	29.10	34.32	34.13	34.21	34.52	33.08	34.82	35.15	35.64	36.42	36.34	36.38
Kaempferol	28.30	33.98	33.82	33.91	34.25	32.72	34.28	34.47	35.33	36.15	35.99	36.07
Beta-sitosterol	51.20	37.50	37.38	37.41	37.19	39.33	36.88	36.51	35.81	35.05	35.06	35.06
Diosgenin	47.30	37.66	37.05	37.47	38.21	40.02	37.60	37.42	36.79	35.93	35.94	35.93
Lycopene	74.10	39.31	40.60	39.89	38.93	39.74	39.76	39.46	38.34	37.47	37.81	37.65
Gallotannins	55.40	68.90	69.46	69.16	68.62	70.93	68.91	67.59	65.97	62.86	63.58	63.25
Caffeic acid	18.80	30.90	30.90	30.81	30.62	29.21	31.10	32.43	31.83	32.66	32.64	32.65
Gallic acid	15.40	30.60	30.47	30.40	30.22	28.83	31.11	32.65	31.61	32.51	32.58	32.55
Ferulic acid	20.70	31.17	31.31	31.19	30.99	29.88	31.19	32.43	32.01	32.81	32.73	32.77
Sesamin	35.70	36.21	36.00	36.49	36.86	37.04	35.04	33.11	36.21	36.32	35.84	36.07
Vitexin	41.10	37.41	37.39	37.47	37.54	37.58	37.15	36.28	37.45	37.57	37.36	37.46
Emodin	27.40	33.68	33.39	33.50	33.59	31.88	33.72	34.01	34.88	35.90	35.72	35.81
Aloe-emodin	27.30	33.32	33.09	33.22	33.69	32.22	33.26	32.88	34.75	35.80	35.47	35.62
Oleanolic acid	53.00	38.57	37.71	37.92	38.22	39.56	39.26	41.05	37.59	36.72	37.24	36.99
Palmitic acid	30.80	32.07	32.91	32.56	32.12	33.58	31.54	30.16	31.39	31.41	31.29	31.35

3.2. Multi linear regression.

The data of the actual physicochemical properties and the corresponding predictions from the multilinear regression analysis calculated using Equation 2 are listed in Table 8.

Table 8. Comparison between actual values and predicted values of the MLRs.

Phyto chemical	P	A	P	A	P	A
	BP		FP		P	
Quercetin	642.4	579.35	248.1	271.24	29.1	26.09
Kaempferol	582.1	555.29	226.1	258.86	28.3	27.35
Beta-sitosterol	501.9	501.90	220.4	220.40	51.2	51.20
Diosgenin	527.1	527.10	272.6	272.60	47.3	47.30
Lycopene	660.9	557.11	350.7	307.32	74.1	61.73
Gallotannins	1103.6	1152.04	363	378.36	55.4	60.60
Caffeic acid	416.8	473.12	220	238.70	18.8	25.35
Gallic acid	501	519.14	271	269.17	15.4	21.77
Ferulic acid	372.3	412.03	150.5	147.56	20.7	20.84
Sesamin	504.4	519.49	212.3	205.61	35.7	37.40
Vitexin	767.7	608.96	273.1	224.04	41.1	26.57
Emodin	586.9	601.31	322.8	289.34	27.4	23.78
Aloe-emodin	568.8	610.04	311.9	307.30	27.3	27.43
Oleanolic acid	553.5	553.50	302.6	302.60	53	53.00
Palmitic acid	351.5	470.52	192	244.00	30.8	45.19

P-Predicted; A-Actual.

The statistical parameters derived from the multiple linear regression models are listed in Tables 9 and 10, respectively.

Table 9. Statistical parameters of the MLR model of CQ.

Physical property	Coefficients of ABC (G)	Coefficients of R (G)	Coefficients of S (G)	Coefficients of GA (G)	Coefficients of M1	Coefficients of M2	Coefficients of H (G)	Coefficients of HM (G)
BP	1327.25	-314.85	32.60	170.45	-303.63	34.49	-390.21	35.87
F	2390.87	-776.90	99.49	660.08	-640.80	62.91	-713.40	69.50

Physical property	Coefficients of ABC (G)	Coefficients of R (G)	Coefficients of S (G)	Coefficients of GA (G)	Coefficients of M1	Coefficients of M2	Coefficients of H (G)	Coefficients of HM (G)
P	211.97	-87.28	14.02	94.22	-66.92	5.23	-62.17	6.50

Table 10. Statistical parameters of the MLR model of CQ.

Physical property	Coefficients of H (G)	Coefficients of HM (G)	Coefficients of ZG3	Coefficients of F (G)	Coefficients of SSD (G)	Intercept	R ²	MSE	RMSE
BP	-390.21	35.87	4.40	-33.10	0.04	426.54	0.86	4292.82	65.52
F	-713.40	69.50	32.95	-56.33	0.52	225.13	0.80	692.26	26.31
P	-62.17	6.50	3.84	-3.96	0.17	16.42	0.81	47.18	6.87

3.3. Random forest regression.

The data of the actual physicochemical properties and the corresponding predictions from the Random Forest Regression analysis calculated from Equation 3 are listed in Table 11.

Table 11. Comparison between actual and predicted values of the RFR.

Phyto chemical	A	P	A	P	A	P	A	P	A	P	A	P
	BP		MP		FP		MV		P		IC ₅₀	
Quercetin	642.4	621.25	316	293.82	248.1	260.76	168	169.56	29.1	28.74	10	332.87
Kaempferol	582.1	588.35	275	272.96	226.1	264.80	169.5	170.25	28.3	28.19	28	371.27
Beta-sitosterol	501.9	525.82	140	177.17	220.4	233.28	424.3	399.08	51.2	50.30	2500	2163.67
Diosgenin	527.1	543.13	203	218.49	272.6	274.21	366.9	373.61	47.3	48.05	6960	4908.60
Lycopene	660.9	672.46	175	202.08	350.7	320.22	604.2	500.55	74.1	63.24	5800	3669.80
Gallotannins	1103.6	954.29	200	208.01	363	335.24	320.4	372.23	55.4	56.77	1000	1538.33
Caffeic Acid	416.8	422.01	225	217.68	220	216.26	121.9	126.76	18.8	18.89	2	1474.80
Gallic Acid	501	471.21	258	239.40	271	251.30	97.3	111.41	15.4	17.09	10000	6916.40
Ferulic Acid	372.3	407.44	168	189.35	150.5	167.34	147.5	138.93	20.7	20.13	6440	5433.53
Sesamin	504.4	531.13	123	172.54	212.3	230.39	255.8	274.22	35.7	40.71	7100	5165.33
Vitexin	767.7	692.84	256	237.34	273.1	262.60	256.3	319.19	41.1	46.10	72	1091.00
Emodin	586.9	583.09	256	258.87	322.8	303.48	170.6	171.00	27.4	27.67	240	478.47
Aloe-emodin	568.8	556.89	222	220.51	311.9	287.54	169.7	180.37	27.3	27.73	1000	1344.87
Oleanolic Acid	553.5	593.24	310	310.00	302.6	302.60	414.9	414.90	53	53.00	90	90.00
Palmitic Acid	351.5	392.85	61.8	105.21	192	244.00	287.3	386.34	30.8	45.19	2600	2795.05

P-Predicted; A-Actual.

The statistical metrics obtained during the random forest regression analysis are listed in Table 12.

Table 12. Statistical parameters of RFR.

Property	R ²	MSE	RMSE
BP	0.92	2377.15	48.76
MP	0.85	705.02	26.55
FP	0.89	378.45	19.45
MV	0.92	1433.74	37.86
P	0.95	12.39	3.52
IC ₅₀	0.83	1825129.95	1350.97

3.4. AdaBoost.

The data of the actual physicochemical properties and the corresponding predictions from the AdaBoost regression analysis calculated from Equation 4 are listed in Table 13.

Table 13. Comparison between actual and predicted values of the AdaBoost.

Phyto chemical	A	P	A	P	A	P	A	P	A	P	A	P
	BP		MP		FP		MV		P		IC ₅₀	
Quercetin	642.4	607.05	316	286.00	248.1	240.50	168	169.70	29.1	29.67	10	939.67
Kaempferol	582.1	605.60	275	270.50	226.1	236.67	169.5	169.70	28.3	29.10	28	939.67

Phyto chemical	A	P	A	P	A	P	A	P	A	P	A	P
	BP		MP		FP		MV		P		IC ₅₀	
Beta-sitosterol	501.9	533.86	140	142.75	220.4	237.79	424.3	419.60	51.2	51.20	2500	2500.00
Diosgenin	527.1	541.82	203	203.00	272.6	257.65	366.9	366.90	47.3	47.30	6960	6380.00
Lycopene	660.9	660.90	175	191.67	350.7	350.70	604.2	604.20	74.1	74.10	5800	5800.00
Gallotannins	1103.6	1103.60	200	200.00	363	356.85	320.4	320.40	55.4	55.40	1000	1000.00
Caffeic Acid	416.8	416.80	225	217.75	220	220.00	121.9	121.90	18.8	19.75	2	2.00
Gallic Acid	501	501.00	258	226.50	271	271.00	97.3	109.60	15.4	15.40	10000	10000.00
Ferulic Acid	372.3	372.30	168	204.00	150.5	150.50	147.5	147.50	20.7	20.70	6440	5740.00
Sesamin	504.4	537.13	123	142.75	212.3	227.88	255.8	256.30	35.7	35.70	7100	6450.00
Vitexin	767.7	767.70	256	256.00	273.1	256.30	256.3	256.30	41.1	41.10	72	872.67
Emodin	586.9	605.60	256	267.90	322.8	311.90	170.6	169.70	27.4	29.10	240	1000.00
Aloe-emodin	568.8	571.10	222	207.80	311.9	311.90	169.7	169.70	27.3	29.10	1000	1765.71
Oleanolic Acid	553.5	553.50	310	310.00	302.6	287.85	414.9	407.60	53	52.49	90	872.67
Palmitic Acid	351.5	372.30	61.8	61.80	192	192.00	287.3	287.30	30.8	29.67	2600	2277.78

P-Predicted; A-Actual.

The statistical metrics obtained during the AdaBoost analysis are listed in Table 14.

Table 14. Statistical parameters of AdaBoost.

Property	R ²	MSE	RMSE
BP	0.99	326.60	18.07
MP	0.94	285.31	16.89
FP	0.97	106.31	10.31
MV	1.00	15.38	3.92
P	1.00	0.64	0.80
IC ₅₀	0.97	364379.73	603.64

3.5. Discussion.

In this study, the properties of CQ extracts were predicted using TIs, and the results were validated using linear regression (LR) and multiple linear regression (MLR) models. We aimed to predict the IC₅₀, polarizability (P), molar volume (MV), flash point (F), melting point (MP), boiling point (BP), and polarizability (P).

The results demonstrated that certain topological indices exhibited strong correlations with specific properties. For example, M₂ and HM(G) showed significant correlations with BP, making LR effective in predicting this property. However, MLR provided superior performance for BP, as it accounted for the combined influence of all indices, yielding more accurate predictions. For MP, LR yielded a near-zero R-squared value, indicating poor predictive capability. In contrast, MLR achieved a substantial R-squared value of 0.86, highlighting its advantage in cases where multiple indices collectively explain variability.

When MLR was applied, the R-squared value increased to 0.68, demonstrating its capacity to model moderate correlations effectively. The regression models demonstrated varying predictive capabilities for different physicochemical properties of *Cissus quadrangularis* compounds, with ensemble methods outperforming linear approaches in most cases.

The key findings of the regression analysis are summarised in the following phrases:

The LR model achieved moderate performance ($R^2=0.76$) with ABC(G) as the most influential index, the MLR showed a improved accuracy ($R^2=0.86$) using 11 indices, notably ABC(G) and R(G). The Random Forest regression(RFR) significantly enhanced predictions ($R^2=0.92$, RMSE = 48.76) by capturing nonlinear relationships. The AdaBoost delivered near-

perfect accuracy ($R^2=0.99$, RMSE = 18.07), reducing errors by 72% compared with MLR for Boiling Point.

The RFR showed improved predictions ($R^2=0.85$, RMSE = 26.55) using SSD(G) and H(G) and AdaBoost achieved the highest accuracy ($R^2=0.94$, RMSE = 16.89) by focusing on the outliers for melting point. Where as for Flash Point (FP), the RFR exhibited reduced RMSE to 19.45 ($R^2=0.89$) and AdaBoost gave optimized predictions ($R^2=0.97$, RMSE = 10.31) leveraging sequential error correction.

Both RFR and AdaBoost models achieved near-perfect fits ($R^2=0.92$ and 0.99 , respectively), with AdaBoost's RMSE (3.92) being 92% lower than that of MLR for Molar Volume (MV). The RFR outperformed linear models ($R^2=0.95$, RMSE = 3.52) using ZG3(G). The AdaBoost provided nearly flawless predictions ($R^2=0.997$, RMSE = 0.80), demonstrating exceptional precision for Polarizability (P). Also the AdaBoost model excelled ($R^2=0.97$, RMSE = 603.64), reducing errors by 89% compared with RFR for IC₅₀ (Bioactivity).

3.5.1. Key advancements revealed by RFR and AdaBoost.

The AdaBoost reduced prediction errors by 72-89% compared to MLR, leveraging sequential error correction across topological indices such as SSD(G) and H(G) for the property Boiling Point. For Bioactivity (IC₅₀), the AdaBoost achieved $R^2=0.97$ for IC₅₀ predictions by resolving nonlinear dose-response patterns through boosted decision trees, outperforming RFR ($R^2=0.83$) and MLR ($R^2=0.40$) and for molecular Polarizability: AdaBoost delivered near-perfect predictions (RMSE = 0.80 vs. MLR's 6.87) by optimizing weak learners on ZG3(G) indices, which correlate with electron distribution patterns.

3.5.2. Practical implications.

In drug discovery, AdaBoost demonstrated remarkable predictive ability for IC₅₀ values, achieving an accuracy with RMSE = 603.64 nM, which enabled the rapid identification of bioactive candidates solely on the basis of structural descriptors. In the field of green chemistry, the model's near-perfect prediction of molar volume with 99% accuracy (RMSE = 3.92) provides a powerful tool for solvent selection without the need for labor-intensive experimental trials. Furthermore, feature importance analysis from the RFR model highlighted ABC(G) and SSD(G) as consistently influential indices, thereby confirming their significance in quantifying molecular branching patterns and surface-related interactions.

This work establishes ensemble methods as essential tools for phytochemical modeling, with AdaBoost particularly suited for properties requiring high precision (polarizability and flash points) and RFR ideal for initial feature selection. Future studies should focus on integrating these models with 3D topological indices to bridge the gap between computational predictions and experimental observations.

4. Conclusions

This study illustrates the substantial enhancement in the predictive accuracy of topological indices for characterizing *Cissus quadrangularis* compounds through the application of advanced ensemble methods in contrast to conventional regression techniques. Multiple linear regression (MLR) laid the groundwork by integrating the effects of indices such as ABC(G) for boiling point prediction. Nevertheless, random forest regression (RFR) and AdaBoost outperformed MLR by directly modeling intricate nonlinear relationships and

refining predictions through iterative enhancements. This research ultimately contributes to the field by offering methodologies that can streamline the identification and development of promising *Cissus quadrangularis*-derived compounds for pharmaceutical applications.

Author Contributions

Conceptualization, A.D.S.; methodology, A.A.; software, A.A.; validation, A.A. and A.D.S.; formal analysis, A.A. and A.D.S.; investigation, A.D.S.; resources, A.A.; data curation, A.A.; writing—original draft preparation, A.A. and A.D.S.; writing—review and editing, A.D.S.; visualization, A.D.S.; supervision, J.B.; project administration, J.B. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Data supporting the findings of this study are available upon reasonable request from the corresponding author.

Funding

There was no external funding for this research.

Acknowledgments

We would like to express my gratitude to my institution, VIT, Chennai, as well as our colleagues, who generated and sustained a deeply rooted awareness in the study.

Conflicts of Interest

None of the authors has any conflicts of interest to declare. This is the writer's exclusive obligation to write and compile this article.

References

1. Payani, S.; Bhaskar, M.; Kumar, G.S.; Pradeepkiran, J.A. A study on antimicrobial and anticancer properties of *Cissus quadrangularis* using lung cancer cell line. *Cancer Treat. Res. Commun.* **2023**, *36*, 100732, <https://doi.org/10.1016/j.ctarc.2023.100732>.
2. Bafna, P.S.; Patil, P.H.; Maru, S.K.; Mutha, R.E. *Cissus quadrangularis* L: A comprehensive multidisciplinary review. *J. Ethnopharmacol.* **2021**, *279*, 114355, <https://doi.org/10.1016/j.jep.2021.114355>.
3. Sundar Roy, G.; Das Sarkar, R.; Bose, A.; Mondal, A.; Singh, A.; Midya, M.; Goswami, R.; Singha, M.; Maity, P.; Debnath, S.; Maiti, S.; Banerjee, S.; Majumder, R. DESIGN AND ANALYSIS OF *CISSUS QUADRANGULARIS* L. BY DESIGN OF EXPERIMENTS AS A NOVEL HERBAL GEL. *Int. J. Adv. Res.* **2023**, *11*, 983–992, <https://doi.org/10.21474/IJAR01/17611>.
4. Sairaman, S.; Nivedhitha, M.S.; Shrivastava, D.; Al Onazi, M.A.; Algarni, H.A.; Mustafa, M.; Alqahtani, A.R.; AlQahtani, N.; Teja, K.V.; Janani, K.; Eswaramoorthy, R.; Sudhakar, M.P.; Alam, M.K.; Srivastava, K.C. Biocompatibility and antioxidant activity of a novel carrageenan based injectable hydrogel scaffold

- incorporated with *Cissus quadrangularis*: an in vitro study. *BMC Oral Health* **2022**, *22*, 377, <https://doi.org/10.1186/s12903-022-02409-6>.
5. Noviandy, T.R.; Maulana, A.; Idrees, G.M.; Emran, T.B.; Tallei, T.E.; Helwani, Z.; Idrees, R. Ensemble Machine Learning Approach for Quantitative Structure Activity Relationship Based Drug Discovery: A Review. *Infolitika J. Data Sci.* **2023**, *1*, 32–41, <https://doi.org/10.60084/ijds.v1i1.91>.
 6. Anuradha, D.S.; Jaganthan, B. Topological Properties of Boron Triangular Sheet for Robotic Finger Flex Motion through Indices. *AIP* **2023**, 2946, <https://doi.org/10.1063/5.0178070>.
 7. Manzoor, S.; Siddiqui, M.K.; Ahmad, S. Degree-based entropy of molecular structure of hyaluronic acid–curcumin conjugates. *Eur. Phys. J. Plus* **2021**, *136*, 15, <https://doi.org/10.1140/epjp/s13360-020-00976-7>.
 8. Rauf, A.; Ishtiaq, M.; Siddiqui, M.K. Topological Study of Hydroxychloroquine Conjugated Molecular Structure Used for Novel Coronavirus (COVID-19) Treatment. *Polycycl. Aromat. Compd.* **2022**, *42*, 3792–3808, <https://doi.org/10.1080/10406638.2021.1873807>.
 9. Julietraja, K.; Venugopal, P.; Prabhu, S.; Arulmozhi, A.K.; Siddiqui, M.K. Structural Analysis of Three Types of PAHs Using Entropy Measures. *Polycycl. Aromat. Compd.* **2022**, *42*, 4101–4131, <https://doi.org/10.1080/10406638.2021.1884101>.
 10. Chu, Y.-M.; Julietraja, K.; Venugopal, P.; Siddiqui, M.K.; Prabhu, S. Degree- and irregularity-based molecular descriptors for benzenoid systems. *Eur. Phys. J. Plus* **2021**, *136*, 78, <https://doi.org/10.1140/epjp/s13360-020-01033-z>.
 11. Yu, G.; Siddiqui, M.K.; Hussain, M.; Hussain, N.; Saddique, Z.; Petros, F.B. On topological indices and entropy measures of beryllonitrene network via logarithmic regression model. *Sci. Rep.* **2024**, *14*, 7187, <https://doi.org/10.1038/s41598-024-57601-1>.
 12. Anuradha, D.S.; Jaganathan, B. Physicochemical Properties of Benzophenone and Curcumin-conjugated PAMAM Dendrimers Using Topological Indices. *Polycycl. Aromat. Compd.* **2024**, *44*, 3419–3441, <https://doi.org/10.1080/10406638.2023.2234542>.
 13. Kuriachan, G.; Parthiban, A. Computation of domination degree-based topological indices using python and QSPR analysis of physicochemical and ADMET properties for heart disease drugs. *Front. Chem.* **2025**, *13*, 1536199, <https://doi.org/10.3389/fchem.2025.1536199>.
 14. Arockiaraj, M.; Godlin, J.J.J.; Radha, S. Comparative study of degree and neighborhood degree sum-based topological indices for predicting physicochemical properties of skin cancer drug structures. *Mod. Phys. Lett. B* **2025**, *39*, 2550106, <https://doi.org/10.1142/S0217984925501064>.
 15. Sorgun, S.; Birgin, K. Vertex-Edge-Weighted Molecular Graphs: A Study on Topological Indices and Their Relevance to Physicochemical Properties of Drugs Used in Cancer Treatment. *J. Chem. Inf. Model.* **2025**, *65*, 2093–2106, <https://doi.org/10.1021/acs.jcim.4c02013>.
 16. Arockiaraj, M.; Jeni Godlin, J.J.; Radha, S.; Aziz, T.; Al-Harbi, M. Comparative study of degree, neighborhood and reverse degree based indices for drugs used in lung cancer treatment through QSPR analysis. *Sci. Rep.* **2025**, *15*, 3639, <https://doi.org/10.1038/s41598-025-88044-x>.
 17. Kuriachan, G.; Angamuthu, P. Quantitative Structure-Property Relationship Analysis of Physical and ADMET Properties of Anticancer Drugs Using Domination Topological Indices. *Int. J. Quantum. Chem.* **2025**, *125*, e27525, <https://doi.org/10.1002/qua.27525>.
 18. Hasani, M.; Ghods, M.; Mondal, S.; Siddiqui, M.K.; Cheema, I.Z. Modeling QSPR for pyelonephritis drugs: a topological indices approach using MATLAB. *J. Supercomput.* **2025**, *81*, 479, <https://doi.org/10.1007/s11227-025-06967-8>.
 19. Asghar, A. QSPR analysis of anti-hepatitis prescription drugs using degree based topological indices through M-polynomial and NM-polynomial. *Chimica Techno Acta* **2025**, *12*, 12206, <https://doi.org/10.15826/chimtech.2025.12.2.06>.
 20. Saeed, F.; Idrees, N. Integrating Structural Modeling and Decision-Making for Anti-psychotic Drugs through Topological Indices. *BioNanoScience* **2024**, *15*, 59, <https://doi.org/10.1007/s12668-024-01732-2>.
 21. Ashraf, T.; Idrees, N. Topological indices based VIKOR assisted multi-criteria decision technique for lung disorders. *Front. Chem.* **2024**, *12*, 1407911, <https://doi.org/10.3389/fchem.2024.1407911>.
 22. Idrees, N.; Noor, E.; Rashid, S.; Agama, F.T. Role of topological indices in predictive modeling and ranking of drugs treating eye disorders. *Sci. Rep.* **2025**, *15*, 1271, <https://doi.org/10.1038/s41598-024-81482-z>.
 23. Ugasini Preetha, P.; Suresh, M.; Tolasa, F.T.; Bonyah, E. QSPR/QSAR study of antiviral drugs modeled as multigraphs by using TI's and MLR method to treat COVID-19 disease. *Sci. Rep.* **2024**, *14*, 13150, <https://doi.org/10.1038/s41598-024-63007-w>.

24. Zaman, S.; Rasheed, S.; Alamer, A. A quadratic regression model to quantify certain latest corona treatment drug molecules based on coindices of M-polynomial. *J. Supercomput.* **2024**, *80*, 26805–26830, <https://doi.org/10.1007/s11227-024-06434-w>.
25. Anuradha, D.S.; Jaganathan, B. Predictive modelling and ranking: *Azadirachta indica* compounds through indices and multi-criteria decision-making techniques. *Front. Chem.* **2025**, *13*, 1580267, <https://doi.org/10.3389/fchem.2025.1580267>.
26. Harsha Vardhan, K.S.; Anuradha, D.S.; Jaganathan, B. Double bond indices and their application: QSAR of polycyclic aromatic hydrocarbons. *Utilitas Mathematica* **2025**, *122*, 3–27, <https://doi.org/10.61091/um122-01>.
27. Umaiya Bharathi, V.; Thambidurai, S. Green synthesized chitosan-coated iron oxide nanocomposite using *Cissus quadrangularis* plant extract for antibacterial, antioxidant and anticancer applications. *Inorganica Chim. Acta* **2024**, *572*, 122293, <https://doi.org/10.1016/j.ica.2024.122293>.
28. Murugan, G.; Julietraja, K.; Alsinai, A. Computation of Neighborhood M-Polynomial of Cycloparaphenylene and Its Variants. *ACS Omega* **2023**, *8*, 49165–49174, <https://doi.org/10.1021/acsomega.3c07294>.
29. Zhang, X.; Saif, M.J.; Idrees, N.; Kanwal, S.; Parveen, S.; Saeed, F. QSPR Analysis of Drugs for Treatment of Schizophrenia Using Topological Indices. *ACS Omega* **2023**, *8*, 41417–41426, <https://doi.org/10.1021/acsomega.3c05000>.
30. Zhang, X.; Bajwa, Z.S.; Zaman, S.; Munawar, S.; Li, D. The study of curve fitting models to analyze some degree-based topological indices of certain anti-cancer treatment. *Chem. Pap.* **2024**, *78*, 1055–1068, <https://doi.org/10.1007/s11696-023-03143-1>.
31. Malik, M.Y.H.; Binyamin, M.A.; Hayat, S. Correlation Ability of Degree-Based Topological Indices for Physicochemical Properties of Polycyclic Aromatic Hydrocarbons with Applications. *Polycycl. Aromat. Compd.* **2022**, *42*, 6267–6281, <https://doi.org/10.1080/10406638.2021.1977349>.
32. Julietraja, K.; Alsinai, A.; Alameri, A. Theoretical Analysis of Superphenalene Using Different Kinds of VDB Indices. *J. Chem.* **2022**, *2022*, 5683644, <https://doi.org/10.1155/2022/5683644>.
33. Baby, A.; Julietraja, K.; Xavier, D.A. On Molecular Structural Characterization of Cyclen Cored Dendrimers. *Polycycl. Aromat. Compd.* **2024**, *44*, 707–729, <https://doi.org/10.1080/10406638.2023.2179641>.
34. Tharmalingam, G.; Ponnusamy, K.; Govindhan, M.; Konsalraj, J. On Certain Degree Based and Bond Additive Molecular Descriptors of Hexabenzocorene. *Biointerface Res. Appl. Chem.* **2023**, *13*, 495, <https://doi.org/10.33263/BRIAC135.495>.
35. Yang, J.; Konsalraj, J.; Raja S., A.A. Neighbourhood Sum Degree-Based Indices and Entropy Measures for Certain Family of Graphene Molecules. *Molecules* **2023**, *28*, 16, <https://doi.org/10.3390/molecules28010168>.
36. Joy, P.A.; Jaganathan, B. Internal Functionalized Drug Delivery Dendrimers: Theoretical Analysis by Descriptors. *Lett. Appl. NanoBioSci.* **2025**, *14*, 65, <https://doi.org/10.33263/LIANBS142.065>.
37. Gutman, I. Geometric Approach to Degree-Based Topological Indices: Sombor Indices. *MATCH Commun. Math. Comput. Chem.* **2021**, *86*, 11–16.
38. Gnanaraj, L.R.M.; Ganesan, D.; Siddiqui, M.K. Topological Indices and QSPR Analysis of NSAID Drugs. *Polycycl. Aromat. Compd.* **2023**, *43*, 9479–9495, <https://doi.org/10.1080/10406638.2022.2164315>.
39. Lal, S.; Bhat, V.K.; Sharma, K.; Sharma, S. T Topological indices of lead sulphide using polynomial technique. *Mol. Phys.* **2024**, *122*, e2249131, <https://doi.org/10.1080/00268976.2023.2249131>.
40. Jude, T.P.; Panchadcharam, E.; Masilamani, K. Topological Indices of Dendrimers used in Drug Delivery. *J. Sci. Res.* **2020**, *12*, 645–655, <https://doi.org/10.3329/jsr.v12i4.45389>.
41. Chamua, M.; Buragohain, J.; Bharali, A.; Essa Nazari, M. Predictive ability of neighborhood degree sum-based topological indices of Polycyclic Aromatic Hydrocarbons. *J. Mol. Struct.* **2022**, *1270*, 133904, <https://doi.org/10.1016/j.molstruc.2022.133904>.
42. Shi, X.; Kosari, S.; Ghods, M.; Kheirkhahan, N. Innovative approaches in QSPR modelling using topological indices for the development of cancer treatments. *PLoS One* **2025**, *20*, e0317507, <https://doi.org/10.1371/journal.pone.0317507>.
43. Buyukkose, S.; Cangul, I.N. Some notes on Randić index. *Bol. Soc. Parana. Mat.* **2022**, *40*, 1–7, <https://doi.org/10.5269/bspm.47213>.
44. Hayat, S.; Arfan, A.; Khan, A.; Jamil, H.; Alenazi, M.J.F. An Optimization Problem for Computing Predictive Potential of General Sum/Product-Connectivity Topological Indices of Physicochemical Properties of Benzenoid Hydrocarbons. *Axioms* **2024**, *13*, 342, <https://doi.org/10.3390/axioms13060342>.

45. Saidi, N.H.A.M.; Husin, M.N.; Ismail, N.B. Zagreb indices and Zagreb coindices of the line graphs of the subdivision graphs. *J. Discrete Math. Sci. Cryptogr.* **2020**, *23*, 1253–1267, <https://doi.org/10.1080/09720529.2020.1816695>.
46. Usman, M.; Javaid, M. Connection-Based Zagreb Indices of Polycyclic Aromatic Hydrocarbons Structures. *Curr. Org. Synth.* **2024**, *21*, 246–256, <https://doi.org/10.2174/1570179421666230823141758>.
47. Rajasekharaiah, G.V.; Murthy, U.P. Hyper-Zagreb indices of graphs and its applications. *J. Algebra Comb. Discrete Struct. Appl.* **2021**, *8*, 9–22, <https://doi.org/10.13069/jacodesmath.867532>
48. Sarkar, P.; De, N.; Pal, A. On some topological indices and their importance in chemical sciences: a comparative study. *Eur. Phys. J. Plus* **2022**, *137*, 195, <https://doi.org/10.1140/epjp/s13360-022-02431-1>.
49. Anuradha, D.S.; Julietraja, K.; Jaganathan, B.; Alsinai, A. Curcumin-Conjugated PAMAM Dendrimers of Two Generations: Comparative Analysis of Physicochemical Properties Using Adriatic Topological Indices. *ACS Omega* **2024**, *9*, 14558–14579, <https://doi.org/10.1021/acsomega.4c00686>.
50. Farjana, H.N.; Valiathan, G.M. *Cissus quadrangularis*: A comprehensive review as an emerging biomaterial for periodontal regeneration. *Oral Res. Rev.* **2025**, *17*, 87-92, https://doi.org/10.4103/jorr.jorr_27_24.
51. Mathangi, R.; Devarajan, N.; Lakshmi, U.D. An Overview of the Osteogenic potential of Indian herb *Cissus quadrangularis* (Veldt Grape). *South East. Eur. J. Public Health* **2024**, *25*, 2210-2214, <https://doi.org/10.70135/seejph.vi.2354>.
52. Padhihary, S.; Pramanik, K. Development of a novel *Cissus quadrangularis* extract loaded sodium alginate/chitosan based 3D printed scaffold for regeneration of cancellous alveolar bone. *J. Drug Deliv. Sci. Technol.* **2025**, *104*, 106440, <https://doi.org/10.1016/j.jddst.2024.106440>.
53. Sai Nivetha, S.; Ranjani, S.; Hemalatha, S. Synthesis and application of silver nanoparticles using *Cissus quadrangularis*. *Inorg. Nano-Met. Chem.* **2022**, *52*, 82–89, <https://doi.org/10.1080/24701556.2020.1862219>.
54. Xu, Y. Deep Neural Networks (DNNs) for QSAR Quantitative structure activity relationship (QSAR). In *Artificial Intelligence in Drug Design*, Heifetz, A., Ed.; Springer US: New York, NY, **2022**; Volume 2390, pp. 233-260, https://doi.org/10.1007/978-1-0716-1787-8_10.
55. Zhang, X.; Govardhana Reddy, H.G.; Usha, A.; Shanmukha, M.C.; Farahani, M.R.; Alaeiyan, M. A study on anti-malaria drugs using degree-based topological indices through QSPR analysis. *Math. Biosci. Eng.* **2023**, *20*, 3594–3609, <https://doi.org/10.3934/mbe.2023167>.
56. Ramapraba, P.S.; Babu, B.R.; Paul, N.R.R.; Sharmila, V.; Babu, V.R.; Ramya, R.; Murugan, S. Implementing cloud computing in drug discovery and telemedicine for quantitative structure-activity relationship analysis. *Int. J. Electr. Comput. Eng.* **2025**, *15*, 1132-1141, <https://doi.org/10.11591/ijece.v15i1.pp1132-1141>.
57. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487, <https://doi.org/10.5194/gmd-15-5481-2022>.
58. Amiri, S.; Zakeri, N.; Yousefi, T. Evaluation of the APSIM common bean model using different cultivars and water-management scenarios. *Technol. Agron.* **2024**, *4*, e017, <https://doi.org/10.48130/tia-0024-0015>.
59. Sahu, S.K.; Ojha, K.K. Applications of QSAR study in drug design of tubulin binding inhibitors. *J. Biomol. Struct. Dyn.* **2024**, *42*, 12806-12821, <https://doi.org/10.1080/07391102.2023.2273437>.
60. Sardar, M.S.; Ali, M.A.; Farahani, M.R.; Alaeiyan, M.; Cancan, M.; Campus, Z. TOPOLOGICAL INDICES AND QSPR/QSAR ANALYSIS OF SOME DRUGS BEING INVESTIGATED FOR THE TREATMENT OF HEADACHES. *Eur. Chem. Bull.* **2023**, *12*, 69–83.
61. Ahmed, W.; Zaman, S.; Asif, E.; Ali, K.; Mahmoud, E.E.; Asheboss, M.A. Exploring the role of topological descriptors to predict physicochemical properties of anti-HIV drugs by using supervised machine learning algorithms. *BMC Chem.* **2024**, *18*, 167, <https://doi.org/10.1186/s13065-024-01266-4>.
62. Koam, A.N.A.; Majeed, M.U.; Zaman, S.; Ahmad, A.; Masmali, I.; Ahmadini, A.A.H. Machine learning approaches for modeling the physicochemical characteristics of polycyclic aromatic hydrocarbons. *Eur. Phys. J. E* **2025**, *48*, 21, <https://doi.org/10.1140/epje/s10189-025-00487-2>.
63. Javed, S.; Ahmad, S.; Sehar, N.; Khalid, S.; Siddiqui, M.K.; Gegbe, B. Enhancing topological index of calcium chloride network through feature selection methods exploration. *Sci. Rep.* **2024**, *14*, 27722, <https://doi.org/10.1038/s41598-024-79040-8>.
64. Shanmugasundar, G.; Vanitha, M.; Čep, R.; Kumar, V.; Kalita, K.; Ramachandran, M. A Comparative Study of Linear, Random Forest and AdaBoost Regressions for Modeling Non-Traditional Machining. *Processes* **2021**, *9*, 2015, <https://doi.org/10.3390/pr9112015>.

65. Kaur, J.; Dhiman, V.; Bhadada, S.; Katare, O.P.; Ghoshal, G. LC/MS Guided Identification of Metabolites of Different Extracts of *Cissus Quadrangularis*. *Food Chemistry Advances* 2022, 1, 100084, doi:<https://doi.org/10.1016/j.focha.2022.100084>.

Publisher's Note & Disclaimer

The statements, opinions, and data presented in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect the views of the publisher and/or the editor(s). The publisher and/or the editor(s) disclaim any responsibility for the accuracy, completeness, or reliability of the content. Neither the publisher nor the editor(s) assume any legal liability for any errors, omissions, or consequences arising from the use of the information presented in this publication. Furthermore, the publisher and/or the editor(s) disclaim any liability for any injury, damage, or loss to persons or property that may result from the use of any ideas, methods, instructions, or products mentioned in the content. Readers are encouraged to independently verify any information before relying on it, and the publisher assumes no responsibility for any consequences arising from the use of materials contained in this publication.