



OPEN A graph-based computational approach for modeling physicochemical properties in drug design

Ibrahim Al-Dayel¹, Meraj Ali Khan¹, Muhammad Faisal Hanif², Muhammad Kamran Siddiqui³, Saba Hanif³ & Brima Gegbe⁴✉

The efficacy and effectiveness of antibiotics and neuropathic drugs are essentially guided by their physicochemical properties governing stability, bioavailability, and therapeutic activity. This work utilises mathematical modelling and quantitative structure-property relationship (QSPR) analysis for predicting important physicochemical properties such as boiling point, enthalpy of vaporisation, flash point, and molar refraction of chosen antibiotics and neuropathic drugs. Modified degree-based topological indices are utilised as molecular descriptors for correlations between physicochemical functionality and molecular structure. Linear and quadratic forms are various forms of regression models employed for improved predictions. The findings exhibit excellent performance of quadratic models across all but one property compared to linear models, highlighted by significant statistical markers like high R^2 values and low error margins. These results highlight the potential use of topological descriptors in combination with sound mathematical frameworks for drug optimisation and early-stage screening.

Keywords Molecular graph, QSPR analysis, Antibiotic compounds, Topological indices, Regression modeling

Over recent years, QSPR modeling has been established as a highly viable technique in pharmaceutical chemistry for its potential in predicting major physicochemical properties from molecular structure directly¹. This predictive property is important for accelerating drug discovery since early evaluation of molecular performance can be conducted without requiring long-term laboratory experiments². Boiling points, enthalpy for vaporization, flash points, and molar refraction are particularly significant due to their impact on influential pharmaceutical parameters such as stability, availability, and formulation safety³. Such considerations are particularly significant when developing antibiotics, where therapeutic efficacy is highly dependent upon molecular traits. A QSPR model provides standardized means for connecting molecular descriptors and such properties, making the rationale-based lead optimization and drugs design possible⁴. For an elaborate review of QSPR methods applied to drug molecules, see⁵. Kumar and Das⁶ compared Generalized Quadratic (GQ) and Quadratic Generalized (QG) indices as molecular descriptors. Their work shows how these indices are effective in modeling structural aspects pertinent for quantum chemical properties. The analysis indicates that these two descriptors hold potential for improved predictive modeling for QSPR research work.

Here, important physicochemical traits are studied for a group of drugs of pharmacological interest, such as Norfloxacin, Ciprofloxacin, Gemifloxacin, Chlorpyrifos, Dimethoate, Phorate, Monocrotophos, Thiram, Azoxystrobin, Atrazine, Chlorothalonil, and Warfarin. Most of these drugs, which are either antibiotics or structurally analogous bioactive molecules, are extensively employed for various therapeutic and agricultural applications and share diverse and pertinent physicochemical requirements⁷. For example, vaporization enthalpy informs about the amount of energy needed for phase boundary transitions, which has direct bearing on drug formulation and storage functions. In a similar vein, data for boiling points for drugs such as Norfloxacin and doxycycline are important for gauging thermal stability for various temperature and pressure conditions⁸. Kumar

¹Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), P.O. Box 65892, 11566 Riyadh, Saudi Arabia. ²Department of Mathematics and Statistics, The University of Lahore, Lahore Campus, Lahore, Pakistan. ³Department of Mathematics, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan. ⁴Department of Mathematics and Statistics, Njala University, Freetown, Sierra Leone. ✉email: bgegbe@njala.edu.sl

and Das⁹ explored the structural sensitivity and chemical applicability of numerous newly developed degree-based topological indices. Their analysis proved how these indices react under different structural changes in molecules. The analysis reaffirmed their potential applicability in QSPR/QSAR modeling by proving effective for correlation with important chemical parameters. In turn, molar refraction is a gauge of the polarizability of a molecule and has predictive value for drug biological environment interactions, whereas flash point measurements are important for determining flammability and for proper handling of volatile drugs¹⁰.

The present research utilizes degree-based topological indices, which are mathematical descriptors of molecular structure^{11,12}. In graph theory, vertex degree noted λ_a refers to the number of connected edges linking a given vertex a ¹³. Degree-based topological indices are highly effective for predicting physicochemical, thermodynamic, and biological characteristics because they capture essential structural information such as molecular branching and network connectivity¹⁴. Kumar and Das¹⁵ compared the performance of Nirmla and Sombor indices in detail, i.e., comparing their applicability, degeneracy, and smoothness. This work compared how each index functions in differentiating between molecular structures. The results underscore the efficacy and limitations of these two indices and offer insights into their possible use for modeling molecules using graphs. Within QSAR and QSPR analysis, degree-based topological indices constitute an important subdomain because vertex degree-based information allows for high correlation between molecular structure and experimentally tangible properties or bioactivities¹⁶.

Jacob and Clement¹⁷ elucidated the structural complexity of zeolite EDI through the application of topological entropy and proved the efficacy of the model to predict the interaction of molecules inside zeolite structures. The method introduced a new quantitative means to interpret host-guest behavior in porous compounds. Gayathri, Santiago, and Govardhan¹⁸ investigated hexagonal fractals via the application of topological indices and the use of fractal dimensions to describe structure-property relationships. Their work showed the capability of the aforementioned mathematical measures to predict the behavior of molecules and predict material properties. Peter et al.¹⁹ created a predictive model of interaction energies between molecules through the application of topological and spectral entropies of zeolite AWW. The article stressed the significance of descriptors of the entropic kind to predict the interaction of molecules inside porous compounds. Shenoy et al.²⁰ carried out statistical analysis of the application of Revan topological indices to drug compounds for the treatment of tuberculosis. Their analysis showed the promise of the indices to analyze and discriminate molecular structures for application to pharmaceutical research. Jacob et al.²¹ carried out network topology and entropy analysis of tetragonal farneseite zeolites in order to study their structural complexity. Their research emphasized the use of measures of entropy to identify complex topological properties of the zeolite structures.

The application of topological indices as structural descriptors has a history dating back a long time in predicting molecular properties. The first contribution was made by Wiener²², who proved that paraffin boiling points could be well estimated through topological indices, setting a precursor to the relationship between molecular structure and thermodynamic property. This was further developed upon by Gutman and Trinajstić²³ when the first Zagreb indices were introduced degree-based topological indices that have been an integral part in investigating the electronic aspect of molecules thereafter. This framework was expanded upon later by Kulli²⁴ when he introduced reverse Zagreb indices that capture reverse vertex degrees and present additional structural insights. Further building upon this was done by Khabyah et al.²⁵, who applied reverse degree-based indices toward extracting knowledge about coronene structures. Being symmetric, two-dimensional, and highly conjugated in nature, coronene is a perfect system for studying molecular symmetry and stability using topological analysis. For this work, a diversity of modified degree based topological indices are applied toward capturing the antibiotic and related molecules' molecular structure. The indices are capable of encoding essential structural features size, branching, and connectivity effectively, enabling quantitative correlation between architectural building blocks and physicochemical behavior²⁶.

Quantitative Structure-Property Relationship (QSPR) models have been widely developed for predicting chemical and physical properties of pharmacologically active compounds, illustrated by comparison analysis between regression models depicted in Figs. 1, 2, 3, and 4. Amongst these studied compounds, Norfloxacin, a synthetic fluoroquinolone antibiotic, is extensively utilized in clinical practice for treating urinary tract infection (UTI), prostatitis, and various gastrointestinal infections caused by susceptible bacteria. Norfloxacin inhibits bacterial DNA gyrase and ultimately causes cell division and DNA replication inhibition, which leads ultimately to cell death in bacteria. This drug is mainly effective in gram-negative organisms like *Escherichia coli* and *Klebsiella pneumoniae* and is administered orally. Nevertheless, in clinical use, its activity is sometimes hindered

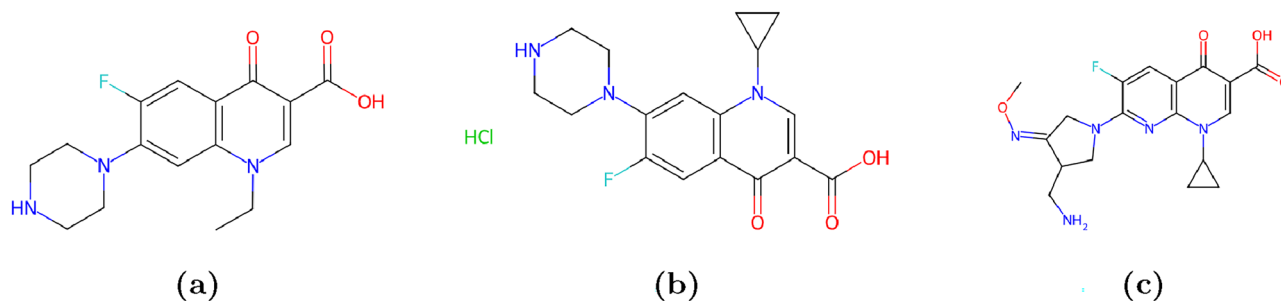


Fig. 1. (a) Norfloxacin (b) Ciprofloxacin (c) Gemifloxacin.

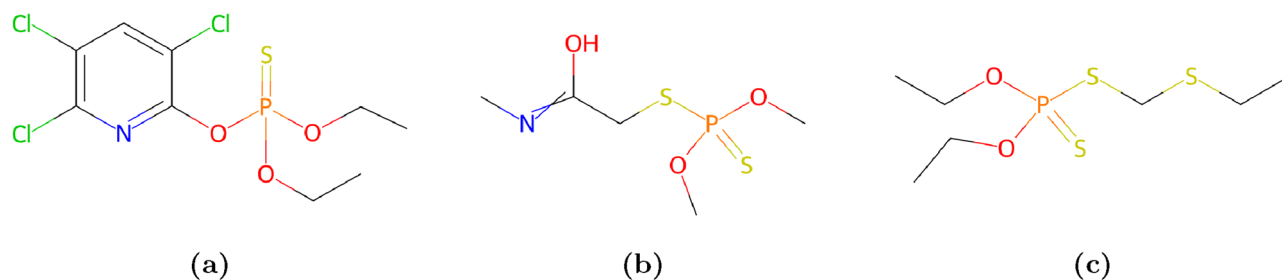


Fig. 2. (a) Chlorpyrifos (b) Dimethoate (c) Phorate.

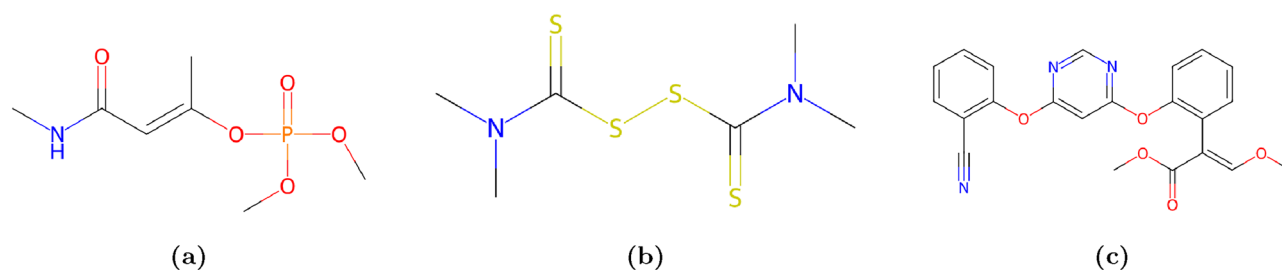


Fig. 3. (a) Monocrotophos (b) Thiram (c) Azoxystrobin.

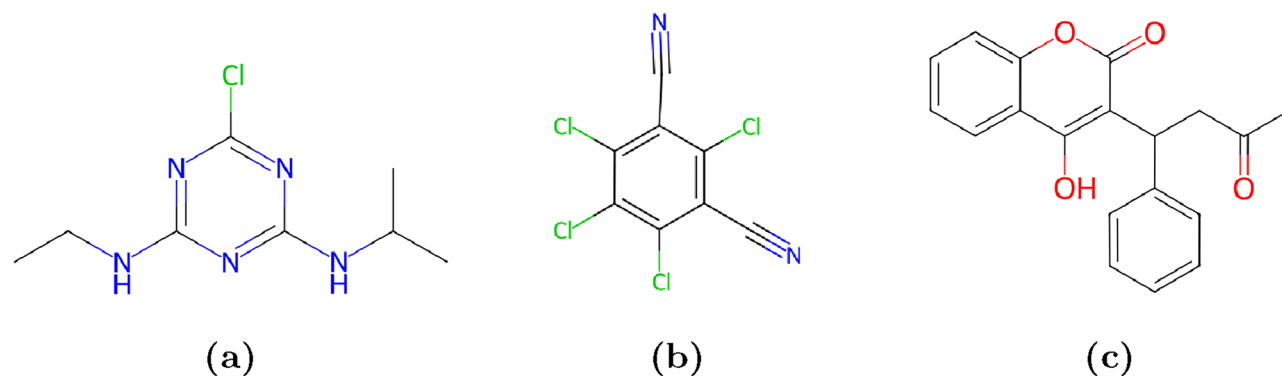


Fig. 4. (a) Atrazine (b) Chlorothalonil (c) Warfarin.

because of side effects like nausea, dizziness, and uncommon but serious tendon issues, particularly in elderly patients. Norfloxacin is contraindicated in infancy and pregnant individuals because of possible adverse cartilage and joint development effects, as depicted in Fig. 1a (source: <http://www.chemspider.com>).

Ciprofloxacin Ciprofloxacin is a fluoroquinolone antibiotic agent that exerts broad-spectrum efficacy against bacterial infections, including respiratory, urinary, gastrointestinal, and skin infections. This agent exerts its antibacterial effects by inhibiting bacterial DNA gyrase and topoisomerase IV, enzymes critical for DNA replication in bacteria^{27,28}. This triggers the death of bacterial cells. Side effects may include nausea, diarrhea, and photosensitivity, risk of tendon ruptures, particularly in the elderly. Because of these possible adverse effects, it is usually avoided in children as well as pregnant women see Fig. 1b.

Gemifloxacin Gemifloxacin is a fluoroquinolone antimicrobial agent indicated for respiratory tract infections, including acute bacterial exacerbation of chronic bronchitis and mild to moderate forms of community-acquired pneumonia^{29,30}. The drug exerts its antibacterial activity by inhibiting DNA gyrase and topoisomerase IV-bacterial enzymes critical for DNA replication. The common side effects include rash, diarrhea, and nausea, while serious ones might include tendonitis and tendon rupture. This is a cautious prescription that should not be dispensed to children or pregnant women see Fig. 1c.

Chlorpyrifos Chlorpyrifos is not an antibiotic but an organophosphate pesticide in common use for the control of insect pests both in agriculture and residentially. It works through the inhibition of acetyl cholinesterase enzyme action in insects, an action which causes overstimulation of the nervous system and eventually death^{31,32}.

It is due to these reasons that it has been restricted or banned in many countries, with studies linking exposures to developmental and neurological problems. see Fig. 2a.

Dimethoate It is not an antibiotic; it is an organophosphate insecticide, widely used in agriculture to control a wide range of pests on crops. It works by inhibiting acetyl cholinesterase, an enzyme essential for insect nerve function, thus disrupting their nervous system and eventually killing them^{33,34}. While effectively applied for pest control, dimethoate had raised health and environmental concerns because of its toxicity, affecting not only non-target species, including beneficial insects, but also humans upon exposure. Due to its associated risks, certain countries restrict its use; recommended guidelines on safety are advised upon application to reduce the impact on the environment and human health see Fig. 2b.

Phorate Phorate is a highly toxic organophosphate insecticide and acaricide. It is not an antibiotic, and its application is only for the safety of crops such as corn, potatoes, and cotton. Phorate acts through the inhibition of acetyl cholinesterase, an enzyme critical to the normal functioning of the nervous system; the insects are then paralyzed and dead^{35,36}. Phorate is extremely toxic and, therefore, particularly dangerous to humans, wild animals, and the environment through exposure. Because of this, its use has been seriously circumscribed when this insecticide is used, and certain countries have banned it. Recommended safety precautions during handling and application are rigorous see Fig. 2c.

Monocrotophos Monocrotophos is not an antibiotic, but it contains a highly toxic organophosphorus insecticide used in large-scale control of various pests in a variety of crops such as cotton, rice, and sugar cane. It acts through the inhibition of acetyl cholinesterase that overstimulates the nervous system and eventually leads to paralysis and death in insects. Monocrotophos is highly toxic, and exposure has a severe threat to human beings, birds, and aquatic life^{37,38}. For this reason, several countries have banned or restricted the use of the chemical, due to which only a High degree of care can be afforded to those areas where the use of the chemical is not prohibited for agricultural purposes see Fig. 3a.

Thiram Thiram is not an antibiotic. It is a chemical type of fungicide, and it finds wide application in agriculture to prevent fungal infections in crops such as fruits, vegetables, and ornamental plants. This effect is through the inhibition of spore germination of fungi, thus protecting the plants from diseases such as blight, rust, and mold^{39,40}. It is injurious to health if inhaled and ingested and can cause irritation to the skin of human beings too. As it is hazardous to the environment, its use has some prescription on the basis of which it has to be handled and applied see Fig. 3b.

Azoxystrobin In fact, it is not an antibiotic but a broad-spectrum fungicide. It finds wide application in agriculture for the protection of various crops against very different fungal diseases such as blight, rust, and powdery mildew. The mode of action consists of the inhibition of mitochondrial respiration in fungi, which halts the growth of fungal cells and stops the proliferation of the infection^{41,42}. This product is taken to be generally safe while being used in the prescribed manner; however, it does have a potential threat to the environment, particularly aquatic life, and should be carried out very carefully near water systems to minimize ecological damage see Fig. 3c.

Atrazine Atrazine is not an antibiotic, but it is one of the most widely used herbicides in the world. It is applied for the control of broadleaf and grassy weeds in a number of crops, including corn, sugarcane, and sorghum^{43,44}. A chemical kill of plants, it inhibits photosynthesis in plants, thus causing death to the weeds and increasing crop yields. Due to the possible influence of the chemical on fauna and suspected endocrine-disrupting activity, its use is either limited or prohibited in some countries, while it remains one of the most widely used herbicides in other countries see Fig. 4a.

Chlorothalonil Chlorothalonil is not an antibiotic. It is a broad-spectrum fungicide. It finds wide application in agriculture and horticulture for the protection of crops such as peanuts, potatoes, tomatoes, and turf grass against fungal diseases like mold, blight, and leaf spot. Where chlorothalonil is effective, there is also a risk to the environment, aquatic organisms in particular, and irritation in man in the case of exposure⁴⁵. Its use is thus regulated in many regions with specific view to minimizing ecological impacts, and safety measures are advised during application with a view to reducing health risks see Fig. 4b.

Warfarin Warfarin is not an antibiotic; it is an anticoagulant or blood thinner used to prevent and treat blood clots that form in the veins of the legs, in the lungs, and with atrial fibrillation. Warfarin requires close monitoring since its efficacy depends on dietary factors, concurrent medications, and patient variables. Blood sampling is frequent to maintain it within a narrow window of safety and efficacy⁴⁶. Adverse effects include bleeding complications, which may well require a patient to be advised about diet and the use of medications cautiously to avoid undesirable interactions see Fig. 4c.

Degree based topological indices

In this section, we provide the initial literature review for topological indices.

Gutman and Polansky introduced the first and second Zagreb index⁴⁷ as:

$$M_1(G) = \sum_{ab \in E(G)} (\hat{\lambda}_a + \hat{\lambda}_b), \quad M_2(G) = \sum_{ab \in E(G)} (\hat{\lambda}_a \times \hat{\lambda}_b)$$

Martinez-Martinez et al.⁴⁸ defined the Harmonic index as follows:

$$H(G) = \sum_{ab \in E(G)} \frac{2}{(\hat{\lambda}_a + \hat{\lambda}_b)}$$

The forgotten index was introduced by Furtula and Gutman⁴⁹ as:

$$F(G) = \sum_{ab \in E(G)} [(\widehat{\lambda}_a)^2 + (\widehat{\lambda}_b)^2]$$

The Shilpa-Shanmukha Index was defined by Zhao et al.⁵⁰ as follows:

$$SS(G) = \sum_{ab \in E(G)} \sqrt{\frac{\widehat{\lambda}_a \times \widehat{\lambda}_b}{\widehat{\lambda}_a + \widehat{\lambda}_b}}$$

The atom bond connectivity index was introduced by Estrada et al.⁵¹ as:

$$ABC(G) = \sum_{ab \in E(G)} \sqrt{\frac{\widehat{\lambda}_a + \widehat{\lambda}_b - 2}{\widehat{\lambda}_a \times \widehat{\lambda}_b}}$$

The Randic index was defined by Randic et al.⁵² as:

$$RI(G) = \sum_{ab \in E(G)} \frac{1}{\widehat{\lambda}_a \times \widehat{\lambda}_b}$$

The sum connectivity index and the geometric arithmetic index was defined by Vukicevic et al.⁵³ as:

$$SC(G) = \sum_{ab \in E(G)} \frac{1}{\widehat{\lambda}_a + \widehat{\lambda}_b}, \quad GA(G) = \sum_{ab \in E(G)} \frac{2\sqrt{\widehat{\lambda}_a \times \widehat{\lambda}_b}}{\widehat{\lambda}_a + \widehat{\lambda}_b}$$

The Hyper Zagreb Index was introduced by Rajasekharaiah et al.⁵⁴ as:

$$HZ(G) = \sum_{ab \in E(G)} (\widehat{\lambda}_a + \widehat{\lambda}_b)^2$$

Main results

In this section, we presented antibiotic drug molecules’ degree-based bond partitions, see Table 1. Moreover using this edge partition for all antibiotic drugs, we have computed degree-based topological indices. For the drug molecule Norfloxacin (G_1), By using first and second Zagreb index and edge partition of G_1 .

$$\begin{aligned} M_1(G) &= \sum_{ab \in E(G)} (\widehat{\lambda}_a + \widehat{\lambda}_b) \\ M_1(G_1) &= (1+2)(2) + (1+3)(5) + (2+2)(3) + (2+3)(8) + (3+3)(7) = 120 \\ M_2(G) &= \sum_{ab \in E(G)} (\widehat{\lambda}_a \times \widehat{\lambda}_b) \\ M_1(G_1) &= (1 \times 2)(2) + (1 \times 3)(5) + (2 \times 2)(3) + (2 \times 3)(8) + (3 \times 3)(7) = 142 \end{aligned}$$

Graphs	Drug molecules	Degree edge classes with their frequencies						
		(1, 2)	(1, 3)	(1, 4)	(2, 2)	(2, 3)	(2, 4)	(3, 3)
G_1	Norfloxacin	2	5	0	3	8	0	7
G_2	Ciprofloxacin	0	4	0	5	10	0	8
G_3	Gemifloxacin	2	4	0	2	12	0	11
G_4	Chlorpyrifos	2	3	1	2	5	3	2
G_5	Dimethoate	3	1	1	1	2	3	0
G_6	Phorate	3	0	1	5	0	3	0
G_7	Monochrotophos	3	2	1	0	4	3	0
G_8	Thiram	0	6	0	1	2	0	2
G_9	Azoxystrobin	2	2	0	9	14	0	4
G_{10}	Atrazine	1	3	0	1	9	0	0
G_{11}	Chlorothalonil	2	4	0	0	2	0	6
G_{12}	Warfarin	0	4	0	7	8	0	6

Table 1. Degree-based bond partitions of antibiotic drug molecules.

By using Harmonic index and edge partition of G_1

$$H(G) = \sum_{ab \in E(G)} \frac{2}{(\hat{\lambda}_a + \hat{\lambda}_b)}$$

$$H(G_1) = \frac{2}{(1+2)}(2) + \frac{2}{(1+3)}(5) + \frac{2}{(2+2)}(3) + \frac{2}{(2+3)}(8) + \frac{2}{(3+3)}(7) = 10.87$$

By using forgotten index and edge partition of G_1

$$F(G) = \sum_{ab \in E(G)} [(\hat{\lambda}_a)^2 + (\hat{\lambda}_b)^2]$$

$$F(G_1) = (1^2 + 2^2)(2) + (1^2 + 3^2)(5) + (2^2 + 2^2)(3) + (2^2 + 3^2)(8) + (3^2 + 3^2)(7) = 314$$

By using Shilpa-Shanmukha Index and edge partition of G_1

$$SS(G) = \sum_{ab \in E(G)} \sqrt{\frac{\hat{\lambda}_a \times \hat{\lambda}_b}{\hat{\lambda}_a + \hat{\lambda}_b}}$$

$$SS(G_1) = \sqrt{\frac{1 \times 2}{(1+2)}}(2) + \sqrt{\frac{1 \times 3}{(1+3)}}(5) + \sqrt{\frac{2 \times 2}{(2+2)}}(3) + \sqrt{\frac{2 \times 3}{(2+3)}}(8) + \sqrt{\frac{3 \times 3}{(3+3)}}(7) = 26.30$$

By using atom bond connectivity index and edge partition of G_1

$$ABC(G) = \sum_{ab \in E(G)} \sqrt{\frac{\hat{\lambda}_a + \hat{\lambda}_b - 2}{\hat{\lambda}_a \times \hat{\lambda}_b}}$$

$$ABC(G_1) = \sqrt{\frac{1+2-2}{(1 \times 2)}}(2) + \sqrt{\frac{1+3-2}{(1 \times 3)}}(5) + \sqrt{\frac{2+2-2}{(2 \times 2)}}(3) + \sqrt{\frac{2+3-2}{(2 \times 3)}}(8) + \sqrt{\frac{3+3-2}{(3 \times 3)}}(7) = 17.94$$

By using Randic index and edge partition of G_1

$$RI(G) = \sum_{ab \in E(G)} \frac{1}{\sqrt{\hat{\lambda}_a \times \hat{\lambda}_b}}$$

$$RI(G_1) = \left(\frac{1}{1 \times 2}\right)(2) + \left(\frac{1}{1 \times 3}\right)(5) + \left(\frac{1}{2 \times 2}\right)(3) + \left(\frac{1}{2 \times 3}\right)(8) + \left(\frac{1}{3 \times 3}\right)(7) = 11.40$$

By using sum connectivity index and edge partition of G_1

$$SC(G) = \sum_{ab \in E(G)} \frac{1}{\hat{\lambda}_a + \hat{\lambda}_b}$$

$$SC(G_1) = \left(\frac{1}{1+2}\right)(2) + \left(\frac{1}{1+3}\right)(5) + \left(\frac{1}{2+2}\right)(3) + \left(\frac{1}{2+3}\right)(8) + \left(\frac{1}{3+3}\right)(7) = 11.59$$

By using geometric arithmetic index and edge partition of G_1

$$GA(G) = \sum_{ab \in E(G)} \frac{2\sqrt{\hat{\lambda}_a \times \hat{\lambda}_b}}{\hat{\lambda}_a + \hat{\lambda}_b}$$

$$GA(G_1) = \frac{2\sqrt{1 \times 2}}{(1+2)}(2) + \frac{2\sqrt{1 \times 3}}{(1+3)}(5) + \frac{2\sqrt{2 \times 2}}{(2+2)}(3) + \frac{2\sqrt{2 \times 3}}{(2+3)}(8) + \frac{2\sqrt{3 \times 3}}{(3+3)}(7) = 24.05$$

By using Hyper Zagreb Index and edge partition of G_1

$$HZ(G) = \sum_{ab \in E(G)} (\hat{\lambda}_a + \hat{\lambda}_b)^2$$

$$HZ(G) = (1+2)^2(2) + (1+3)^2(5) + (2+2)^2(3) + (2+3)^2(8) + (3+3)^2(7) = 598$$

Similarly, in accordance with what was described above, this computation process was applied to all of the entire dataset, and the calculated values correspondingly for all 12 antibiotic drugs. The ERC's governing instruments are compiled with Table 2.

Drug molecules	M_1	M_2	H	F	SS	ABC	RI	SC	GA	HZ
Norfloxacin	120	142	10.87	314	26.30	17.94	11.40	11.59	24.05	598
Ciprofloxacin	134	164	11.17	354	29.22	19.21	11.56	12.24	26.26	682
Gemifloxacin	156	195	12.80	420	33.71	21.91	13.29	14.01	30.11	810
Chlorpyrifos	86	97	7.90	234	18.52	13.13	8.41	8.38	17.01	428
Dimethoate	50	53	5.20	136	10.86	8.05	5.58	5.30	10.28	242
Phorate	70	78	6.90	192	15.27	10.77	7.24	7.13	14.29	348
Monocrotophos	60	64	6	164	12.92	9.57	6.47	6.19	12.11	292
Thiram	50	52	4.97	130	10.84	8.35	5.45	5.21	10.16	234
Azoxystrobin	144	166	13.77	356	32.60	21.98	14.12	14.55	30.33	688
Atrazine	64	69	6.27	160	14.27	10.23	6.61	6.60	13.36	298
Chlorothalonil	68	82	6.13	184	14.64	10.09	6.54	6.50	13.31	348
Warfarin	120	142	10.70	308	26.58	17.87	11.08	11.53	24.30	592

Table 2. Degree based topological indices for antibiotics drug molecules.

G_i	Drug molecules	Boiling point (°C)(BP)	Enthalpy of vaporization (kJ/mol)(ENP)	Flash point (°C)(FP)	Molar refractivity (cm ³)(MR)	Polarizability (cm ³)(POL)	Molar weight (g/mol)(MW)	Heavy atom count (HAC)	Complexity (Com)	Molar volume
G_1	Norfloxacin	555.8	88.1	289.9	80.7	32	319.33	22	452	237.4
G_2	Ciprofloxacin	581.8	91.5	305.6	83.3	33	331.34	24	528	226.8
G_3	Gemifloxacin	638.9	99.2	340.2	94.9	37.6	389.4	28	584	236.3
G_4	Chlorpyrifos	375.9	59.9	181.1	77.2	30.6	350.6	20	378	236.8
G_5	Dimethoate	310.3	63.9	141.4	53.7	21.3	229.26	12	232	169
G_6	Phorate	296	51.4	132.8	68.1	27.0	260.35	13	208	216.8
G_7	Monocrotophos	315.1	55.6	144.4	50.1	19.9	223.16	11	216	186.7
G_8	Thiram	307.4	54.8	139.7	67.9	26.9	240.4	10	160	180
G_9	Azoxystrobin	581.3	86.9	305.3	106.2	42.1	403.4	27	512	300.1
G_{10}	Atrazine	368.5	61.5	176.7	58.5	23.2	215.7	14	209	169.9
G_{11}	Chlorothalonil	350.5	59.5	153.8	55.2	21.9	265.9	14	322	154.6
G_{12}	Warfarin	515.2	82.9	188.8	84.4	33.5	308.33	22	426	235.8

Table 3. Antibiotics' physicochemical characteristics.

Exploring topological indices and physicochemical properties via regression models

Quantitative Structure-property Relationship (QSPR) models allow for predictions of a compound's various physical and chemical properties solely from its molecular structure⁵⁵. Quantitative Structure-Activity Relationship (QSAR) models provide a mathematical approach for investigating complex interdependencies between a molecule's structure and its physicochemical or biochemical activities⁵⁶. Topological descriptors are particularly significant in QSPR investigations, as they contain useful connections with a variety of physicochemical properties pertinent to drugs employed for treating various medical conditions⁵⁷. Both linear and quadratic regression models were employed in this study for analyzing antibiotic drugs' physicochemical properties. The experimental property values for these antibiotics are given in Table 3.

Table 4 compares between linear and quadratic regression models against various physicochemical properties prediction based on the model $M_1(G)$. Quadratic models generally outperform the linear one with a higher value of R^2 and lower SE: properties like BP, ENP, and HAC provide the best fitting with the quadratic model with $R^2 = 0.962$. The quadratic model for HAC is the most accurate in predicting the property with ($R = 0.980$), ($F = 115.763$). Both models perform badly for molar volume and MV since $R^2 \leq 0.7$. Overall, a quadratic model provides the best predictions for most properties in this dataset. The graphical prediction is shown in Fig. 5.

Table 5 compares the performance of linear and quadratic regression models for the prediction of physicochemical properties using $M_2(G)$. A quadratic model tends to make better predictions, higher R^2 values and lower standard errors (SE). For properties like boiling point (BP), enthalpy of vaporization (ENP), and heavy atom count (HAC), quadratic models show excellent performance, with . The best prediction is observed for BP, with $R^2 = 0.957$ and high statistical significance $p < 10^{-6}$. The flash point (FP) and molar refraction (MR) are also well described by quadratic models. Molar volume (MV) yields poorer correlations, with $R^2 < 0.7$. In general, both models normally fit below 0.7, while the quadratic models are the best for most of the properties. The graphical prediction is shown in Fig. 6.

Property	Models	Equations	R	R^2	S_E	F	p -value
BP	Linear	$126.4892 + 3.2788 \times M_1$	0.977	0.955	0.224	213.946	4.4527E-08
	Quadratic	$181.3617 + 2.0014 \times M_1 + 0.0064 \times M_1^2$	0.977	0.957	26.826	101.284	6.75386E-07
ENP	Linear	$32.7574 + 0.4119 \times M_1$	0.939	0.881	0.047	74.579	5.98915E-06
	Quadratic	$52.0149 + -0.0365 \times M_1 + 0.0022 \times M_1^2$	0.939	0.897	5.457	39.189	3.61206E-05
FP	Linear	$32.6748 + 1.8784 \times M_1$	0.933	0.871	0.228	67.571	9.27521E-06
	Quadratic	$139.9623 + -0.6192 \times M_1 + 0.0125 \times M_1^2$	0.933	0.893	25.461	37.770	4.1908E-05
MR	Linear	$35.1149 + 0.4089 \times M_1$	0.906	0.821	0.060	45.939	4.87181E-05
	Quadratic	$41.1357 + 0.2688 \times M_1 + 0.0007 \times M_1^2$	0.906	0.822	7.368	20.872	0.000416716
POL	Linear	$13.9414 + 0.1619 \times M_1$	0.906	0.821	0.023	46.065	4.81598E-05
	Quadratic	$16.2534 + 0.1081 \times M_1 + 0.0003 \times M_1^2$	0.906	0.822	2.914	20.918	0.000413333
MW	Linear	$153.6176 + 1.5096 \times M_1$	0.898	0.807	0.232	42.072	7.0158E-05
	Quadratic	$147.1422 + 1.6603 \times M_1 + -0.0008 \times M_1^2$	0.898	0.808	28.527	18.947	0.000594365
HAC	Linear	$2.8147 + 0.1633 \times M_1$	0.980	0.960	0.010	246.038	2.27386E-08
	Quadratic	$0.4159 + 0.2191 \times M_1 + -0.0003 \times M_1^2$	0.980	0.962	1.249	115.763	3.79194E-07
Com	Linear	$9.5081 + 3.6657 \times M_1$	0.965	0.932	0.312	137.303	3.6534E-07
	Quadratic	$-31.1430 + 4.6120 \times M_1 + -0.0047 \times M_1^2$	0.965	0.933	38.100	62.683	5.2094E-06
MV	Linear	$131.1677 + 0.8700 \times M_1$	0.811	0.658	0.197	19.318	0.001344888
	Quadratic	$77.2248 + 2.1258 \times M_1 + -0.0063 \times M_1^2$	0.811	0.678	23.547	9.515	0.006020919

Table 4. Statistical parameters and regression models for $M_1(G)$.

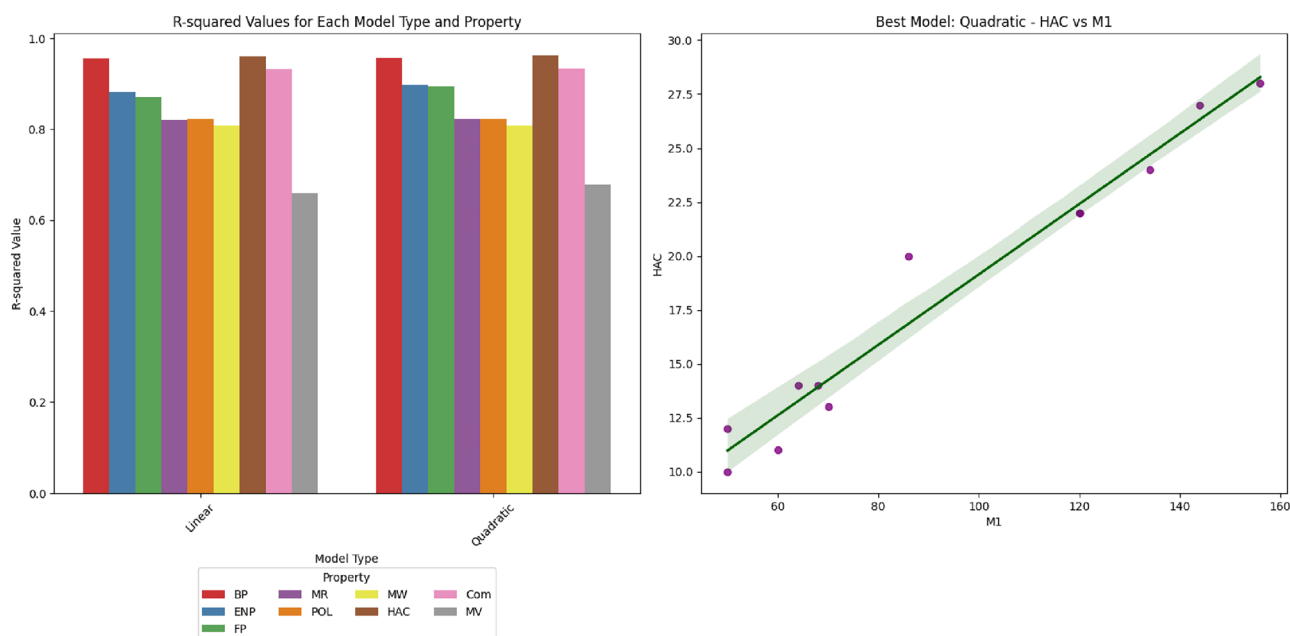


Fig. 5. Comparison of fits of models for HAC prediction based on $M_1(G)$.

Table 6 presents a comparison of linear versus quadratic regression models against each other for the prediction of physicochemical properties using $H(G)$. The quadratic model gives better performance with considerably high values of R^2 and lower values of standard error (SE) in the case of most of the properties. The boiling point and heavy atom count are best fitted by quadratic models with $R^2 = 0.917$ and $R^2 = 0.972$, respectively, with high statistical significance $p < 10^{-5}$. Quadratic equations also provide very good predictions for properties such as molar refraction and polarization. The opposite holds for the enthalpy of vaporization, ENP; both have fair correlations, $R^2 = 0.831$. Globally, quadratic models are the best predictors for the data. The graphical explanation of the comparison is shown in Fig. 7.

Property	Models	Equations	R	R ²	S _E	F	p-value
BP	Linear	157.6296 + 2.5346 × M ₂	0.977	0.955	0.172	216.954	4.16424E-08
	Quadratic	182.7138 + 2.0215 × M ₂ + 0.0022 × M ₂ ²	0.977	0.957	27.046	99.566	7.27026E-07
ENP	Linear	36.4880 + 0.3200 × M ₂	0.944	0.891	0.035	82.224	3.8679E-06
	Quadratic	49.2737 + 0.0585 × M ₂ + 0.0011 × M ₂ ²	0.944	0.904	5.268	42.391	2.62732E-05
FP	Linear	50.5209 + 1.4520 × M ₂	0.933	0.871	0.176	67.859	9.103E-06
	Quadratic	120.2526 + 0.0255 × M ₂ + 0.0060 × M ₂ ²	0.933	0.889	25.978	36.104	5.02221E-05
MR	Linear	39.8439 + 0.3083 × M ₂	0.884	0.781	0.051	35.830	0.00013462
	Quadratic	37.5926 + 0.3544 × M ₂ + −0.0002 × M ₂ ²	0.884	0.782	8.165	16.158	0.001
POL	Linear	15.8151 + 0.1221 × M ₂	0.884	0.782	0.020	35.887	0.0001
	Quadratic	14.8643 + 0.1416 × M ₂ + −0.0001 × M ₂ ²	0.884	0.782	3.230	16.188	0.001
MW	Linear	169.4657 + 1.1531 × M ₂	0.888	0.789	0.188	37.463	0.0001
	Quadratic	143.4164 + 1.6860 × M ₂ + −0.0023 × M ₂ ²	0.888	0.792	29.637	17.222	0.0008
HAC	Linear	4.4529 + 0.1254 × M ₂	0.974	0.949	0.009	187.427	8.38367E-08
	Quadratic	1.6039 + 0.1837 × M ₂ + −0.0002 × M ₂ ²	0.974	0.953	1.390	92.565	9.94658E-07
Com	Linear	42.3828 + 2.8515 × M ₂	0.971	0.944	0.218	170.113	1.32945E-07
	Quadratic	−7.7873 + 3.8779 × M ₂ + −0.0043 × M ₂ ²	0.971	0.947	33.878	80.470	1.81026E-06
MV	Linear	143.0209 + 0.6395 × M ₂	0.772	0.596	0.166	14.761	0.003
	Quadratic	85.2153 + 1.8221 × M ₂ + −0.0050 × M ₂ ²	0.772	0.638	24.975	7.959	0.010

Table 5. Statistical parameters and regression models for M₂(G).

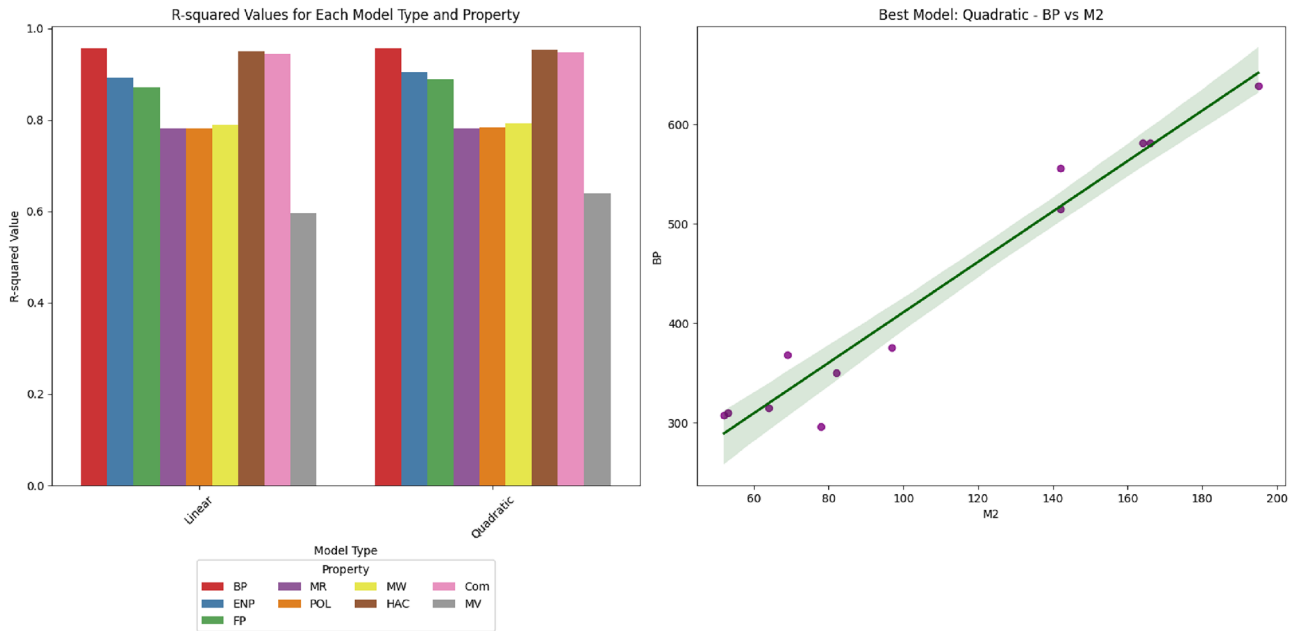


Fig. 6. Comparison of fits of models for BP prediction based on M₂(G).

This Table 7 collects some statistical parameters and also the regression model in the different properties as functions of F(G). Each property is modeled by a linear and quadratic model. The best model between these two is the one that realizes higher R^2 and the lowest p -values. For the most important properties, the quadratic model resulted in better fitted (higher R and R^2), indicating a nonlinear relation. Properties like BP, ENP, MR, POL, and Com give good (R^2) values greater than 0.9 for a quadratic model, indicating that the predictive accuracy is fairly good. The P -values of such models are also very low, hence statistically significant. In the scatter plot of highlighted properties, the quadratic model comes out to be the best for many properties, something as HAC. The graphical explanation of the comparison is shown in Fig. 8.

Property	Models	Equations	R	R^2	S_E	F	p -value
BP	Linear	$91.1873 + 39.9538 \times H$	0.957	0.917	3.796	110.724	9.94653E-07
	Quadratic	$49.1682 + 50.2834 \times H + -0.5646 \times H^2$	0.957	0.917	37.255	50.350	1.29763E-05
ENP	Linear	$28.7006 + 4.9746 \times H$	0.911	0.831	0.707	49.419	3.58412E-05
	Quadratic	$ENP = 29.2871 + 4.8304 \times H + 0.0079 \times H^2$	0.911	0.831	6.976	22.240	0.0003
FP	Linear	$12.3942 + 22.8961 \times H$	0.914	0.836	3.198	51.257	3.07024E-05
	Quadratic	$71.7364 + 8.3079 \times H + 0.7974 \times H^2$	0.914	0.841	31.103	23.826	0.0002
MR	Linear	$28.6535 + 5.2236 \times H$	0.930	0.866	0.648	64.839	1.11214E-05
	Quadratic	$50.8435 + -0.2314 \times H + 0.2982 \times H^2$	0.930	0.878	6.096	32.560	7.57427E-05
POL	Linear	$11.3792 + 2.0690 \times H$	0.931	0.867	0.256	65.265	1.08065E-05
	Quadratic	$20.1197 + -0.0796 \times H + 0.1174 \times H^2$	0.931	0.879	2.407	32.754	7.39874E-05
MW	Linear	$133.2252 + 18.8787 \times H$	0.903	0.817	2.825	44.646	5.48747E-05
	Quadratic	$158.0742 + 12.7701 \times H + 0.3339 \times H^2$	0.903	0.818	27.771	20.240	0.0004
HAC	Linear	$0.8442 + 2.0147 \times H$	0.972	0.945	0.152	174.128	1.19004E-07
	Quadratic	$-4.2766 + 3.2736 \times H + -0.0688 \times H^2$	0.972	0.950	1.437	86.329	1.34103E-06
Com	Linear	$-26.1946 + 44.2280 \times H$	0.936	0.877	5.230	71.513	7.21829E-06
	Quadratic	$-192.8238 + 85.1908 \times H + -2.2389 \times H^2$	0.936	0.887	49.479	35.335	5.4731E-05
MV	Linear	$112.8041 + 11.6532 \times H$	0.874	0.764	2.046	32.421	0.0001
	Quadratic	$121.5129 + 9.5123 \times H + 0.1170 \times H^2$	0.874	0.764	20.162	14.616	0.001

Table 6. Statistical parameters and regression models for $H(G)$.

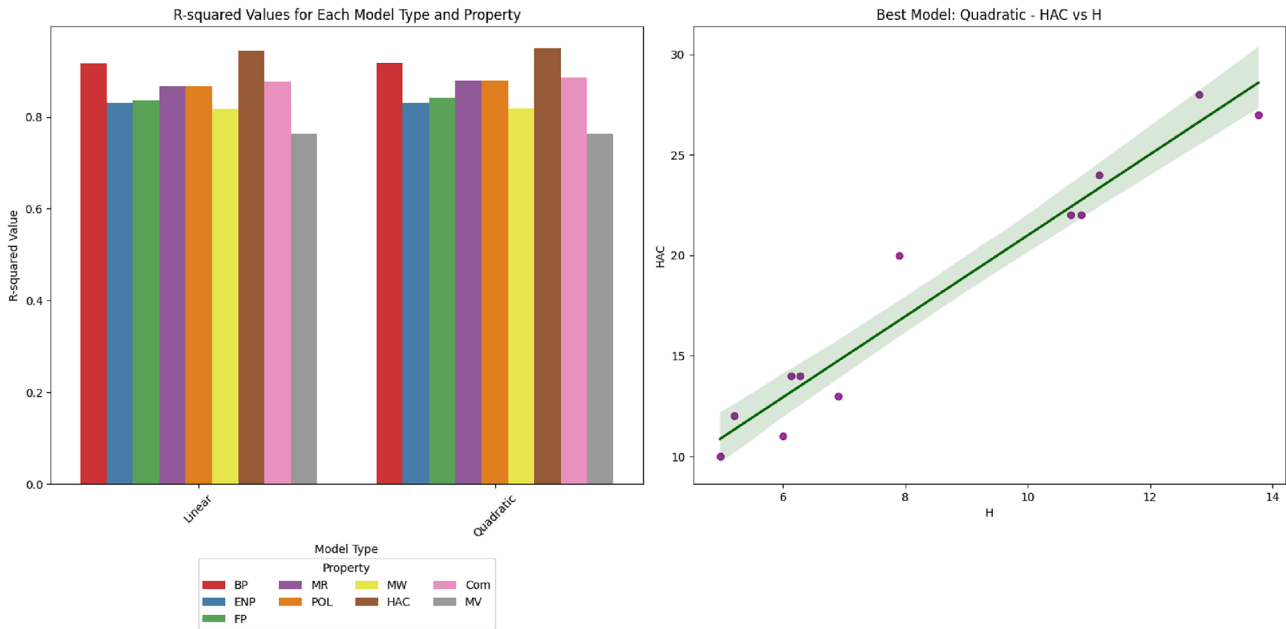


Fig. 7. Comparison of fits of models for HAC prediction based on $H(G)$.

The Table 8 below shows the statistical parameters and regression models for different properties as functions of $SS(G)$, for which both linear and quadratic models have been fitted, considering that the best fit always correlates with higher R^2 values with lower p -values. Most of the properties have a better fit, as by R^2 values closer to 1, with quadratic models like BP, FP, POL, and MW. On the other hand, for properties such as HIAC, ENP, and MR, the quadratic model is highly predictive with large p -values, indicating thereby how appropriate this approach is. In general, the quadratic models do better for this data set. The graphical comparison is shown in Fig. 9.

Table 9 above shows, for a number of properties, regression models as functions of the ABC index. In general, quadratic models are better fits, as measured by the higher p -value. The quadratic model for BP has

Property	Models	Equations	R	R^2	S_E	F	p -value
BP	Linear	$121.5524 + 1.2663 \times F$	0.969	0.939	0.101	156.558	1.96982E-07
	Quadratic	$174.9928 + 0.7983 \times F + 0.0009 \times F^2$	0.969	0.942	31.245	73.479	2.6641E-06
ENP	Linear	$31.9948 + 0.1596 \times F$	0.934	0.873	0.019	69.298	8.29744E-06
	Quadratic	$50.7944 + -0.0050 \times F + 0.0003 \times F^2$	0.934	0.890	5.620	36.697	4.70458E-05
FP	Linear	$28.9742 + 0.7290 \times F$	0.930	0.865	0.090	64.326	1.15156E-05
	Quadratic	$127.8348 + -0.1368 \times F + 0.0016 \times F^2$	0.930	0.887	26.159	35.543	5.34636E-05
MR	Linear	$35.0983 + 0.1555 \times F$	0.885	0.783	0.025	36.142	0.000130037
	Quadratic	$31.6974 + 0.1853 \times F + -0.0001 \times F^2$	0.885	0.783	8.135	16.314	0.001015862
POL	Linear	$13.9361 + 0.0616 \times F$	0.885	0.783	0.010	36.196	0.00012926
	Quadratic	$12.5102 + 0.0741 \times F + -0.0000 \times F^2$	0.885	0.784	3.218	16.344	0.001009175
MW	Linear	$150.3524 + 0.5870 \times F$	0.897	0.805	0.091	41.543	7.39097E-05
	Quadratic	$114.1775 + 0.9038 \times F + -0.0006 \times F^2$	0.897	0.810	28.365	19.215	0.000564693
HAC	Linear	$2.5279 + 0.0632 \times F$	0.974	0.950	0.004	191.872	7.49683E-08
	Quadratic	$-0.6319 + 0.0909 \times F + -0.0001 \times F^2$	0.974	0.953	1.388	92.836	9.82213E-07
Com	Linear	$-0.9150 + 1.4356 \times F$	0.971	0.943	0.111	165.779	1.50252E-07
	Quadratic	$-50.6254 + 1.8710 \times F + -0.0008 \times F^2$	0.971	0.944	34.623	76.852	2.20184E-06
MV	Linear	$131.6511 + 0.3287 \times F$	0.787	0.620	0.081	16.349	0.002348309
	Quadratic	$44.1036 + 1.0954 \times F + -0.0015 \times F^2$	0.787	0.681	23.443	9.640	0.005785411

Table 7. Statistical parameters and regression models for $F(G)$.

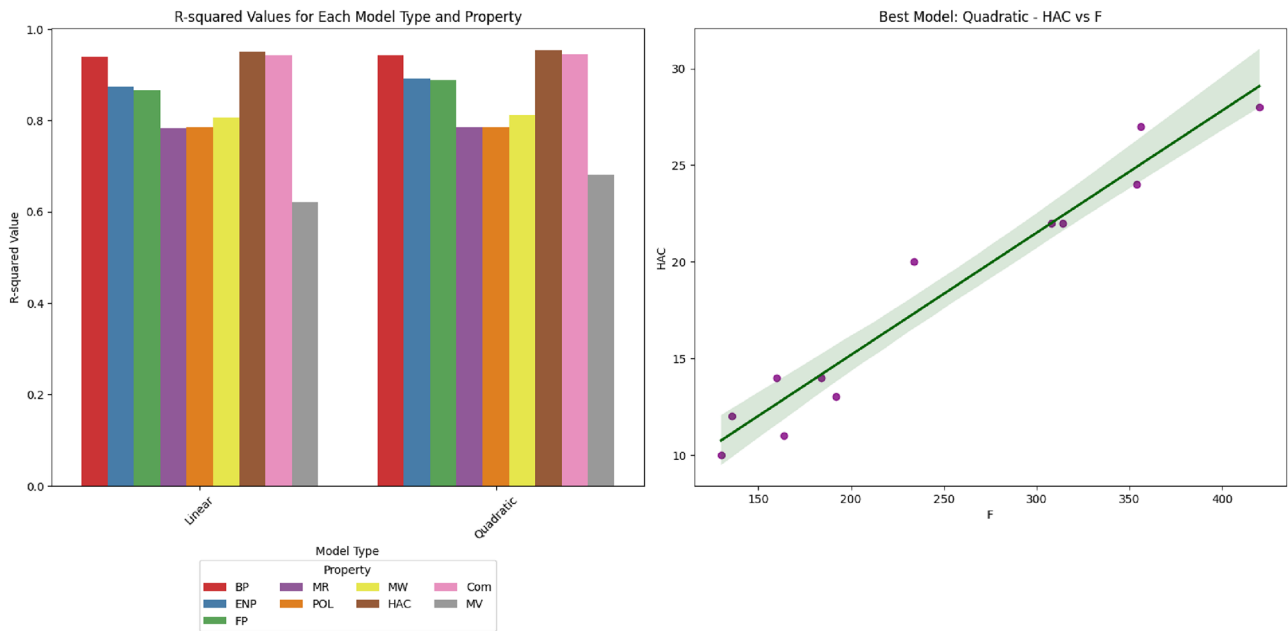


Fig. 8. Comparison of fits of models for HAC prediction based on $F(G)$.

an $R^2 = 0.947$ with a $p = 1.68 \times 10^{-6}$, where as the quadratic model for ENP has an $R^2 = 0.863$ with a $p = 1.82 \times 10^{-5}$. The quadratic model systematically demonstrates an outstanding fit for HIAC with $R^2 = 0.960$ and $p = 5.25 \times 10^{-7}$; hence, it was the best model for this property. Most of the other properties, such as FP, POL, and MW, also presented high R^2 values in their respective quadratic models, confirming their accuracy and suitability. The graphical comparison is shown in Fig. 10.

The Table 10 below depicts the quality of fit tests utilizing linear and quadratic regression models for several properties with $RI(G)$. From the rather higher values, it is obvious that in most cases the quadratic model fits best. Most importantly, the best fit is manifested by HAC with $R^2 = 0.948$, followed by BP with $R^2 = 0.920$,

Property	Models	Equations	R	R^2	S_E	F	p -value
BP	Linear	$131.0827 + 14.7467 \times SS$	0.976	0.953	1.033	203.639	5.64079E-08
	Quadratic	$172.1350 + 10.3668 \times SS + 0.0998 \times SS^2$	0.976	0.954	27.783	94.123	9.25859E-07
ENP	Linear	$33.4085 + 1.8488 \times SS$	0.936	0.876	0.219	70.869	7.51366E-06
	Quadratic	$49.4323 + 0.1392 \times SS + 0.0390 \times SS^2$	0.936	0.886	5.720	35.272	5.51201E-05
FP	Linear	$35.6829 + 8.4300 \times SS$	0.930	0.865	1.051	64.263	1.15646E-05
	Quadratic	$133.7198 + -2.0296 \times SS + 0.2384 \times SS^2$	0.930	0.884	26.574	34.305	6.15824E-05
MR	Linear	$35.2972 + 1.8583 \times SS$	0.914	0.836	0.259	51.148	3.09804E-05
	Quadratic	$46.7619 + 0.6351 \times SS + 0.0279 \times SS^2$	0.914	0.841	6.964	23.899	0.000250914
POL	Linear	$14.0129 + 0.7360 \times SS$	0.914	0.836	0.102	51.334	3.05072E-05
	Quadratic	$18.4904 + 0.2582 \times SS + 0.0109 \times SS^2$	0.914	0.841	2.754	23.964	0.000248375
MW	Linear	$155.7071 + 6.7907 \times SS$	0.898	0.806	1.052	41.661	7.30489E-05
	Quadratic	$165.1202 + 5.7864 \times SS + 0.0229 \times SS^2$	0.898	0.806	28.630	18.777	0.000614092
HAC	Linear	$3.0307 + 0.7351 \times SS$	0.980	0.960	0.047	242.616	2.43266E-08
	Quadratic	$0.7372 + 0.9798 \times SS + -0.0056 \times SS^2$	0.980	0.961	1.260	113.633	4.10941E-07
Com	Linear	$16.4441 + 16.3988 \times SS$	0.959	0.920	1.527	115.184	8.28581E-07
	Quadratic	$-33.3993 + 21.7166 \times SS + -0.1212 \times SS^2$	0.959	0.921	41.253	52.806	1.06544E-05
MV	Linear	$130.9741 + 3.9821 \times SS$	0.825	0.680	0.862	21.329	0.000952867
	Quadratic	$101.1432 + 7.1647 \times SS + -0.0725 \times SS^2$	0.825	0.686	23.253	9.873	0.005376017

Table 8. Statistical parameters and regression models for SS(G).

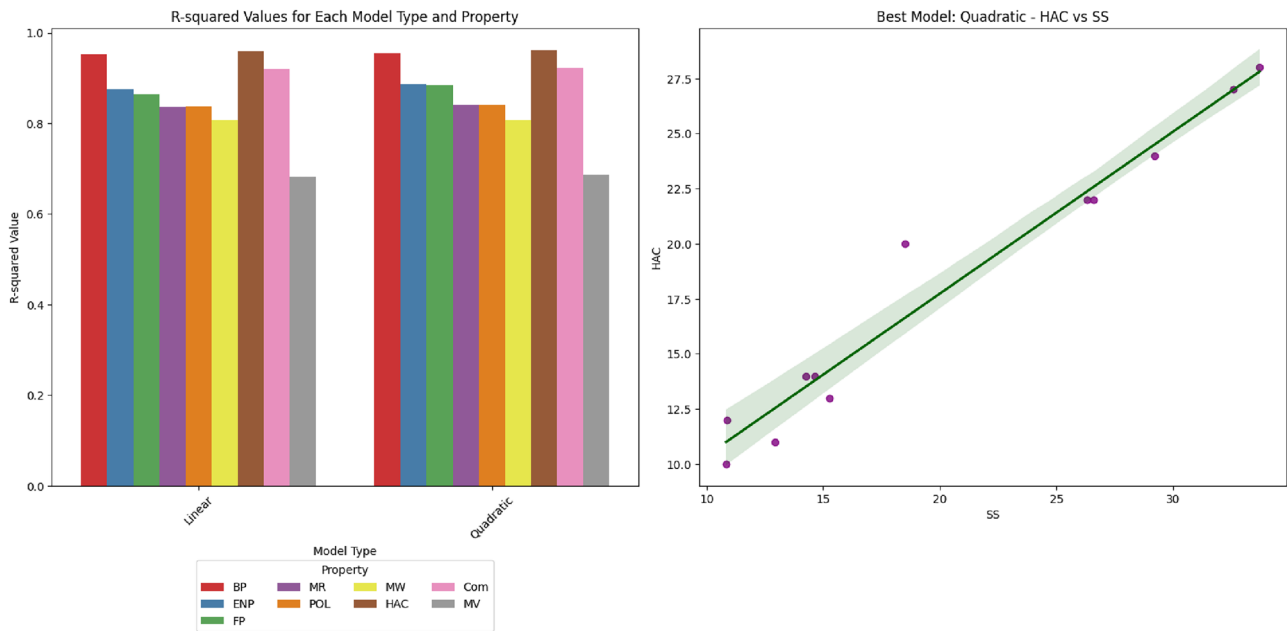


Fig. 9. Comparison of fits of models for HAC prediction based on SS(G).

and POL with $R^2 = 0.868$. The results point to how well the quadratic model fits to catch nonlinear relationships and thus serves as the best predictive model across the data. The graphical comparison is shown in Fig. 11.

The Table 11 shows quadratic regression models fitting best for most properties of SS(G), as evidenced by high values of R^2 . The best fittings are for HAC, with $R^2 = 0.957$, MW, with $R^2 = 0.977$, and BP, with $R^2 = 0.984$. This suggests that the quadratic model does a better job of capturing nonlinear trends in the data than the linear model. The graphical comparison is shown in Fig. 12.

Table 12 presents the regression models together with their statistical parameters for different properties versus GA(G). Most of the properties were tested both for linear and quadratic models, and the Table 12 gives values of the equation, correlation coefficient R , determination coefficient R^2 , standard error S_E , and p -values.

Property	Models	Equations	R	R ²	S _E	F	p-value
BP	Linear	97.8098 + 23.7906 × ABC	0.973	0.947	1.776	179.413	1.03236E-07
	Quadratic	140.5329 + 17.3140 × ABC + 0.2162 × ABC ²	0.973	0.947	29.664	82.009	1.66982E-06
ENP	Linear	29.3842 + 2.9721 × ABC	0.929	0.864	0.371	63.941	1.18225E-05
	Quadratic	47.7947 + 0.1812 × ABC + 0.0932 × ABC ²	0.929	0.873	6.055	30.988	9.20529E-05
FP	Linear	16.5105 + 13.6107 × ABC	0.928	0.861	1.727	62.084	1.34481E-05
	Quadratic	134.3638 + −4.2550 × ABC + 0.5964 × ABC ²	0.928	0.877	27.290	32.295	7.82283E-05
MR	Linear	30.4808 + 3.0422 × ABC	0.925	0.855	0.394	59.413	1.62837E-05
	Quadratic	47.7026 + 0.4315 × ABC + 0.0871 × ABC ²	0.925	0.862	6.478	28.326	0.000130757
POL	Linear	12.1047 + 1.2049 × ABC	0.925	0.856	0.155	59.685	1.5963E-05
	Quadratic	18.8590 + 0.1810 × ABC + 0.0342 × ABC ²	0.925	0.863	2.560	28.430	0.000128908
MW	Linear	138.9778 + 11.0552 × ABC	0.903	0.8160	1.659	44.361	5.6352E-05
	Quadratic	160.5555 + 7.7842 × ABC + 0.1092 × ABC ²	0.903	0.816	27.868	20.068	0.000481673
HAC	Linear	1.3250 + 1.1892 × ABC	0.979	0.959	0.076	238.599	2.63638E-08
	Quadratic	−1.2836 + 1.5847 × ABC + −0.0132 × ABC ²	0.979	0.960	1.276	110.735	4.59532E-07
Com	Linear	−19.1974 + 26.3594 × ABC	0.952	0.907	2.658	98.324	1.71745E-06
	Quadratic	−84.2710 + 36.2241 × ABC + −0.3293 × ABC ²	0.952	0.909	44.386	45.0007	2.05922E-05
MV	Linear	119.3617 + 6.6106 × ABC	0.846	0.716	1.315	25.259	0.000517268
	Quadratic	96.7676 + 10.0357 × ABC + −0.1143 × ABC ²	0.846	0.718	22.048	11.486	0.00333078

Table 9. Statistical parameters and regression models for ABC.

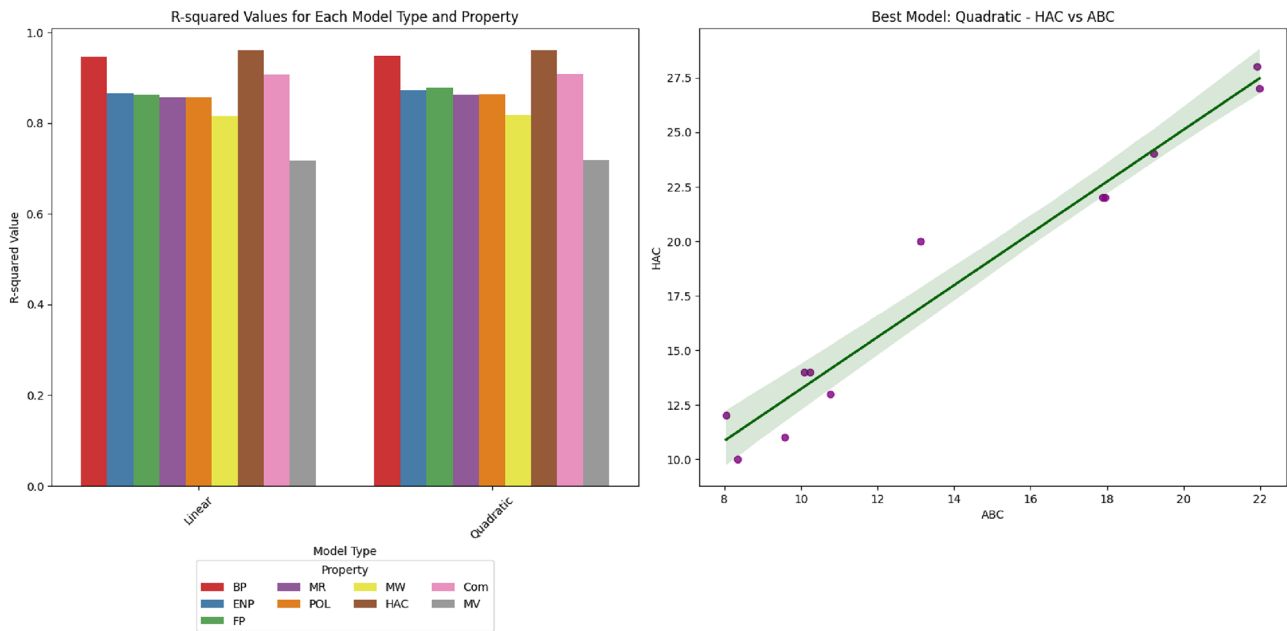


Fig. 10. Comparison of fits of models for HAC prediction based on ABC(G).

Quadratic models generally exhibit better R^2 values compared to linear models, hence a better fit of the model. For example, the quadratic model for HAC has a high value of $R^2 = 0.978$, compared to $R^2 = 0.897$ for the linear model. All the models have very low values of p, hence they are statistically significant. Notably, properties such as ENP and POL also exhibit excellent fits with R^2 values surpassing 0.9. The graphical comparison is shown in Fig. 13.

Table 13 gives statistical parameters and regression models of different properties concerning relation HZ(G). Quadratic models generally provide higher values R^2 than linear models, which express better fittings. For example, a quadratic model of the HAC property yields an R^2 of 0.951 against that of the linear model, 0.917, indicating that the performance is better. The computed values in all the models were much less, indicating

Property	Models	Equations	R	R^2	S_E	F	p -value
BP	Linear	$74.0206 + 39.9856 \times RI$	0.959	0.920	3.710	116.114	7.9826E-07
	Quadratic	$42.8556 + 47.2418 \times RI + -0.3795 \times RI^2$	0.959	0.921	36.546	52.499	1.09156E-05
ENP	Linear	$26.5630 + 4.9786 \times RI$	0.913	0.834	0.700	50.577	3.24946E-05
	Quadratic	$30.3521 + 4.0964 \times RI + 0.0461 \times RI^2$	0.913	0.835	6.903	22.808	0.000299283
FP	Linear	$2.2366 + 22.9500 \times RI$	0.917	0.842	3.136	53.532	2.55132E-05
	Quadratic	$81.9137 + 4.3986 \times RI + 0.9703 \times RI^2$	0.917	0.848	30.338	25.273	0.000202887
MR	Linear	$26.4929 + 5.2184 \times RI$	0.930	0.866	0.647	64.976	1.10189E-05
	Quadratic	$50.5894 + -0.3920 \times RI + 0.2935 \times RI^2$	0.930	0.878	6.111	32.387	7.73542E-05
POL	Linear	$10.5236 + 2.0670 \times RI$	0.931	0.867	0.255	65.392	1.07144E-05
	Quadratic	$20.0105 + -0.1419 \times RI + 0.1155 \times RI^2$	0.931	0.878	2.413	32.572	7.56313E-05
MW	Linear	$124.7720 + 18.9318 \times RI$	0.907	0.823	2.771	46.645	4.57033E-05
	Quadratic	$150.0256 + 13.0520 \times RI + 0.3075 \times RI^2$	0.907	0.824	27.289	21.121	0.00039877
HAC	Linear	$-0.0096 + 2.0150 \times RI$	0.973	0.948	0.149	182.666	9.47723E-08
	Quadratic	$-5.2643 + 3.2385 \times RI + -0.0640 \times RI^2$	0.973	0.952	1.414	89.385	1.15549E-06
Com	Linear	$-45.6506 + 44.3138 \times RI$	0.939	0.882	5.107	75.262	5.75043E-06
	Quadratic	$-221.6903 + 85.3013 \times RI + -2.1439 \times RI^2$	0.939	0.891	48.537	36.897	4.60345E-05
MV	Linear	$108.0125 + 11.6385 \times RI$	0.874	0.764	2.045	32.385	0.000200829
	Quadratic	$111.6924 + 10.7817 \times RI + 0.0448 \times RI^2$	0.874	0.764	20.183	14.577	0.001503624

Table 10. Statistical parameters and regression models for $RI(G)$.

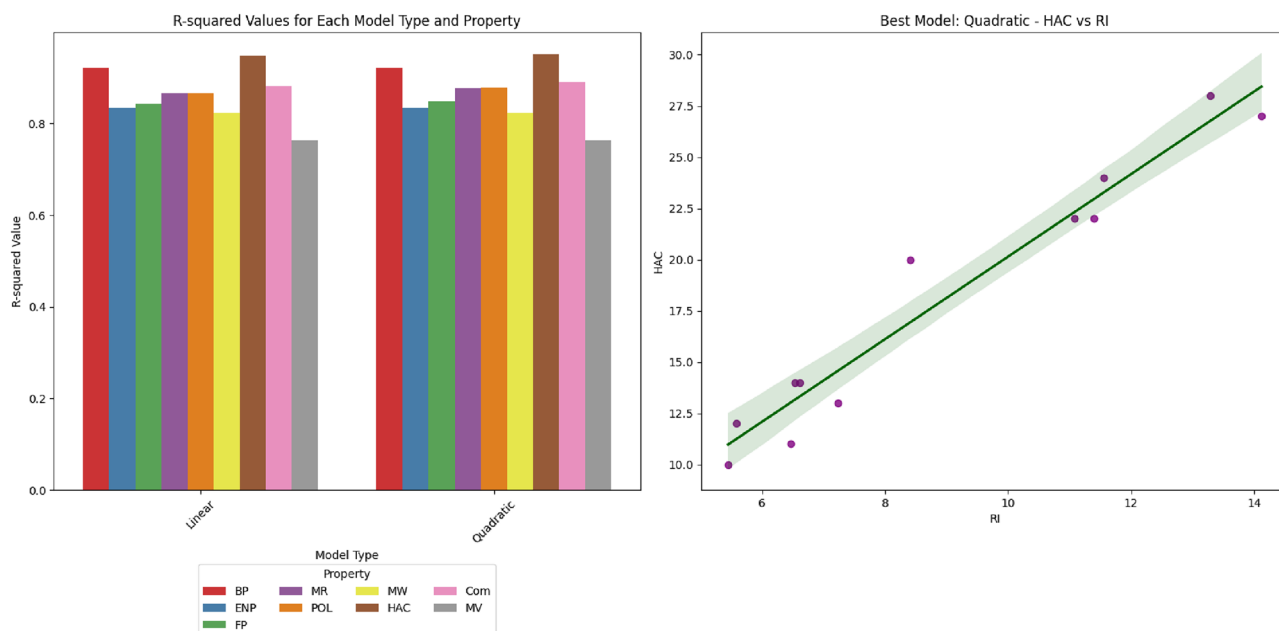


Fig. 11. Comparison of fits of models for HAC prediction based on $RI(G)$.

that they are statistically significant. In general, the quadratic models provide a much better explanation for the properties with much higher R^2 values for most properties and smaller standard errors. In other words, the quadratic models can normally explain property $HZ(G)$ more effectively. The graphical comparison is shown in Fig. 14.

Comparison of actual and prediction values of drugs properties

Table 14 compares the predicted boiling point, enthalpy of vaporization and Flash point from various models linear and quadratic of different types of drug molecules to real values. In each case, the quadratic model provided the closest yield on boiling point and enthalpy of vaporization to real values, especially for complex

Property	Models	Equations	<i>R</i>	<i>R</i> ²	<i>S_E</i>	<i>F</i>	<i>p</i> -value
BP	Linear	101.7231 + 36.4005 × <i>SC</i>	0.966	0.934	3.051	142.273	3.09181E-07
	Quadratic	104.3140 + 35.7944 × <i>SC</i> + 0.0312 × <i>SC</i> ²	0.966	0.934	33.330	64.026	4.76533E-06
ENP	Linear	29.9218 + 4.5421 × <i>SC</i>	0.922	0.850	0.601	57.110	1.93205E-05
	Quadratic	39.0251 + 2.4126 × <i>SC</i> + 0.1096 × <i>SC</i> ²	0.922	0.853	6.513	26.174	0.000177399
FP	Linear	18.6871 + 20.8318 × <i>SC</i>	0.922	0.850	2.766	56.720	1.99E-05
	Quadratic	105.2521 + 0.5814 × <i>SC</i> + 1.0424 × <i>SC</i> ²	0.922	0.860	29.195	27.648	0.000143623
MR	Linear	30.5754 + 4.6992 × <i>SC</i>	0.927	0.860	0.598	61.708	1.38082E-05
	Quadratic	51.0181 + −0.0830 × <i>SC</i> + 0.2462 × <i>SC</i> ²	0.927	0.871	6.271	30.528	9.7622E-05
POL	Linear	12.1413 + 1.8612 × <i>SC</i>	0.928	0.861	0.236	62.049	1.34805E-05
	Quadratic	20.1819 + −0.0197 × <i>SC</i> + 0.0968 × <i>SC</i> ²	0.928	0.872	2.477	30.669	9.58714E-05
MW	Linear	139.7475 + 17.0301 × <i>SC</i>	0.903	0.815	2.557	44.335	5.6493E-05
	Quadratic	165.5208 + 11.0009 × <i>SC</i> + 0.3103 × <i>SC</i> ²	0.903	0.817	27.839	20.119	0.000477199
HAC	Linear	1.4517 + 1.8272 × <i>SC</i>	0.977	0.954	0.125	210.371	4.82729E-08
	Quadratic	−2.1712 + 2.6747 × <i>SC</i> + −0.0436 × <i>SC</i> ²	0.977	0.957	1.336	100.527	6.97588E-07
Com	Linear	−14.8111 + 40.3253 × <i>SC</i>	0.946	0.895	4.365	85.328	3.27266E-06
	Quadratic	−126.6197 + 66.4809 × <i>SC</i> + −1.3463 × <i>SC</i> ²	0.946	0.899	46.613	40.383	3.19925E-05
MV	Linear	118.4792 + 10.3310 × <i>SC</i>	0.858	0.737	1.950	28.053	0.000349277
	Quadratic	115.1736 + 11.1042 × <i>SC</i> + −0.0398 × <i>SC</i> ²	0.858	0.737	21.301	12.627	0.002442658

Table 11. Statistical parameters and regression models for SC(G).

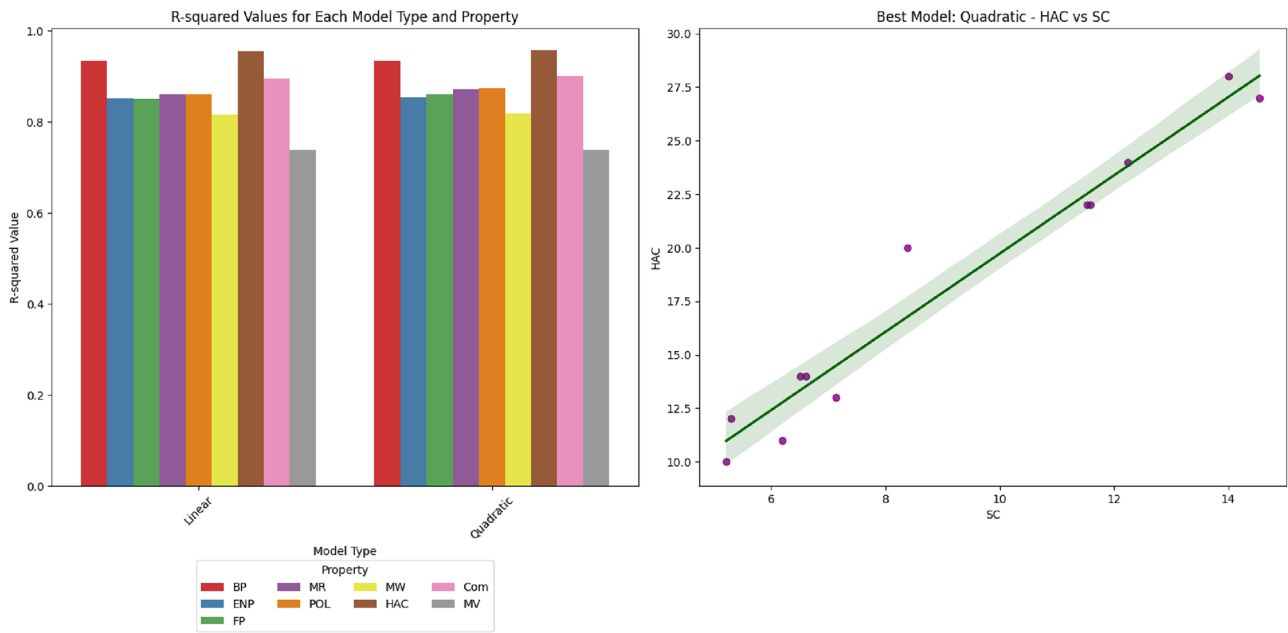


Fig. 12. Comparison of fits of models for HAC prediction based on SC(G).

drug molecules . This means that the quadratic model handles the representation of drug molecular complexity better than the other models. The graphical comparison is shown in Fig. 15.

The Table 15 below shows the comparison of the predicted and experimental values of Molar refractivity, Polarizability and Molar weight of various drug molecules using different models: linear, and quadratic. The quadratic model generally shows closer alignment with the actual values of both properties, especially for the most complex molecules. This trend in general points toward the better performance of the quadratic model in correct property prediction, although some deviations remain, more precisely concerning molar refraction. The graphical comparison is shown in Fig. 16.

Property	Models	Equations	R	R ²	S _E	F	p-value
BP	Linear	123.7566 + 16.4544 × GA	0.972	0.945	1.254	172.148	1.25646E-07
	Quadratic	146.7483 + 13.8036 × GA + 0.0659 × GA ²	0.972	0.945	30.384	77.962	2.07164E-06
ENP	Linear	32.5776 + 2.0582 × GA	0.930	0.864	0.257	64.065	1.17227E-05
	Quadratic	44.9872 + 0.6275 × GA + 0.0355 × GA ²	0.930	0.870	6.117	30.275	0.000100858
FP	Linear	31.5685 + 9.4023 × GA	0.925	0.857	1.213	60.069	1.5525E-05
	Quadratic	120.9621 + −0.9042 × GA + 0.2561 × GA ²	0.925	0.871	28.014	30.417	9.90279E-05
MR	Linear	33.8986 + 2.0988 × GA	0.921	0.849	0.279	56.542	2.01714E-05
	Quadratic	49.9067 + 0.2531 × GA + 0.0459 × GA ²	0.921	0.858	6.580	27.313	0.000150561
POL	Linear	13.4582 + 0.8312 × GA	0.922	0.850	0.110	56.801	1.97789E-05
	Quadratic	19.7413 + 0.1068 × GA + 0.0180 × GA ²	0.922	0.858	2.601	27.410	0.000148517
MW	Linear	151.5244 + 7.6201 × GA	0.899	0.808	1.172	42.273	6.87981E-05
	Quadratic	171.5341 + 5.3132 × GA + 0.0573 × GA ²	0.899	0.809	28.406	19.146	0.000572193
HAC	Linear	2.6195 + 0.8227 × GA	0.978	0.957	0.054	227.828	3.29252E-08
	Quadratic	−0.0736 + 1.1332 × GA + −0.0077 × GA ²	0.978	0.959	1.295	107.402	5.2441E-07
Com	Linear	9.3340 + 18.2426 × GA	0.952	0.906	1.849	97.313	1.80062E-06
	Quadratic	−62.9248 + 26.5736 × GA + −0.2070 × GA ²	0.952	0.909	44.322	45.144	2.0325E-05
MV	Linear	126.9924 + 4.5498 × GA	0.841	0.707	0.924	24.224	0.000603237
	Quadratic	113.6293 + 6.0905 × GA + −0.0383 × GA ²	0.8413	0.708	22.422	10.958	0.003874299

Table 12. Statistical parameters and regression models for GA(G).

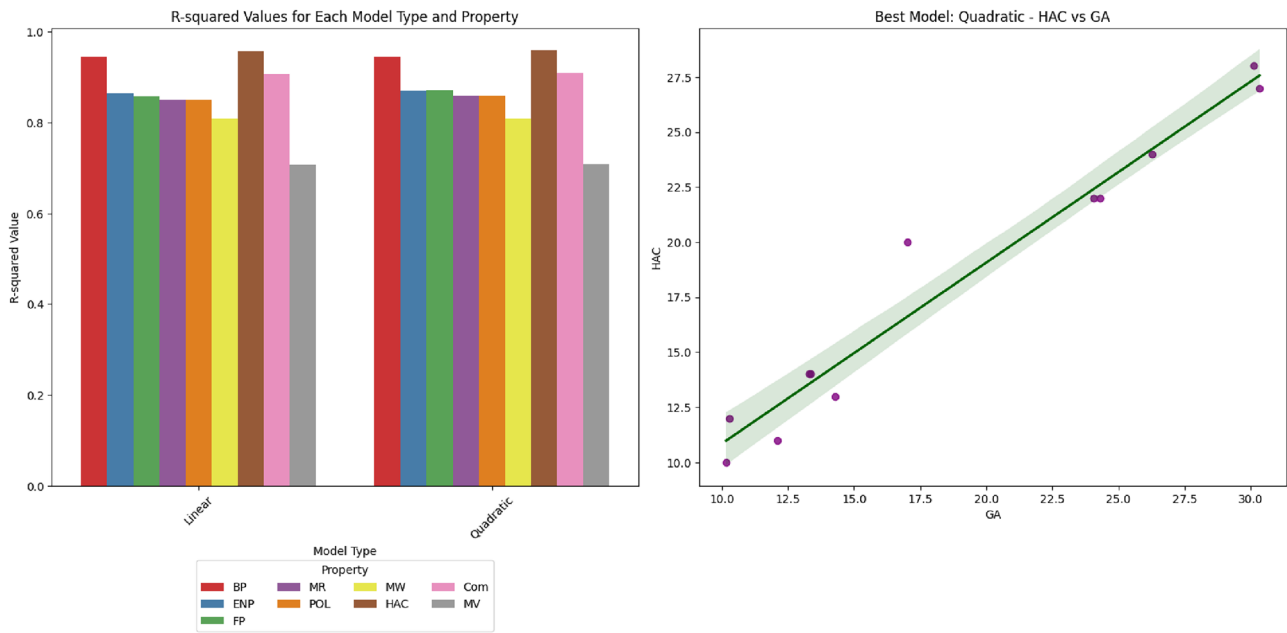


Fig. 13. Comparison of fits of models for HAC prediction based on GA(G).

Table 16 shows that, overall, the quadratic model gives the closest predictions to real values in respect to polarization and molar weight for different drug molecules. While the model generally predicts polarization very well for most drugs considered, significant deviations become apparent for Abamectin, which documents the limitations of the model for some compounds. The linear and logarithmic models tend to underestimate polarization, especially if high values are considered. This is so true, especially for molar weight, since all models' results keep close to the actual values, therefore indicating simpler predictive needs versus polarization. Lastly, according to both properties in this dataset, the quadratic model gives more accurate results. The graphical comparison is shown in Fig. 17.

Property	Models	Equations	<i>R</i>	<i>R</i> ²	<i>S_E</i>	<i>F</i>	<i>p</i> -value
BP	Linear	139.3030 + 0.6340 × <i>HZ</i>	0.974	0.948	0.046	185.514	8.8038E-08
	Quadratic	176.4656 + 0.4589 × <i>HZ</i> + 0.0002 × <i>HZ</i> ²	0.974	0.950	29.009	85.965	1.36546E-06
ENP	Linear	34.2030 + 0.0800 × <i>HZ</i>	0.939	0.883	0.009	75.889	5.54122E-06
	Quadratic	49.7770 + 0.0066 × <i>HZ</i> + 0.0001 × <i>HZ</i> ²	0.939	0.897	5.431	39.617	3.45713E-05
FP	Linear	39.6109 + 0.3641 × <i>HZ</i>	0.932	0.869	0.044	66.531	9.93113E-06
	Quadratic	122.8744 + −0.0283 × <i>HZ</i> + 0.0004 × <i>HZ</i> ²	0.932	0.888	26.008	36.011	5.07411E-05
MR	Linear	37.4476 + 0.0775 × <i>HZ</i>	0.885	0.783	0.012	36.141	0.00013005
	Quadratic	34.4667 + 0.0915 × <i>HZ</i> + −0.0000 × <i>HZ</i> ²	0.885	0.783	8.135	16.311	0.001016468
POL	Linear	14.8663 + 0.0307 × <i>HZ</i>	0.885	0.783	0.005	36.197	0.00012925
	Quadratic	13.6167 + 0.0366 × <i>HZ</i> + −0.0000 × <i>HZ</i> ²	0.885	0.784	3.218	16.342	0.001009636
MW	Linear	159.8652 + 0.2911 × <i>HZ</i>	0.893	0.798	0.046	39.589	9.00062E-05
	Quadratic	128.0136 + 0.4413 × <i>HZ</i> + −0.0002 × <i>HZ</i> ²	0.893	0.802	28.941	18.280	0.000676754
HAC	Linear	3.4807 + 0.0315 × <i>HZ</i>	0.975	0.950	0.002	193.233	7.24812E-08
	Quadratic	0.3749 + 0.0462 × <i>HZ</i> + −0.0000 × <i>HZ</i> ²	0.975	0.954	1.373	94.997	8.89804E-07
Com	Linear	20.4974 + 0.7160 × <i>HZ</i>	0.971	0.944	0.054	170.776	1.3051E-07
	Quadratic	−31.8463 + 0.9627 × <i>HZ</i> + −0.0002 × <i>HZ</i> ²	0.971	0.946	33.9368	80.178	1.83852E-06
MV	Linear	137.3365 + 0.1623 × <i>HZ</i>	0.780	0.608	0.041	15.560	0.002753684
	Quadratic	64.7385 + 0.5044 × <i>HZ</i> + −0.0003 × <i>HZ</i> ²	0.780	0.661	24.186	8.785	0.007661374

Table 13. Statistical parameters and regression models for *HZ*(*G*).

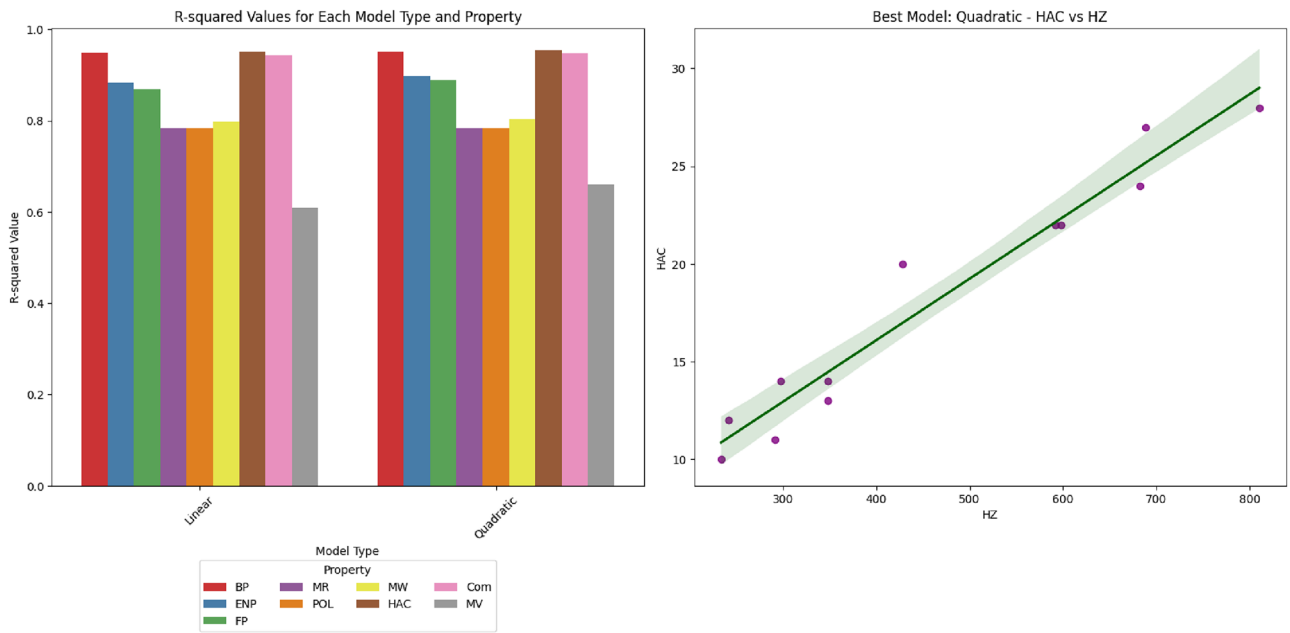


Fig. 14. Comparison of fits of models for *HAC* prediction based on *HZ*(*G*).

Conclusion

Our study presents a systematic QSPR analysis of antibiotic drugs, utilizing degree-based topological indices to predict critical physicochemical properties. The findings highlight the effectiveness of reverse degree-based indices in capturing molecular structural attributes and their correlation with drug characteristics. The application of regression models, particularly quadratic models, demonstrates high predictive accuracy, indicating the potential for structural tailoring in drug development. Future research can extend this approach to a broader range of pharmaceuticals, incorporating machine learning techniques for enhanced predictive performance. These findings highlight the application of the use of QSPR models as a valuable drug development aid that enables scientists to design the required drug candidates at an early development phase. This computational

Properties	Boiling point			Enthalpy of vapourization			Flash point		
	Linear	Quadratic	Actual	Linear	Quadratic	Actual	Linear	Quadratic	Actual
Norfloxacin	519.945	513.69	555.8	82.185	79.315	88.1	258.083	245.658	289.9
Ciprofloxacin	565.848	564.468	581.8	87.952	86.627	91.5	284.38	281.44	305.6
Gemifloxacin	637.982	649.33	638.9	97.014	99.86	99.2	325.705	347.567	340.2
Chlorpyrifos	408.466	400.816	375.9	68.181	65.147	59.9	194.217	179.161	181.1
Dimethoate	290.429	297.432	310.3	53.352	55.69	63.9	126.595	140.252	141.4
Phorate	356.005	352.82	296	61.59	60.24	51.4	164.163	157.868	132.8
Monocrotophos	323.217	324.486	315.1	57.471	57.745	55.6	145.379	147.81	144.4
Thiram	290.429	297.432	307.4	53.352	55.69	54.8	126.595	140.252	139.7
Azoxystrobin	598.636	602.274	581.3	92.071	92.378	86.9	303.164	309.998	305.3
Atrazine	336.332	335.666	368.5	59.119	58.69	61.5	152.892	151.534	176.7
Chlorothalonil	349.448	347.05	350.5	60.767	59.706	59.5	160.406	155.657	153.8
Warfarin	519.945	513.69	515.2	82.185	79.315	82.9	258.083	245.658	188.8

Table 14. Comparison of actual values against predictions for Boiling Point, Enthalpy of Vapourization and Flash point.

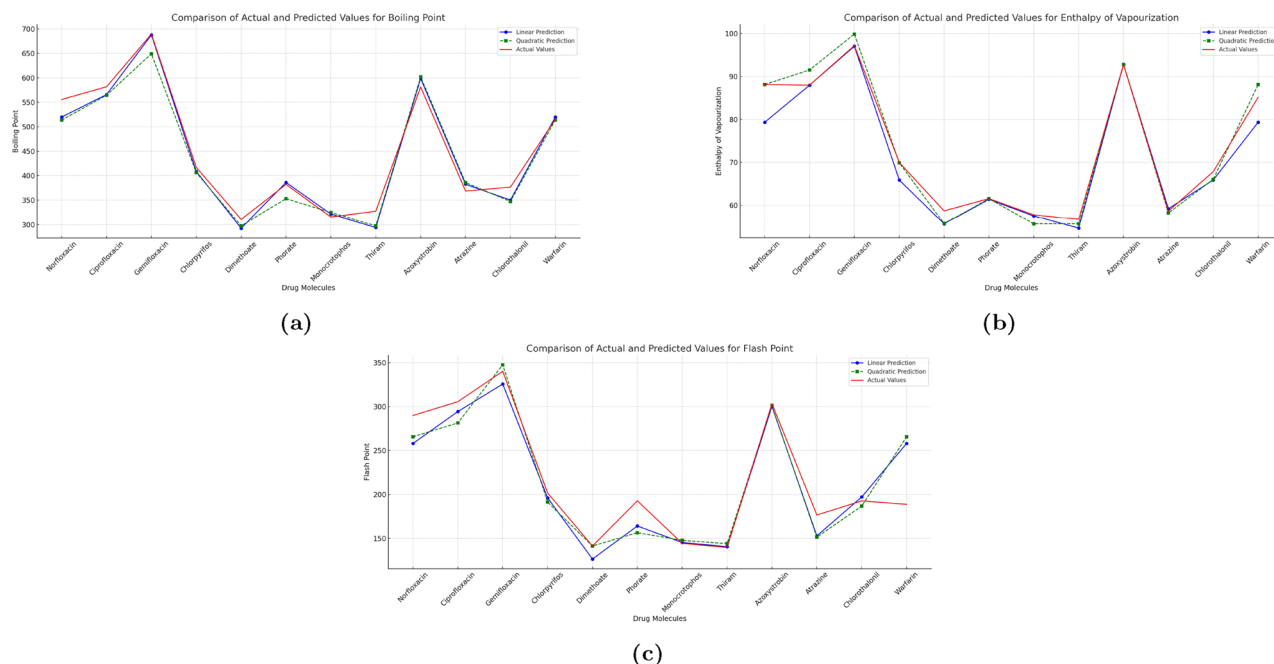


Fig. 15. (a) Comparison of actual and predicted values for BP, (b) Comparison of actual and predicted values for ENV, (c) Comparison of actual and predicted values for FP.

approach not only saves the effort and the cost of laboratory screening but also provides a superior insight into the molecular factors affecting the behavior of the drug.

Future work

For applied drug discovery purposes, merging bioactivity information (QSAR) into QSPR descriptors has the potential for producing multifaceted predictive models enabling lead optimization early in drug development. Further, affiliating these topological descriptors with ADMET (absorption, distribution, metabolism, excretion, and toxicity) traits would lead to a comprehensive predictive system, enhancing new drug candidates' pharmacokinetic profiling. Finally, building an automation pipeline or an analogous software package using this methodology would help medicinal chemists screen drug-like molecules virtually in high-throughput, speeding up drug development.

Properties	Molar refractivity			Polarizability			Molar weight		
	Linear	Quadratic	Actual	Linear	Quadratic	Actual	Linear	Quadratic	Actual
Norfloxacin	84.183	83.472	80.7	33.369	33.545	32	334.77	334.858	319.33
Ciprofloxacin	89.908	89.724	83.3	35.636	36.126	33	355.904	355.258	331.34
Gemifloxacin	98.903	100.104	94.9	39.198	40.418	37.6	389.115	386.68	389.4
Chlorpyrifos	70.28	69.43	77.2	27.864	27.769	30.6	283.443	284.011	350.6
Dimethoate	55.56	56.326	53.7	22.036	22.408	21.3	229.098	228.157	229.26
Phorate	63.738	63.382	68.1	25.274	25.29	27	259.29	259.443	260.35
Monocrotophos	59.649	59.784	50.1	23.655	23.819	19.9	244.194	243.88	223.16
Thiram	55.56	56.326	67.9	22.036	22.408	26.9	229.098	228.157	240.4
Azoxystrobin	93.996	94.358	106.2	37.255	38.041	42.1	371	369.637	403.4
Atrazine	61.284	61.206	58.5	24.303	24.401	23.2	250.232	250.125	215.7
Chlorothalonil	62.92	62.651	55.2	24.95	24.991	21.9	256.27	256.343	265.9
Warfarin	84.183	83.472	84.4	33.369	33.545	33.5	334.77	334.858	308.33

Table 15. Comparison of actual values against predictions for Molar refractivity, Polarizability, and Molar weight.

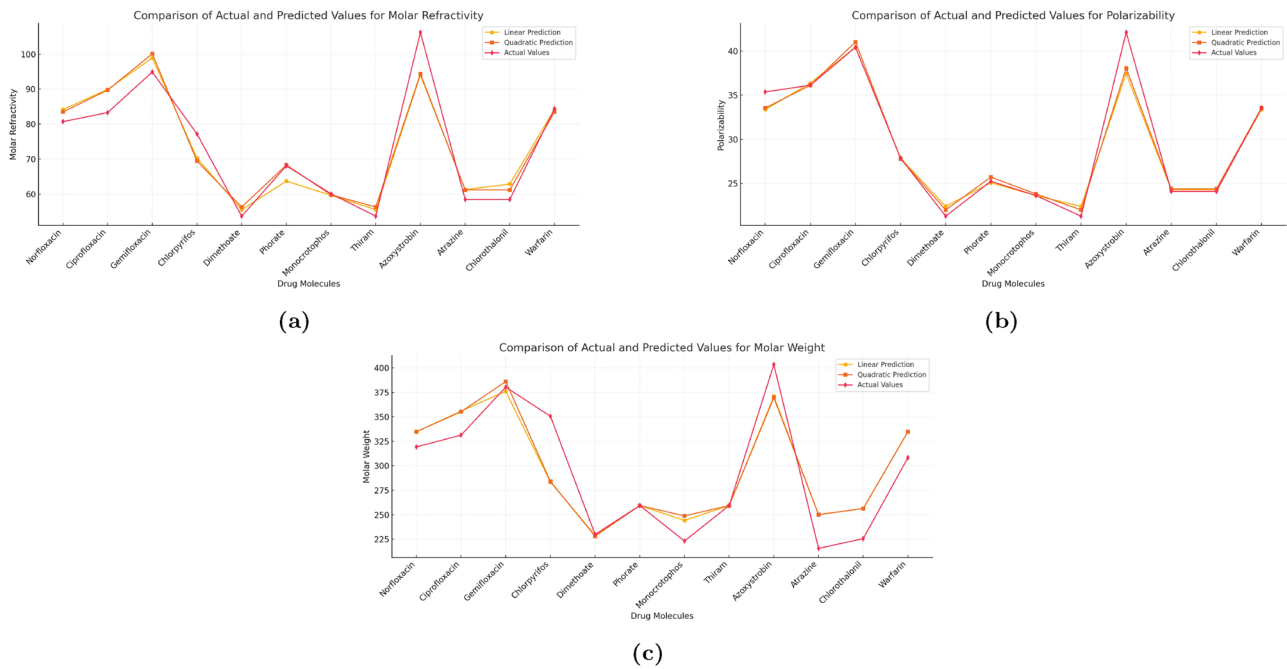


Fig. 16. (a) Comparison of actual and predicted values for Molar refractivity, (b) Comparison of actual and predicted values for Polarizability, (c) Comparison of actual and predicted values for Molar weight.

Properties	Heavy atom count			Complexity			Molar volume		
	Linear	Quadratic	Actual	Linear	Quadratic	Actual	Linear	Quadratic	Actual
Norfloxacine	22.411	33.545	22	449.392	518.874	452	235.568	241.601	237.4
Ciprofloxacin	24.697	36.126	24	500.712	566.729	528	247.748	248.959	226.8
Gemifloxacin	28.29	40.418	28	581.357	638.207	584	266.888	255.533	236.3
Chlorpyrifos	16.858	27.769	20	324.758	394.985	378	205.988	213.449	236.8
Dimethoate	10.98	22.408	12	192.793	251.964	232	174.668	167.765	169
Phorate	14.246	25.29	13	266.107	332.924	208	192.068	195.161	216.8
Monocrotophos	12.613	23.819	11	229.45	292.914	216	183.368	182.093	186.7
Thiram	10.98	22.408	10	192.793	251.964	160	174.668	167.765	180
Azoxystrobin	26.33	38.041	27	537.369	599.783	512	256.448	252.703	300.1
Atrazine	13.266	24.401	14	244.113	309.031	209	186.848	187.471	169.9
Chlorothalonil	13.919	24.991	14	258.776	324.997	322	190.328	192.648	154.6
Warfarin	22.411	33.545	22	449.392	518.874	426	235.568	241.601	235.8

Table 16. Comparison of actual values against predictions for Heavy atom count, Complexity and Molar volume.

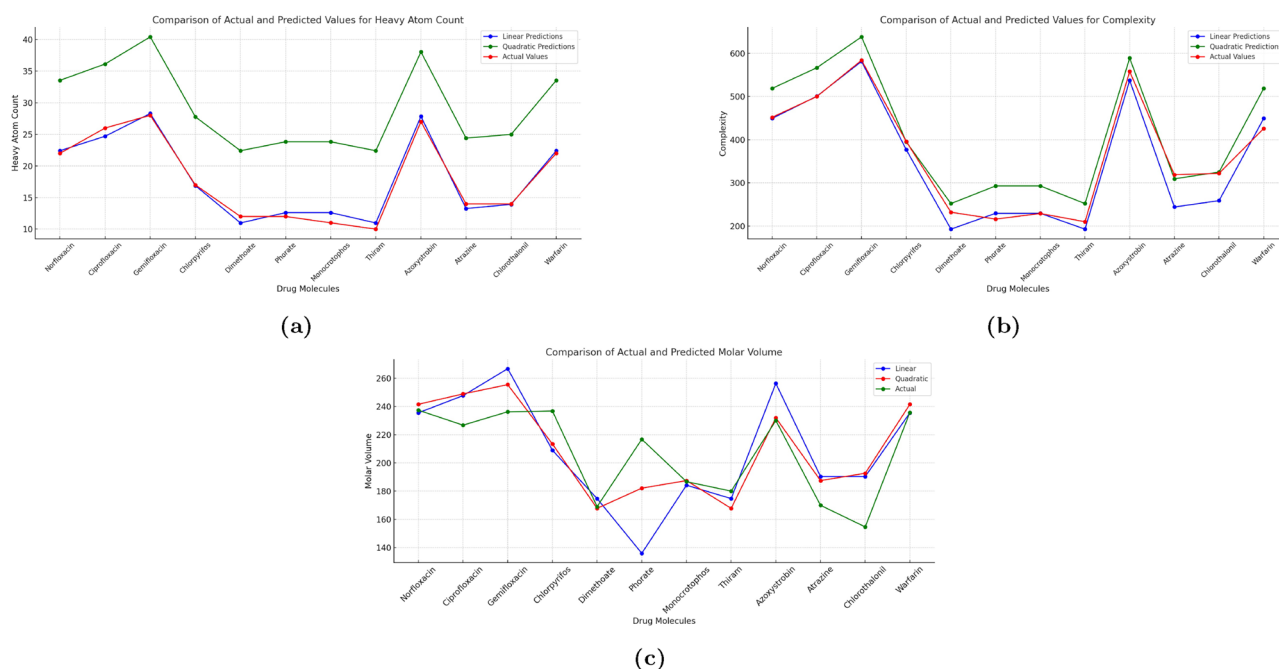


Fig. 17. (a) Comparison of actual and predicted values for Heavy atom count, (b) Comparison of actual and predicted values for Complexity, (c) Comparison of actual and predicted values for Molar volume.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 25 March 2025; Accepted: 10 June 2025

Published online: 01 July 2025

References

- Qin, H. et al. On QSPR analysis of pulmonary cancer drugs using python-driven topological modeling. *Sci. Rep.* **15**(1), 3965 (2025).
- Kanabur, R. & Shigehalli, V. QSPR analysis of degree-based topological indices with physical properties of benzenoid hydrocarbons. *Gen. Lett. Math.* **2**(3), 150–169 (2017).
- Zaman, S., Yaqoob, H. S. A., Ullah, A. & Sheikh, M. QSPR analysis of some novel drugs used in blood cancer treatment via degree based topological indices and regression models. *Polycycl. Aromat. Compd.* **44**(4), 2458–2474 (2024).
- Das, P., Mondal, S., Some, B. & Pal, A. Extension of adjacency matrix in QSPR analysis. *Chemom. Intell. Lab. Syst.* **243**(105024), 15–25 (2023).

5. Qin, H. et al. A python approach for prediction of physicochemical properties of anti-arrhythmia drugs using topological descriptors. *Sci. Rep.* **15**(1), 1742 (2025).
6. Kumar, V. & Das, S. Comparative study of GQ and QG indices as potentially favorable molecular descriptors. *Int. J. Quantum Chem.* **124**(3), e27334 (2024).
7. Mahboob, A., Rasheed, M. W., Dhiaa, A. M., Hanif, I. & Amin, L. On quantitative structure-property relationship (QSPR) analysis of physicochemical properties and anti-hepatitis prescription drugs using a linear regression model. *Heliyon* **10**(4), 1–18 (2024).
8. Huang, R., Mahboob, A., Rasheed, M. W., Alam, S. M. & Siddiqui, M. K. On molecular modeling and QSPR analysis of lyme disease medicines via topological indices. *Eur. Phys. J. Plus* **138**(3), 243–249 (2023).
9. Kumar, V. & Das, S. On structure sensitivity and chemical applicability of some novel degree-based topological indices. *MATCH Commun. Math. Comput. Chem.* **92**(1), 165–203 (2024).
10. Zhou, H. et al. On QSPR analysis of molecular descriptor and thermodynamic features of narcotic drugs. *Polycycl. Aromat. Compd.* **44**(5), 3079–3099 (2024).
11. Ghazwani, H., Koam, A. N., Nadeem, M. F. & Ahmad, A. Topological Insights into nanostar Dendrimers by computing the augmented Zagreb index. *Comb. Chem. High Throughput Screen.* **23**, 14–24 (2024).
12. Alhulwah, K. H., Hussain, M., Almohanna, N. E., Hanif, M. F. & Siddiqui, M. K. On physical analysis of cadmium bismuth sulfide using quadratic regression approach. *Chem. Pap.* **79**, 1–20 (2025).
13. Wei, J., Hanif, M. F., Mahmood, H., Siddiqui, M. K. & Hussain, M. QSPR analysis of diverse drugs using linear regression for predicting physical properties. *Polycycl. Aromat. Compd.* **44**(7), 4850–4870 (2024).
14. Nadeem, M. F. et al. Topological aspects of metal-organic structure with the help of underlying networks. *Arab. J. Chem.* **14**(6), 103–123 (2021).
15. Kumar, V. & Das, S. Comparative study between Nirmala and Sombor indices based on their applicability, degeneracy and smoothness. *Eur. Phys. J. Plus* **139**(11), 1–22 (2024).
16. Delahi, K., Ahmad, A., Asim, M. A. & Hasni, R. Computation of topological indices of binary and ternary trees using algorithmic approach. *Iran. J. Math. Chem.* **15**(2), 107–115 (2024).
17. Jacob, K. & Clement, J. Topological entropy characterization of zeolite EDI and its application in predicting molecular interactions. *Eur. Phys. J. Plus* **139**(2), 161 (2024).
18. Gayathri, K. B., Santiago, R. & Govardhan, S. Hexagonal fractals: Topological indices, fractal dimensions, structure-property modeling and its applications. *Curr. Org. Synth.* **15**(4), 112 (2025).
19. Peter, P., Clement, J., Arockiaraj, M. & Jacob, K. Predictive modeling of molecular interaction energies using topological and spectral entropies of zeolite AWW. *Front. Chem.* **13**, 1543588 (2025).
20. Shenoy, B. G., Kumar, V., Acharya, D., HM, N. & Poojary, P. Statistical analysis of Revan topological indices for drug compounds used in treatment of tuberculosis. *Phys. Scr.* **100**(3), 035203 (2025).
21. Jacob, K., Clement, J., Arockiaraj, M., Peter, P. & Balasubramanian, K. Network topology and entropy analysis of tetragonal farnesite zeolites. *Sci. Rep.* **15**(1), 14896 (2025).
22. Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **69**(1), 17–20 (1947).
23. Gutman, I. & Trinajstić, N. Graph theory and molecular orbitals. Total electron energy of alternant hydrocarbons. *Chem. Phys. Lett.* **17**(4), 535–538 (1972).
24. Kulli, V. R. Reverse Zagreb and reverse hyper-Zagreb indices and their polynomials of rhombus silicate networks. *Ann. Pure Appl. Math.* **16**(1), 47–51 (2018).
25. Huang, L., Alhulwah, K. H., Hanif, M. F., Siddiqui, M. K. & Ikram, A. S. On QSPR analysis of glaucoma drugs using machine learning with XGBoost and regression models. *Comput. Biol. Med.* **187**, 109731 (2025).
26. Mufti, Z. S., AlQadi, H., Tabraiz, A., Hanif, M. F. & Fiidow, M. A. Fuzzy and crisp computational analysis of certain graphs structures via machine learning techniques. *Sci. Rep.* **15**(1), 5995 (2025).
27. Gao, M. et al. Removal of ciprofloxacin by PAA-PAM hydrogel: Adsorption performance and mechanism studies. *J. Water Process Eng.* **71**, 107361 (2025).
28. Feng, X., Ma, Z., Yu, C. & Xin, R. MRNDR: Multihead attention-based recommendation network for drug repurposing. *J. Chem. Inf. Model.* **64**(7), 2654–2669 (2024).
29. Li, X. et al. Screening for primary aldosteronism on and off interfering medications. *Endocrine* **83**(1), 178–187 (2024).
30. Zhang, R. et al. MvMRL: A multi-view molecular representation learning method for molecular property prediction. *Brief. Bioinform.* **25**(4), bbae298 (2024).
31. Huang, C., Liu, X., Wang, W. & Guo, Z. Exploring the mechanism of centipeda minima in treating nasopharyngeal carcinoma based on network pharmacology. *Curr. Comput. Aided Drug Des.* **48**(1) (2024).
32. Hui, Z. et al. Mechanisms and therapeutic potential of chinonin in nervous system diseases. *J. Asian Nat. Prod. Res.* **26**(12), 1405–1420 (2024).
33. Saleem, N. et al. Hematological changes in the blood of experimental male and female albino rats on exposure to pesticide, dimethoate. *PLoS One* **20**(5), e0321848 (2025).
34. Lou, Y. et al. Effects of the CYP3A inhibitors, voriconazole, itraconazole, and fluconazole on the pharmacokinetics of osimertinib in rats. *PeerJ* **11**, e15844 (2023).
35. Yang, J., Li, G. & Yang, Y. Identification of indigenous phorate-degrading pathways during natural attenuation of contaminated soils. *Environ. Res.* **278**, 121754 (2025).
36. Li, W. et al. Puerarin-loaded PEG-PE micelles with enhanced anti-apoptotic effect and better pharmacokinetic profile. *Drug Deliv.* **25**(1), 827–837 (2018).
37. Renu, V. V. & Vardhanan, Y. S. Transcriptomic insights into the effects of monocrotophos on *Chrysomya megacephala*: Forensic implications of altered pupation time. *Mol. Biol. Rep.* **52**(1), 378 (2025).
38. Lin, S. et al. A single-dose, randomized, open-label, four-period, crossover equivalence trial comparing the clinical similarity of the proposed biosimilar rupatadine fumarate to reference Wystamm in healthy Chinese subjects. *Front. Pharmacol.* **15**, 1328142 (2024).
39. Ge, Z. et al. Three-dimensional urchin-like K₂Ti₈O₁₇/Ag NPs composite as a SERS substrate for detecting folic acid and thiram. *Talanta* **292**, 127926 (2025).
40. Zhao, X. et al. Comparison of the efficacy and safety of low-dose antihypertensive combinations in patients with hypertension: Protocol for a systematic review and network meta-analysis. *BMJ Open* **14**(10), e086323 (2024).
41. Wan, C. et al. Distribution and accumulation dynamics of fungicide azoxystrobin in the soil-plant system. *Environ. Res.* **274**, 121287 (2025).
42. Feng, C. et al. Precisely tailoring molecular structure of doxorubicin prodrugs to enable stable nanoassembly, rapid activation, and potent antitumor effect. *Pharmaceutics* **16**(12), 1582 (2024).
43. Ahmad, F., Tahir, S., Wali, A., Khan, M. I. & Shanableh, A. A review article on the photocatalytic degradation of atrazine by potential catalysts. *Next Mater.* **8**, 100534 (2025).
44. Zhou, X. et al. Cascade reaction of isocyanides with carboxylic acid and CD₃SSO₃Na: Toward S-CD₃ thiocarbamates. *Org. Lett.* **27**(18), 4742–4746 (2025).
45. Rinkevich, F. D., Dodge, D. & Egnew, N. Minimal toxicological impact of chlorothalonil on adult honey bees (*Apis mellifera* L.). *Pestic. Biochem. Physiol.* **208**, 106300 (2025).

46. Lassoued, A. et al. An experimental study to assess the ecotoxicity of warfarin and tinzaparin on meiobenthic amphipods: Original taxonomic data from Saudi Arabia and computational modeling. *Toxics* **13**(4), 264 (2025).
47. Gutman, I. & Polansky, O. E. *Mathematical concepts in organic chemistry* (Springer, UK, 2012).
48. Martinez-Martinez, C. T., Mendez-Bermudez, J. A., Rodriguez, J. M. & Sigarreta, J. M. Computational and analytical studies of the harmonic index on Erdos–Renyi models. *MATCH-Commun. Math. Comput. Chem.* **85**(2), 395 (2021).
49. Furtula, B. & Gutman, I. A forgotten topological index. *J. Math. Chem.* **53**(4), 1184–1190 (2015).
50. Zhao, W., Shanmukha, M. C., Usha, A., Farahani, M. R. & Shilpa, K. C. Computing SS index of certain dendrimers. *J. Math.* **2021**(1), 7483508 (2021).
51. Estrada, E., Torres, L., Rodriguez, L. & Gutman, I. *An atom-bond connectivity index: Modelling the enthalpy of formation of alkanes* (1998).
52. Randic, M. Characterization of molecular branching. *J. Am. Chem. Soc.* **97**(23), 6609–6615 (1975).
53. Vujoevic, S., Popivoda, G., Vukicevic, K., Furtula, B. & Krekovski, R. Arithmetic geometric index and its relations with geometric arithmetic index. *Appl. Math. Comput.* **391**, 125706 (2021).
54. Rajasekharaiah, G. V. & Murthy, U. P. Hyper-Zagreb indices of graphs and its applications. *J. Algebra Comb. Discret. Struct. Appl.* **8**(1), 9–22 (2021).
55. Hosamani, S., Perigidad, D., Jamagoud, S., Maled, Y. & Gavade, S. QSPR analysis of certain degree based topological indices. *J. Stat. Appl. Prob.* **6**(2), 361–371 (2017).
56. Bokhary, S. A. U. H., Adnan, Siddiqui, M. K. & Cancan, M. On topological indices and QSPR analysis of drugs used for the treatment of breast cancer. *Polycycl. Aromat. Compd.* **42**(9), 6233–6253 (2022).
57. Havare, Q. C. Topological indices and QSPR modeling of some novel drugs used in the cancer treatment. *Int. J. Quantum Chem.* **121**(24), e26813 (2021).

Acknowledgements

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2502).

Author contributions

Ibrahim Al-Dayel worked on carrying out investigations, analyzing the data, designing tests and developing them. Meraj Ali Khan deals with financing sources, computation, analysis of the data and verifying calculations. Muhammad Faisal Hanif acquired funding and performed the formal analysis of the study. Aside from project overseeing, Muhammad Kamran Siddiqui planned the approach, scheduled it, found the sources, and wrote the introduction of the paper. Saba Hanif contributed to the Matlab and Maple graph elevation of the project. Brima Gegbe handled the development of the software and validation of the study. The final report of the project is reviewed and approved by each of the authors.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to B.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025