

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2026; 11(1): 44-54
© 2026 Stats & Maths
<https://www.mathsjournal.com>
Received: 08-10-2025
Accepted: 07-12-2025

Hanumantha Reddy DT
Department of Mathematics
Government College for Women,
Chintamani, Karnataka, India

K Raja Sekhara Reddy
Department of Mathematics
Government First Grade College,
Gauribidanur, Karnataka, India

Sunilkumar M Hosamani
Department of Mathematics
Rani Channamma University,
Belagavi, Karnataka, India

QSPR-analysis of amino acids with fuzzy topological indices

Hanumantha Reddy DT, K Raja Sekhara Reddy and Sunilkumar M Hosamani

DOI: <https://www.doi.org/10.22271/math.2026.v11.i1a.2240>

Abstract

Topological indices play vital role in mathematical chemistry. The first entire Zagreb index (FEZI) is one among them. This parameter has proven to be essential in various real-life scenarios, such as networking businesses and traffic management on roads. In this paper, we explore the first entire Zagreb index for various classes of generalized fuzzy transformation graphs. This research paper also deals with the bounds for FEZI for generalized fuzzy transformation graphs in terms of elements of a fuzzy graph G .

Keywords: First Zagreb index, second Zagreb index, first entire Zagreb index, AMS subject classification: 05C09, 05C72

1. Introduction

Let $G = (V, E)$ be a graph with $|V| = n$ and $|E| = m$. The open neighborhood of a vertex $v \in V$ is $N(v) = \{u/ uv \in E\}$. The degree of a vertex v is $d_v = |N(v)|$. A molecular graph is a labeled graph in which the atoms and bonds of the molecule are represented by the vertices and edges, respectively. Chemical graph theory is a branch of mathematical chemistry, which is found to be useful in study of QSAR (Quantitative Structure-Activity Relationships) and QSPR (Quantitative Structure-Property Relationship) analysis. Numerous studies have been made relating to the above mentioned fields by using what are called molecular descriptors.

There are mainly three types of descriptors:

1. Degree based molecular descriptors
2. Distance based molecular descriptors
3. Eigenvalue based molecular descriptors

A pair of molecular descriptors known as the first Zagreb index $M_1(G)$ and the second Zagreb index $M_2(G)$, are the first degree-based molecular descriptors and they first appeared in the topological formula for the total π -electron energy of conjugated molecules that has been derived by Gutman and Trinajstić^[7], the first and second Zagreb indices are defined as:

$$M_1(G) = \sum_{uv \in E(G)} (d_u + d_v) \quad M_2(G) = \sum_{uv \in E(G)} (d_u \cdot d_v)$$

Nowadays, due to various applications of fuzzy graph theory, a huge number of researchers are working on topological indices^[2, 3, 5, 6, 9-11, 15, 16]. Kalathian *et al.*^[12] have studied the first index for fuzzy graphs and is defined as follows:

$$M_1(G) = \sum_{i=1}^n \omega(u) \deg_G(v_i)^2 \quad (1)$$

To maintain the similarity of the first Zagreb index for crisp graphs, Islam and Pal^[10] redefined the first Zagreb index in the following way.

Corresponding Author:
Hanumantha Reddy DT
Department of Mathematics
Government College for Women,
Chintamani, Karnataka, India

$$FZ_1(G) = \sum_{i=1}^n [\omega(u)deg_G(v_i)]^2 \quad (2)$$

The second fuzzy Zagreb index ^[3] is defined as:

$$FZ_2(G) = \sum_{uv \in E} [\omega(u)\omega(v)deg_G(u)deg_G(v)] \quad (3)$$

The forgotten Zagreb index for fuzzy graph was put forward in ^[15]. It is defined as:

$$FF_1(G) = \sum_{i=1}^n [\omega(u)deg_G(v_i)]^3 \quad (4)$$

Amino Acids are the organic compounds that combine to form peptide and pro-teins, hence they are referred to as the building blocks for peptide and proteins. Amino acids are organic compounds containing the basic amino groups ($-NH_2$) and carboxyl groups ($-COOH$). Altogether, there are twenty amino acids, which are involved in the Construction of proteins.

Hellberg ^[13], developed three principal properties for 20 amino acids and variation in amino acid sequence within sets of peptides is described by three principal properties hydrophilicity (z_1), bulk (z_2) and electronic properties (z_3). These are derived from principal components analysis (PCA) of a matrix of 29 physicochemical variables for 20 amino acids.

Apart from this two descriptor are derived from the side chain surface area and atomic charge. One descriptor is isotropic surface area (ISA), proposed and developed by Grigoros ^[6] as the portion of the solute molecule which is accessible to nonspecific interactions with the solvent, water. The other one is Electronic charge index (ECI), ECI is the sum of absolute value of CNDO/2 charges of the side chain atoms. ISA and ECI were derived and QSAR methodologies were used to correlate structural variation with variation in activity for the sets of peptides. One physical property of amino acid is molecular weight (MW), molecular weight is a measure of the sum of the atomic weight values of the atoms in a molecule. Molecular weight is used in chemistry to determine stoichiometry in chemical reactions and equations. Molecular weight is commonly abbreviated by M.W. or MW. Molecular weight is either unitless or expressed in terms of atomic mass units (amu) or Daltons (Da).

The following molecular descriptors are considered for the study.

2. Data set

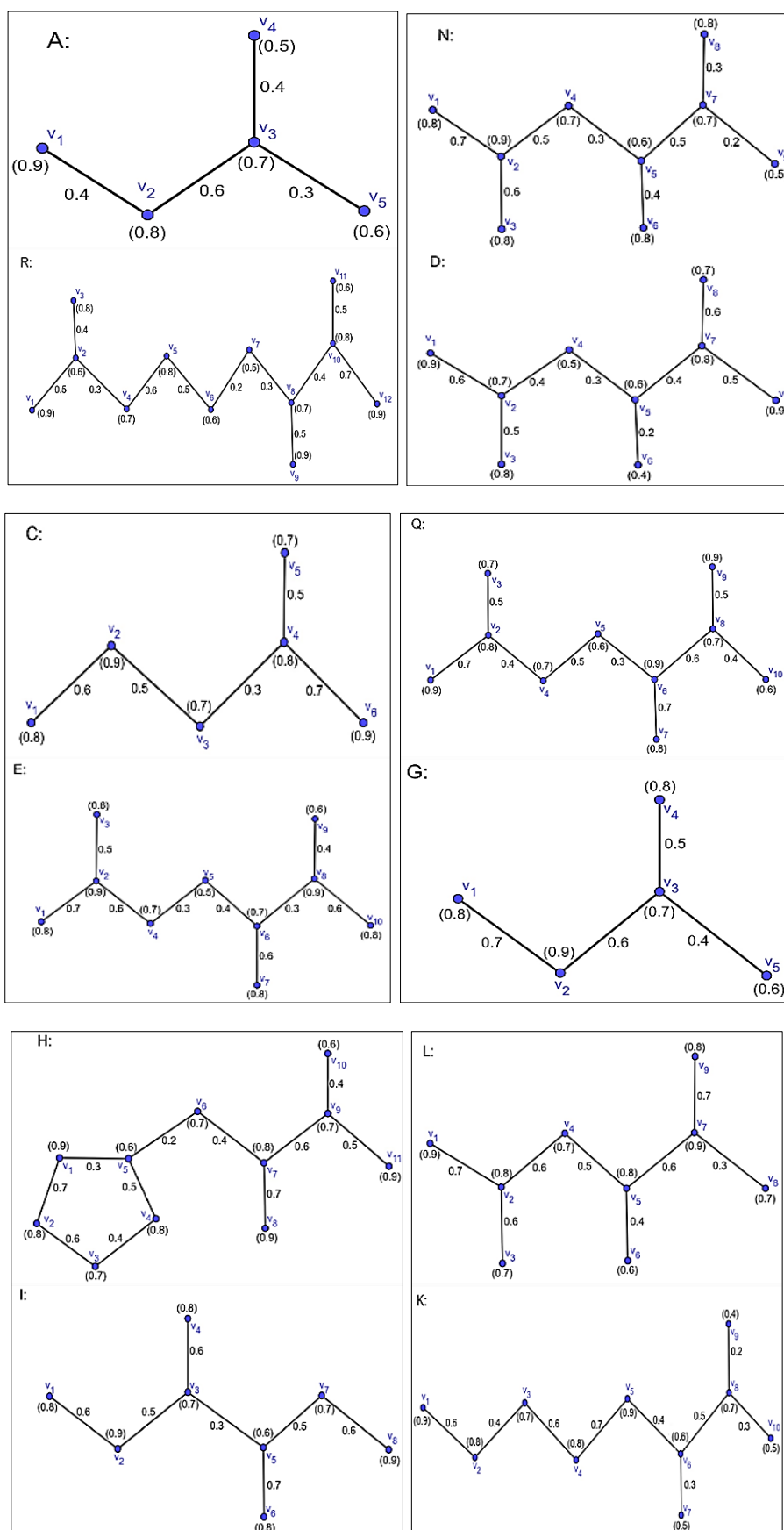
Table 1: Physical properties of Amino acids

S. No	Amino acid	Abbreviations	z_1	z_2	z_3	ISA	ESI	MW
1.	Alanine	A	0.07	-1.73	0.09	62.9	0.05	89.0
2.	Valine	V	-2.69	-2.53	-1.29	120.91	0.07	117
3.	Leucine	L	-4.19	-1.03	-0.98	154.35	0.01	131.0
4.	Isoleucine	I	-4.44	-1.68	-1.03	149.77	0.09	131.0
5.	Proline	P	-1.22	0.88	2.23	122.35	0.16	115.0
6.	Phenylalanine	F	-4.92	1.3	0.45	189.42	0.14	165.0
7.	Tryptophan	W	-4.75	3.65	0.85	179.16	1.08	204.0
8.	Methionine	M	-2.49	-0.27	-0.41	132.22	0.34	149.0
9.	Lysine	K	2.84	1.41	-3.14	102.78	0.53	147.0
10.	Arginine	R	2.88	2.52	-3.44	52.98	1.69	175.0
11.	Histidine	H	2.41	1.74	1.11	87.38	0.56	155.0
12.	Glycine	G	2.23	-5.36	0.3	19.93	0.02	75.0
13.	Serine	S	1.96	-1.63	0.57	19.75	0.56	105.0
14.	Threonine	T	0.92	-2.09	-1.4	59.44	0.65	119.0
15.	Cysteine	C	0.71	-0.97	4.13	78.51	0.15	121.0
16.	Tyrosine	Y	-1.39	2.32	0.01	132.16	0.72	181.0
17.	Asparagine	N	3.22	1.45	0.84	17.87	1.31	132.0
18.	Glutamine	Q	2.18	0.53	-1.14	19.53	1.36	146.0
19.	Aspartic	D	3.64	1.13	2.36	18.46	1.25	132.0
20.	Glutamic	E	3.08	0.39	-0.07	30.19	1.31	146.0

Table 2: Amino acids with their molecular formula

Amino acid	Abbreviations	Molecular formula	Linear formula
Alanine	A	$C_3H_7NO_2$	$CH_3-CH(NH_2)-COOH$
Arginine	R	$C_6H_{14}N_4O_2$	$HN=C(NH_2)-NH-(CH_2)_3-CH(NH_2)-COOH$
Asparagine	N	$C_4H_8N_2O_3$	$H_2N-CO-CH_2-CH(NH_2)-COOH$
Aspartic acid	D	$C_4H_7NO_4$	$HOOC-CH_2-CH(NH_2)-COOH$
Cysteine	C	$C_3H_7NO_2S$	$HS-CH_2-CH(NH_2)-COOH$
Glutamine	Q	$C_5H_{10}N_2O_3$	$H_2N-CO-(CH_2)_2-CH(NH_2)-COOH$
Glutamic acid	E	$C_5H_9NO_4$	$HOOC-(CH_2)_2-CH(NH_2)-COOH$
Glycine	G	$C_2H_5NO_2$	NH_2-CH_2-COOH
Histidine	H	$C_6H_9N_3O_2$	$NH-CH=N-CH-CH_2-CH(NH_2)-COOH$
Isoleucine	I	$C_6H_{13}NO_2$	$CH_3-CH_2-CH(CH_3)-CH(NH_2)-COOH$
Leucine	L	$C_6H_{13}NO_2$	$(CH_3)_2-CH-CH_2-CH(NH_2)-COOH$
Lysine	K	$C_6H_{14}N_2O_2$	$H_2N-(CH_2)_4-CH(NH_2)-COOH$

Methionine	M	$C_5H_{11}NO_2S$	$CH_3-S-(CH_2)_2-CH(NH_2)-COOH$
Phenylalanine	F	$C_9H_9NO_2$	$Ph-CH_2-CH(NH_2)-COOH$
Proline	P	$C_5H_9NO_2$	$NH-(CH_2)_3-CH-COOH$
Serine	S	$C_3H_7NO_3$	$HO-CH_2-CH(NH_2)-COOH$
Threonine	T	$C_4H_9NO_3$	$CH_3-CH(OH)-CH(NH_2)-COOH$
Tryptophan	W	$C_{11}H_{12}N_2O_2$	$Ph-NH-CH=C-CH_2-CH(NH_2)-COOH$
Tyrosine	Y	$C_9H_9NO_3$	$HO-Ph-CH_2-CH(NH_2)-COOH$
Valine	V	$C_5H_{11}NO_2$	$(CH_3)_2-CH-CH(NH_2)-COOH$



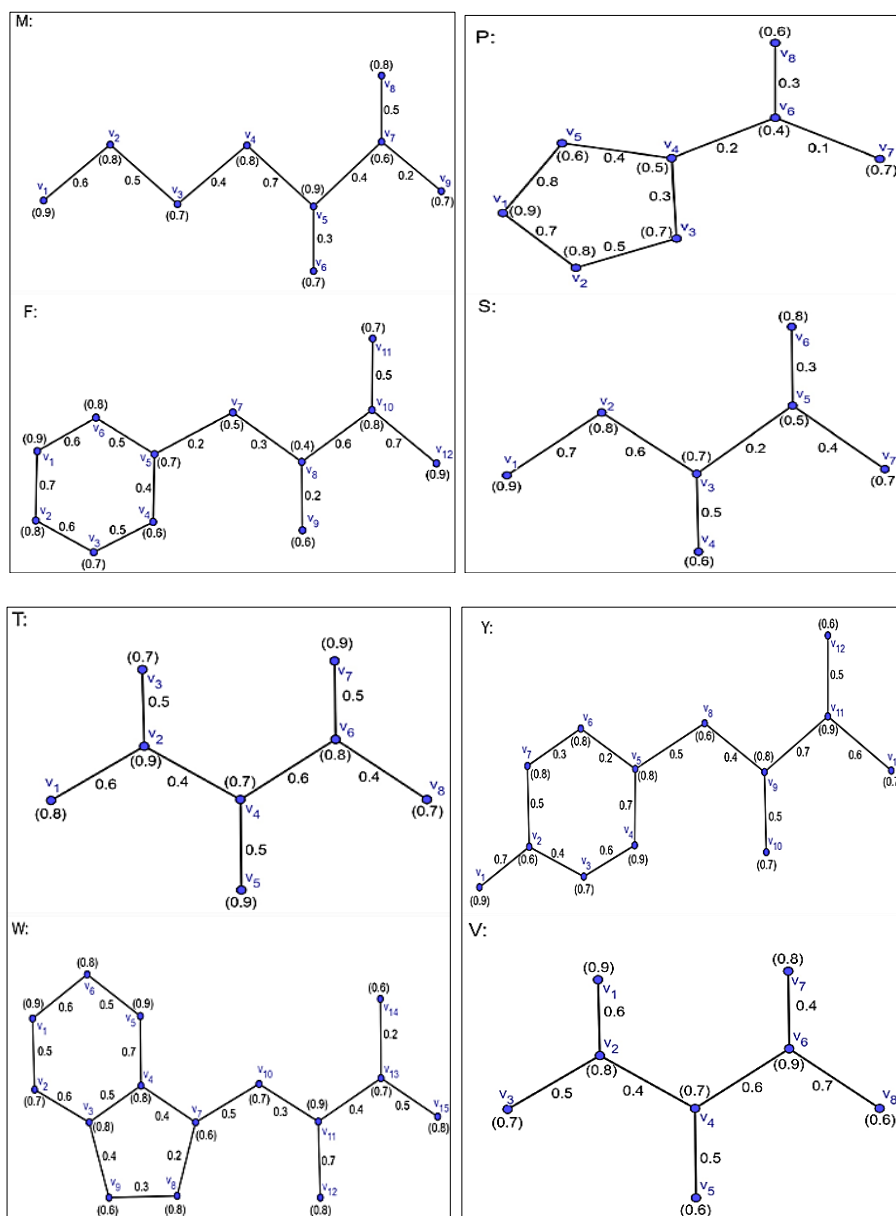


Fig 1: Molecular fuzzy graphs on amino acids

Table 3: Values of molecular descriptors of Amino acids

S. No.	Amino acid	Abbreviations	First fuzzy Zagreb Index	Second fuzzy Zagreb Index	F- fuzzy Index
1.	Alanine	Ala (A)	1.6701	1.2218	1.326059
2.	Arginine	Arg (R)	5.2669	4.8906	4.576401
3.	Asparagine	Asn (N)	4.6604	3.9676	5.477192
4.	Aspartic acid	Asp (D)	3.7935	3.2787	3.473153
5.	Cysteine	Cys (C)	3.4835	2.8776	3.277429
6.	Glutamine	Gln (Q)	6.2838	6.0973	6.649278
7.	Glutamic acid	Glu (E)	6.2628	5.2967	7.337258
8.	Glycine	Gly (G)	3.0026	2.5557	3.012678
9.	Histidine	His (H)	7.0455	6.9725	6.888277
10.	Isoleucine	Ile (I)	4.4094	4.4106	3.651288
11.	Leucine	Leu (L)	7.4055	6.7816	9.205161
12.	Lysine	Lys (K)	4.5431	4.0770	3.831137
13.	Methionine	Met (M)	4.4842	4.1454	4.134332
14.	Phenylalanine	Phe (F)	7.5658	6.9565	7.859922
15.	Proline	Pro (P)	3.8882	3.5496	4.007500
16.	Serine	Ser (S)	2.7351	2.5181	2.282383
17.	Threonine	Thr (T)	5.2013	5.1465	5.703669
18.	Tryptophan	Trp (W)	10.3463	11.3585	10.492621
19.	Tyrosine	Tyr (Y)	10.4955	9.6726	12.549463
20.	Valine	Val (V)	5.6663	5.5959	6.801397

3. Interrelation analysis between molecular descriptors

The intercorrelation between 03 molecular descriptors (FZ_1 , $FZ_2(G)$ and $FF_1(G)$) is carried out in table 6.4 and the values for pair of indices is judged by correlation coefficient r . For $r > 0.97$ it is highly intercorrelated, for $r > 0.9$ it is appreciably intercorrelated, for $r < 0.9$ it is weakly intercorrelated and for $r < 0.5$ it is not intercorrelated^[14].

3.1 Interrelation of molecular descriptors

Table 4: Interrelation of molecular descriptors.

	FZ_1	$FZ_2(G)$	$FF_1(G)$
FZ_1	1	0.974	0.972
$FZ_2(G)$	0.974	1	0.984
$FF_1(G)$	0.972	0.984	1

For physical property z_1 , the regression models are:

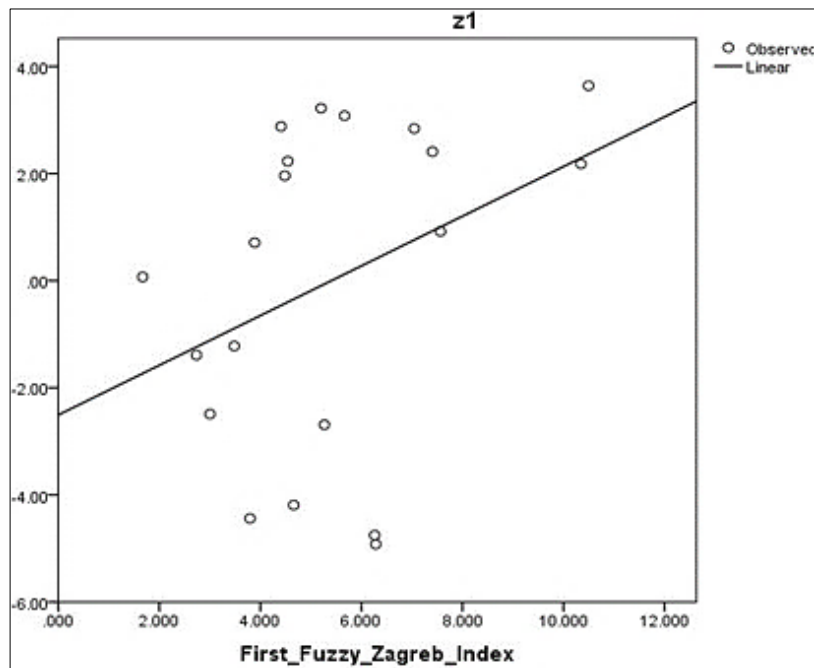


Fig 1: The linear regression model for z_1 with $FFZI$

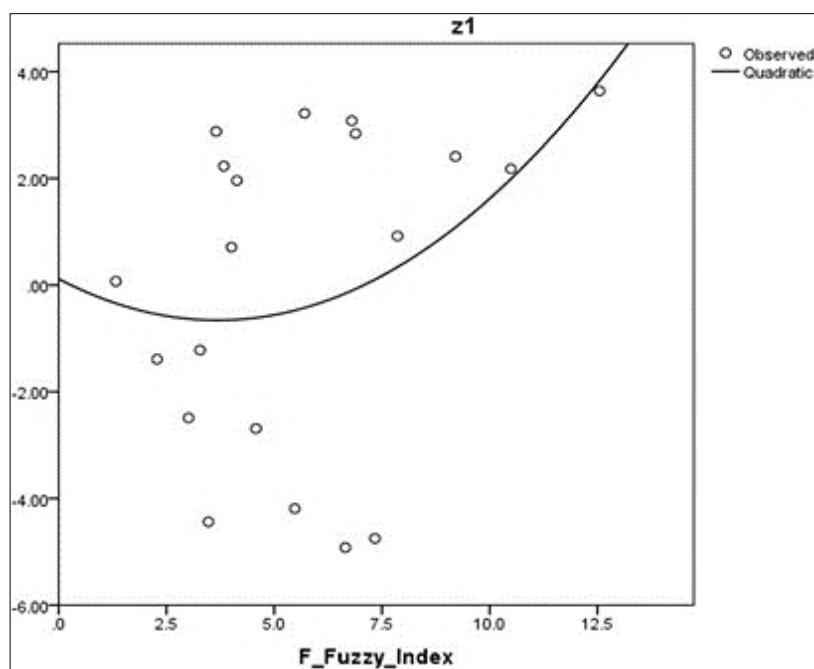


Fig 2: The quadratic regression model for z_1 with FFI

The linear regression for z_1 is

$$z_1 = (0.464)(FFZI) + (-2.507)$$

$$n = 20 \quad r = 0.860 \quad F = 2.646^\circ\text{C}$$

The quadratic regression for z_1 is

$$z_1 = (-0.421)(FFI)^2 + (0.057)(FFI) + (0.115)$$

$$n = 20 \quad r = 0.921 \quad F = 1.398^\circ\text{C}$$

For physical property z_2 , the regression models are:

The quadratic regression for z_2 is

$$z_2 = (0.475)(SFZI)^2 + (-0.022)(SFZI) + (-1.725)$$

$$n = 20 \quad r = 0.897 \quad F = 0.566^\circ\text{C}$$

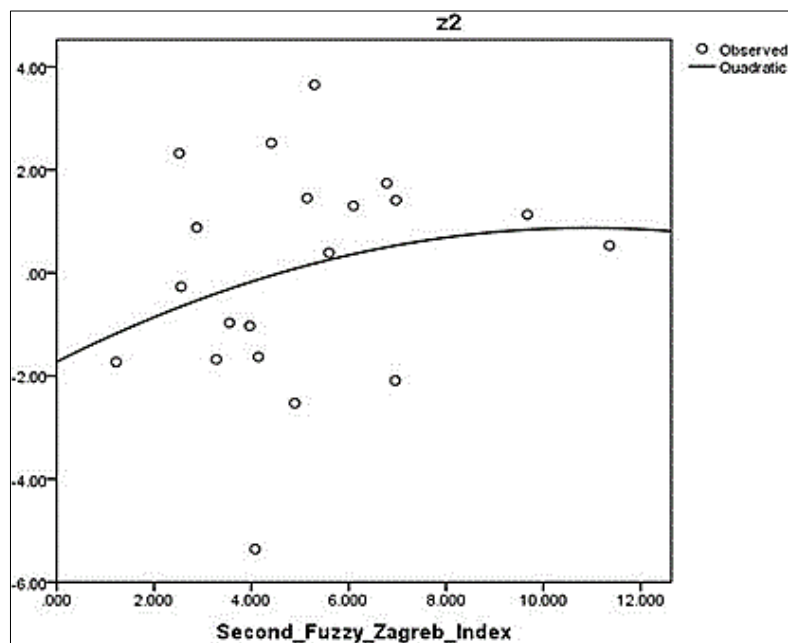


Fig 3: The quadratic regression model for z_2 with SFZI

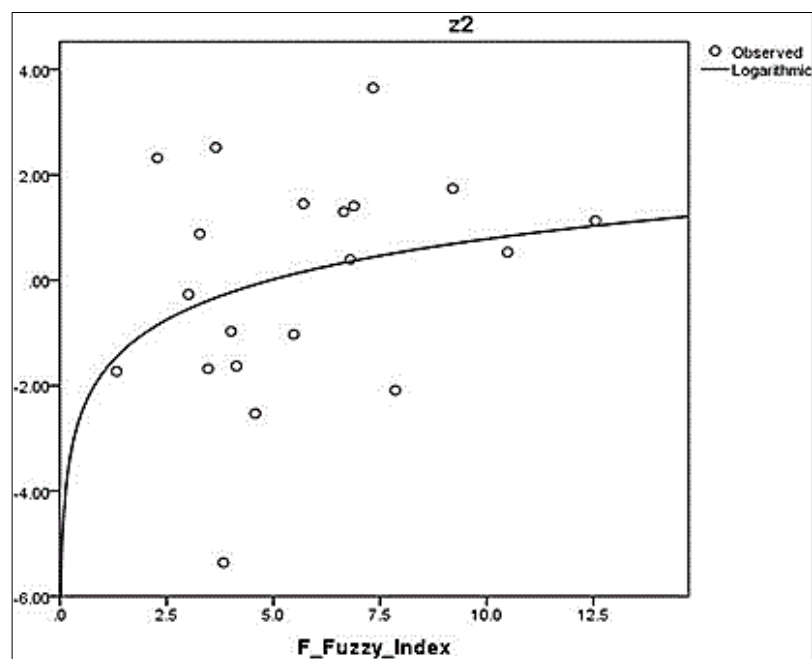


Fig 4: The logarithmic regression model for z_2 with FFI

The logarithmic regression for z_2 is

$$z_2 = (1.106)\ln(FFI) + (-1.765)$$

$$n = 20 \quad r = 0.950 \quad F = 1.559^\circ\text{C}$$

For physical property z_3 , the regression models are:

The logarithmic regression for z_3 is

$$z_3 = (-0.289)\ln(FFZI) + (0.464)$$

$$n = 20 \quad r = 0.881 \quad F = 0.096^\circ\text{C}$$

The linear regression for z_3 is $z_3 = (-0.073)(SFZI) + (0.370)$

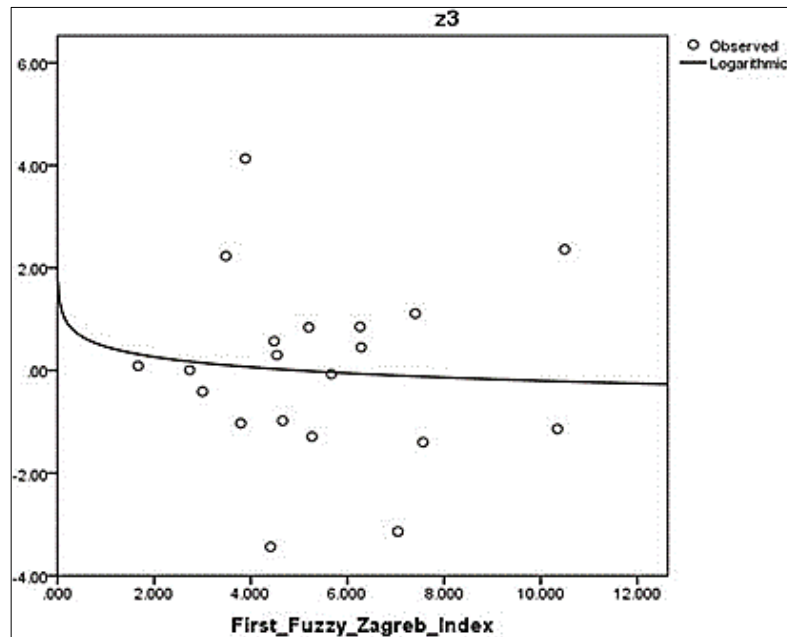


Fig 5: The logarithmic regression model for z_3 with $FFZI$

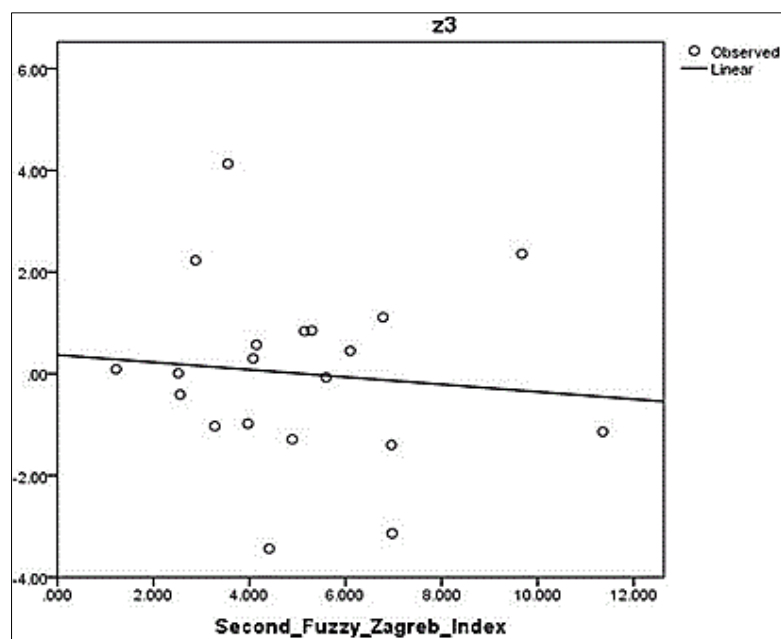


Fig 6: The linear regression model for z_3 with SFI

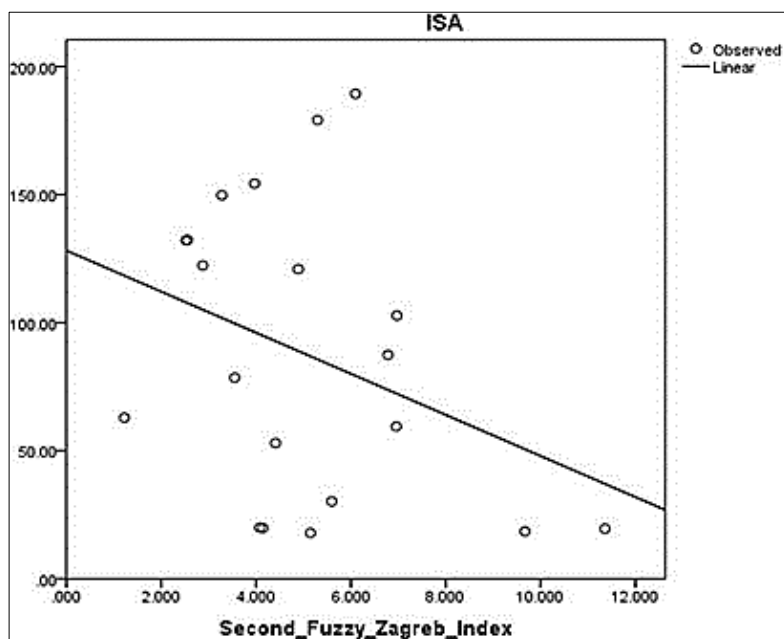


Fig 7: The linear regression model for *ISA* with *SFZI*

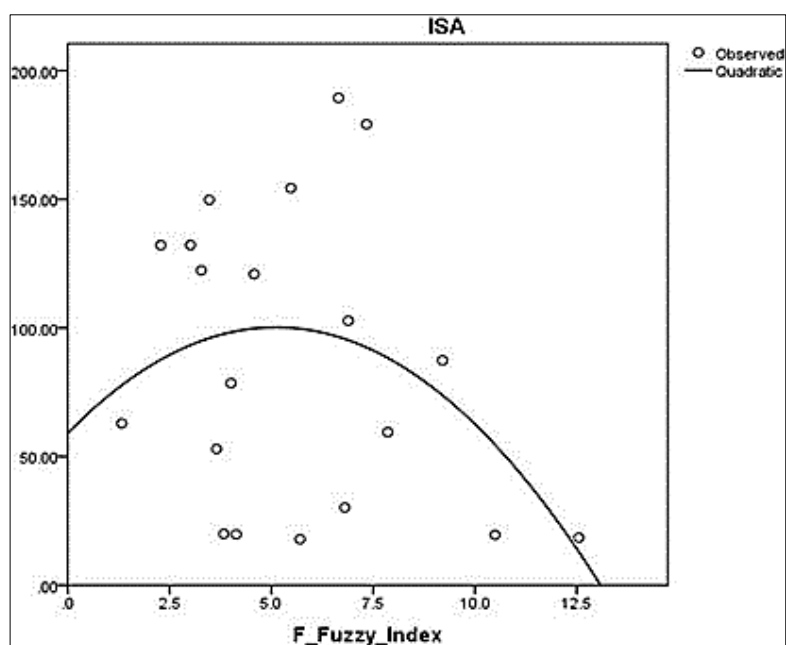


Fig 8: The linear regression model for *ISA* with *FFI*

$$n = 20 \quad r = 0.903 \quad F = 0.182 \text{ } ^\circ\text{C}$$

For physical property *ISA*, the regression models are: The linear regression for *ISA* is

$$ISA = (-8.024)(SFZI) + (128.172)$$

$$n = 20 \quad r = 0.947 \quad F = 2.365 \text{ } ^\circ\text{C}$$

The quadratic regression for *ISA* is $ISA = (16.090)(FFI)^2 + (-1.575)(FFI) + (59.126)$

$$n = 20 \quad r = 0.855 \quad F = 1.271 \text{ } ^\circ\text{C}$$

For physical property *ECI*, the regression models are:

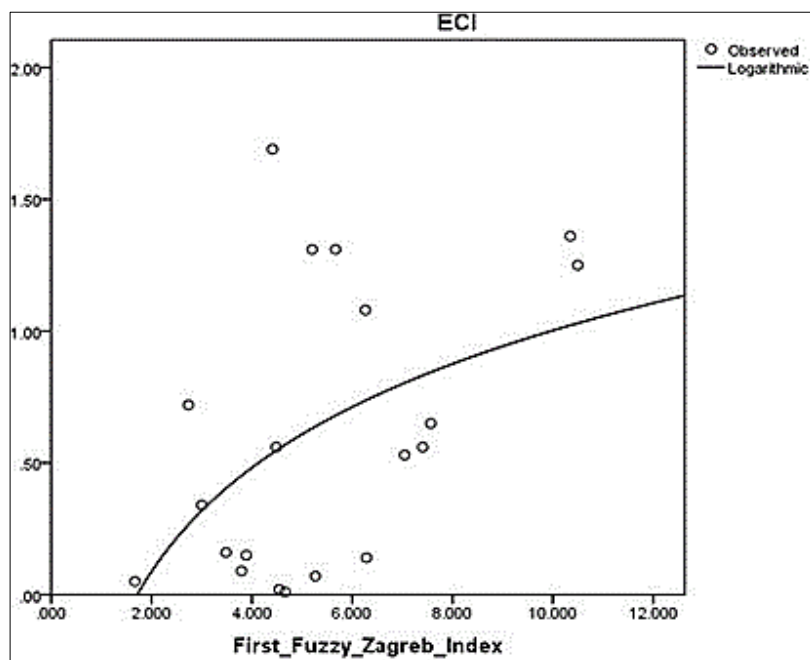


Fig 9: The logarithmic regression model for *ECI* with *FFZI*

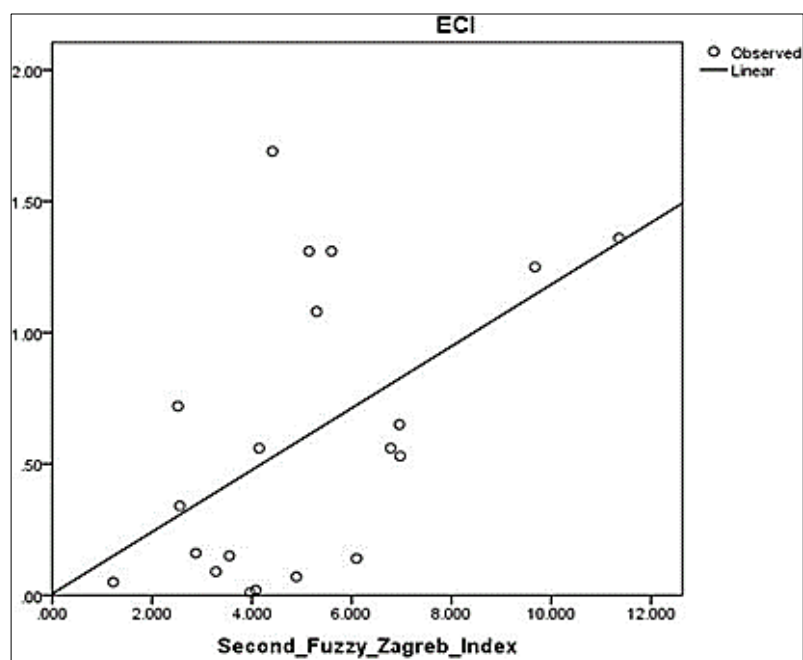


Fig 10: The linear regression model for *ECI* with *SFZI*

The logarithmic regression for *ECI* is $ECI = (0.568) \ln(FFZI) + (-0.306)$

$n = 20$ $r = 0.932$ $F = 4.960^\circ\text{C}$

The linear regression for *ECI* is

$ECI = (0.118)(SFZI) + (0.006)$ $n = 20$ $r = 0.874$ $F = 6.964^\circ\text{C}$

For physical property *MW*, the regression models are: The logarithmic regression for *MW* is $MW = (17.423)\ln(FFZI) + (108.892)$

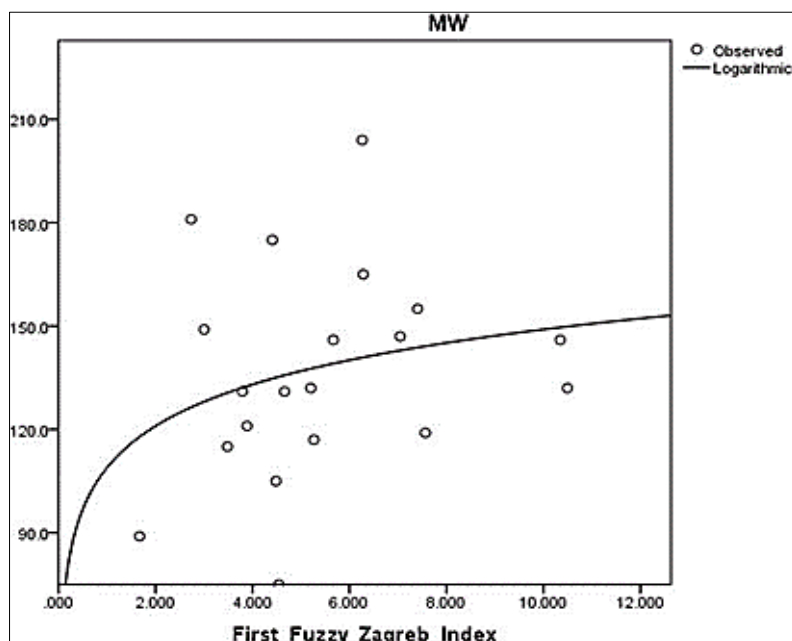


Fig 11: The logarithmic regression model for *MW* with *FFZI*

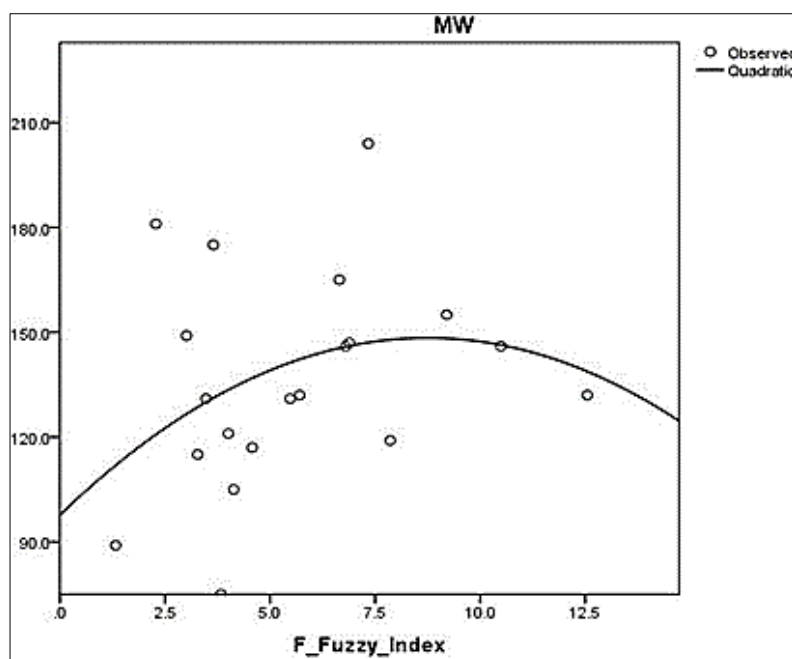


Fig 12: The quadratic regression model for *ECI* with *FFI*

$$n = 20 \quad r = 0.863 \quad F = 1.216^{\circ}\text{C}$$

The quadratic regression for *MW* is

$$MW = (11.590) (FFI)^2 + (-0.663) (FFI) + (97.695)$$

$$n = 20 \quad r = 0.925 \quad F = 1.001^{\circ}\text{C}$$

A list of conclusions are made from the following analysis:

- FFZI and FFI indices are intercorrelated with principal property z_1 , the linear regression model of FFZI and quadratic regression model of FFI gives appreciable intercorrelation with correlation coefficient $r = 0.860$ and $r = 0.921$.
- SFZI and FFI indices are intercorrelated with principal property z_2 , the quadratic
- regression model of SFZI and logarithmic regression model of FFI gives appreciable intercorrelation with correlation coefficient $r = 0.897$ and $r = 0.950$.
- FFZI and SFZI indices are intercorrelated with principal property z_3 , the logarithmic regression model of FFZI and linear regression model of SFZI gives appreciable intercorrelation with correlation coefficient $r = 0.881$ and $r = 0.903$.
- SFZI and FFI indices are intercorrelated with principal property *ISA*, the linear regression model of SFZI and quadratic regression model of FFI gives appreciable intercorrelation with correlation coefficient $r = 0.947$ and $r = 0.855$.

- FFZI and SFZI indices are intercorrelated with principal property *ECI*, the linear regression model of FFZI and quadratic regression model of SFZI gives appreciable intercorrelation with correlation coefficient $r = 0.932$ and $r = 0.874$.
- FFZI and FFI indices are intercorrelated with principal property *MW*, the linear regression model of FFZI and quadratic regression model of FFI gives appreciable intercorrelation with correlation coefficient $r = 0.863$ and $r = 0.925$.

References

1. Bindusree AR, Cangul IN, Lokesha V, Cevik AS. Zagreb polynomials of three graph operators. *Filomat*. 2016;30:1979-1986.
2. Zhou B, Trinajstić N. On a novel connectivity index. *Journal of Mathematical Chemistry*. 2009;46:1252-1270.
3. Furtula B, Gutman I, Ediz S. On difference of Zagreb indices. *Discrete Applied Mathematics*. 2014;178:83-88.
4. Vukićević D, Gašperov M. Bond additive modeling 1. Adriatic indices. *Croatian Chemical Acta*. 2010;83:243-260.
5. Collantes ER, Dunn WJ III. Amino acid side chain descriptors for quantitative structure-activity relationship studies of peptide analogues. *Journal of Medicinal Chemistry*. 1995;38:2705-2713.
6. Grigoraş S. A new semiempirical method for the calculation of the partition coefficient [PhD thesis]. Chicago (IL): University of Illinois at Chicago; 1985.
7. Gutman I, Trinajstić N. Graph theory and molecular orbitals. Total π -electron energy of alternate hydrocarbons. *Chemical Physics Letters*. 1972;17:535-538.
8. Gutman I. Topological indices and irregularity measures. *Bulletin of the International Mathematical Virtual Institute*. 2018;8:469-475.
9. Ahmad M, Hussain S, Parveen U, Zahid I, Sultan M, Qayyum A. On degree-based topological indices of Petersen subdivision graph. *European Journal of Mathematical Analysis*. 2023;3:20.
10. Trinajstić N, Nikolić S, Basak SC, Lukovits I. Distance indices and their hyper-counterparts: intercorrelation and use in the structure-property modeling. *SAR and QSAR in Environmental Research*. 2001;12:31-54.
11. Kanabura R, Giregola SK, Nagaral A, Badave P, Sannakkai R, Sasalatti R. QSPR analysis of chemical graph theory. *General Letters in Mathematics*. 2020;9(1):9-16.
12. Ediz S. On the reduced first Zagreb index of graphs. *Pacific Journal of Applied Mathematics*. 2016;8(2):99-102.
13. Hellberg S, Sjöström M, Skagerberg B, Wold S. Peptide quantitative structure-activity relationships: a multivariate approach. *Journal of Medicinal Chemistry*. 1987;30:1126-1135.
14. Réti T, Sharafzadeh R, Drégelyi-Kiss Á, Haghbin H. Graph irregularity indices used as molecular descriptors in QSPR studies. *MATCH Communications in Mathematical and in Computer Chemistry*. 2018;79:509-524.
15. Islam TU, Mufti ZS, Ameen A, Aslam MN, Tabraiz A. On certain aspects of topological indices. *Journal of Mathematics*. 2021;2021:1-20.
16. Marrero-Ponce Y, Medina-Marrero R, Castillo-Garit JA, Romero-Zaldivar V, Torrens F, Castro EA. Protein linear indices of the macromolecular pseudograph α -carbon atom adjacency matrix in bioinformatics. Part 1: Prediction of protein stability effects of a complete set of alanine substitutions in Arc repressor. *Bioorganic and Medicinal Chemistry*. 2005;13:3003-3015.
17. Promega Corporation. Amino acid chart: amino acid structure [Internet]. Madison (WI): Promega; c2024. <https://www.promega.in/resources/tools/amino-acid-chart-amino-acid-structur>