Research Journal of Physical Sciences ISSN 2320–4796
Vol. **14(2),** 14-24, August **(2026)** *Res. J. Physical Sci.*

Correlation and Regression Analysis of VL(G) and Its Transformed Forms for Predicting Physicochemical Properties of Breast Cancer Drugs

Bharathi Shettahalli Nagaraje Urs[1]* and Srikumar Krishnamurthy[2]
[1]Department of Mathematics, Government Science College, Chitradurga, Karnataka, India
[2]Department of Mathematics, School of Engineering, Dayananda Sagar University, Bangalore, Karnataka, India
bharathihemanth2022@gmail.com

Available online at: www.isca.in, www.isca.me
Received 20th May 2026, revised 10th June 2026, accepted 30th July 2026

Abstract

As one of the commonly diagnosed cancers worldwide, breast cancer necessitates efficient computational approaches for drug characterization and property prediction. In this study, the predictive ability of the topological descriptor VL(G) and its transformed forms, namely $\sqrt{VL(G)}$, $[VL(G)]^2$, $1/VL(G)$, *and* $1/\sqrt{VL(G)}$, *was investigated QSPR analysis of selected breast cancer drugs. The molecular structures were represented as molecular graphs, and the selected physicochemical properties were examined through correlation and simple linear regression analysis. The correlation study revealed that VL(G),* $\sqrt{VL(G)}$, *and [VL(G)]² exhibited strong positive relationships with the considered physicochemical properties, whereas the reciprocal forms 1/VL(G) and* $1/\sqrt{VL(G)}$ *showed inverse correlations with comparatively weaker predictive strength. Consequently, detailed regression analysis was performed only for VL(G),* $\sqrt{VL(G)}$, *and [VL(G)]², producing statistically significant models with high coefficients of determination for most properties. Comparative assessment indicated that VL(G) and [VL(G)]² demonstrated comparatively stronger predictive efficiency for several physicochemical parameters, while* $\sqrt{VL(G)}$ *also exhibited satisfactory performance. The findings establish that VL(G) and its transformed forms effectively capture structural information of breast cancer drugs and can serve as reliable molecular descriptors in QSPR investigations, thereby contributing to graph-theoretical modelling and computational prediction of physicochemical properties in medicinal chemistry.*

Keywords: Breast cancer drugs; QSPR analysis; Topological index; VL(G); Correlation analysis; Linear regression; Molecular graph; Physicochemical properties.

Introduction

Breast cancer remains a major health challenge worldwide owing to its high incidence and mortality, thereby posing a significant healthcare burden[1]. It develops due to the uncontrolled growth of abnormal cells in breast tissue, leading to tumour formation that may spread to surrounding tissues or distant organs if not detected at an early stage. Although breast cancer mainly occurs in women, men can also be affected. In spite of advancements in screening methods, diagnostic techniques, and therapeutic strategies, it continues to be a major cause of cancer-related deaths among women[1].

Advances in medical technologies, including mammography, molecular diagnostics, hormone therapy, chemotherapy, targeted therapy, and immunotherapy, have significantly improved disease management and patient survival. Nevertheless, the identification of efficient therapeutic compounds and the prediction of their physicochemical behaviour continue to remain important challenges in medicinal and pharmaceutical research[1].

The physicochemical characteristics of drug molecules play significantly influence molecular stability, reactivity, transport, absorption, and pharmacological effectiveness. Experimental determination of these properties generally involves extensive laboratory investigations, considerable expenditure, and significant time. Hence, computational and mathematical methods for estimating molecular properties have gained increasing importance in pharmaceutical chemistry and drug discovery[2,3].

Quantitative Structure-Property Relationship (QSPR) studies offer a systematic approach for elucidating the relationship between molecular structure and physicochemical properties[4]. QSPR modelling employs numerical molecular descriptors to estimate chemical and biological characteristics without relying entirely on experimental procedures. Such approaches reduce experimental cost and effort while facilitating the screening of compounds for potential therapeutic applications[5]. Closely related to QSPR studies, Quantitative Structure-Activity Relationship (QSAR) analysis is also widely employed to evaluate biological activity based on structural characteristics of molecules[6].

Graph theory has emerged as an important mathematical tool in QSPR/QSAR investigations because molecular structures can naturally be represented in the form of graphs[7-10]. Molecular graphs describe chemical structures by assigning vertices to atoms and edges to the bonds between them[8,10,11]. Numerical quantities derived from graph structures, commonly known as topological indices, have been successfully employed for predicting physicochemical, biological, and pharmaceutical properties of compounds[12]. In particular, degree-based and distance-based topological descriptors have attracted substantial attention owing to their applications in mathematical chemistry and cheminformatics[13].

Several researchers have investigated the applicability of topological indices in medicinal chemistry and QSPR modelling. Bokhary et al. analysed topological indices for breast cancer drugs and established their significance in QSPR studies[1]. Shanmukha et al., examined degree-based topological indices of anticancer drugs and reported meaningful correlations with physicochemical parameters[14]. Recent studies have also focused on graph-theoretical descriptors and their mathematical properties, including the investigation of the $VL(G)$index and related topological measures[15,16].

The present work focuses on the $VL(G)$topological index and its mathematical transformations, namely $\sqrt{VL(G)}$, $[VL(G)]^2$, $1/VL(G)$, and $1/\sqrt{VL(G)}$, to investigate their effectiveness in predicting the physicochemical characteristics of breast cancer drugs. Since mathematical transformations of molecular descriptors often reveal hidden structural relationships and may improve predictive performance, transformed forms of $VL(G)$were incorporated to examine whether stronger associations with physicochemical properties could be achieved[17].

To evaluate the strength and direction of association among the selected topological descriptors and physicochemical properties, correlation analysis was performed. Correlation coefficients provide useful information regarding the dependence between variables and help identify descriptors with predictive significance. However, correlation analysis alone is insufficient to establish predictive mathematical relationships.

Therefore, simple linear regression analysis was employed to construct predictive models and determine the extent to which variations in physicochemical properties can be explained through the selected descriptors. Regression analysis plays an important role in QSPR studies by enabling property prediction and assessing model adequacy through statistical measures such as the coefficient of determination (R^2), correlation coefficient (R), F-statistics, and significance levels (p).

In the present investigation, the predictive capability of $VL(G)$ and its transformed forms, namely $\sqrt{VL(G)}$, $[VL(G)]^2$, $1/VL(G)$, and $1/\sqrt{VL(G)}$, was examined for selected physicochemical properties of breast cancer drugs. Correlation analysis was employed to evaluate the degree of association among the selected descriptors and physicochemical parameters, while simple linear regression analysis was utilized to establish predictive models and assess descriptor efficiency in QSPR modeling. The study aims to identify suitable graph-theoretical descriptors for reliable prediction of physicochemical properties and to evaluate the applicability of transformed forms of $VL(G)$in medicinal chemistry.

For a simple connected chemical graph G, the topological index $VL(G)$ is[18],

$$VL(G) = \frac{1}{2}\sum_{xy \in E(G)}[d(x) + d(y) + d(x)d(y)] \quad (1)$$

Where: $d(x)$denotes the degree of vertex x in G.

Figure-1 illustrates the two-dimensional molecular structures of the selected breast cancer drugs considered in the present study.

Materials and Methods

Collection of Molecular Structures and Physicochemical Data: The chemical structures of the selected breast cancer medications (Figure-1) were obtained from previously published literature on molecular descriptors and QSPR analysis of breast cancer medications. The molecular graphs were adopted from the reported study by Bokhary et al.[1], for topological descriptor computation. Physicochemical properties, BP, MP, E, FP, MR, MV and P, were collected from published scientific sources and authenticated chemical databases. Missing data were retrieved from the ChemSpider database to maintain dataset consistency.

Molecular Graph Representation and Descriptor Computation: The selected drugs were represented as molecular graphs $G(V,E)$, where vertices correspond to atoms and edges denote chemical bonds[14-17]. Based on these graphs, the topological descriptor $VL(G)$ and its mathematical variants were calculated for each breast cancer drug. The investigated forms include $VL(G)$, $\sqrt{VL(G)}$, $[VL(G)]^2$, $1/VL(G)$, and $1/\sqrt{VL(G)}$, to evaluate the effect of descriptor modification on physicochemical property prediction[12,18].

Correlation Analysis: To investigate the association between selected topological descriptors and the physicochemical characteristics of breast cancer drugs, Pearson's correlation analysis was carried out. The correlation coefficient (R) was considered for evaluating both the strength and direction of the relationship, and analysis was carried out using Microsoft Excel. Descriptor forms showing strong positive associations were selected for detailed predictive modeling, whereas weaker inverse correlations were included for comparative assessment.

Regression Analysis: To examine the predictive performance of the selected topological descriptors for physicochemical properties of breast cancer drugs, simple linear regression

analysis was employed[13,18]. Following the correlation analysis, an extensive regression study was undertaken for $VL(G)$,$\sqrt{VL(G)}$, and $[VL(G)]^2$, while reciprocal transformations $1/VL(G)$ and $1/\sqrt{VL(G)}$ were excluded from detailed modeling due to lower predictive strength. Model adequacy was assessed using the coefficient of determination (R^2), correlation coefficient (R), F-statistic, and significance level (p-value).

Computational Tools and Graphical Representation: Correlation analysis was performed using Microsoft Excel, while regression statistics were obtained using Excel-assisted tools and online statistical resources. Python programming was employed for graphical visualization and regression plots.

Results and Discussion

The computed physicochemical characteristics of the selected breast cancer medications are presented in Table-1, while the calculated values of the topological descriptor and its transformed forms are summarized in Table-2. To evaluate the predictive efficiency of the selected descriptors, correlation analysis was carried out between the physicochemical parameters and five variants of the descriptor, namely VL(G), $\sqrt{\text{VL(G)}}$, [VL(G)]², 1/VL(G), and $1/\sqrt{\text{VL(G)}}$. The obtained correlation coefficients are reported in Table-3.

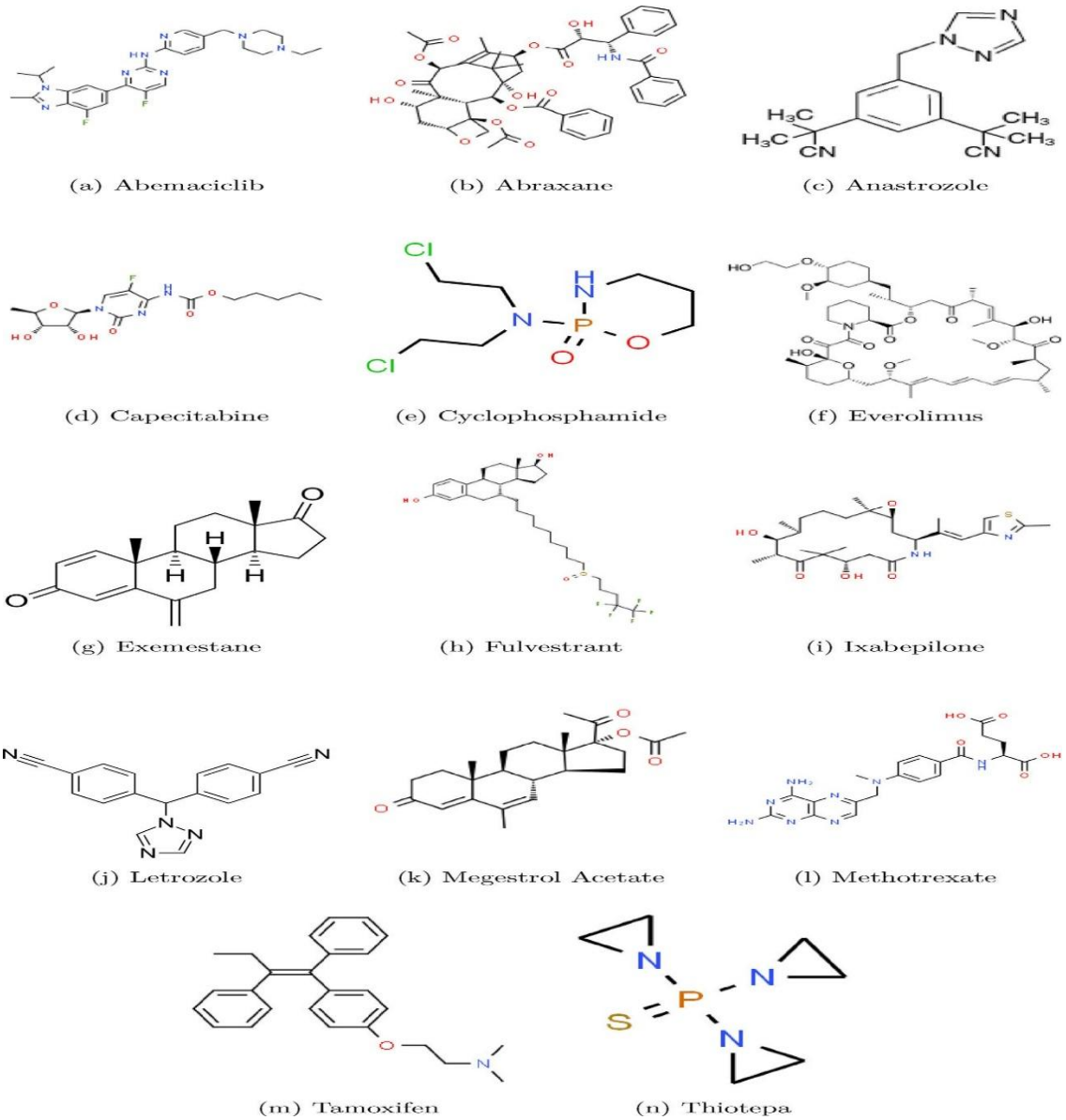

Figure-1: Two-dimensional chemical structures of breast cancer drugs.

Table-1: Physicochemical Properties of Selected Breast Cancer Drugs.

Drugs	BP	MP	E	FP	MR	MV	P
Abemaciclib	689.3	-	101	370.7	140.4	382.3	55.7
Abraxane	957.1	-	146	532.6	219.3	610.6	86.9
Anastrozole	469.7	81.5	73.2	237.9	90	270.3	35.7
Capecitabine	517	115.5	-	m	82.3	240.5	32.6
c	336.1	51	57.9	157.1	58.1	195.7	23
Everolimus	998.7	998.7	165.1	557.8	257.7	811.2	102.2
Exemestane	453.7	155.13	71.3	169	85.8	260.6	34
Fulvestrant	674.8	104	104.1	361.9	154	505.1	61.1
Ixabepilone	697.8	-	107.3	375.8	140.1	451.6	55.5
Letrozole	563.5	181	84.7	294.6	87.1	234.5	34.5
Megestrol Acetate	507.1	214	77.7	77.7	106.4	333.4	42.2
Methotrexate	-	192	-	-	119	295.7	47.2
Tamoxifen	482.3	96	74.7	140	118.9	118.9	47.1
Thiotepa	270.2	51.5	50.8	117.2	49.1	125.8	19.5

Table-2: Computed Values of VL(G) and Its Transformed Forms for Selected Breast Cancer Drugs.

Medicines	VL(G)	$\sqrt{VL(G)}$	$[VL(G)]^2$	$\frac{1}{VL(G)}$	$\frac{1}{\sqrt{VL(G)}}$
Abemaciclib	219	14.79865	47961	0.004566	0.067574
Abraxane	371	19.26136	137641	0.002695	0.051917
Anastrozole	113	10.63015	12769	0.00885	0.094072
Capecitabine	129.5	11.37981	16770.25	0.007722	0.087875
Cyclophosphamide	68	8.246211	4624	0.014706	0.121268
Everolimus	363	19.05256	131769	0.002755	0.052486
Exemestane	182	13.49074	33124	0.005495	0.074125
Fulvestrant	230.5	15.18223	53130.25	0.004338	0.065866
Ixabepilone	203	14.24781	41209	0.004926	0.070186
Letrozole	112.5	10.6066	12656.25	0.008889	0.094281
Megestrol Acetate	190.5	13.80217	36290.25	0.005249	0.072452
Methotrexate	175	13.22876	30625	0.005714	0.075593
Tamoxifen	148	12.16553	21904	0.006757	0.082199
Thiotepa	75.5	8.689074	5700.25	0.013245	0.115087

Correlation Analysis of VL(G) and Its Variants: To evaluate the predictive ability of the selected descriptor and its transformed forms, correlation analysis was performed among the physicochemical characteristics and the molecular descriptors VL(G), $\sqrt{VL(G)}$, $[VL(G)]^2$, 1/VL(G), and $1/\sqrt{VL(G)}$. The obtained correlation coefficients are reported in Table-3.

The findings reveal that VL(G), $\sqrt{VL(G)}$, and $[VL(G)]^2$ exhibit strong positive correlations with all selected physicochemical properties. In particular, VL(G) showed comparatively stronger correlations for BP (R = 0.9435), MR (R = 0.9615), MV (R = 0.9170), and P (R = 0.9614), whereas $[VL(G)]^2$ demonstrated slightly improved correlations for MP (R = 0.9352), E (R = 0.9461), and FP (R = 0.8429). The transformed descriptor √VL(G) also displayed substantial positive correlations across all properties, confirming its predictive relevance.

In contrast, the reciprocal transformations 1/VL(G) and $1/\sqrt{VL(G)}$ produced negative correlation values for all physicochemical parameters, indicating inverse relationships and comparatively weaker predictive performance. Therefore, detailed regression modeling was restricted to VL(G), $\sqrt{VL(G)}$, and $[VL(G)]^2$, while reciprocal forms were considered only for comparative correlation assessment.

Regression Analysis of VL(G) and Its Mathematical Transformations with Physicochemical Properties: To evaluate the predictive efficiency of the selected molecular descriptor, simple linear regression analysis was carried out using VL(G) and its transformed forms, namely $\sqrt{VL(G)}$ and $[VL(G)]^2$, against various physicochemical properties. These three forms were considered due to their strong positive correlations with the studied descriptors, indicating substantial predictive capability in QSPR modelling. In contrast, the reciprocal transformations, namely 1/VL(G) and $1\sqrt{VL(G)}$, exhibited negative correlations with all physicochemical properties and comparatively weaker association strengths. Therefore, detailed regression analysis for these reciprocal indices was not considered, and the study primarily focused on VL(G), $\sqrt{VL(G)}$, and $[VL(G)]^2$ as efficient predictors of physicochemical behaviour.

Regression Analysis of VL(G) with Physicochemical Properties: As the first molecular descriptor considered in the present study, VL(G) was analyzed to evaluate its predictive ability toward physicochemical properties.

The relationship between boiling point (BP) and VL(G) is expressed as:

$$BP = 194.7138 + 2.1143\,[VL(G)] \quad (2)$$

The model exhibited strong predictive performance (R^2 = 0.8903, R = 0.9435, F(1,11) = 89.2552, p < 0.001), revealing a strong positive association between VL(G) and boiling point. These results indicate that the VL(G) index is an effective molecular descriptor for estimating the boiling point of the selected breast cancer drugs.

The relationship between melting point (MP) and VL(G) is represented by:

$$MP = -247.1796 + 2.7744\,[VL(G)] \quad (3)$$

The linear modeldemonstrated a statistically meaningfulpositive relationship ($R^2 = 0.7306, R = 0.8547, F(1,9) = 24.4046, p < 0.001$, demonstrating that 73.06% of the variation in MP is described by VL(G).

The relationship between energy (E) and VL(G) is expressed as:

$$E = 30.5674 + 0.3282\,[VL(G)] \quad (4)$$

Table-3: Correlations coefficients of TI's with Molecular properties of Breast Cancer Drugs.

Physio-chemical properties/TI's	VL(G)	$\sqrt{VL(G)}$	$[VL(G)]^2$	$\frac{1}{VL(G)}$	$\frac{1}{\sqrt{VL(G)}}$
BP	0.943546517	0.935402744	0.925340195	-0.826889851	-0.873357852
MP	0.854737412	0.789774833	0.935159835	-0.555043176	-0.631192112
E	0.945210688	0.923797799	0.946064832	-0.779348609	-0.835160236
FP	0.825941404	0.796401402	0.842854565	-0.640725534	-0.697489081
MR	0.961455968	0.94529689	0.951444194	-0.801229368	-0.86005985
MV	0.917023226	0.894868897	0.912916426	-0.72646668	-0.793510685
P	0.961435654	0.945300833	0.951365926	-0.801262568	-0.860090476

The model exhibited strong predictive performance ($R^2 = 0.8934, R = 0.9452, F(1,10) = 83.8291, p < 0.001$), revealing a strong positive association between VL(G) and E.

Similarly, the relationship between flash point (FP) and VL(G) is represented by:

$$FP = 28.7414 + 1.3389\ [VL(G)] \quad (5)$$

The regression model showed statistically significant predictive ability ($R^2 = 0.6822, R = 0.8259, F(1,10) = 21.4643, p < 0.001$), indicating satisfactory predictive accuracy.

For molar refractivity (MR), the regression equation is given by:

$$MR = 9.7458 + 0.6091[VL(G)] \quad (6)$$

The model demonstrated excellent predictive efficiency ($R^2 = 0.9244, R = 0.9615, F(1,12) = 146.7251, p < 0.001$). These results indicate that the VL(G) index is an effective molecular descriptor for predicting the molar refractivity of the selected breast cancer drugs.

Likewise, the relationship between molar volume (MV) and VL(G) is expressed as:

$$MV = -9.0344 + 1.9231[VL(G)] \quad (7)$$

The regression model exhibited strong predictive performance ($R^2 = 0.8409, R = 0.9170, F(1,12) = 63.4392, p < 0.001$).

For polarizability (P), the regression equation is represented by:

$$P = 3.8542 + 0.2415\ [VL(G)] \quad (8)$$

The model showed excellent predictive capability ($R^2 = 0.9244, R = 0.9614, F(1,12) = 146.6431, p < 0.001$). Residual diagnostics confirmed model adequacy, indicating normally distributed residuals and absence of influential outliers.

The linear plots showing the association between VL(G) and the physicochemical characteristics of breast cancer drugs is illustrated in Figure-2.

Regression Analysis of $\sqrt{VL(G)}$ with Physicochemical Properties: As the second molecular descriptor, the square root transformation of the index, $\sqrt{VL(G)}$, was examined to assess its predictive capability.

The relationship between boiling point (BP) and $\sqrt{VL(G)}$ is represented by:

$$BP = -186.475 + 58.5328\sqrt{VL(G)} \quad (9)$$

The model exhibited strong predictive efficiency ($R^2 = 0.8750, R = 0.9354, F(1,11) = 76.9847, p < 0.001$), indicating that 87.50% of the variation in BP is explained by √VL(G).

The relationship between melting point (MP) and $\sqrt{VL(G)}$ is expressed as:

$$MP = -655.9703 + 69.288\sqrt{VL(G)} \quad (10)$$

The linear model demonstrated a statistically reliable relationship ($R^2 = 0.6237, R = 0.7898, F(1,9) = 14.9199, p = 0.004$).

For enthalpy (E):

$$E = -26.4369 + 8.9344\sqrt{VL(G)} \quad (11)$$

The model showed strong predictive performance ($R^2 = 0.8534, R = 0.9238, F(1,10) = 58.2139, p < 0.001$).

Similarly, for flash point (FP):

$$FP = -197.2849 + 35.9593\sqrt{VL(G)} \quad (12)$$

The model demonstrated satisfactory predictive capability ($R^2 = 0.6343, R = 0.7964, F(1,10) = 17.3415, p = 0.002$).

For molar refractivity (MR):

$$MR = -98.7991 + 16.7299\sqrt{VL(G)} \quad (13)$$

The model revealed excellent predictive strength ($R^2 = 0.8936, R = 0.9453, F(1,12) = 100.7673, p < 0.001$).

For molar volume (MV):

$$MV = -346.5415 + 52.4283\sqrt{VL(G)} \quad (14)$$

The regression model showed good predictive ability ($R^2 = 0.8008, R = 0.8949, F(1,12) = 48.2380, p < 0.001$).

For polarizability (P):

$$P = -39.1887 + 6.634\sqrt{VL(G)} \quad (15)$$

The model demonstrated excellent predictive performance ($R^2 = 0.8936, R = 0.9453, F(1,12) = 100.7752, p < 0.001$). Residual analysis supported model reliability, although the FP model showed slight deviation from normality ($p = 0.0311$).

The linear plots showing the relationship between $\sqrt{VL(G)}$ and the physicochemical characteristics of breast cancer drugs are presented in Figure-3.

Regression Analysis of [VL(G)]² with Physicochemical Properties: As the third molecular descriptor, the squared form [VL(G)]² was investigated to determine its predictive efficiency toward physicochemical properties.

The relationship between boiling point (BP) and [VL(G)]² is represented by:

$$BP = 391.5237 + 0.00455[VL(G)]^2 \quad (16)$$

The model demonstrated strong predictive ability ($R^2 = 0.8563, R = 0.9253, F(1,11) = 65.5241, p < 0.001$).

The relationship between melting point (MP) and $[VL(G)]^2$ is expressed as:

$$MP = -25.1869 + 0.007005[VL(G)]^2 \quad (17)$$

The regression model exhibited strong predictive performance ($R^2 = 0.8745, R = 0.9352, F(1,9) = 62.7268, p < 0.001$).

For energy (E):

$$E = 60.4293 + 0.0007214[VL(G)]^2 \quad (18)$$

The model showed excellent predictive capability ($R^2 = 0.8950$, R=0.9461,F(1,10)=85.2732,p*<0.001.*

Similarly, for flash point (FP):

$$FP = 147.9815 + 0.003[VL(G)]^2 \quad (19)$$

The model demonstrated statistically significant predictive performance ($R^2 = 0.7104, R = 0.8429, F(1,10) = 24.5308$, p*<0.001.*

For molar refractivity (MR):

$$MR = 66.7824 + 0.001319[VL(G)]^2 \quad (20)$$

The model exhibited excellent predictive performance ($R^2 = 0.9052$,R=0.9514,F(1,12)=114.6438,p*<0.001.*

For molar volume (MV):

$$MV = 170.0075 + 0.00419[VL(G)]^2 \quad (21)$$

The regression model demonstrated strong predictive ability ($R^2 = 0.8334, R = 0.9129, F(1,12) = 60.0359, p < 0.001$).

For polarizability (P):

$$P = 26.4720 + 0.000523[VL(G)]^2 \quad (22)$$

The model exhibited excellent predictive efficiency ($R^2 = 0.9051$,R=0.9514,F(1,12)=114.4451,p*<0.001.* Residual analysis supported model adequacy with normally distributed residuals and no significant outliers.

The linear plots showing the relationship between $[VL(G)]^2$and the physicochemical characteristics of breast cancer drugs are presented in Figure-4.

Overall, the regression findings indicate that $VL(G)$, $\sqrt{VL(G)}$, and $[VL(G)]^2$ exhibited strong predictive ability toward all selected physicochemical properties.

A comparative assessment revealed that $[VL(G)]^2$ demonstrated comparatively stronger correlations for MP (R = 0.9352), E (R = 0.9461), and FP (R = 0.8429), whereas VL(G) exhibited superior correlation with BP (R = 0.9435), MR (R = 0.9615), MV (R = 0.9170), and P (R = 0.9614).

The transformed index $\sqrt{VL(G)}$ also showed substantial predictive capability, maintaining strong positive correlations across all studied descriptors.

In contrast, the reciprocal transformations 1/VL(G) and $1\sqrt{VL(G)}$ displayed negative correlations with all physicochemical properties, with correlation coefficients ranging from −0.5550 to −0.8734 and −0.6312 to −0.8734, respectively.

Owing to their inverse relationships and comparatively lower predictive strength, detailed regression analysis for these indices was not undertaken. Hence, VL(G), $\sqrt{VL(G)}$ and $[VL(G)]^2$ serve as reliable molecular descriptors for QSPR modeling of the considered physicochemical properties.

Comparison with Existing Literature and Significance of the Present Work: Previous studies on breast cancer drugs have investigated QSPR behaviour using various established topological indices to predict physicochemical properties.

In particular, the work of Bokhary et al.[1] focused on the application of selected topological descriptors for QSPR analysis of breast cancer drugs. However, the predictive efficiency of the descriptor VL(G) and its transformed variants was not examined in their investigation.

The present work differs from earlier studies by introducing a systematic analysis of VL(G) together with its mathematical transformations, namely $\sqrt{VL(G)}$, $[VL(G)]^2$, $1/VL(G)$, and $1/\sqrt{VL(G)}$, for modeling physicochemical properties of breast cancer drugs.

The motivation behind considering transformed forms of a descriptor lies in the fact that molecular properties often exhibit nonlinear structural behavior, and suitable mathematical transformations can enhance correlation strength and predictive accuracy in QSPR models.

A comparative analysis was therefore necessary to determine whether transformed forms of VL(G) provide improved predictive performance over the original descriptor.

The results revealed that VL(G), $\sqrt{VL(G)}$, and $[VL(G)]^2$ exhibited strong positive correlations and statistically significant regression performance toward physicochemical characteristics such as BP, MP, E, FP, MR, MV, and P.

In several cases, transformed forms demonstrated stronger predictive capability than the original descriptor, indicating that descriptor modification plays an important role in improving QSPR efficiency.

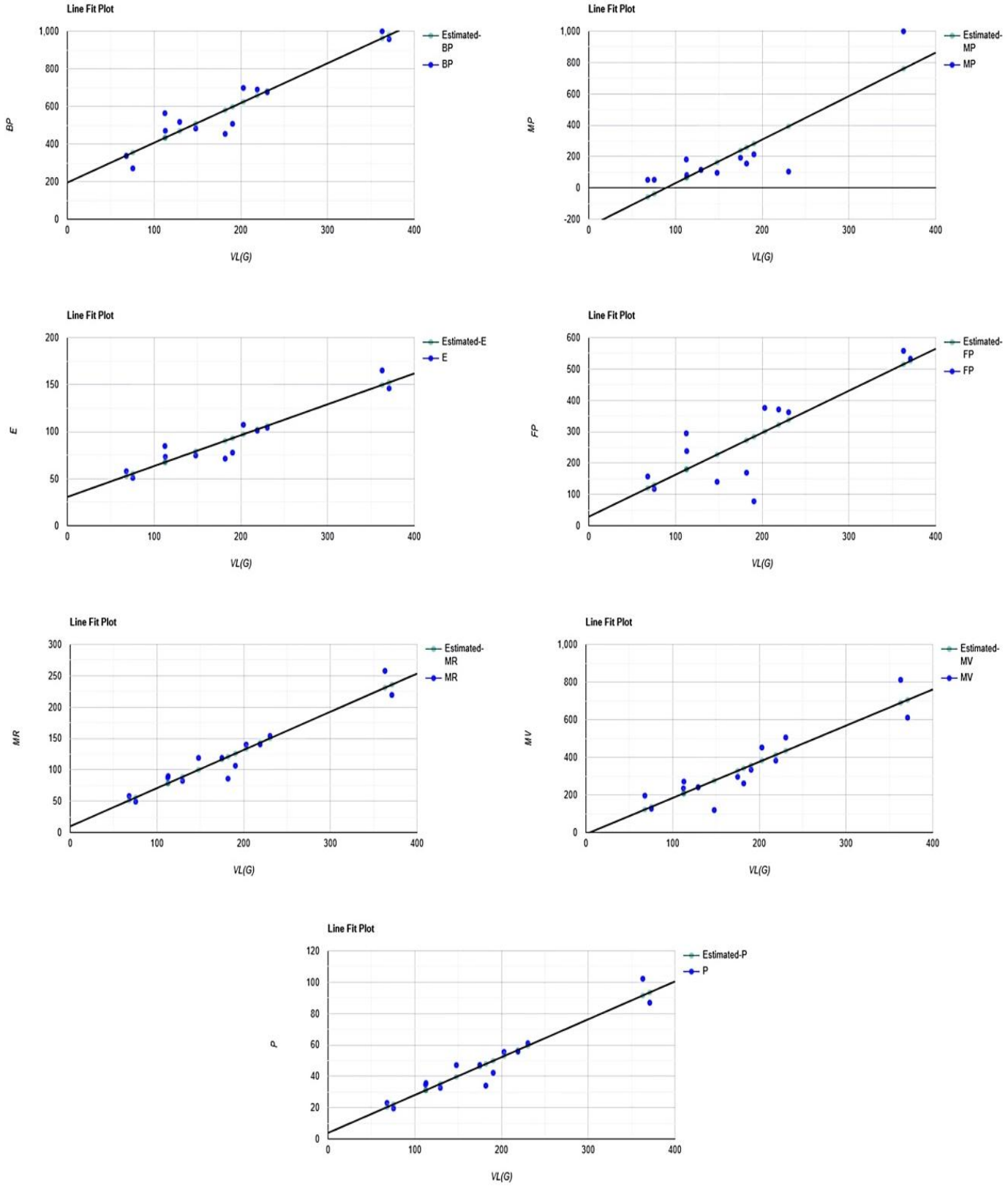

Figure-2: Linear plots of VL(G)against the physicochemical attributes of breast cancer drugs.

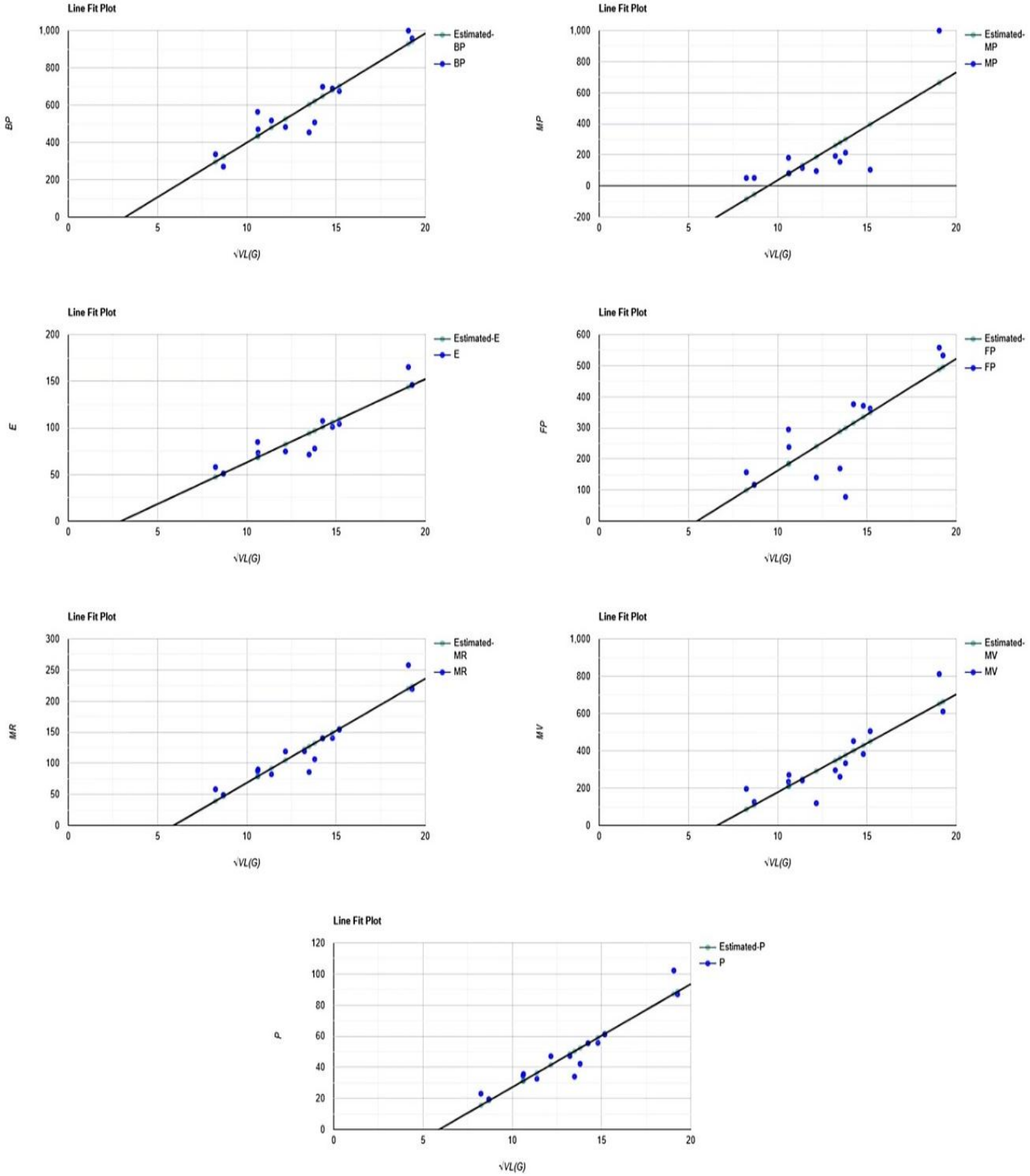

Figure-3: Linear plots of $\sqrt{\mathrm{VL(G)}}$against the physicochemical attributes of breast cancer drugs

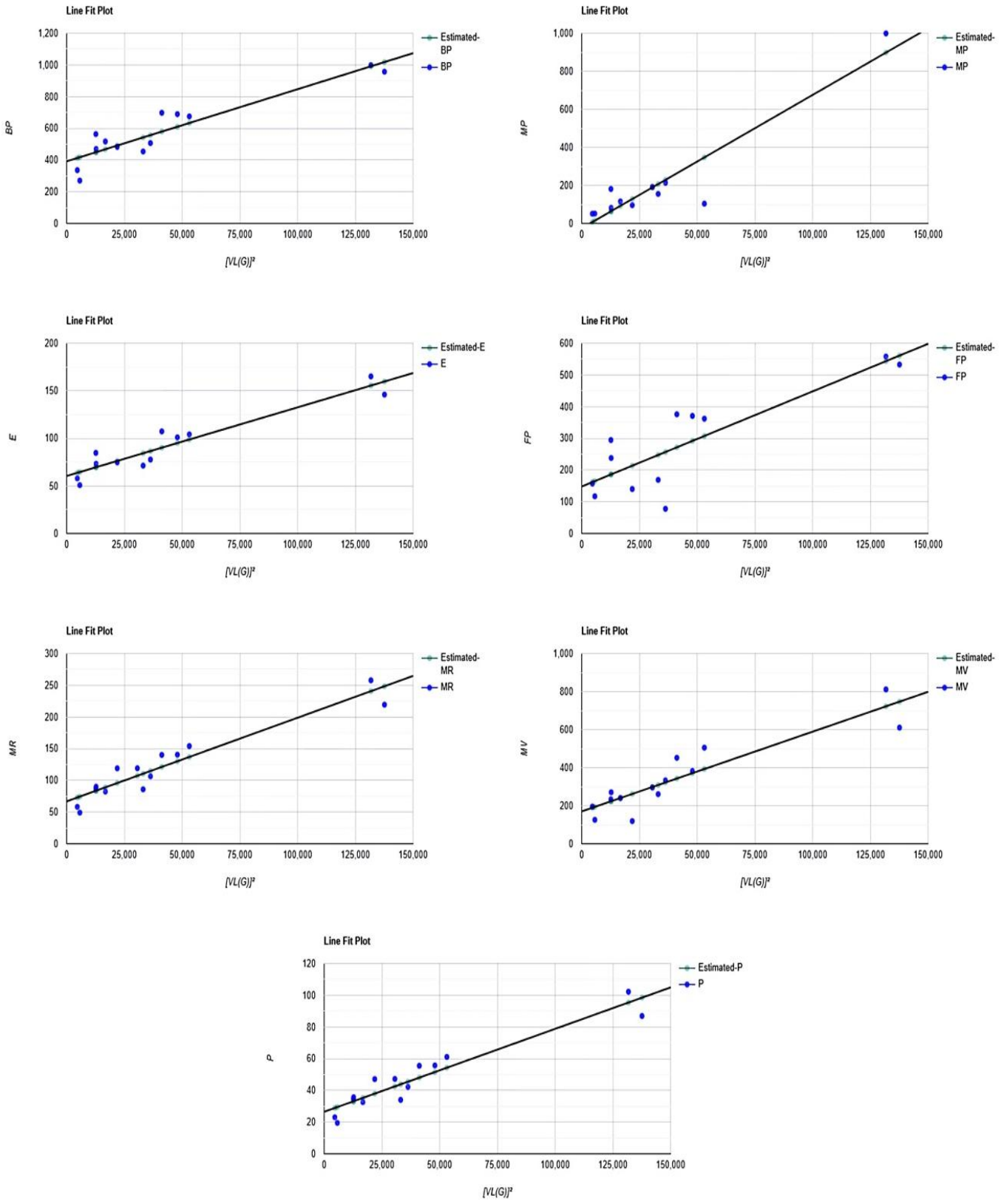

Figure-4: Linear plots of $[\mathrm{VL(G)}]^2$against the physicochemical attributes of breast cancer drugs.

In contrast, reciprocal transformations $1/\mathrm{VL(G)}$ and $1/\sqrt{\mathrm{VL(G)}}$ exhibited inverse and comparatively weaker correlations, indicating lower predictive suitability for the selected dataset. This comparative observation itself constitutes an important outcome, as it identifies which mathematical forms of the descriptor are more reliable for physicochemical property estimation. Therefore, the significance of the present study lies not in repeating an existing QSPR investigation, but in evaluating the effectiveness of a new graph-theoretical descriptor and its transformed variants for breast cancer drugs. The study provides insight into the suitability of VL(G)-based descriptors for QSPR modeling and establishes a comparative framework for selecting efficient molecular descriptors in medicinal chemistry and drug-property prediction.

Conclusion

The present study investigated the predictive efficiency of VL(G) and its transformed forms, namely $\sqrt{\mathrm{VL(G)}}$, and $[\mathrm{VL(G)}]^2$, toward the physicochemical characteristics of selected breast cancer drugs employing correlation and linear regression analysis. The statistical findings revealed strong and significant positive relationships between the selected descriptors and the studied physicochemical parameters, confirming their applicability in QSPR modelling. Among the investigated descriptors, VL(G) demonstrated comparatively stronger correlations with boiling point (BP, R = 0.9435), molar refractivity (MR, R = 0.9615), molar volume (MV, R = 0.9170), and polarizability (P, R = 0.9614), whereas $[\mathrm{VL(G)}]^2$ exhibited superior predictive performance for melting point (MP, R = 0.9352), energy (E, R = 0.9461), and flash point (FP, R = 0.8429). The transformed descriptor $\sqrt{\mathrm{VL(G)}}$ also showed substantial predictive capability, with correlation coefficients ranging from 0.7898 to 0.9453. Furthermore, the regression models produced satisfactory coefficients of determination (R^2) ranging from 0.6237 to 0.9244, indicating that a substantial proportion of variability in the physicochemical properties was explained by the selected descriptors. In contrast, the reciprocal descriptors $1/\mathrm{VL(G)}$ and $1/\sqrt{\mathrm{VL(G)}}$ exhibited inverse correlations, ranging from −0.5550 to −0.8734 and −0.6312 to −0.8734, respectively, indicating comparatively weaker predictive behaviour, therefore, these descriptors were excluded from detailed regression modelling. Overall, the findings establish VL(G), $\sqrt{\mathrm{VL(G)}}$, and $[\mathrm{VL(G)}]^2$ as reliable molecular descriptors for predicting physicochemical characteristics of breast cancer drugs and provide a useful computational basis for QSPR studies and future drug property prediction.

References

1. Bokhary, S. A. U. H., Adnan, M., Siddiqui, M. K., & Cancan, M. (2022). On topological indices and QSPR analysis of drugs used for the treatment of breast cancer. *Polycyclic Aromatic Compounds*, 42(10), 5773–5793.

2. ChemSpider Database. (2026). Royal Society of Chemistry. Retrieved May 2026, from https://www. chemspider.com

3. Todeschini, R. & Consonni, V. (2000). *Handbook of molecular descriptors*. Wiley-VCH.

4. Devillers, J. & Balaban, A. T. (1999). *Topological indices and related descriptors in QSAR and QSPR*. Gordon and Breach Science Publishers.

5. Kier, L. B., & Hall, L. H. (1986). *Molecular connectivity in structure–activity analysis*. Research Studies Press.

6. Hansch, C., Leo, A., & Hoekman, D. (1995). *Exploring QSAR: Fundamentals and applications in chemistry and biology*. American Chemical Society.

7. Gutman, I., & Polansky, O. E. (1986). *Mathematical concepts in organic chemistry*. Springer-Verlag.

8. Wiener, H. (1947). Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1), 17–20.

9. Hosoya, H. (1971). Topological index: A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bulletin of the Chemical Society of Japan*, 44(9), 2332–2339.

10. Trinajstić, N. (1992). *Chemical graph theory* (2nd ed.). CRC Press.

11. Randić, M. (1975). On characterization of molecular branching. *Journal of the American Chemical Society*, 97(23), 6609–6615.

12. Deepika, T. (2021). VL index and bounds for tensor products of F-sum graphs. *TWMS Journal of Applied and Engineering Mathematics*, 11(2), 374–385.

13. Shanmukha, M. C., Basavarajappa, N. S., Shilpa, K. C., & Usha, A. (2020). Degree-based topological indices on anticancer drugs with QSPR analysis. *Heliyon*, 6, e04235.

14. Estrada, E. (2011). *The structure of complex networks: Theory and applications*. Oxford University Press.

15. Urs, B. S. N., & Kuntal, R. S. (2024). Lower bounds and upper bounds for the topological indices of ξ-graph. *AIP Conference Proceedings*, 3149, 140033.

16. Urs, B. S. N. & Krishnamurthy S. (2025). Figuring of some degree-based topological indices of graphene. *AIP Conference Proceedings*, 3258, 020002.

17. Urs, B. S. N., Krishnamurthy, S. & Maji, S. (2026). Topological indices of certain derived graphs of triglycerides. *Boletim da Sociedade Paranaense de Matemática*, 44, 1–11.

18. Urs, B. S. N., Krishnamurthy, S., Kamalakaran, A. S., & Togarichetu, D. (2025). Evaluating degree-based topological indices in QSPR modeling of anticancer drugs using linear and multilinear regression. *International Journal of Quantum Chemistry*, 125, e70123.