

Effect of 3D parameters on Antifungal Activities of Some Heterocyclic Compounds

Anita K*¹, Vijay K. Agrawal², Bashirulla Shaik², Shweta Sharma³

^{1,3}Dept. of Chemistry, Career College, Bhopal, Madhya Pradesh, India

²Dept. of Applied Science, National Institute of Technical Teachers Training and Research, Shamla Hills, Bhopal-462002, Madhya Pradesh, India

Abstract: Quantitative Structure Activity Relationships (QSAR) of some heterocyclic compounds was studied using some 3D parameters. The QSAR models indicated that Dipole Y, Dipole mag., Y length and some indicator parameters are very effective in describing the antifungal activities of these compounds against *Candida albicans* in the training and external test set. The multiple regression analysis have produced well predictive statistically significant and cross validated QSAR models which help to explore some expectedly potent compounds.

Key words: QSAR, Dipole Y, Dipole mag, Y length, indicator parameters, cross validation

I. Introduction

Despite of tremendous progress on medicinal chemistry, infectious diseases have dramatically increased and become a major threat to human health. Infections caused by fungal species are common in immuno-compromised patients and carry significant treatment cost and mortality [1]. Antibacterial resistance is a growing problem which makes the discovery of potent antibiotics with activity against resistant strains very essential [2].

Derivatives of benzoxazoles, benzimidazoles and oxazolopyrimidines and benzothiazoles have been found to possess inhibitory properties against *Candida albicans*. We know that *Candida* causes infections which are yeast like fungus and in medical terms commonly known as candidiasis [3]. Studies reveal that most fungi are completely resistant to antimicrobial agents; therefore new agents are required which possess inhibitory effect and must be relatively less toxic. The above mentioned compounds and their derivatives are known for their antifungal activities and also they have less toxicity.

QSAR is one of the important tools of drug designing used in medicinal chemistry. QSAR models are mathematical equations which correlates chemical structure to physico-chemical and biological properties. The desired relationship between molecular descriptors and activity is used to estimate the property of other molecules or to calculate the important parameters affecting the biological activity [4].

Antifungal activities of various compounds have been modeled in recent past by many workers. They have showed that Topological descriptors can be used for estimating antifungal activities of the compounds. In the present study we have selected 68 such compounds. During our study we found that Dipole mag., Dipole Y, Y length and indicator parameters are highly co-related with the antifungal activities of these compounds.

The magnitude of dipole moment has been used as a descriptor in QSAR studies. The magnitudes of one or more of the vector components along X, Y, and Z axis have been found to be a good descriptor in QSAR studies. Dipole moment can influence electrostatic interactions of biological macromolecules. Electric dipoles aligned themselves parallel to external electric field but in the opposite direction resulting in decrease of the magnitude of the field. Thus the magnitude of dipole moment could be a measure of the ability of a molecule to diminish the electric field across lipid membranes which in turn will influence membrane transport.

In this study we have also used few indicator parameters IP₁, IP₂, IP₃ which are related to substituent X, R₁, and Z. These parameters show the significance of a particular group or substituent in the given series of drug. They account for the increase or decrease of a given pharmacological activity at any specific site in the drug molecule [5]. The general structure of antifungal compounds which are active against *Candida albicans* is shown below:

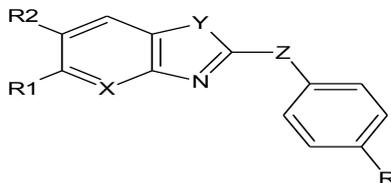


Fig.1. General Structure of antifungal compound used in the present study

II. Experimental

Quantitative Structure Activity Relationship (QSAR) modeling establishes a quantitative correlation between chemical structure and biological activity [6]. In the present study we have considered 68 compounds having antifungal activities against *Candida albicans* which have been represented in terms of $\log 1/C$ values, where C is the molar concentration of a compound and $\log 1/C$ is the dependent variable that defines the biological parameter for QSAR equations. Activities of these compounds have been taken from literature [7]. The structural details of compounds used in the present study are given in Table 1.

In order to obtain a validated QSAR model for meaningful predictions, the data set should be divided into training and test sets. The dataset is divided into a training set and a test set molecules on the basis of chemical and biological diversity using the random selection method. Accelry's Discovery Studio software has been used for calculation of the descriptors. These calculated descriptors include: the energy value of LUMO (Lowest Unoccupied Molecular Orbital), Dipole mag., Dipole Y, Jurs FNSA 3, PMI X, Shadow Y length, Shadow YZ. From the descriptors calculated useful descriptors were selected by variable selection procedure of multiple regression analysis using NCCS software [8]. A relationship between independent and dependent variables (calculated parameters and biological activities) were determined statistically using regression analysis. Linear regression is achieved by fitting a best-fit straight line to the data using the least square method. The quality of fit for a regression equation was assessed relative to its correlation coefficient and standard deviation. The F value represents the level of statistical significance of the regression. Quality of selected models was further ascertained to select the best model from cross-validated squared correlation coefficient (r^2).

The model validation is done to test the stability and predictive ability of the QSAR models. The developed models were validated by internal validation using leave-one-out method [9]. For calculating r^2 each molecule in the training set was eliminated once and the activity of the eliminated molecule was predicted by using the model developed by the remaining molecules. The predictive power of the selected model was also confirmed by external cross validation of test set compounds which is denoted by R^2_{pred} .

III. Results and Discussion

The structural details of the compounds used in the present study along with $\log 1/C$ values of compounds have been presented in Table 1. Correlation matrix showing relatedness among parameters and activity is presented in Table 2. A close look at this table reveals that the dipole Y is the only parameter which is suitable for modeling the anti-fungal activity. The calculated parameters also contain three indicator parameters IP_1 , IP_2 , and IP_3 . IP_1 has been given a value of 1 when N is present at X otherwise its value is assigned as zero. IP_2 has been assigned a value of unity if R_1 contains $-NO_2$ otherwise its value is zero and similarly IP_3 has given a value of 1 if Z contains $-CH_2$ otherwise its value is zero. The correlation matrix indicated good correlation of antifungal activity with indicator parameter IP_2 . However, this correlation has a lower magnitude than the previous one. The data was subjected to regression analysis and statistically significant correlations obtained have been reported in Table 3.

One variable model (model 6, Table 3)

$$\log 1/C = 0.0923 (\pm 0.0088) D.Y + 4.0999 \quad [1]$$

$$N = 55, R^2 = 0.6744, R^2_A = 0.6683, Se = 0.0308, F = 109.779, Q = 26.6629$$

The best one parametric model contains Dipole Y for which the R^2 value comes to be 0.6744.

Two variable model (model 10, Table 3)

$$\log 1/C = 0.0866 (\pm 0.0078) D.Y + 0.1197 (\pm 0.0282) Y \text{ length} + 3.2002 \quad [2]$$

$$N = 55, R^2 = 0.7580, R^2_A = 0.7487, Se = 0.0268, F = 81.438, Q = 32.4863$$

The systematic addition of Y length and indicator parameters gave five bi-parametric correlations which were found to be statistically better than the mono parametric correlations. However the best bi parametric model contains dipole Y and Y length. The R^2 for this model comes out to be 0.7580. The adjusted R^2 for this model is 0.7487 which clearly indicates that addition of Y length is justified.

Three variable model (model 15, Table 3)

$$\log 1/C = 0.0590 (\pm 0.0085) D.Y + 0.1946 (\pm 0.0408) IP_2 + 0.1703 (\pm 0.0335) IP_3 + 4.0516 \quad [3]$$

$$N = 55, R^2 = 0.8242, R^2_A = 0.8138, Se = 0.0231, F = 79.691, Q = 39.3011$$

Four variable model (model 19, Table 3)

$$\log 1/C = 0.0462 (\pm 0.0078) D.Y + 0.1532 (\pm 0.0346) IP_1 + 0.2556 (\pm 0.0376) IP_2 + 0.1974 (\pm 0.0293) IP_3 + 4.0217 \quad [4]$$

$$N = 55, R^2 = 0.8736, R^2_A = 0.8635, Se = 0.0197, F = 86.409, Q = 47.4450$$

Five variable model (model 21, Table 3)

$$\log 1/C = 0.0483(\pm 0.0068)D.Y + 0.1960(\pm 0.0318)IP_1 + 0.2146(\pm 0.0341)IP_2 + 0.1508(\pm 0.0278)IP_3 + 0.0935(\pm 0.0225)Y \text{ length} + 3.3253 \quad [5]$$

N = 55, R² = 0.9065, R²_A = 0.8970, Se = 0.0172, F = 95.030, Q = 55.3548

Six variable model (model 22, Table 3)

$$\log 1/C = -0.0254(\pm 0.0054)D. \text{ mag} + 0.0449(\pm 0.0057)D.Y + 0.2289(\pm 0.0276)IP_1 + 0.2869(\pm 0.0325)IP_2 + 0.1553(\pm 0.0233)IP_3 +$$

$$0.1249(\pm 0.0200)Y \text{ length} + 3.1396 \quad [6]$$

N = 55, R² = 0.9358, R²_A = 0.9278, Se = 0.0144, F = 116.576, Q = 67.1783

Further improvement in R² and adjusted R² is observed in the three, four, five and six parametric models. The R² and adj. R² values for the six parametric model containing dipole mag., dipole Y, Y length, IP₁, IP₂, and IP₃ have come out to be 0.9358 and 0.9278 respectively and the Q value [10-11] has come out to be 67.18. From these values we can infer that the six parametric model is the best for modeling the present set of compounds.

The observed activity has been plotted against estimated activity and such comparison is demonstrated in Figure 1. The predictive power of the model has come out to be 0.9359 suggesting that the model can explain 94% of the activity values. Further confirmation is obtained by estimating the log 1/C values of present set of compounds using model 22 (Table 4). The predicted log 1/C values are in good agreement with observed values showing that the proposed model is best suited for estimating log 1/C values of present set of compounds. On the basis of cross validated parameters (Table 5) the six parametric model discussed above is found to be the best model for which the cross validated R² value comes out to be 0.9314 (Table 5).

Lowest value of PSE, S_{PRESS} and PRESS/SSY further confirmed our findings. The ratio PRESS/SSY can be used to calculate approximate confidence intervals of prediction of new compounds. To be a reasonable and significant QSAR model the ratio PRESS/SSY should be less than 0.4 (PRESS/SSY < 0.4) and the value of this ratio smaller than 0.1 indicates an excellent model. From Table 5 it is clear that the six parametric model (model no. 22, eqn. 6) have the PRESS/SSY less than 0.1 indicating that this model has the best predicting capacity. R²_{CV} is the cross validated squared correlation coefficient. The highest R²_{CV} value for model 22 (eqn. 6) further confirms our predictions. The two important cross-validation parameters Uncertainty in prediction (S_{PRESS}) and Predictive squared error (PSE) were also calculated. There is no co linearity among the used parameters which has been established by ridge analysis as well as variance inflation factor calculated from the model (Table 6, Fig. 2 & Fig. 3). The parameters whose VIF value is greater than 10 will show co linearity. A perusal of this table shows that in all the cases VIF are less than 10 which mean that all the proposed models reported by us are free from co linearity. Similarly if λ (Eigen value) is found to be greater than 5 then the model will suffer from co linearity. The Table 6 shows that for all the parameters λ is less than 5. Therefore, from this point of view also proposed models are free from the defect of co linearity. Another test for co linearity is condition number (k) if its value is found to be >100 then the co linearity exists but the values reported in Table 6 indicates that the values are <100. Tolerance value (T) equal to 1 or less indicates absence of co linearity. Table 6 indicates that all the above mentioned parameters or models discussed in the study are free from multi co linearity.

External cross-validation of log 1/C value of test set is done to validate the best model obtained from training set. The results are reported in Table 5. It is clearly visible from the table that estimated activities are in good agreement with the observed activities. Therefore model 22 can be used for modeling the log 1/C activity of the present set of compounds. The correlation potential for the test set is 0.934 which is reported in Figure 1.

IV. Conclusions

On the basis of above discussion following conclusions can be drawn.

1. Dipole Y is a parameter which can model the anti-fungal activity of present set compounds along with indicator parameters, Y length and dipole mag..
2. Coefficient for Dipole Y and Y length is positive suggesting that a higher value of dipole Y and Y length will favour the log 1/C activity.
3. Negative coefficient of dipole mag suggests that it has retarding effect towards the log 1/C values, hence in future designing of potent compounds their lower values will give better results.
4. Since all the indicator parameters have positive coefficients presence of N at X, presence of -NO₂ at R₁ and presence of -CH₂ at Z is essential for exhibition of anti-fungal activity of the present set of compounds

Therefore, while designing new antifungal agents of this series they should be retained.

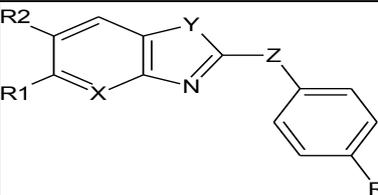
Acknowledgement

One of the authors (Anita K.) is very thankful to University Grant Commission New Delhi for providing financial support (F.No.Ms-132/102003/11-12 CRO).

References

- [1]. K.G. Yu, C.H. Li, Zhou, J.L. Diao, (2009), Study on the synthesis and antimicrobial activity of novel cationic porphyrins Chin. Chem. Lett. 20, 411-414.
- [2]. I.R. Ezabadi, C. Camoutsis, P. Zoumpoulakis, A. Geronikaki, M. Socovic, J. Clamocilija, A. Ciric, (2008), Sulfonamide-1,2,4-triazole derivatives as antifungal and antibacterial agents. Synthesis, Biological evaluation, Lipophilicity and Conformational Studies, Bioorg. Med. Chem. 16, 1150-1161.
- [3]. <http://www.emedicine.com/emerg/byname/candidiasis.htm>.
- [4]. C. Hanch, T.J. Fujita, (1964), A method for the correlation of biological activity and chemical Structure, Am. Chem. Soc. 86, 1616-1626.
- [5]. V.K. Agrawal, K. Mishra, R. Sharma, P.V. Khadikar, (2004), Topological estimation of cytotoxic activity of some anti- HIV Agents: HEPT Analogues, J. Chem. Sci., 116 (2), 93.
- [6]. Oleg, A. Costescu, M.V. Diudea, B. Parv, (2006), QSAR modeling of antifungal activity of some heterocyclic compounds, CROATICA CHEMICA ACTA, 79, (3), 483, 17-20.
- [7]. X.Y. Yao, J.P. Panaye, J.P. Doucet, R.S. Zang, H.F. Chen, N.C. Liu, Hu Z.D., B.T. Fan, (2004), Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks and multiple linear regression, J. Chem. Inf. Comput. Sci., 44, 1257.
- [8]. NCCS, Kaysville Utah, www.nccs.com
- [9]. S. Chatterjee, A.S. Hadi, B. Price, (2000), Regression Analysis by Examples, 3rd Ed. Wiley: New York.
- [10]. L. Pogliani, (1994), Structure property relationships of amino acids and some dipeptides, Amino Acids, 6, 141-153.
- [11]. L. Pogliani, (1996), Modeling with Special Descriptors Derived from a Medium-Sized Set of Connectivity Indices, J. Phys. Chem., 100, 18065-18077.

Table 1. The structural details of compounds used in the present study and their experimental logI/C values

Compd. No.	X	Y	Z	R	R ₁	R ₂		logI/C
1.	CH	O	-	H	H	H	3.89	
2.	CH	O	-	C(CH ₃) ₃	H	H	4.00	
3.	CH	O	-	NH ₂	H	H	3.92	
4.	CH	O	-	NH(CH ₃)	H	H	3.95	
5. *	CH	O	-	C ₂ H ₅	Cl	H	4.01	
6.	CH	O	-	NHCOCH ₃	Cl	H	4.06	
7.	CH	O	-	NHCH ₃	Cl	H	4.02	
8.	CH	O	-	Cl	Cl	H	4.02	
9.	CH	O	-	NO ₂	Cl	H	4.04	
10. *	CH	O	-	H	NO ₂	H	4.28	
11.	CH	O	-	CH ₃	NO ₂	H	4.31	
12.	CH	O	-	C(CH ₃) ₃	NO ₂	H	4.38	
13.	CH	O	-	NH ₂	NO ₂	H	4.31	
14.	CH	O	-	Cl	NO ₂	H	4.34	
15. *	CH	O	-	Br	NO ₂	H	4.41	
16.	CH	O	-	C ₂ H ₅	NH ₂	H	3.98	
17.	CH	O	-	F	NH ₂	H	3.96	
18.	CH	O	-	N(CH ₃) ₂	CH ₃	H	4.01	
19.	CH	O	-	CH ₃	CH ₃	H	3.95	
20. *	CH	O	-	C ₂ H ₅	CH ₃	H	3.98	
21.	CH	O	-	OCH ₃	CH ₃	H	3.98	
22.	CH	O	-	F	CH ₃	H	3.96	
23.	CH	O	-	NHCOCH ₃	CH ₃	H	4.03	
24.	CH	O	-	NHCH ₃	CH ₃	H	3.98	
25. *	CH	O	-	N(CH ₃) ₂	H	H	4.00	
26.	N	O	-	CH ₃	H	H	4.23	
27.	N	O	-	C ₂ H ₅	H	H	4.25	
28.	N	O	-	OCH ₃	H	H	4.26	
29.	N	O	-	OC ₂ H ₅	H	H	4.28	
30. *	N	O	-	NH ₂	H	H	4.23	
31.	N	O	-	NO ₂	H	H	4.29	
32.	CH	O	CH ₂	H	H	H	4.22	
33.	CH	O	CH ₂	OCH ₃	H	H	4.28	
34.	CH	O	CH ₂	Cl	H	H	4.29	
35. *	CH	O	CH ₂	NO ₂	H	H	4.31	
36.	CH	O	CH ₂	H	Cl	H	4.29	
37.	CH	O	CH ₂	OCH ₃	Cl	H	4.34	
38.	CH	O	CH ₂	Br	Cl	H	4.41	

39.	CH	O	CH ₂	NO ₂	Cl	H	4.36
40. *	CH	O	CH ₂	H	NO ₂	H	4.61
41.	CH	O	CH ₂	OCH ₃	NO ₂	H	4.66
42.	CH	O	CH ₂	Br	NO ₂	H	4.73
43.	CH	O	CH ₂	Cl	NO ₂	H	4.66
44.	CH	O	CH ₂	NO ₂	NO ₂	H	4.68
45. *	CH	O	CH ₂ O	H	CH ₃	H	3.98
46.	CH	O	CH ₂ O	H	H	NO ₂	3.73
47.	CH	O	CH ₂ O	H	Cl	NO ₂	3.79
48.	CH	O	CH ₂ O	Cl	Cl	NO ₂	3.83
49.	CH	O	CH ₂ S	H	NO ₂	H	4.36
50. *	CH	O	CH ₂ S	H	CH ₃	H	4.01
51.	N	O	CH ₂ O	H	H	H	4.26
52.	N	O	CH ₂ O	Cl	H	H	4.32
53.	CH	NH	CH ₂ O	Cl	CH ₃	H	4.04
54.	CH	NH	CH ₂ S	H	NO ₂	H	4.36
55. *	CH	NH	CH ₂ S	H	CH ₃	H	4.01
56.	CH	O	CH ₂ O	H	COOCH ₃	H	4.05
57.	CH	O	CH ₂ O	Cl	COOCH ₃	H	4.10
58.	CH	NH	CH ₂ O	Cl	COOCH ₃	H	4.10
59.	CH	NH	CH ₂ S	H	COOCH ₃	H	4.08
60. *	CH	O	C ₂ H ₄	H	NO ₂	H	4.33
61.	N	O	C ₂ H ₄	H	H	H	4.25
62.	CH	O	-	Br	NH ₂	H	4.11
63.	CH	O	CH ₂	Br	H	H	4.36
64.	CH	O	CH ₂ O	Cl	H	H	4.02
65. *	CH	NH	CH ₂ O	H	NO ₂	H	4.28
66.	CH	NH	CH ₂ O	Cl	H	H	4.02
67.	CH	NH	CH ₂ S	H	Cl	H	4.04
68.	CH	NH	C ₂ H ₄	H	H	H	4.08

*Test Set

Table 2. Correlation matrix

	logI/C	LUMO	D.mag	D.Y	FNSA3	PMI.X	Y length	YZ	IP1	IP2	IP3
logI/C	1.00										
LUMO	-0.49	1.00									
D.mag	0.43	-0.88	1.00								
D.Y	0.78	-0.31	0.37	1.00							
FNSA3	-0.51	0.95	-0.90	-0.41	1.00						
PMI.X	0.48	-0.49	0.51	0.42	-0.50	1.00					
Ylength	0.26	-0.43	0.45	0.22	-0.45	0.83	1.00				
YZ	0.02	-0.06	0.04	0.17	0.00	0.35	0.15	1.00			
IP1	0.18	-0.02	-0.06	0.12	0.09	-0.28	-0.42	-0.40	1.00		
IP2	0.65	-0.66	0.61	0.58	-0.72	0.35	0.32	0.17	-0.21	1.00	
IP3	0.47	-0.14	0.16	0.34	-0.09	0.44	0.31	0.09	0.12	-0.01	1.00

LUMO - **LUMO Eigen value VAMP**
D. mag - **Dipole mag.**
D.Y - **Dipole Y**
FNSA3 - **Jurs FNSA 3**
PMI.X - **PMI X**
Y length - **Shadow Y length**
YZ - **Shadow YZ**
IP₁, IP₂, IP₃ - **Indicator parameters.**
 If X=N, IP₁=1, R₁=NO₂, IP₂=1, and Z=CH₂, IP₃=1

Table 3. Regression parameters and quality of correlation of training set

Model No.	Parameters Used	Se	R ²	R ² _A	F-ratio	Q=R/Se
1	IP ₁	0.0530	0.0347	0.0165	1.906	3.5147
2	D. mag	0.0507	0.1197	0.0990	6.936	6.8240
3	Y length	0.0488	0.1818	0.1662	11.766	8.7373
4	IP ₃	0.0459	0.2758	0.2621	20.185	11.4415
5	IP ₂	0.0404	0.4403	0.4297	41.686	16.4245
6	D.Y	0.0308	0.6744	0.6683	109.779	26.6629
7	D. mag D.Y	0.0305	0.6865	0.6744	56.930	27.1657
8	D.Y IP ₂	0.0280	0.7351	0.7249	72.149	30.6207
9	D.Y IP ₃	0.0275	0.7458	0.7360	76.270	31.4035
10	D.Y Y length	0.0268	0.7580	0.7487	81.438	32.4863
11	D.Y IP ₂ Y length	0.0255	0.7842	0.7715	61.783	34.7275
12	IP ₁ IP ₂ IP ₃	0.0255	0.7851	0.7724	62.093	34.7474
13	D.Y IP ₃ Y length	0.0254	0.7870	0.7745	62.816	34.9264
14	D.Y IP ₁ Y length	0.0244	0.8031	0.7915	69.349	36.7278
15	D.Y IP ₂ IP ₃	0.0231	0.8242	0.8138	79.691	39.3011
16	D.Y IP ₁ IP ₃ Y length	0.0228	0.8309	0.8174	61.430	39.9797
17	D.Y IP ₂ IP ₃ Y length	0.0226	0.8341	0.8208	62.846	40.4111
18	D.Y IP ₁ IP ₂ Y length	0.0215	0.8505	0.8386	71.121	42.8942
19	D.Y IP ₁ IP ₂ IP ₃	0.0197	0.8736	0.8635	86.409	47.4450
20	D. mag D.Y IP ₁ IP ₂ IP ₃	0.0191	0.8837	0.8718	74.459	49.2174
21	D.Y IP ₁ IP ₂ IP ₃ Y length	0.0172	0.9065	0.8970	95.030	55.3548
22	D. mag D.Y IP ₁ IP ₂ IP ₃ Y length	0.0144	0.9358	0.9278	116.576	67.1783

Table 4. Observed and estimated values of log1/C of Training and Test setusing model no. 22(Table 3)

Compd. No	Observed log1/C	Estimated log1/C	Residual
1	3.89	4.01	-0.12
2	4.00	4.01	-0.01
3	3.92	4.00	-0.07
4	3.95	4.01	-0.06
5*	4.01	4.06	-0.05
6	4.06	4.03	0.03
7	4.02	4.02	-0.01
8	4.02	4.03	0.00
9	4.04	3.91	0.13
10*	4.28	4.34	-0.06
11	4.31	4.32	-0.01
12	4.38	4.36	0.02
13	4.31	4.34	-0.03
14	4.34	4.35	-0.01
15*	4.41	4.35	0.05
16	3.98	3.99	-0.01
17	3.96	3.97	-0.01
18	4.01	3.99	0.01
19	3.95	4.01	-0.06
20*	3.98	3.99	-0.01
21	3.98	3.97	0.02
22	3.96	3.95	0.01
23	4.03	3.95	0.08
24	3.98	4.03	-0.06
25*	4.00	4.00	0.01
26	4.23	4.24	-0.01
27	4.25	4.29	-0.04
28	4.26	4.26	0.00
29	4.28	4.35	-0.06
30*	4.23	4.23	0.00
31	4.29	4.31	-0.03
32	4.22	4.19	0.03
33	4.28	4.18	0.11
34	4.29	4.31	-0.02
35*	4.31	4.37	-0.06
36	4.29	4.35	-0.06
37	4.34	4.28	0.06
38	4.41	4.39	0.02
39	4.36	4.42	-0.06
40*	4.61	4.69	-0.08
41	4.66	4.64	0.01
42	4.73	4.73	-0.01
43	4.66	4.67	-0.01
44	4.68	4.64	0.04
45*	3.98	4.14	-0.16
46	3.73	3.81	-0.08
47	3.79	3.83	-0.05
48	3.83	3.85	-0.02
49	4.36	4.34	0.02
50*	4.01	4.01	0.00
51	4.26	4.26	0.00
52	4.32	4.22	0.10
53	4.04	4.02	0.02
54	4.36	4.38	-0.02
55*	4.01	4.09	-0.08
56	4.05	4.02	0.03
57	4.10	3.99	0.11
58	4.10	4.07	0.03
59	4.08	4.13	-0.06
60*	4.33	4.38	-0.08
61	4.25	4.20	0.05
62	4.11	4.15	-0.04

63	4.36	4.19	0.17
64	4.02	3.94	0.08
65*	4.28	4.35	-0.06
66	4.02	4.08	-0.06
67	4.04	4.11	-0.07
68	4.08	4.08	0.00

Table 5. Cross validation parameters of proposed models (Table 3)

Model No.	Parameters used	PRESS/SSY	R ² _{CV}	S _{PRESS}	PSE
6	D.Y	0.4828	0.5172	0.1283	0.1259
10	D.Y,Y length	0.3192	0.6808	0.1116	0.1085
15	D.Y, IP ₂ , IP ₃	0.2133	0.7867	0.0961	0.0925
19	D.Y, IP ₁ IP ₂ ,IP ₃	0.1447	0.8553	0.0823	0.0784
21	D.Y,IP ₁ ,IP ₂ ,IP ₃ , Y length	0.1031	0.8969	0.0715	0.0675
22	D.mag,D.Y,IP ₁ ,IP ₂ , IP ₃ ,Y length	0.0686	0.9314	0.0598	0.0559

Table 6. Ridge analysis for the six variable model (model 22)

Independent variable	VIF	T	λ	k
D mag	1.74	0.57	2.43	1.00
D.Y	1.95	0.51	1.29	1.88
Y length	1.80	0.55	0.99	2.33
IP ₁	1.45	0.69	0.68	3.56
IP ₂	2.41	0.42	0.36	6.68
IP ₃	1.43	0.70	0.22	10.64

VIF = Variance inflation factor

λ = Eigen Values

T = Tolerance

k = Condition number

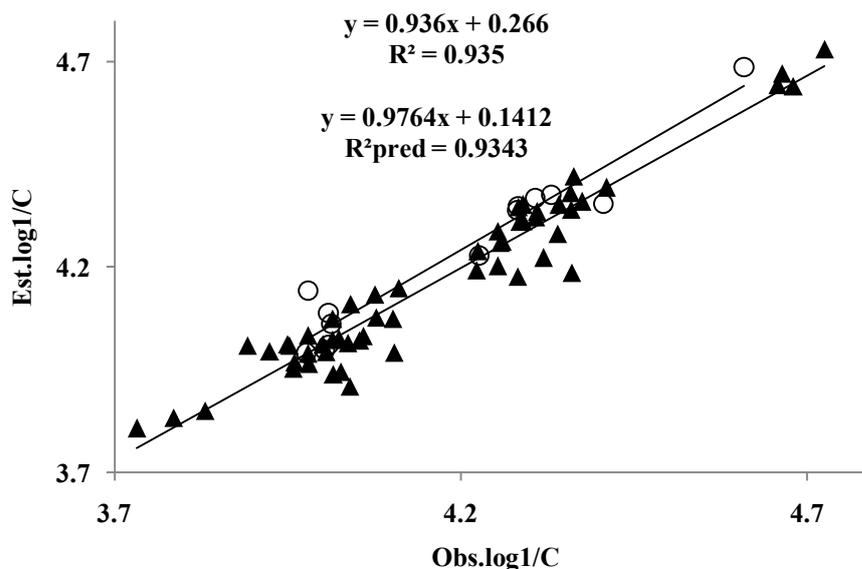


Fig. 1. Correlation between observed and estimated log1/C values using model 22

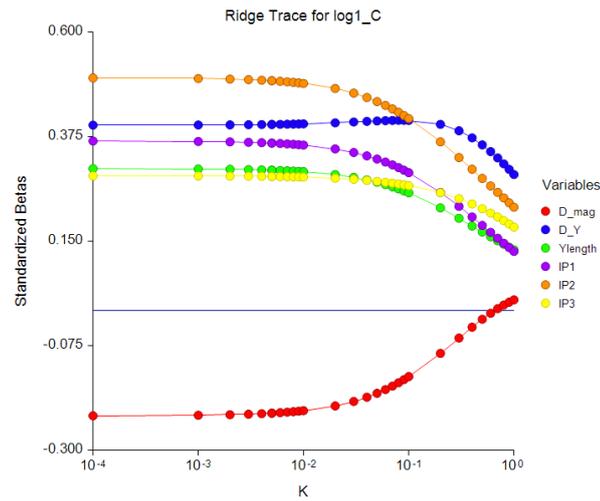


Fig. 2. Ridge Trace for six variable model (model-22)

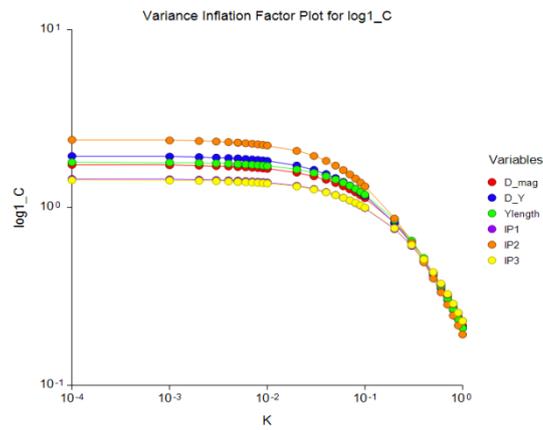


Fig. 3. VIF plot for six variable model (model-22)