International Journal of Research in BioSciences
Volume 7 Issue 1, pp. (77-86), January 2018
Available online at http://www.ijrbs.in
ISSN 2319-2844

Research Paper

Novel topological modeling of 'log P' value of few aromatic derivative compounds

Shanno Pathan and S. M. Ali
Department of Chemistry, Islamia Karimia College, Indore (M.P.), INDIA

(Received June 24, 2017, Accepted October 25, 2017)

Abstract

The present work, effort have been made to model of lipophilicity term log p of the set of aromatic compounds to identify the quantitative structural activity relationship between biological activity and structure of aromatic derivatives by using multiple regression analysis with various descriptor the best model is obtained and validated by some cross validation parameters. The results were discussed on the basis of maximum R^2 value which indicates the six parametric model is the most significant model for the biological activity log P.

Keywords: QSAR, Log P, Indicator parameters, Topological modeling & lipophilicity

Introduction

Few arenes[1] or aryl hydrocarbons[2] containing conjugated planner ring of carbon atoms and delocalized pi-electron clouds have been selected for this study. Topological approach employs different method for the characterization and quantization of molecular structure. Recently these models have been extensively used in the prediction of activities of organic compounds acting as a drug[3]. QSAR (Quantitative structure activity relationship) a mathematical approach that allows the identification of the quantitative link between structure of the molecule and biological activity of investigated compounds[4]. QSAR approach developed starting with the information obtained the 2D and 3D structure of a chemical compounds lead to introduction of molecular descriptors on the structural activity relationship method[5]. It is a well established fact that lipophilicity affects pharmaco-dynamic profile of the drug, it seems to have the greatest impact on ADMET (.i.e absorption distribution, metabolism, excretion, and toxicity) properties. Since, lipophilicity is a physico-chemical property, which describes a partitioning equilibrium of solute molecules between water and an immiscible organic solvent[6-9] it plays an important role in drug development. Many researches have been carried out to show that the toxicity of organic compounds depended on their lipophilicity and expressed by logarithmic partition coefficient log P.

Materials and Methods

According to Hansch[10-11] the drug action depends upon two process:

(i) The journey from the point of entry in the body to the site of action.
(ii) The interaction with the receptor site.

He has suggested that the biological activity of the set of the compounds can be modeled by using few independent parameters and the equation can be written as follows

$Y = m_1x_1 + m_2x_2 + m_3x_3 \ldots\ldots\ldots\ldots m_nx_n + C$ ……..……..(1)

Where, Y is the biological activity and x_1,x_2, x_3 and x_n are independent parameters, m is slope of the correlation expression and C is the intercept of the straight line correlation expression.

Data Set

The data for the present study of a set of 37 aromatic compounds carrying –COOH, –OH, $–NH_2$ functional groups along with their biological activity in term of log P were taken from the literature and is given in Table -1. The aim of this study is to develop a QSAR model to correlate the structural features of these set of aromatic compounds with their biological activity using topological descriptors. A perusal of Table -1 reveals that these compounds can be arranged in decreasing order of lipophilicity (log P) as according to the Compound number given below:

7>4>6>24=25>5>2=3>22=23>18>13>11=12>8>20=21>34>9>35>
32>26=27>1>17>33>29=30=31>14=15=16>19>10>36>37>28

This means that 4-Bromo benzoic acid has highest lipophilicity (log P) while aniline has lowest value of lipophilicity (log P). These sequences do not give any structure-activity relationship. No one is to one correlation is seen in the activity and structure of the compounds. Therefore, statistical analysis is required to obtain a model for modeling the log P of present set of compounds. The structures of the compounds were drawn from ACD labs Chem. Sketech soft ware[12].

Table 1: List of 37 Aromatic compounds along with their Biological Activity (Log P)

Comp. No.	Name of Compounds	Log P	Comp. No.	Name of Compounds	Log P
1	Benzoic Acid	1.885	20	m-Cresol	2.124
2	3-Methyl Benzoic Acid	2.534	21	p-Cresol	2.124
3	4-Methyl Benzoic Acid	2.534	22	3-Chloro Phenol	2.485
4	3-Chloro Benzoic Acid	2.696	23	4-Chloro Phenol	2.485
5	4-Chloro Benzoic Acid	2.596	24	3-Bromo Phenol	2.635
6	3-Bromo Benzoic Acid	2.686	25	4-Bromo Phenol	2.635
7	4-Bromo Benzoic Acid	2.846	26	3-Flouro Phenol	1.915
8	3-Flouro Benzoic Acid	2.126	27	4-Flouro Phenol	1.915
9	4-Flouro Benzoic Acid	2.036	28	Aniline	0.915
10	Phenyl Acetic Acid	1.414	29	o-Toluedine	1.564
11	3-Chloro Phenyl Acetic Acid	2.127	30	m-Toluedine	1.564
12	4-Chloro Phenyl Acetic Acid	2.127	31	p-Toluedine	1.564
13	3-Bromo Phenyl Acetic Acid	2.277	32	3-Chloro-Aniline	1.930
14	2-Flouro Phenyl Acetic Acid	1.557	33	4-Chloro-Aniline	1.730
15	3-Flouro Phenyl Acetic Acid	1.557	34	3-Bromo Aniline	2.080
16	4-Flouro Phenyl Acetic Acid	1.557	35	4-Bromo-Aniline	2.020
17	3-Phenyl Propionic Acid	1.863	36	3-Fluoro-Aniline	1.360
18	4-Phenyl Butanoic Acid	2.392	37	4-Fluoro-Aniline	1.260
19	Phenol	1.475			

Used Descriptors

We have used following topological descriptors for modeling lipophilicity of present set of aromatic compounds.

1. First order Valence connectivity index ($^1\chi^V$):
2. Balaban-type Index[13] from Z weighted distance matrix ($Jhet_Z$)
3. Xu Index(X_u)
4. Balaban type Index (F)

All the topological descriptors employed in the present study were calculated using hydrogen suppressed graph[14-16] by DRAGON Software[17].It is pertinent to mention that we calculated more than 30 descriptors from the DRAGON Software, but variable selection suggested that only $^1\chi^V$**,** $Jhet_Z$ and X_u are useful in modeling the log P activity of present set of compounds**.** We also calculated **F** index for these compounds. The calculated parameters are reported in Table 2.

Indicator Parameters

To improve the QSAR models two dummy indicator parameters have been used which are also called *denovo* constants to account for the functional groups. These indicator parameters I_1 and I_2 are reported in Table 2.

I_1 is given a value of 1 if **– COOH** group is present in the structure otherwise its value is zero.

I_2 is given a value of 1 if **– NH_2** group is present in the compound otherwise its value is zero.

Table 2: Indicator Parameters and Calculated Values of Descriptors

Comp. No.	I_1	I_2	$^1\chi^V$	$Jhet_Z$	X_u	F
1	1	0	2.588	3.089	8.860	3.076
2	1	0	2.999	3.162	9.861	3.174
3	1	0	2.999	3.091	9.950	3.150
4	1	0	3.066	3.339	9.861	3.256
5	1	0	3.066	3.254	9.950	3.284
6	1	0	3.481	3.395	9.861	2.692
7	1	0	3.481	3.305	9.950	2.734
8	1	0	2.688	3.249	9.861	3.342
9	1	0	2.688	3.172	9.950	2.962
10	1	0	3.046	2.743	10.163	3.106
11	1	0	3.523	2.972	11.135	3.102
12	1	0	3.523	2.906	11.226	3.090
13	1	0	3.938	3.009	11.135	2.290
14	1	0	3.151	2.987	11.059	3.524
15	1	0	3.145	2.912	11.135	3.572
16	1	0	3.145	2.851	11.226	3.044
17	1	0	3.546	2.501	11.226	3.044
18	1	0	4.046	2.325	12.668	3.690
19	0	0	2.136	3.141	6.529	4.246
20	0	0	2.545	3.173	7.614	3.986
21	0	0	2.545	3.124	7.663	3.864
22	0	0	2.612	3.455	7.614	4.044
23	0	0	2.612	3.392	7.663	3.364
24	0	0	3.027	3.550	7.614	3.346
25	0	0	3.027	3.480	7.663	3.520
26	0	0	2.234	3.299	7.614	3.662
27	0	0	2.234	3.254	7.663	3.480
28	0	1	2.199	3.089	6.529	4.246
29	0	1	2.616	3.191	7.576	4.558
30	0	1	2.610	3.131	7.614	4.462
31	0	1	2.610	3.084	7.663	4.384
32	0	1	2.677	3.406	7.614	4.642
33	0	1	2.677	3.345	7.663	4.684
34	0	1	3.092	3.345	7.614	4.384
35	0	1	3.092	3.432	7.663	4.384
36	0	1	2.299	3.264	7.614	4.462
37	0	1	2.299	3.211	7.663	4.384

Results and Discussion

We have tried variable selection regression to get multi-parametric correlations and the results are reported in Table 3. The indicator parameter I_2 (presence of -NH_2 functional group) plays appreciable role in modeling of Log P of present set of compounds. Similarly, X_u Index also plays a dominating role in six parametric model using maximum R^2 method. The above table shows that, I_2, J_{hetz}, $^1\chi^V$, I_1, F, X_u are suitable parameters for modeling log P of present set of compounds.

Table 3: Selection Results Section with Variable Names

Model Size	R- Squared	Change in R- Square value	Parameters
1	0.290311	0.290311	I_2
2	0.571359	0.281048	I_2, J_{hetz}
3	0.765580	0.194220	I_2, J_{hetz}, $^1\chi^V$
4	0.817232	0.051653	I_2, J_{hetz}, $^1\chi^V$, I_1
5	0.826859	0.009627	I_2, J_{hetz}, $^1\chi^V$, I_1, F
6	0.830513	0.003654	I_2, J_{hetz}, $^1\chi^V$, I_1, F, X_u

The suitable parameters in variable selection process are shown in Table 3. This table shows that the drastic change in R^2 value is seen from mono to six-parametric correlation. R^2 is variance square of correlation constant. The graph is plotted in between R-squared and variables for selection regression in Figure No.1.

Plots Section

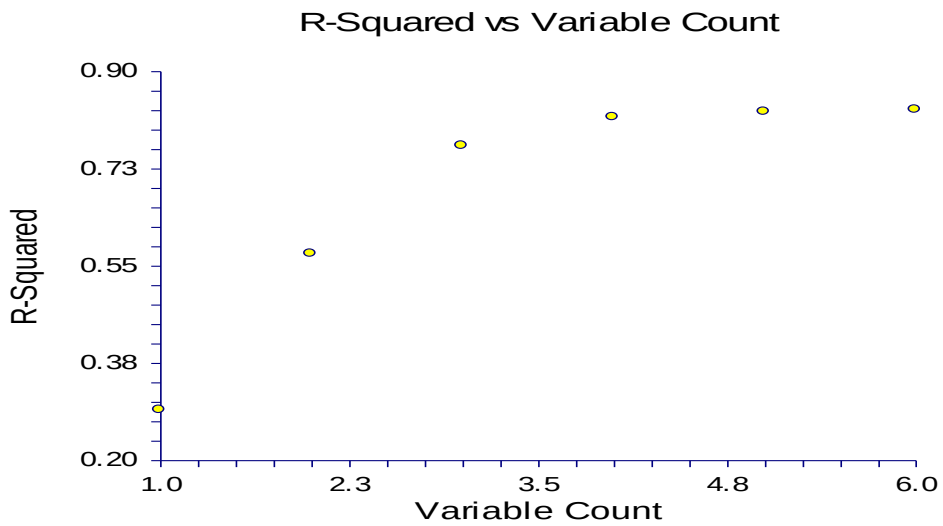

Figure 1: Graph between the R-squared and variables for selection regression of set of 37 aromatic compounds

The data was subjected to statistical analysis. We calculated correlation matrix of various parameters which considered for modeling the Log P lipophilicity value of aromatic compounds used in the present study. The derived correlation matrix is showing inter correlation among all the parameters is shown in the Table 4.

Table 4: Correlation Matrix

	I_2	J_{hetz}	$^1\chi^V$	I_1	F	X_u	Log P
I_2	1.0000						
J_{hetz}	0.2333	1.0000					
$^1\chi^V$	-0.3625	-0.4528	1.0000				
I_1	-0.0620	0.4740	0.1437	1.0000			
F	0.8074	0.2542	-0.5930	-0.1647	1.0000		
X_u	-0.5419	-0.6740	0.8166	0.0941	-0.6991	1.0000	
Log P	-0.5388	0.3011	0.4818	0.2709	-0.5201	0.2565	1.0000

A close look at this table gives following information:

1. Log P shows best correlation with indicator parameter I_2.
2. Indicator parameter I_1 is also correlated with Log P.
3. Valence connectivity indices $^1\chi^V$ also show high correlation.
4. J_{hetz}, F and X_u are the best parameters to be used in multi-parametric modeling.
5. Indicator parameter I_1which accounts for the presence of carboxylic functional group (-COOH) may be a suitable parameter in multi-parametric analysis.

On the basis of above, the data was subjected to regression analysis[18] and the obtained models are reported in Table 5. This table includes some statistically significant models which have been discussed in subsequent paragraphs.

Table 5: Regression Parameters and Quality of Correlations Obtained Models

Model No.	Parameters Used	A_i (i=1…..6)	B	Se	R^2	R^2A	F	Q=R/Se
1.	I_2	-0.5718(±0.1511)	2.1705	0.0786	0.2903	0.2700	14.317	6.854
2.	I_2	-0.6836(±0.1346)	-0.4312	0.7340	0.4830	0.4526	15.881	0.9467
	J_{hetz}	0.8350(±0.2346)						
3.	I_2	-0.5036(±0.0963)	-3.6999	0.7213	0.7656	0.7443	35.924	1.212
	J_{hetz}	0.6218(±0.0986)						
	$^1\chi^v$	1.2846(±0.1755)						
4.	I_2	-0.2873(±0.0955)	-5.0296	0.7835	0.8172	0.7944	35.771	1.1537
	J_{hetz}	-0.5239(±0.0866)						
	$^1\chi^v$	0.7464(±0.0976)						
	I_1	1.6528(±0.1993)						
5.	I_2	-0.2773(±0.0948)	-5.8044	0.9740	0.8269	0.7969	29.609	0.9336
	J_{hetz}	-0.6702(±0.1405)						
	$^1\chi^v$	0.8180(±0.1109)						
	I_1	1.6690(±0.1975)						
	F	0.1514(±0.1153)						
6.	I_2	-0.3274(±0.1139) -	-6.8434	1.6213	0.8305	0.7966	24.501	0.5620
	J_{hetz}	0.6577(±0.1422)						
	$^1\chi^v$	0.7507(±0.1394)						
	I_1	1.8686(±0.3179)						
	F	0.0557(±0.0693)						
	Xu	0.1884(±0.1248)						

Mono-parametric model

The mono-parametric model with highest R^2 value is 0.2903 with I_2 indicator parameter. The model is as under :

Log P = -0.5718(±0.1511)I_2+ 2.1705 ……..(2)

N = 37, Se = 0.0786, R^2 = 0.2903

R^2A = 0.2700, F = 14.317, Q = 6.854

Here and here onwards 'N' is total the number of compounds, R^2 is variance square of correlation constant, R^2A is adjusted Regression coefficient, Se is standard error of estimation, F is Fischer's Ratio and 'Q' is the Pogliani's quality factor[19], which is a ratio of R/Se.

Bi-parametric model

When $Jhet_Z$ is added to the mono-parametric model discussed above, a bi – parametric model is obtained. A drastic improvement in R^2 is observed. The R^2value changes from 0.2903 to 0.4830. The adjusted R^2A also changes from 0.2700 to 0.4526, which clearly indicates that the addition of $Jhet_Z$ is significant and it has fair share in the model. The model is as under

log P = -0.6836(±0.1346) I_2+0.8350(±0.2346)J_{hetz} –0.4312…….(3)

N =37, Se = 0.7340, R^2 = 0.4830
R^2A = 0.4526, F = 15.881, Q = 0.9467

Tri-parametric model

A tri-parametric model is obtained when $^1\chi^v$ added to the above model. For this model a better improvement is observed in R^2 value. The value of R^2 changes from 0.4830 to 0.7656 and R^2A also increases 0.4526 to 0.7443. Hence the addition of $^1\chi^v$ is significant and has fair share in the model. The model is an under

Log P=-0.5036(±0.0963)I_1+0.6218(±0.0986)J_{hetz}+1.2846(±0.1755)$^1\chi^V$–3.6999…….. (4)

N =37,	Se = 0.7213	R^2= 0.7656
R^2A = 0.7443,	F = 35.924,	Q = 1.212

Four-parametric model

To get a better model we tried for a higher parametric model. For this I_1 is added to the above model which resulted into a four-parametric model with R^2 equal to 0.8172 and value of R^2A comes to be 0.7944. This clearly shows that the addition of I_1 is favourable and has a significant contribution in the model. The model is as under

$$\log P = -0.2873(\pm 0.0955)\, I_1 - 0.5239\,(\pm 0.0866)\, J_{hetz} + 0.7464(\pm 0.0976)\,{}^1\chi^V + 1.6528(\pm 0.1993) I_1 - 5.0296 \quad \ldots\ldots\ldots\ldots (5)$$

N =37	Se = 0.7835,	R^2 = 0.8172
R^2A = 0.7944,	F = 35.771,	Q = 1.1537

In this model the R^2 changes from 0.7656 to 0.8172 which shows that the added indicator parameter I_1 has its fair share in the model.

Five Parametric model

When Balaban type 'F' parameter is added to above model, R^2 slightly increases from 0.8172 to 0.8269. Similarly, R^2A also shows significant improvement. . The model is as below

$$\log P = -0.2773(\pm 0.0948)\, I_2 - 0.6702(\pm 0.1405) J_{hetz} + 0.8180(\pm 0.1109)\,{}^1\chi^v + 1.6690(\pm 0.1975) I_1 + 0.1514(\pm 0.1153) F - 5.8044 \quad \ldots\ldots (5)$$

N = 37,	Se = 0.9740,	R^2 = 0.8269
R^2A = 0.7969,	F = 29.609,	Q = 0.9336

Six-Parametric model

Finally, a six parameter model with I_2, J_{hetz}, ${}^1\chi^v$, I_1,F and X_u The R^2 value comes out to be 0.8305 as compare to 5-parametric model mentioned above with 0.8269. This model explains more than 83% variance of the data. This model is the best among all the above proposed models. The model is as below

$$\text{LogP} = -0.3274(\pm 0.1139) I_2 - 0.6577(\pm 0.1422) \text{Jhetz} + 0.7507(\pm 0.1394)\,{}^1\chi^v + 1.8686 (\pm 0.3179) I_1 + 0.05571(\pm 0.0693) F + 0.1884(\pm 0.1248) X_u - 6.8434 \quad \ldots\ldots\ldots\ldots\ldots (7)$$

N =37,	Se = 1.6213,	R^2= 0.8305
R^2A = 0.7966,	F = 24.501,	Q = 0.5620

To support the above finding we calculated the log P lipophilicity value of the compounds used in the present study using the best six-parametric model. Such values are reported in the Table 6. These values are in good agreement with the observed values. A comparison between observed and estimated values are demonstrated in figure 2. The potential of the model comes out to be 0.830 which further suggests that the six-parametric model is the best suited for modeling the log P value of the present set of the aromatic compounds.

Cross validation parameters also support our findings. The cross validated parameters for different models are reported in Table 7. The lowest PRESS/SSY value 0.3211 and higher R^2_{CV} value 0.7203 for the six parametric model confirms our findings. The PSE value 0.2491 for this model is lowest and S_{PRESS} comes out to be 0.2767, further verifies our results. The ridge traces is recorded in figure 3.

In order to explain whether or not the proposed model is free from the defect of collinearity, we have calculated VIF(variance inflation factor), Eigen value (λi), condition No (k) and tolerance (T) for all the independent parameters used in proposed models. These values are reported in Table -8. All the four ridge parameters are within the permissible range. Therefore, from this point view also proposed six-parametric model is free from the defect of collinearity. The VIF plot is recorded in figure 4.

Table 6: Residual Report of Observed and Estimated value of Log P

Compd. N	Obs. Log	Est. Log	Residual
1	1.885	1.945	-0.060
2	2.534	2.464	0.070
3	2.534	2.331	0.203
4	2.696	2.533	0.163
5	2.596	2.384	0.212
6	2.686	2.843	-0.157
7	2.846	2.687	0.159
8	2.126	2.097	0.029
9	2.036	1.887	0.149
10	1.414	1.720	-0.306
11	2.127	2.232	-0.105
12	2.127	2.112	0.015
13	2.277	2.460	-0.183
14	1.557	2.056	-0.499
15	1.557	1.925	-0.368
16	1.557	1.716	-0.159
17	1.863	1.691	0.172
18	2.392	1.939	0.453
19	1.475	1.793	-0.318
20	2.124	2.171	-0.047
21	2.124	2.059	0.065
22	2.485	2.432	0.053
23	2.485	2.189	0.296
24	2.635	2.790	-0.155
25	2.635	2.694	-0.059
26	1.915	1.785	0.130
27	1.915	1.669	0.246
28	0.915	1.085	-0.170
29	1.564	1.706	-0.142
30	1.564	1.574	-0.010
31	1.564	1.474	0.090
32	1.930	1.844	0.086
33	1.730	1.741	-0.011
34	2.080	1.993	0.087
35	2.020	2.158	-0.138
36	1.360	1.261	0.099
37	1.260	1.150	0.110

Table 7: Cross Validation Parameters for Obtained Models

Model no.	Parameter Used	PRESS/SSY	R^2cv	S_{PRESS}	PSE
1.	I_2	0.5014	0.2123	0.4645	0.4182
2.	I_2,J_{hetz}	0.4949	0.3139	0.4334	0.3902
3.	$I_2,J_{hetz},{}^1\chi^V$	0.3616	0.6722	0.2996	0.2698
4.	$I_2,J_{hetz},{}^1\chi^V,I_1$	0.3550	0.6767	0.2976	0.2679
5.	$I_2,J_{hetz},{}^1\chi^V,I_1,F$	0.3421	0.6912	0.2908	0.2618
6.	$I_2, J_{hetz},{}^1\chi^V,I_1,F,X_u$	0.3211	0.7203	0.2767	0.2491

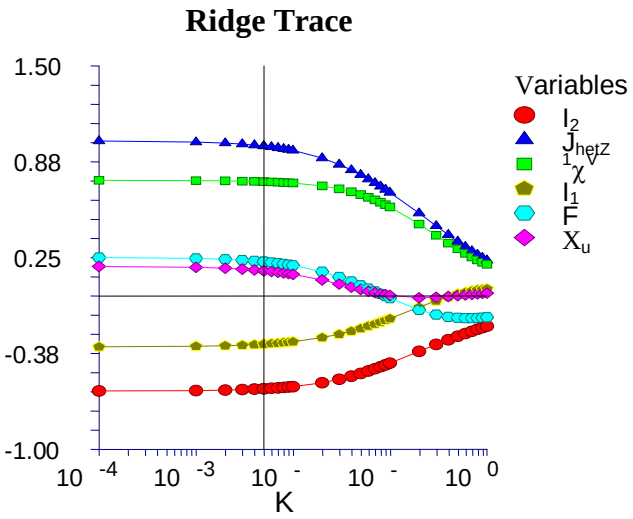

Figure 3: Ridge Trace for Best six parametric Model

Table 8: Ridge Regression Parameters for Obtained Models

Model no.	Parameter Used	VIF	T	λi	K
1.	I_2	0.9901	1.0000	1.0000	1.00
2.	I_2	1.0459	0.9456	1.2333	1.00
	J_{hetz}	1.0459	0.9456	0.7666	1.61
3.	I_2	1.1446	0.8625	0.7066	1.00
	$J_{hetz,}$	1.2479	0.7894	0.7748	2.20
	$^{1}{}_{v}\chi$	1.3561	0.7251	0.5185	3.29
4.	I_2	1.1515	0.8573	1.7481	1.00
	$J_{hetz,}$	1.9708	0.4917	1.2962	1.35
	$^{1}{}_{v}\chi$	1.6381	0.5946	0.6869	2.54
	I_1	1.5961	0.6099	0.2686	6.51
5.	I_2	2.9616	0.3185	2.3960	1.00
	J_{hetz}	1.9766	0.4898	1.4421	1.66
	$^{1}{}_{v}\chi$	2.1297	0.4508	0.7441	3.22
	I_1	1.6072	0.6060	0.2792	8.58
	F	3.9364	0.2378	0.1385	17.30
6.	I_2	2.9989	0.3147	3.2256	1.00
	J_{hetz}	4.6044	0.1913	1.4488	2.23
	$^{1}{}_{v}\chi$	3.2213	0.2884	0.8242	3.91
	I_1	2.1752	0.4284	0.3005	10.73
	F	4.4868	0.2054	0.1404	22.96
	Xu	8.6951	0.0987	0.0603	53.44

VIF=Variance Inflation Factor, T=Tolerance, λi=Eigen values, k=Condition Number

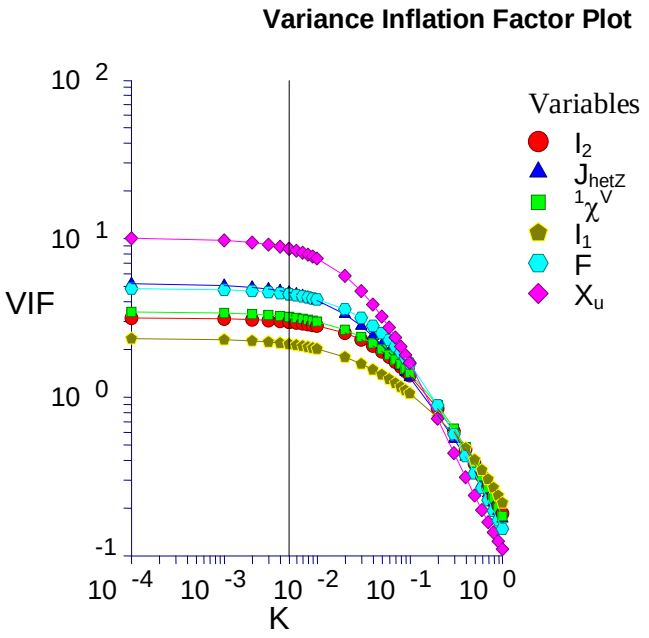

Figure 4: VIF plot for best six parametric model

Conclusion

$LogP= -0.3274(\pm0.1139)I_2-0.6577(\pm0.1422)J_{hetz}+0.7507(\pm0.1394)^1\chi^v +1.8686 (\pm0.3179)I_1+0.05571(\pm0.0693)\ F+0.1884(\pm0.1248)X_u -6.8434$

On the basis of our study following conclusions may be drawn.

- The J_{hetz} and I_2 have negative coefficient indicating that these parameter will have negative impact towards activity.
- The $^1\chi^v$, I_1, F and Xu have positive coefficient in the best model suggesting that these parameters have positive role towards the lipophilicity.
- Since I_1 has a positive coefficient, presence of -COOH functional group is very effective in modeling the activity of present set of aromatic compounds.
- In designing the new compounds functional group with lower values of the I_2 is preferred.

Acknowledgement

Author is thankful to Prof. Vijay Kumar Agrawal, Head, Department of Chemistry, APS University, Rewa (M.P.) for sharing the knowledge and the help which resulted into the above paper. She is also thankful to late Prof. P.V. Khadikar well known topologist, who initiated this work.

References

1. Muetterties L.E. and Bleeke R.J.,Catalytic hydrogenation of aromatic hydrocarbons: Accounts of chemical research, 12(9): 324-331 **(1979)**

2. Backlund M., Mechanism of Activation of the Aryl Hydrocarbon Receptor: Institute of Environmental medicine, Karolinska **(2007)**

3. Manikpuri A.D. and Khadikar P.V., Chemical shift in NMR as a molecular descriptors in computer Aided Drug designing: Topology and NRM, National Academy Science Letter **(2009)**

4. Hammet L.P., Some Relations between Reaction Rates and Equilibrium Constants: Chemical reviews.17(1): 125 -136 **(1935)**

5. Jantschi L.L., Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology, Leonardo Electronic Journal of practices and technologies, 4(6): 76 -98 **(2005)**

6. Raevsky O.A., Schaper K.J., Quantitative estimation of hydrogen bond contribution to permeability and absorption processes of some chemicals and drugs, 33(10): 799-807 **(1998)**

7. Shcaper K.J., Zhang H. and Raevsky O.A., pH-Dependent Partitioning of Acidic and Basic Drugs into Liposomes—A Quantitative Structure-Activity Relationship Analysis- Quantitative Sturcture Activity relationship. 20(1): 46-54 **(2001)**

8. Khadikar P.V., Singh. S. and Shrivastava A., On the topological evidences for modelling lipophilicity On the topological evidences for modelling lipophilicity :Bioorganic and Medicinal chemistry, 10: 3981-3996 **(2002)**

9. Agrawal V.K., Singh J., Bruno L., Joshi S., Joshi A., and Khadikar P.V., Topology of molecule and its lipophilicity. Current computer aided drug design, 2(4): 1-35 **(2006)**

10. Hansch C., Leo A. and Hockman D., Exploring QSAR: Hydrophobic, Electronic and Steric constants. ACS Professional Reference Book, Washington page 110 (**1995**)

11. Hansch C., The QSAR paradigm in the design of less toxic molecules. Drug Metab., Rev., 15: 1279-1294 (**1984**)

12. Chem. Sketch ACD /Labs Software:(11.01 version) wwwl.acdlabs.com

13. Balaban A.T., Khadikar P.V., Supuran C.T., Thakur A. and Thakur M., Orbital Symmetry in QSAR: Some Schiff's Base Inhibitors of Carbonic Anhydrase: Bioorganic and Medicinal Chemistry Letters, 15: 3966-3973 **(2005)**

14. Todeschini R. and Cosonnin V., Handbook of molecular Descriptor, Wiely-VCH: Weinheim **(2000)**

15. Trinajstic N., Chemical Graph Theory, CRC Press: Boca Raton Florida **(1992)**

16. Karelson M., Molecular Descriptors in QSAR/QSPR, John Wiely & Sons, New York **(2000)**

17. DRAGON Software : Roberto Todeschini, Milano chemo metrics and QSAR research group, University of Milano-Bicocea, P.za della scienza,1-20126' Milano, Italy. Website: http://www.disat.unimibit/chem/ .

18. Chatarjee S., Hadi A.S. and Price, B. Regression Analysis by examples, 3rd Ed., Wiley: New York (**2000**)

19. Pogliani L.J., Physical Chemistry, 100: 18065 – 18077 **(1996)**