Topological Modeling of Lipophilicity of Some Alkanes

SHANNO PATHAN, S.M. ALI, SALMA BEE KHAN and MEENAL SHRIVASTAVA

Department of Chemistry, Islamia Karimia College, Indore (M.P.)
Department of Chemistry, Govt. Holkar Science College, Indore (M.P.)
Email of Corresponding Author: shannopathan355@gmail.com
http://dx.doi.org/10.22147/juc/120301

Acceptance Date 27th Sep., 2016, Online Publication Date 2nd Nov. 2016

Abstract

In this work a study has been done to model the lipophilicity of few alkanes using topological indices. The result says that the indicator parameters play dominating role in predicting the activity of alkanes.

Key words: Topological modeling, QSPR, lipophilicity, log P, indicator parameters.

Introduction

Lipophilicity in terms of log P is an important parameter in drug development. Log P has been modeled by Khadikar and Agrawal for many compounds/drugs using topological indices. These indices include Randic connectivity, Kier & hall valence connectivity, Wiener index, Balaban & Balaban type indices. Sz index, Padmakar – Ivan index etc. It has also been shown that in case of derivatives the contribution of functional groups can be studied by adding indicator parameters which are termed as de novo constants. These can be abbreviated as I_1, I_2, etc. They are usually assigned a value of 1 for their presence and if the group is absent they are given value of zero. e.g. if in a series a functional group –CH_3 is present in few of the compounds then I_1 will be assigned value of 1 for the compound with –CH_3 group and others will be given zero.

Methodology

In fact Hansch has suggested that the biological activity of a series of compounds can be modeled by using few independent parameters and the equation for this can be—

$$Y = m_1 x_1 + m_2 x_2 + ....... + m_i x_i + C$$ (1)

Where, Y is biological activity & x_1, x_2, .... are independent parameters C is intercept of the straight line. For obtaining the model one can use regression analysis for which standard softwares are available.

Calculation of Xu and F:

Xu index is a topological molecular descriptor based on adjacency and distance matrices. It was calculated using the following expression:

$$Xu = \sqrt{n \log L} = \sqrt{n \log \frac{\sum \xi_i \alpha_j}{\sum \xi_i \alpha_j}}$$ (2)

Where, ‘n’ is the number of atoms and ‘L’ represents the valence average topological distance calculated by vertex degree \(\delta\) and vertex distance.
degree $\sigma$ of all the atoms. It was proposed as a particularly high discriminate molecular descriptor accounting for molecular size and branching.

For those properties where not only the “shape” but also the size of graph influences the property or activity, other Balaban index $F_{16}$ has been developed. It defines as:

$$F = J(R+1) = \frac{E}{S_{\text{alledges}}[\text{di. dj}]^{1/2}}$$  \hspace{1cm} (3)

Where, $J$ is Balaban index and $R$ is cyclomatic number, this index is able to separate graph size, cyclicity and branching. For the calculation of Balaban index ‘$J$’ and cyclomatic number ‘$R$’ following expressions were used.

$$J = \frac{E}{(R+1)} = \frac{S_{\text{alledges}}[\text{di. dj}]^{1/2}}{2}$$  \hspace{1cm} (4)

Where, $E$ is the number of bonds in a graph, $R$ is the cyclomatic number.

$$R = E - N + 1$$  \hspace{1cm} (5)

Where, $E$ is number of edges in graph and $N$ is number of vertices.

In the present paper we have taken 36 alkanes including cyclo compounds & -chloro, -nitro, -methyl derivatives. These alkanes and their derivatives are reported in Table- 1. This Table also includes their lipophilicity values which have been expressed in the form of logP. These compounds can be arranged in decreasing order of activity as below:

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

Table1. List of 36 Alkanes along with their biological activity (log P)

| S.No. | Name of compounds  | Log p  | S.No. | Name of compounds  | Log p  |
|-------|--------------------|--------|-------|--------------------|--------|
| 1     | n-Butane           | 2.81   | 19    | 1-Chloro Propane   | 1.994  |
| 2     | n-Pentane          | 3.339  | 20    | 1-Chloro Butane    | 2.523  |
| 3     | n-Hexane           | 3.868  | 21    | 1-Chloro Pentane   | 3.052  |
| 4     | n-Heptane          | 4.397  | 22    | 1-Chloro Hexane    | 3.581  |
| 5     | n-Octane           | 4.926  | 23    | 1-Chloro Heptane   | 4.11   |
| 6     | n-Nonane           | 5.455  | 24    | Carbon Tetra Chloride | 2.875 |
| 7     | n-Decane           | 5.984  | 25    | 1,2-Di chloro Ethane | 1.458 |
| 8     | n-Undecane         | 6.513  | 26    | 1,1,1-Tri Chloro Ethane | 2.481 |
| 9     | 2-Methyl Butane    | 3.209  | 27    | 1,1,2,2-Tetra Chloro Ethane | 2.644 |
| 10    | 2-Methyl Pentane   | 3.738  | 28    | Penta Chloro Ethane | 3.627  |
| 11    | 2-Methyl Hexane    | 4.267  | 29    | 1,2-DiBromo Ethane | 1.738  |
| 12    | 2,2,4-Tri Methyl Pentane | 4.536 | 30    | Flouro Tri Chloro Methane | 2.435 |
| 13    | Cyclopentane       | 2.795  | 31    | Nitro Methane     | -0.284 |
| 14    | Cyclo hexane       | 3.354  | 32    | NitroEthane       | 0.245  |
| 15    | Cyclo Heptane      | 3.913  | 33    | 1-Nitro Propane   | 0.774  |
| 16    | Cyclo Octane       | 4.472  | 34    | 1-Nitro Butane    | 1.303  |
| 17    | Adamantane         | 3.982  | 35    | 1-Nitro Pentane   | 1.832  |
| 18    | 1-Chloroethane     | 1.465  | 36    | 2-Nitro Propane   | 0.554  |

No one is to one correlation is seen in the activity and structure of the compounds. The structures of the molecules were drawn from Chem.Sketch software which is freely available from ACD Labs. The molecules so drawn were used for the calculation of topological indices for which mol files are needed, therefore hydrogen depleted graphs were drawn and these mol files were used for the calculation of topological indices using DRAGON software. More than 50 topological indices were calculated, but variable selection suggested that only $X_u$ index is useful in modeling the log P of present set of compounds. We also calculated $F$ index for these compounds. These calculated parameters are reported in Table-2. The data was subjected to statistical analysis. We calculated correlation matrix of various parameters which are considered for modeling the log P of compounds used in the present study.
Table 2. Indicator parameters and calculated values of descriptors

| Comp. No. | I₁ | I₂ | Xu  | F     |
|-----------|----|----|-----|-------|
| 1         | 0  | 0  | 3.161 | 1.974 |
| 2         | 0  | 0  | 4.593 | 2.19  |
| 3         | 0  | 0  | 5.969 | 2.339 |
| 4         | 0  | 0  | 7.297 | 2.447 |
| 5         | 0  | 0  | 8.584 | 2.53  |
| 6         | 0  | 0  | 9.835 | 2.595 |
| 7         | 0  | 0  | 11.053 | 2.647 |
| 8         | 0  | 0  | 12.242 | 2.692 |
| 9         | 0  | 0  | 4.301 | 2.52  |
| 10        | 0  | 0  | 5.699 | 2.62  |
| 11        | 0  | 0  | 7.056 | 2.678 |
| 12        | 0  | 0  | 7.752 | 3.388 |
| 13        | 0  | 0  | 4.006 | 4.166 |
| 14        | 0  | 0  | 5.382 | 4.000 |
| 15        | 0  | 0  | 6.574 | 4.082 |
| 16        | 0  | 0  | 7.842 | 8.000 |
| 17        | 0  | 0  | 9.32  | 9.485 |
| 18        | 1  | 0  | 1.655 | 1.632 |
| 19        | 1  | 0  | 3.161 | 1.974 |
| 20        | 1  | 0  | 4.593 | 2.19  |
| 21        | 1  | 0  | 5.699 | 2.339 |
| 22        | 1  | 0  | 7.297 | 2.447 |
| 23        | 1  | 0  | 8.584 | 2.53  |
| 24        | 1  | 0  | 3.972 | 3.024 |
| 25        | 1  | 0  | 3.161 | 1.975 |
| 26        | 1  | 0  | 3.972 | 3.024 |
| 27        | 1  | 0  | 5.395 | 2.993 |
| 28        | 1  | 0  | 6.367 | 3.541 |
| 29        | 1  | 0  | 3.161 | 1.975 |
| 30        | 1  | 0  | 3.972 | 3.023 |
| 31        | 0  | 1  | 2.894 | 2.324 |
| 32        | 0  | 1  | 4.301 | 2.54  |
| 33        | 0  | 1  | 5.699 | 2.627 |
| 34        | 0  | 1  | 7.056 | 2.678 |
| 35        | 0  | 1  | 8.37  | 2.716 |
| 36        | 0  | 1  | 5.395 | 2.993 |

The derived correlation matrix is showing inter-correlation among all the parameters is shown in the Table-3.

Table 3. Correlation Matrix showing inter-correlation among all the parameters

|       | I₁  | I₂  | X_u | F    | Log p |
|-------|-----|-----|-----|------|-------|
| I₁    | 1.0000 |     |     |      |       |
| I₂    | -0.3362 | 1.0000 |     |      |       |
| X_u   | -0.3979 | -0.0687 | 1.0000 |     |       |
| F     | -0.2369 | -0.1876 | 0.3038 | 1.0000 |       |
| Log p | -0.2139 | -0.6696 | 0.7475 | 0.2411 | 1.0000 |
A close look at this table gives following information.

1. No auto correlation exists with any of the independent parameters.
2. Log P shows strong correlation with $X_u$ parameter.
3. Indicator parameter $I_2$ is also correlated with log P.
4. Therefore, the only mono-parametric model which is statistically acceptable may be $X_u$.

5. Indicator parameter $I_2$ which accounts for the presence of—NO2, functional group may be a suitable parameter in multi-parametric analysis.

On the basis of above, the data was subjected to regression analysis & the models obtained are reported in Table 4. This table includes few statistical parameters which have been discussed in subsequent paragraphs.

Table 4. Regression Parameters and Quality of correlation of obtained Models

| Model no. | Parameter Used | $A_i$ (i=1…4) | B   | Se  | $R^2$ | $R^2_A$ | F     | Q     |
|----------|----------------|----------------|------|-----|-------|--------|-------|-------|
| 1        | Xu             | 0.4792(±0.0730)| 0.1842 | 0.4717 | 0.5587 | 0.5458 | 43.052 | 1.584 |
| 2        | F              | 0.2894(±0.1997) | 2.2405 | 0.6181 | 0.0581 | 0.304  | 2.099  | 0.389 |
| 3        | Xu             | 0.5045(±0.0800) | -0.0827 | 0.5812 | 0.5670 | 0.5408 | 21.607 | 1.295 |
|          | $I_1$          | 0.3195(±0.4023) |       |      |       |        |       |       |
| 4        | Xu             | -0.2596(±0.1734) | 0.7780 | 0.1770 | 0.9428 | 0.9393 | 271.718 | 5.485 |
|          | $I_2$          | 0.0268(±0.4518) |       |      |       |        |       |       |
| 5        | F              | -0.5354(±0.5525) | 2.5668 | 0.7043 | 0.0842 | 0.0287 | 1.517  | 0.412 |
|          | $I_1$          | 0.2421(±0.2058) |       |      |       |        |       |       |
| 6        | F              | 0.1437(±0.560) | 3.0983 | 0.5045 | 0.4622 | 0.4299 | 14.178 | 1.347 |
|          | $I_2$          | 0.1437(±0.1560) |       |      |       |        |       |       |
| 7        | Xu             | 0.4762(±0.778) | 0.1500 | 0.5486 | 0.5590 | 0.5322 | 20.911 | 1.362 |
|          | F              | 0.0186(±0.186) |       |      |       |        |       |       |
| 8        | F              | -0.1636(±0.4935) | 4.3615 | 0.4935 | 0.6670 | 0.6358 | 21.363 | 1.654 |
|          | $I_1$          | -0.0459(±0.1318) |       |      |       |        |       |       |
|          | $I_2$          | -3.5198(±0.4704) |       |      |       |        |       |       |
| 9        | Xu             | 0.3974(±0.0215) | 1.3887 | 0.1669 | 0.9713 | 0.9687 | 361.577 | 5.905 |
|          | $I_1$          | -0.6472(±0.1145) |       |      |       |        |       |       |
|          | $I_2$          | -0.8843(±0.1357) |       |      |       |        |       |       |
| 10       | Xu             | 0.4999(±0.08366) | -0.1567 | 0.6698 | 0.5677 | 0.5272 | 14.010 | 1.124 |
|          | F              | 0.3222(±0.4119) |       |      |       |        |       |       |
|          | $I_1$          | 0.0344(±0.1477) |       |      |       |        |       |       |
| 11       | Xu             | 0.4709(±0.0621) | 1.0191 | 0.1917 | 0.9519 | 0.9474 | 210.928 | 0.783 |
|          | F              | -0.1221(±0.0496) |       |      |       |        |       |       |
|          | $I_2$          | -2.5064(±0.1640) |       |      |       |        |       |       |
| 12       | Xu             | -0.7676(±0.0703) | 1.8570 | 0.1171 | 0.9901 | 0.9888 | 772.544 | 8.497 |
|          | F              | -0.4153(±0.0839) |       |      |       |        |       |       |
|          | $I_1$          | -0.4153(±0.0131) |       |      |       |        |       |       |
|          | $I_2$          | -0.1796(±0.0235) |       |      |       |        |       |       |
Result and Discussion

The best mono parametric with $R^2 = 0.5587$ is with $X_u$ and is as below:

$$\text{Log } P = 0.4792 \pm 0.0730 \times X_u + 0.1842$$  \hspace{1cm} (6)

$n = 36$, $Se = 0.4717$, $R^2 = 0.5587$, $R^2A = 0.5458$, $F = 43.052$, $Q = 1.584$

Here & here onwards ‘n’ is total number of compounds, $R^2$ is variance square of correlation constant, $R^2A$ is adjusted $R^2$, $Se$ is standard error of estimation, $F$ is Fischer’s ratio & ‘Q’ is Pogliani’s quality factor which is a ratio $R/Se$.

When $I_1$ is added to $X_u$ parameter we obtained a bi-parametric correlation with slightly improved statistical values. The model shows slight improvement in $R^2$ value which changes from 0.5587 to 0.5670. The model is as under—

$$\text{Log } P = 0.5045 \pm 0.0800 \times X_u + 0.3195 \pm 0.4023 \times I_1 - 0.0827$$  \hspace{1cm} (7)

$n = 36$, $Se = 0.5812$, $R^2 = 0.5670$, $R^2A = 0.5408$, $F = 21.607$, $Q = 1.295$

The decrease in $R^2A$ from 0.5458 to 0.5408 clearly shows that the added parameter $I_1$ is not contributing in the model. In fact if the value of $R^2A$ increases only the added parameter shows its contribution. Similarly the error in the coefficient of $I_1$ is also more than the coefficient. Hence, this bi-parametric model consisting of $X_u$ & $I_1$ is not acceptable.

Out of five bi-parametric correlations one with $X_u$ and $I_2$ gave best $R^2$ value equal to 0.9428. The increase is $R^2$ from 0.5587 to 0.9428 is due to the addition of an indicator parameter $I_2$. The adjusted $R^2A$ value also shows dramatic improvement. The Q value shows a drastic increase which changes from 1.584 to 5.485. The obtained model is reported below:

$$\text{Log } P = -0.2596 \pm 0.1734 \times X_u + 0.0268 \pm 0.4518 \times I_2 + 0.7780$$  \hspace{1cm} (8)

$n = 36$, $Se = 0.1770$, $R^2 = 0.9428$, $R^2A = 0.9393$, $F = 271.718$, $Q = 5.485$

Unfortunately in this model the error in the coefficient of $I_2$ is more than the value of the coefficient which is statistically not acceptable.

To get a better model $I_1$, $I_2$ and $X_u$ are taken together which resulted into a three-parametric model with $R^2 = 0.9713$. The derived model is as follows:

$$\text{Log } P = 0.3974 \pm 0.0215 \times X_u - 0.6472 \pm 0.1145 \times I_1 - 0.8843 \pm 0.1357 \times I_2 + 1.3887$$  \hspace{1cm} (9)

$n = 36$, $Se = 0.1669$, $R^2 = 0.9713$, $R^2A = 0.9687$, $F = 361.577$, $Q = 5.905$

In this model the $R^2$ changes from 0.9393 to 0.9687 which shows that the added $I_1$ parameter has its fair share in the model.

Through another three-parametric model ($X_u$, $F$, $I_2$) obtained is statistically significant but, it has lower $R^2$ value equal to 0.9519 than the three parametric model discussed above. Hence the model with $X_u$, $F$, $I_2$ is also discarded.

Finally, a four-parametric model with $X_u$, $F$, $I_1$ and $I_2$ as correlating parameter is obtained with $R^2 = 0.9901$. The model is discussed below:

$$\text{Log } P = -0.7676 \pm 0.0703 \times X_u - 0.4153 \pm 0.0839 \times F - 0.4153 \pm 0.0131 \times I_1 - 0.1796 \pm 0.0233 \times I_2 + 1.8570$$  \hspace{1cm} (10)

$n = 36$, $Se = 0.1171$, $R^2 = 0.9901$, $R^2A = 0.9888$, $F = 772.544$, $Q = 8.497$

For four-parametric model with $X_u$, $F$, $I_1$ and $I_2$, the $R^2$ value comes out to be 0.9901 as compared to three-parametric model with $X_u$, $I_1$, $I_2$ (0.9713), which is certainly better. The change in $R^2A$ from 0.9687 to 0.9888 clearly indicates that the added parameter $F$ has significant role and its fair share in the model. The $Q$ value also changes from 5.905 to 8.497 and is in favour of the above model.

Here the four parametric models with $X_u$, $F$, $I_1$ & $I_2$ is the best for modeling for log P value of the present set of the compounds.

To confirm above finding log P values have been estimated using the best four parametric model such values are reported in Table – 5. These values are in good agreement with the observed activities.
Table 5 Observed and Estimated values of log p for using model no. 12

| Comp. No. | Observed log P | Estimated log P | Residual |
|-----------|----------------|----------------|----------|
| 1         | 2.810          | 2.815          | -0.005   |
| 2         | 3.339          | 3.371          | -0.032   |
| 3         | 3.868          | 3.916          | -0.048   |
| 4         | 4.397          | 4.448          | -0.051   |
| 5         | 4.926          | 4.967          | -0.041   |
| 6         | 5.455          | 5.475          | -0.020   |
| 7         | 5.984          | 5.972          | 0.012    |
| 8         | 6.513          | 6.457          | 0.056    |
| 9         | 3.209          | 3.190          | 0.019    |
| 10        | 3.738          | 3.753          | -0.015   |
| 11        | 4.267          | 4.306          | -0.039   |
| 12        | 4.536          | 4.468          | 0.068    |
| 13        | 2.795          | 2.772          | 0.023    |
| 14        | 3.354          | 3.374          | -0.020   |
| 15        | 3.913          | 3.854          | 0.059    |
| 16        | 4.472          | 4.395          | 0.077    |
| 17        | 3.982          | 4.024          | -0.042   |
| 18        | 1.465          | 1.383          | 0.082    |
| 19        | 1.994          | 2.109          | -0.115   |
| 20        | 2.523          | 2.642          | -0.119   |
| 21        | 3.052          | 3.175          | -0.123   |
| 22        | 3.581          | 3.700          | -0.119   |
| 23        | 4.110          | 4.215          | -0.105   |
| 24        | 2.875          | 2.284          | 0.591    |
| 25        | 1.458          | 1.859          | -0.401   |
| 26        | 2.481          | 2.384          | 0.097    |
| 27        | 2.644          | 2.787          | -0.143   |
| 28        | 3.627          | 3.196          | 0.431    |
| 29        | 1.738          | 1.865          | -0.127   |
| 30        | 2.435          | 2.384          | 0.051    |
| 31        | -0.284         | -0.341         | 0.057    |
| 32        | 0.245          | 0.204          | 0.041    |
| 33        | 0.774          | 0.824          | -0.050   |
| 34        | 1.303          | 1.325          | -0.022   |
| 35        | 1.832          | 1.832          | 0.000    |
| 36        | 0.554          | 0.580          | -0.026   |

A graph has been drawn using observed and estimated log P values which are demonstrated in Figure – 1. The predictive power of the model comes out to be 0.990, which shows that this model explains 99% variance of the data used in the present study.

![Fig. 1 Correlation between observed and estimated log P using model no. 12](image1)

The model was tested by evaluating cross validated parameters. These parameters for different models are reported in Table – 6. The lowest PRESS/SSY value 0.279 for the four parametric model and higher Rcv value which is 0.9874 confirms the finding for the four-parametric model with Xu, F, I1, and I2. The PSE value for this model is 0.173 which is lowest and SPRESS which comes out to be 0.1868 further verifies our result.

The four-parametric model is free from any defect. For this we have performed Ridge analysis. The Ridge trace is reported in Figure – 2. All the parameters are with the permissible limits.

![Fig. 2. Ridge Trace for Best four parametric Model no. 12](image2)
### Table 6. Cross Validated parameters for various models.

| Model no. | Parameter Used | PRESS/SSY | $R^2_{cv}$ | $S_{PRESS}$ | PSE  |
|-----------|----------------|-----------|------------|-------------|------|
| 1         | Xu             | 1.726     | 0.5148     | 1.217       | 1.176|
| 2         | F              | 2.203     | 0.000      | 1.752       | 1.706|
| 3         | Xu,I$_1$       | 1.371     | 0.5001     | 1.142       | 1.094|
| 4         | Xu,I$_2$       | 0.484     | 0.9342     | 0.414       | 0.394|
| 5         | F,I$_1$        | 2.227     | 0.000      | 1.770       | 1.694|
| 6         | F,I$_2$        | 1.564     | 0.3556     | 1.297       | 1.242|
| 7         | Xu,F           | 1.630     | 0.3797     | 1.273       | 1.218|
| 8         | F,I$_1$,I$_2$  | 1.223     | 0.6008     | 1.037       | 0.977|
| 9         | Xu,I$_1$,I$_2$ | 0.460     | 0.9655     | 0.307       | 0.289|
| 10        | Xu,F,I$_1$     | 1.594     | 0.3325     | 1.317       | 1.242|
| 11        | Xu,F,I$_1$,I$_2$ | 0.507 | 0.9285     | 0.439       | 0.413|
| 12        | Xu,F,I$_1$,I$_2$ | 0.279 | 0.9874     | 0.1868      | 0.173|

### Table 7 Ridge Regression Parameters for the obtained model

| Model no. | Parameter Used | VIF   | T     | i     | K   |
|-----------|----------------|-------|-------|-------|-----|
| 1         | Xu             | 0.9901| 1.0000| 1.0000| 1.00|
| 2         | F              | 0.9901| 1.0000| 1.0000| 1.00|
| 3         | Xu             | 1.1719| 0.8417| 1.3979| 1.00|
|           | I$_1$          | 1.1719| 0.8417| 0.6020| 2.32|
| 4         | Xu             | 0.9947| 0.9953| 1.0686| 1.00|
|           | I$_1$          | 0.9947| 0.9953| 0.9313| 1.15|
| 5         | F              | 1.0477| 0.9439| 1.2369| 1.00|
|           | I$_1$          | 1.0477| 0.9439| 0.7630| 1.62|
| 6         | F              | 1.0255| 0.9648| 1.876  | 1.00|
|           | I$_2$          | 1.0255| 0.9648| 0.8123| 1.46|
| 7         | Xu             | 1.0477| 0.9439| 1.2369| 1.00|
|           | F              | 1.0477| 0.9439| 0.7630| 1.62|
| 8         | F              | 1.1424| 0.8633| 1.3420| 1.00|
|           | I$_1$          | 1.2404| 0.9737| 1.1705| 1.15|
|           | I$_2$          | 1.2140| 0.8113| 0.4874| 2.75|
| 9         | Xu             | 1.2373| 0.7954| 1.4883| 1.00|
|           | I$_1$          | 1.3850| 0.7089| 1.0676| 1.39|
|           | I$_2$          | 1.1755| 0.8382| 0.4439| 3.35|
| 10        | Xu             | 1.2394| 0.7952| 1.6305| 1.00|
|           | F              | 1.1080| 0.8972| 0.7772| 2.10|
|           | I$_1$          | 1.1929| 0.8262| 0.5922| 2.75|
|           | I$_2$          | 1.2522| 0.7855| 0.6715| 2.44|
|           | I$_3$          | 1.4564| 0.6733| 0.4226| 3.87|
| 11        | Xu             | 1.0887| 0.9076| 1.3902| 1.00|
|           | F              | 1.1224| 0.8798| 0.9390| 1.48|
|           | I$_1$          | 1.0256| 0.9646| 0.6706| 2.07|
| 12        | Xu             | 1.2783| 0.7699| 1.6374| 1.00|
|           | F              | 1.1803| 0.8356| 1.2682| 1.29|
|           | I$_1$          | 1.2522| 0.7855| 0.6715| 2.44|
|           | I$_2$          | 1.4564| 0.6733| 0.4226| 3.87|
The VIF plots (Figure-3) also shows that all the parameters are below 10 hence, the used parameters in the four-parametric model is free from the defect of the colearinity or chance.

The calculated ridge parameters are reported in Table -7. The VIF value for Xu parameter in all the models reported are less than 10. Similar T is also with the Limit of one. The value of li and k are in favour of proposed model hence the four parametric models is acceptable on the basis of ridge analysis.

Conclusion

\[ \text{Log P} = -0.7676 \pm 0.0703 \times \text{Xu} - 0.4153 \pm 0.0839 \times \text{F} - 0.4153 \pm 0.0131 \times \text{I}_1 - 0.1796 \pm 0.0235 \times \text{I}_2 + 1.8570 \]

On the basis of our study following conclusion may be drawn:
1. Topological indices Xu and F are capable of modelling Log P value of alkanes along with indicator parameters.
2. The indicator parameters are important for prediction of biological activity.
3. Indicator parameter I₁ is very effective in modelling the activity of presence set of alkanes.
4. The four parametric models with Xu, F, I₁ & I₂ is the best model which is statistically significant for predicting and estimating lipophilicity (Log P) of presence set of compounds.
5. The best parametric models proposed in free from any defect.
6. Xu, F alone are not capable of modeling the lipophilicity, therefore use of indicator parameters are essential in the study.
7. The coefficients of all the four parameters are negative suggesting that their lesser value will favour the activity (log P).

Acknowledgment

Authors are thankful to late. Prof. P.V. Khadikar a world known topologist, who initiated this work. They also want to place on record the help rendered by Prof. Vijay K. Agrawal, Head, Department of Chemistry, APS University, Rewa (M.P).

References

1. Shaik, Basheerulla; Singh, Jyoti; Singh, Neelu; Sharma Mamta, Singh, Satyendra; Agrawal, Vijay K.; “Modeling lipophilicity of 2-Furylethylene Derivatives using physico-chemical parameters; International J. pure Appl. chem. 6, 93-99 (2011).
2. Agrawal, Vijay K.; Dubey Vinay K.; Shaik, Singh Jyoti; Singh Kamana: Khadikar P.V.; “Modeling of lipophilicity of some organic compounds using structural and topological indices” J. Indian chem. Soc. 86, 337-345 (2009).
3. Singh Jyoti, Dubey Vinay K.; Agrawal Vijay K.; Khadikar P.V.; “QSAR study on octanal-water partitioning : Dominating role of equalized electronegativity” oxid commun., 1, 27-43 (2008).
4. Agrawal, Vijay K.; Singh J.; Louis, B. and Khadikar P.V.; “The topology of molecule and its Lipophilicity” Current Computer–Aided Drug Design, (CCADD), 2, 369-403 (2006).
5. Todeschini, R.; Consonni, V.; Hand book of molecular descriptors, Wiley – VCH, Weinheim (GER) (2000).
6. Karelson, M., Molecular descriptors in QSAR / QSAR, J. Wiley & Sons, New York (2000).
7. Trinajstic, N.; Chemical Graph Theory 2nd revised ed. CRC Press Boca Raton Fl, Chapter 10, 225-273 (1992).
8. Diudea, M.V.; (Ed.) QSAR / QSPR studies by molecular descriptors, Babes-Bolyai University Cluj Romania, (2000).
9. Agrawal, V.K.; Singh, J.; Khadikar, P. V.; Bioorg. Med. Chem., 10, 3981 (2002).
10. Khadikar, P.V.; Karmakar S.; Gour, K.; Agrawal,
11. Agrawal, V.K.; Bano, S.; Khadikar P.V.; Bioorg. Med. Chem., 11, 4039-4047 (2003).
12. Hansch, C.; Leo, A.; Hockman, D.; Eds. Exploring QSAR: Hydrophobic, Electronic and Steric constants; ACS Professional Reference Book, Washington, D.C. P. 110 (1995).
13. Hansch, C.; The QSAR paradigm in the design of less toxic molecules, Drug Metab. Rev., 15, 1279-1294 (1984).
14. Chatarjee, S.; Hadi, A.S.; Price, B. Regression Analysis by examples, 3rd Ed., Wiley: New York, (2000).
15. Ren, B.; J. Chem. inf. Comput. Sci., 39, 139-143, 21 (1999).
16. Balaban, A.T.; Khadikar, P.V.; Supuran, C.T.; Thakur, A.; Thakur, M.; Bio org Med. Chem. Lett., 15, 3966-3973 (2005).
17. Chem. Sketch ACD/Labs Software; (11.01 version) www.acdlabs.com
18. DRAGON Software: Roberto Todeschini, Milano chemo metrics and QSAR research group, University of Milano-Bicocca, Pza della scienza, 1-20126 Milano, Italy. Website: http://www.disat.unimib.it/chm/
19. Pogliani, L.J.; Phys. Chem., 100, 18065 – 18077 (1996).