**Global Journal of Engineering Science and Researches**

**ESTIMATION OF LIPOPHILICITY OF SOME ALCOHOLS USING TOPOLOGICAL INDICES QSAR STUDIES**

**Asmita Sharma**

Department of Chemistry, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India

**ABSTRACT**

In This paper, I used different topological indices for modeling of lipophilicity of a series of alcohols. A wide variety of indices like the Weiner(W), The Padmakar Iwan (PI) index, Kier and Hall valence connectivity indices , Randic connectivity indices and Balaban and Balaban type indices were used for obtaining statistically significant model. The statistically significant models are governed by a variety of statistical parameters .The regression analysis has shown that out of pool of topological indices used, the topological indices W and PI in combination with connectivity indices given an excellent result. The results indicate that lipophilicity of given series of alcohols can be successfully modeled by using topological indices W and PI in combination with connectivity indices as correlating parameters.

The best model has excellent statistic as well as predictive power. The predictive power of these proposed models was discussed on the basis of cross-validation parameters.

***Keywords:*** *Topological indices, Lipophilicity , QSAR.*

1. **AIM & BACKGROUND**

In the last decades, several scientific researchers have been focused on studying how to catch and convert by a theoretical pathway the information encoded in the molecular structure into numbers called molecular descriptors. These are used to establish quantitative relationships between structures and properties, biological activities and other properties i.e. QSAR/QSPR. A graph theoretical approach to QSAR is based on the use of topological indices for encoding the structural information 1-5. Topological indices are numerical descriptors of molecular graph and are sensitive to size, shape, symmetry and heterogencity of atomic environments in the molecule. There is a recent upsurge of interest in the use of topological indices in QSAR studies. These are quite useful in the development of QSAR and capable of predicting the pharmacological as well as toxic properties of bioactive molecules 6. The use of these indices in risk assessment of chemicals and toxicology is described by Basak(1999) 7,8. Randic and co workers have shown that graph theoretical techniques could also be used to obtain the chemical shift of nuclei 9. Devenbeck(1995) has discussed topological approach to develop models for the prediction of 13C NMR chemical shift 10 Khadikar and coworkers(2002) have discussed the use of PI, W and Sz indices for the prediction of 13C NMR chemical shifts () in alkanes and cycloalkanes 11

In QSAR studies no other physiochemical property has attracted as much interest as lipophilicity 12,13 This is due to its direct relationship to stability in aqueous phases, to membrane permeation and its entropic contribution to binding.

In view of the above, we have undertaken the present investigation in which I have modeled lipophilicity (log p) of 32 alcohols using topological indices. Our aim is to construct mathematical models for predicting lipophilicity (log p) of alcohols by taking different combination of topological indices.

1. **MATERIAL AND METHODS**

Lipophilicity: 32 alcohols are used in the study. Their lipophilicity (log P) indices are taken from the previous work reported in Literature 14.

Topological indices : A set of topological indices as given below are used in the

investigation.

Weiner Index : W

Padmakar Ivan Index : PI

Randic Connectivity Indices : 0χ, 1χ, 2χ, 3χ

Balaban Indices : J, Jhet p, Jhet v, Jhet e, Jhet m, J het z

There indices are calculated using DRAGON Software 15. The structure optimation is made using ACD labs 16. The expressions used for the calculation of these indices are available in the literature. Regression Analysis: I have adopted maximum R2 method. The models giving significant R2 values were selected using NCSS software 17

1. **RESULT AND DISCUSSIONS**

The values of Lipophilicity and topological indices of 32 alcohols are shown in Table I. The results obtained by regression analysis of the data are discussed below.

**Modeling log P using W, PI and 0χ**

A stepwise regression analysis using the above parameters is done. Models having R = 0.49 or higher were selected by NCSS software, out of which a biparametric model consisting of 0χ and PI is statistically more significant. The biparametric model is given as

logP = -2.3426 (±) 0.5142) – 0.0217 (± 0.0121) PI +0.7990 (± 0.1516) 0χ

n = 32, Se = 0.3444, R = 0.9199, F = 79.8397, Q = 2.6710

Here and thereafter n is number of compounds used, Se is standard error of estimation R is multiple correlation coefficient, F is Fisher’s statistics and Q is Poglian’s quality factor.

**Modeling log P using W, PI, 0χ, 1χ, 2χ**

Five models selected by NCSS software are shown in Table II. The triparemetric model using W, PI and 1χ has the values of R2 and R2A as 0.8859 and 0.8737 respectively.

In pentaparametric model also there is a decline in the value of R2A considering this the triparametric model is supposed to be the best:

log P = -1.8178 - 0.0442 (± 0.0080) w + 0.0695 (±0.0146) PI + 1.1198

(± 0.2072) 1χ

n = 32, Se = 0.3036, R = 0.9406, F = 71.5831, Q = 3.0981

**Modeling log P using W, PI and Balaban indices**

The value of R2A goes on increasing up to the IV th model and then declines, considering all this the tetraparametric model consisting of W, PI, J and Jhetm is found to be good model.

log P = 1.5339 (±0.6057) – 0.0419 (±0.0154)W + 0.1039 (±0.0319) PI

+ 5.6050 (±0.2307)J – 5.7634 (±1.3354) JhetM

n = 32, Se = 0.3265, R = 0.9335, F = 45.7350, Q = 2.8591

1. **VALIDATION**

In statistics and chemometrics several validation techniques have been proposed in the last few decades in order to estimate the model prediction capabilities. A model with good statistics does not necessarily mean that it will have good predictive power too. Both the qualities good statistics and good predictive power are necessary for a perfect model. The predictive power of the model can be obtained by calculating Pogliani’s quality factor Q.The higher the value of R and lower the value of Se, the better will be the predictive power of the model. The values of Q for all the models are shown in table IX. By considering the values of Q, the models can be ranked (from the best to the worst) with the following order,2,3,1, The same ranking can be obtained from the values of R.

Another parameter used for validation purposes is PRESS i.e. Predictive Error of Sum of Squares. It is the sum of the squared difference between the experimental response and the response predicted by the regression model. It is one of the most important cross validated parameters.

PRESS should be smaller than SSY (Sum of squares of deviations of each activity). The ratio smaller than 0.4, indicates statistically significant model. In the present case the model numbers 2 have values around 0.13 indicating their excellent predictive power. Higher the value of R2cv higher the predictive power of the model. Once again R2cv is in favor of models 2.

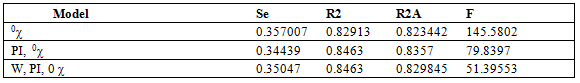
1. **CONCLUSION**

From the aforementioned results and discussion, I conclude that lipophilicity (logP) of alcohols can be successfully modeled by using topological indices W and PI in combination with 0χ, 1χ and 2χas correlating parameters. This Tri parametric model has excellent statistics as well as predictive power.

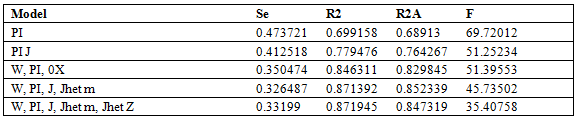
***Table 1: The values of Lipophilicity and topological indices of alcohols***

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Compound | log P | 13C NMR  shift | W | PI | 0X | 1X | 2X | 0XV | 1XV | 2XV | 3X | 3XV | J | JhetZ | JhetM | Jhetv | Jhete | Jhetp |
| methanol | -0.764 | 49 | 1 | 0 | 2 | 1 | 0 | 1.4472 | 0.4472 | 0 | 0 | 0 | 1 | 1.333 | 1.332 | 0.512 | 1.327 | 0.455 |
| ethanol | -0.235 | 57 | 4 | 2 | 2.7071 | 1.4142 | 0.7071 | 2.1543 | 1.0233 | 0.3162 | 0 | 0 | 1 | 1.333 | 1.332 | 0.512 | 1.327 | 0.455 |
| propanol | 0.294 | 63.6 | 10 | 6 | 3.4142 | 1.9142 | 1 | 2.8614 | 1.5233 | 0.7236 | 0.5 | 0.224 | 1.975 | 2.122 | 2.122 | 1.57 | 2.12 | 1.492 |
| butanol | 0.823 | 61.4 | 20 | 12 | 4.1213 | 2.4142 | 1.3536 | 3.5685 | 2.0233 | 1.0772 | 0.707 | 0.512 | 2.191 | 2.29 | 2.29 | 1.886 | 2.289 | 1.822 |
| pentanol | 1.352 | 61.8 | 35 | 20 | 4.8284 | 2.9142 | 1.7071 | 4.2756 | 2.5233 | 1.4307 | 0.957 | 0.762 | 2.339 | 2.411 | 2.411 | 2.106 | 2.41 | 2.055 |
| hexanol | 1.881 | 61.9 | 56 | 30 | 5.5355 | 3.4142 | 2.0607 | 4.9827 | 3.0233 | 1.7843 | 1.207 | 1.012 | 2.447 | 2.501 | 2.501 | 2.266 | 2.501 | 2.224 |
| isopropanol | 0.154 | 63.4 | 9 | 6 | 3.5774 | 1.7321 | 1.7321 | 3.0246 | 1.4129 | 1.0937 | 0 | 0 | 2.324 | 2.538 | 2.537 | 1.775 | 2.534 | 1.675 |
| 2-butanol | 0.603 | 68.7 | 18 | 12 | 4.2845 | 2.2701 | 1.8021 | 3.7317 | 1.9509 | 1.2573 | 0.816 | 0.591 | 2.54 | 2.682 | 2.682 | 2.127 | 2.68 | 2.044 |
| 2-pantanol | 1.132 | 67 | 32 | 20 | 4.9916 | 2.7701 | 2.1825 | 4.4388 | 2.4509 | 1.6377 | 0.866 | 0.706 | 2.627 | 2.724 | 2.724 | 2.326 | 2.723 | 2.261 |
| 2-haxanol | 1.661 | 67.2 | 52 | 30 | 5.6987 | 3.2701 | 2.5361 | 5.1459 | 2.9509 | 1.9912 | 1.135 | 0.975 | 2.678 | 2.747 | 2.747 | 2.453 | 2.746 | 2.402 |
| 3-pentanol | 1.132 | 73.8 | 31 | 20 | 4.9916 | 2.8081 | 1.9217 | 4.4388 | 2.4889 | 1.4703 | 1.394 | 0.942 | 2.754 | 2.864 | 2.864 | 2.419 | 2.863 | 2.348 |
| 3-haxanol | 1.661 | 72.3 | 50 | 30 | 5.6987 | 3.3081 | 2.3021 | 5.1459 | 2.9889 | 1.8507 | 1.478 | 1.093 | 2.832 | 2.913 | 2.913 | 2.573 | 2.912 | 2.516 |
| 3-heptanol | 2.19 | 72.6 | 76 | 42 | 6.4058 | 3.8081 | 2.6556 | 5.853 | 3.4889 | 2.2043 | 1.747 | 1.362 | 2.862 | 2.923 | 2.923 | 2.662 | 2.922 | 2.616 |
| 4-heptanol | 2.19 | 70.6 | 75 | 42 | 6.4058 | 3.8081 | 2.6825 | 5.853 | 3.4889 | 2.2312 | 1.563 | 1.244 | 2.92 | 2.985 | 2.984 | 2.708 | 2.984 | 2.66 |
| 4-octanol | 2.68 | 70.9 | 108 | 56 | 7.1129 | 4.3081 | 3.0361 | 6.5601 | 3.9889 | 2.5847 | 1.832 | 1.513 | 2.955 | 3.006 | 3.006 | 2.784 | 3.005 | 2.745 |
| 5-nananol | 1.572 | 71.1 | 149 | 72 | 7.82 | 4.8081 | 3.3896 | 7.2672 | 4.4889 | 2.9383 | 2.101 | 1.782 | 2.998 | 3.041 | 3.041 | 2.855 | 3.04 | 2.822 |
| isobutanol | 0.805 | 68.9 | 18 | 12 | 4.2845 | 2.2701 | 1.8021 | 3.7317 | 1.8792 | 1.5764 | 0.816 | 0.365 | 2.54 | 2.674 | 2.673 | 5.141 | 2.672 | 2.059 |
| t-butanol | 0.532 | 68.4 | 16 | 12 | 4.5 | 2 | 3 | 3.9472 | 1.7236 | 2.1708 | 0 | 0 | 3.024 | 3.228 | 3.228 | 2.458 | 3.225 | 2.348 |
| neopentanol | 1.664 | 72.6 | 28 | 20 | 5.2071 | 2.5601 | 2.9142 | 4.6543 | 2.1698 | 2.7188 | 1.061 | 0.474 | 3.168 | 3.3 | 3.3 | 2.76 | 3.298 | 2.673 |
| 2-me-pentanol | 0.693 | 66.9 | 50 | 30 | 5.6987 | 3.3081 | 2.3021 | 5.1459 | 2.9172 | 2.0764 | 1.478 | 1.093 | 2.832 | 2.905 | 2.905 | 2.589 | 2.904 | 2.535 |
| 3-me-butanol | 1.28 | 60.2 | 32 | 20 | 4.9916 | 2.7701 | 2.1825 | 4.4388 | 2.3792 | 1.9061 | 0.866 | 0.706 | 2.627 | 2.717 | 2.716 | 2.341 | 2.715 | 2.278 |
| 3-me-2-butanol | 1.28 | 72 | 29 | 20 | 5.1547 | 2.6427 | 2.488 | 4.6019 | 2.3236 | 1.9846 | 1.333 | 0.965 | 2.993 | 3.118 | 3.118 | 2.613 | 3.116 | 2.533 |
| 4-me2-butanol | 1.687 | 65.2 | 32 | 20 | 4.9916 | 2.7701 | 2.1825 | 4.4388 | 2.4509 | 1.6377 | 0.866 | 0.706 | 2.627 | 2.724 | 2.724 | 2.326 | 2.723 | 2.261 |
| 4-me-3-pantanol | 1.687 | 77.3 | 46 | 30 | 5.8618 | 3.1807 | 2.6295 | 5.309 | 2.8616 | 2.2196 | 1.782 | 1.188 | 3.144 | 3.243 | 3.243 | 2.832 | 2.242 | 764 |
| 3,3di me-butanol | 1.808 | 58.9 | 46 | 30 | 5.9142 | 3.0607 | 3.3107 | 5.3614 | 2.6698 | 3.0343 | 1 | 0.862 | 3.154 | 3.242 | 3.242 | 2.865 | 3.241 | 2.8 |
| 2,3di me-2-butanol | 1.529 | 72.2 | 42 | 30 | 6.0774 | 2.9434 | 3.5207 | 5.5246 | 2.667 | 2.8084 | 1.732 | 1.413 | 3.541 | 3.665 | 3.665 | 3.156 | 3.664 | 3.073 |
| 3,3 di me-2-butanol | 1.48 | 74.8 | 42 | 30 | 6.0774 | 2.9434 | 3.5207 | 5.5246 | 2.6242 | 3.042 | 1.732 | 1.253 | 3.541 | 3.66 | 3.66 | 3.166 | 3.658 | 3.084 |
| 4,4 di-me-3-butanol | 2.154 | 80.9 | 46 | 30 | 5.8618 | 3.1807 | 2.6295 | 5.309 | 2.8616 | 2.2196 | 1.782 | 1.188 | 3.144 | 3.243 | 3.243 | 2.832 | 3.242 | 2.764 |
| 2,4,di me 3-pantanol | 2.148 | 80.4 | 65 | 42 | 6.7321 | 3.5534 | 3.3472 | 6.1793 | 3.2343 | 2.9786 | 2.103 | 1.366 | 3.464 | 3.553 | 3.553 | 3.178 | 3.552 | 3.115 |
| 2,3,3tr-me-2-butanol | 1.996 | 74.1 | 58 | 42 | 7 | 3.25 | 4.5 | 6.4472 | 2.9736 | 3.809 | 2.25 | 1.853 | 4.02 | 4.136 | 4.136 | 3.652 | 4.135 | 3.57 |
| 2,4,4 t-me-3pentanol | 2.615 | 82.8 | 86 | 56 | 7.6547 | 3.8541 | 4.3987 | 7.1019 | 3.5349 | 4.0549 | 2.366 | 1.519 | 3.878 | 3.963 | 3.963 | 3.6 | 3.962 | 3.537 |
| 2,2,4,4tetrame-3pentanol | 3.082 | 84.7 | 111 | 72 | 8.5774 | 4.1547 | 5.4537 | 8.0246 | 3.8355 | 5.1346 | 2.598 | 1.641 | 4.231 | 4.312 | 4.312 | 3.966 | 4.311 | 3.906 |

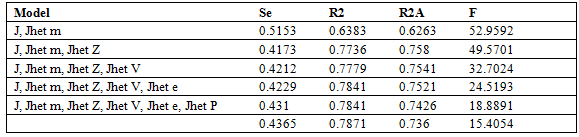
***Table 2 : Model using W, PI , & 0X (Model-1)***

**

***Table 3 : Model using W,PI,0X,1X and 2X (Model-2)***



***Table IV : Model using Balaban Indices (Model-3)***

******

***Table 5 : Cross validation parameters of the Three models selected for estimating log P of alcohols***

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model No.** | **R** | **Q** | **PRESS** | **SSY** | **Press/SSY** | **R2cv** | **PE** |
| 1 | 0.9199 | 2.670 | 3.4397 | 20.8959 | 0.1646 | 0.8353 | 0.0179 |
| 2 | 0.9406 | 3.0981 | 2.5532 | 91.7772 | 0.1291 | 0.8709 | 0.0133 |
| 3 | 0.9335 | 2.8591 | 2.8782 | 19.448 | 0.1480 | 0.8520 | 0.0150 |

**REFERENCES**

1. *Trinajstic, N.;(****1983****) Chemical Graph Theory, CRC Boca Raton, FL, ; vol. II chapter* ***4***
2. *Sabljic, A.; rinajstic, N.; (****1981)*** *Acta Pharm. Jugosl., 31, 189*
3. *Balaban, A.T.; Motoc. I. Bonchev, D.; Mekenyan, O.;(* ***1983*** *)Topics Curr. Chem., 114, 21*
4. *Hansen, P.J.; (****1988 )****Jurs, P.C. J.Chem. Educ., 65, 575*
5. *Randic, M.;(* ***1990)*** *J. Math. Chem., 4, 157*
6. *Khadikar P.V.; Phadnis, A.; Shrivastava, A.;(****2002)*** *''QSAR study on Toricity to Aqueous organisms using the PI Index'', Biorganic and Medicinal chemistry,101181 – 1188*
7. *Gate, B.D.; Basak, S.C.; (****1997)*** *SAR and QSAR in Environ. Res., 7, 117*
8. *Basak, S.C.; Bertelsen, S.; Grunwald, G.D. ;(****1995)*** *Toxicology Letters, 79, 239.*
9. *Randic, M. ;(****1983 )****Int. J. Quantum Chem., 23, 1707*
10. *Duvenbeck, C.;(****1995)''****Topological and Geometrical Approach to Develop Models for Prediction of 13C NMR shifts'', Bochum: GER*
11. *Khadikar, P.V.; Bajaj, A.V.; Mandloi, D.;(****2002****,) ''Prediction of 13C nuclear magnetic resonance chemical shifts (Σ13 cn) in alkanes and cycloalkanes.'' Indian J. Chem. 41A, 2065-2067*
12. *K.A. Dill,;(****1990 ),****Science, 250-297*
13. *P.J. Taylor,;(* ***1990*** *Quantitative drug design, Vol. 4 of comprehensive Medicinal chemistry, Pergamon Press, Oxford, PP 241-294*
14. *Khadikar, P.V.; Sharma, V.; Verma, R.G.; (****2005),****''Novel estimation of lipophilicity using 13C NMR chemical shifts as molecular descriptor''. Bioorganic and medicinal chemistry letters 15, 421-425*
15. *DRAGON: http:// disat. Unimib. It./chem./dragon.com*
16. *ACD labs: http://www.acdlabs.com*
17. *NCSS: NCSS,* [*http://www.ncss.com*](http://www.ncss.com/)
18. *Selassie, C.; Verma, R.P., (2010)History of quantitative structure–activity relationships, Burger's Medicinal Chemistry, Drug Discovery and Development.*
19. *Martin Michalik,;(2016)'' The validation of quantum Chemical lipophilicity prediction of alcohols '' Acta Chimica solvaca 9(2),89-94.*
20. *Saadi saaidpour ,; (2014)'' Prediction of Drug lipophilicity using back propogation artificial neural network modeling'' (30)2.*
21. *Bernard Tesla,;Pieme Alain corncept,;Patric Gaillard,;'' Lipophilicity of Molecular Modeling '' (2017),Pharmaceutical Research 13(3), 335-343.*