

# PREDICTION OF PHYSICAL AND THERMODYNAMIC PROPERTIES OF ALIPHATIC ETHERS FROM MOLECULAR STRUCTURES BY MULTIPLE LINEAR REGRESSION

F. SHAFIEI

Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran

## ABSTRACT

The interrelation of topological indices with enthalpy of vaporization at standard conditions ( $\Delta H^\circ_{\text{vap}}$ ), and normal temperature of boiling point ( $T_{\text{boil}}/K$ ) for series of aliphatic ethers has been investigated. For obtaining model for predicted target properties we have used multiple linear regression (MLR) techniques and followed Back ward regression analysis. The results have shown that combining the four descriptors (J, Wp,  $^1X$ , H) are included, with values of the correlation coefficient ( $r=0.979$ ), the standard error ( $s=1.343 K$ ), the Fisher  $-ratio(F=304.992)$  could be used successfully for modeling and predicting the normal temperature of boiling points of 30 aliphatic ethers. The best model for estimating the enthalpy of vaporization of compounds are included two descriptors (WW,  $^1X$ ), with values of the correlation coefficient ( $r=0.994$ ), the standard error ( $s=5.900 kJ/mol$ ) and the Fisher  $-ratio(F=558.789)$ .

**Keywords:** Quantitative structure–property relationships (QSPR); topological indices; aliphatic ethers; graph theory; multiple linear regressions (MLR).

## 1. INTRODUCTION

Symmetrical or simple aliphatic ethers are an important class of organic compounds characterized by an oxygen atom bonded to two alkyl or aryl groups. Ethers are important in medicine and pharmacology, especially for use as anesthetics and they are useful as insecticides, gasoline additive, solvents and plasticizers for fats, oils, waxes, perfumes, resins, dyes, gums, and hydrocarbons.

As a rule, the molecular properties are characterized by physical values, such as, enthalpy, entropy, heat capacity, heat of formation, etc. All physicochemical values are characterized by numerical values, that is have a quantitative characteristic. As regards the quantitative measure of the structure of organic molecules, there has been yet no conventional view of it [1]. But nevertheless, the numerical method, allowing to define the structure of organic molecules does exist. This is a so-called method of topological indices (TI) [2-5]. This method allows to describe a wide spectrum of molecular properties, establishing the correlation dependences between the properties and values of topological indices [6-9].

In this study, the relationship analysis between the enthalpy of vaporization at standard condition ( $\Delta H^\circ_{\text{vap}}/kJ.mol^{-1}$ ), normal temperature of boiling point ( $T_{\text{boil}}/K$ ) and topological indices was done by multiple linear regressions (MLR) analysis.

The normal boiling point of a substance is defined as the temperature at which a liquid boils at 1 atmosphere of pressure (760 mm Hg). It is a function of a number of molecular properties that control the ability of a molecule to escape from the surface of a liquid into the vapor phase.

Enthalpy of vaporization or heat of vaporization is the change in enthalpy (energy) when a liquid is converted to a gas at constant pressure.

Comparison of the molecular topological indices with molecular surface areas and aqueous solubilities for the aliphatic alcohols, ethers, ketones, aldehydes, carboxylic acids, esters, and hydrocarbons has been proposed [10].

The relationship study of Wiener's topological index and Bonchev's informational index and quantum-chemical parameters with thermodynamic properties and boiling point of a number of aliphatic and aromatic compounds has been investigation [11].

The interaction of a topological distance feature location index (T) and quantum-chemical descriptors with the physicochemical properties of polybrominated diphenyl ethers was proposed [12].

The relations of activity of aliphatic hydrocarbons and ethers with molecular connectivity index ( $^1X$ ) and the various topological indices have been developed [13].

The critical pressures, critical temperatures of aliphatic symmetrical ethers have been measured [14].

A model for predicting boiling points of aliphatic hydrocarbons by using valance molecular connectivity indices, Kappa shape index, dipole moment, and molecular weight has been reported [15].

The additional descriptors for the prediction of boiling points of 298 organic compounds containing O, N, Cl and Br was investigated [16].

The main aim of this study is to illustrate the usefulness of topological indices by MLR study of the enthalpy of vaporization at standard condition ( $\Delta H^\circ_{\text{vap}}/kJ.mol^{-1}$ ) and temperature of boiling points ( $T_{\text{boil}}/K$ ) of 30 aliphatic ethers. As far as we are aware, this is the first quantitative structure-property relationship (QSPR) study for prediction of normal boiling points and enthalpy of vaporization of aliphatic ethers derivatives using topological indices.

## 2. Materials, Mathematical Method and Graphs

The enthalpy of vaporization at standard condition ( $\Delta H^\circ_{\text{vap}}/kJ.mol^{-1}$ ) and normal temperature of boiling points ( $T_{\text{boil}}/K$ ) of 30 ethers are taken from national institute of standards and technology (NIST) chemistry web book [17].

All values are given in SI units and are for a standard-state pressure of 100 000 Pa (1 bar).

Studied ethers and their enthalpy of vaporization and temperature of boiling points are listed in Table 1.

## 3. Topological indices

All the used topological indices, the Randić' ( $^1X$ ) [18], Balaban (J) [19], Harary (H) [20], Wiener (W) [21], hyperWiener (WW) [21,4] and Wiener polarity (Wp) [22,23] were calculated using all hydrogen suppressed graph by deleting all the carbon hydrogen. The descriptors were calculated with chemicalize program [24]. Six topological indices tested in the present study are recorded in Table 1.

## 4. Statistical analysis

Structure- Property models (MLR models) are generated using the multiple linear regression procedure of SPSS version 16. The enthalpy of vaporization at standard conditions ( $\Delta H^\circ_{\text{vap}}/kJ.mol^{-1}$ ) and normal temperature of boiling point ( $T_{\text{boil}}/K$ ) are used as the dependent variable and  $^1X$ , J, Sz, H, Wp and WW indices as the independent variables. The models are assessed with correlation coefficient (r), squared multiple correlation coefficient ( $r^2$ ), adjusted correlation coefficient ( $r^2_{\text{adj}}$ ), standard error of estimate (s), Durbin-Watson value (D), significance (Sig) and Fisher ratio (F).

Distribution of the dependent variable against the independent variable for 30 chemicals employed in developing quantitative structure- Properties relationship. For obtaining appropriate structure- Property model we have used maximum  $r^2$  method and followed Back ward regression analysis. The predictive ability of the model is discussed on the basis of predictive correlation coefficient.

## 4.1. Structure- Property models for the enthalpy of vaporization ( $\Delta H^\circ_{\text{vap}}/kJ.mol^{-1}$ )

Initial regression analysis indicated that combination of six topological indices plays a dominating role in modeling the enthalpy of vaporization. In Table 2 are given the regression parameters and quality of correlation of the proposed models for the enthalpy of vaporization of 30 ethers.

**Table 1.** Aliphatic ethers and their boiling points, enthalpies of vaporization at standard conditions and topological indices used in present study.

compounds	No.	T <sub>boil</sub> /K	ΔH° vap	<sup>1</sup> X	J	H	W	Wp	WW
Dimethyl ether	1	249	19.30	1.41	1.63	2.50	4	0.00	5
Methyl ethyl ether	2	283	24.50	1.91	1.97	4.33	10	1.00	15
Methyl propyl ether	3	311	27.70	2.41	2.19	6.42	20	2.00	35
Methyl isopropyl ether	4	323.75	26.78	2.27	2.54	6.67	18	3.00	28
Methyl butyl ether	5	344.2	32.50	2.91	2.34	8.70	35	3.00	70
Methyl isobutyl ether	6	331.7	30.31	2.77	2.63	9.00	32	4.00	58
Methyl sec-butyl ether	7	332.2	30.41	2.81	2.75	9.08	31	3.00	54
Methyl tert-butyl ether	8	328.2	30.00	2.56	3.17	9.50	28	6.00	9.5
Di butyl ether	9	415	45.00	4.41	2.60	16.46	120	2.00	330
Diethyl ether	10	307.2	27.10	2.41	2.19	6.42	20	3.00	35
Ethyl propyl ether	11	336.8	31.40	2.91	2.34	8.70	35	3.00	70
Ethyl isopropyl ether	12	326	30.32	2.77	2.63	9.00	32	4.00	58
Ethyl butyl ether	13	364.7	36.41	3.41	2.45	11.15	56	4.00	126
Ethyl sec-butyl ether	14	348	34.20	3.31	2.83	11.62	50	5.00	99
Ethyl tert-butyl ether	15	346.1	33.10	3.06	3.15	12.08	46	4.00	84
Ethyl iso-butyl ether	16	354.2	34.14	3.27	2.68	11.48	52	5.00	108
Ethyl pentyl ether	17	392.7	41.06	3.91	2.53	13.74	84	7.00	210
Ethyl tert-pentyl ether	18	375.2	39.20	3.62	3.37	15.03	67	4.00	131
Dipropyl ether	19	363	35.72	3.41	2.45	11.15	56	4.00	126
Propyl isopropyl ether	20	356.2	33.95	3.27	2.68	11.48	52	4.00	108
Di-isopropyl ether	21	341.4	32.26	3.13	2.95	11.83	48	2.00	91
Metyl sec-butyl ether	22	332.2	30.41	2.81	2.75	9.08	31	4.00	54
Di iso-butyl ether	23	395.7	41.20	4.13	2.91	17.22	108	6.00	265
Di sec-butyl ether	24	394	40.88	4.2	3.22	17.63	100	8.00	224
Metyl pentyl ether	25	372.7	36.91	3.41	2.45	11.15	56	4.00	126
Methyl hexyl ether	26	399.2	42.10	3.91	2.53	13.74	84	5.00	210
Methyl heptyl ether	27	423.65	40.10	4.41	2.60	16.46	120	6.00	330
Methyl octyl ether	28	445.2	42.30	4.91	2.65	19.29	165	7.00	495
Methyl nonyl ether	29	468.15	41.40	5.41	2.69	22.22	220	8.00	715
Methyl decyl ether	30	488.35	43.40	5.91	2.73	25.24	286	9.00	1001

**Table 2.** Regression parameters and quality of correlation of the proposed models for the enthalpy of vaporization.

Model	independent variables	r	r <sup>2</sup>	r <sup>2</sup> adj	F	s
1	Wp, J, WW, <sup>1</sup> X, H, W	0.981	0.962	0.952	97.477	1.374
2	Wp, J, WW, <sup>1</sup> X, W	0.981	0.962	0.954	121.968	1.346
3	J, WW, <sup>1</sup> X, W	0.981	0.962	0.956	158.708	1.319
4	J, WW, <sup>1</sup> X	0.980	0.960	0.956	210.696	1.321
5	WW, <sup>1</sup> X	0.979	0.958	0.954	304.992	1.343

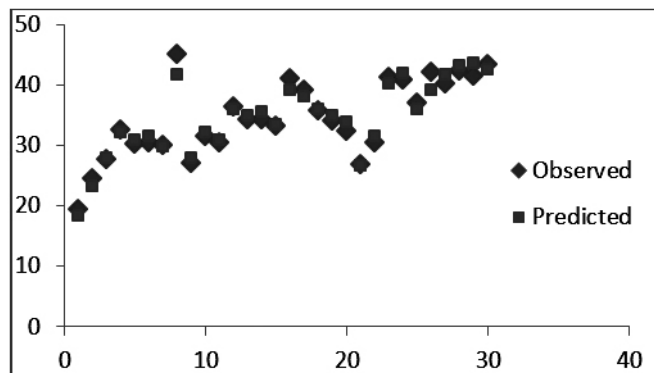
It turns out that the enthalpy of vaporization has a highly correlation with all descriptors as well as with a combination of the two parameters, namely, the Randić' (<sup>1</sup>X) and hyperWiener (WW) indices. Due to the regression parameters can be seen that the model (4.1.5) is more suitable than other models. This model has a minimum of independent variables and maximum of Fisher- ratio (F), compared to the other models.

Model 4.1.5

$$\Delta H^\circ \text{ vap} = 4.197 + 10.101^1X - 0.021WW \quad (1)$$

N= 30, r= 0.979, r<sup>2</sup> = 0.958, r<sup>2</sup> adj= 0.954, s= 1.343 kJ/mol, F= 304.992, Sig= 0.000 D= 1.926

Figure 1 shows the linear correlation between the observed and the predicted enthalpy of vaporization values obtained using eq. (1).



**Figure 1.** Comparison between the predicted and observed values of enthalpy of vaporization by MLR (eq. 1).

#### 4.2. Structure- Property models for the normal temperature of boiling point ( $T_{\text{boil}}/K$ )

In Table 3 are given the regression parameters and quality of correlation of the proposed models for the normal temperature of boiling point ( $T_{\text{boil}}/K$ ).

**Table3:** Regression parameters and quality of correlation of the proposed models for the normal temperature of boiling point ( $T_{\text{boil}}/K$ ).

Model	independent variables	r	r <sup>2</sup>	r <sup>2</sup> <sub>adj</sub>	F	s
1	Wp, J, WW, 'X, H, W	0.995	0.991	0.988	400.319	5.696
2	Wp, J, 'X, H, W	0.995	0.989	0.987	445.271	5.913
3	Wp, J, 'X, H	0.994	0.989	0.987	558.789	5.900

Regression parameters of models (4.2.1-4.2.3) such as correlation coefficient (r), squared multiple correlation coefficient (r<sup>2</sup>), adjusted correlation coefficient (r<sup>2</sup><sub>adj</sub>) and standard error of estimate (s) are close together, but model (4.2.3) with having the lowest number of descriptors and highest of Fisher-ratio (F) is better than other models.

The tetra parametric model containing four descriptors (Wp, J, 'X, H) is found as below:

Model 4.2.3

$$T_{\text{boil}} = 119.176 + 79.231'X + 16.814 J - 3.666 H - 5.692 Wp \quad (2)$$

$$N = 30, r = 0.994, r^2 = 0.989, r^2_{\text{adj}} = 0.987, s = 5.900, F = 558.789$$

$$\text{Sig} = 0.000, D = 1.906$$

Figure 2 shows the linear correlation between the observed and the predicted normal temperature of boiling point values obtained using eq. (2).

## 5. RESULT AND DISCUSSION

We studied the relationship between topological indices to the enthalpy of vaporization ( $\Delta H^\circ \text{vap}$ ) and normal temperature of boiling points ( $T_{\text{boil}}$ ) of 30 aliphatic ethers.

In this study, to find the best model for predict the properties mentioned, we will use the following points:

### 5.1. Verification and Validity of models

In this section for verification and validity of the regression models, we will focus on the Durbin-Watson statistic and unstandardized predicted and residual values.

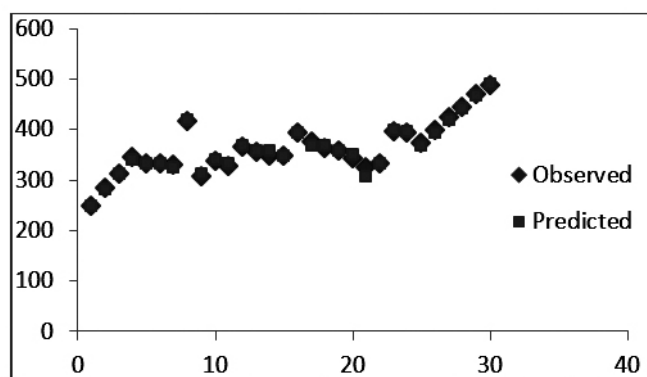
### 5.2. Test for autocorrelation by using the Durbin-Watson statistic

The Durbin-Watson statistic ranges in value from 0 to 4. A value near 2 indicates non-autocorrelation; a value toward 0 indicates positive autocorrelation; a value toward 4 indicates negative autocorrelation. Therefore the value of Durbin-Watson statistic is close to 2 if the errors are uncorrelated. In our all models, the value of Durbin-Watson statistic is less than 2 (See eq. 1, 2) and hence the errors are uncorrelated.

### 5.3. Regular residuals

The residual that is the difference between the observed value of the dependent variable (y) and the predicted value ( $\hat{y}$ ). Comparison between predicted and observed values of enthalpy of vaporization ( $\Delta H^\circ \text{vap}$ ) and normal temperature of boiling point ( $T_{\text{boil}}$ ) show in Table 4. Plot the residuals, and use other diagnostic statistics, to determine whether our model is adequate and the assumptions of regression are met. The residuals can also identify how much a model explains the variation in the observed data.

The residuals values of enthalpy of vaporization ( $\Delta H^\circ \text{vap}$ ) and normal temperature of boiling point ( $T_{\text{boil}}$ ) expressed by eqs. (1, 2) shows a fairly random pattern. This random pattern indicates that a linear model provides a decent fit to the data.



**Figure 2.** Comparison between the predicted and observed values of normal temperature of boiling point by MLR (eq.2).

## 6. CONCLUSION

The character of interrelation of Randic', Balaban, Harary, Wiener, Wiener polarity and hyper Wiener topological indices with the enthalpies of vaporization at standard conditions ( $\Delta H^\circ \text{vap}/\text{kJ}\cdot\text{mol}^{-1}$ ) and temperature of boiling points ( $T_{\text{boil}}/K$ ) of 30 aliphatic ethers using MLR method has been investigated.

Analysis of the obtained results indicates that the MLR model can well represent the structure-property relationships of aliphatic ethers, and that, we can use only two topological indices, namely, the Randic' ('X) and hyper Wiener (WW) for predicting enthalpy of vaporization at standard conditions ( $\Delta H^\circ \text{vap}/\text{kJ}\cdot\text{mol}^{-1}$ ) and four topological indices, namely, the Randic' ('X), Harary (H), Balaban (J) and Wiener polarity (Wp) for predicting temperature of boiling points ( $T_{\text{boil}}/K$ ) of studied compounds.

**Table 4.** Comparison between predicted and observed values of temperature of boiling point and enthalpy of vaporization at standard conditions of respect aliphatic ethers.

Comp. No.	Observed (T <sub>boil</sub> /K)	Predicted (T <sub>boil</sub> /K)	Residual	Observed ( $\Delta H^{\circ}\text{vap}/\text{kJ}\cdot\text{mol}^{-1}$ )	Predicted ( $\Delta H^{\circ}\text{vap}/\text{kJ}\cdot\text{mol}^{-1}$ )	Residual
1	249.00	249.13	-0.13	19.30	18.33	0.97
2	283.00	282.07	0.93	24.50	23.17	1.33
3	311.00	312.03	-1.03	27.70	27.79	-0.09
4	344.20	340.11	4.09	32.50	32.10	0.40
5	331.70	332.80	-1.10	30.31	30.94	-0.63
6	332.20	332.00	0.20	30.41	31.43	-1.02
7	328.20	323.41	4.79	30.00	29.85	0.15
8	415.00	417.81	-2.81	45.00	41.70	3.30
9	307.20	312.03	-4.83	27.10	27.79	-0.69
10	336.80	340.11	-3.31	31.40	32.10	-0.70
11	326.00	332.80	-6.80	30.32	30.94	-0.62
12	364.70	366.91	-2.21	36.41	35.95	0.46
13	354.20	358.47	-4.27	34.14	34.92	-0.78
14	348.00	357.96	-9.96	34.20	35.52	-1.32
15	346.10	347.54	-1.44	33.10	33.31	-0.21
16	392.70	392.68	0.02	41.06	39.21	1.85
17	375.20	367.71	7.49	39.20	37.97	1.23
18	363.00	366.91	-3.91	35.72	35.95	-0.23
19	356.20	358.47	-2.27	33.95	34.92	-0.97
20	341.40	350.64	-9.24	32.26	33.87	-1.61
21	323.75	305.90	17.85	26.78	26.53	0.25
22	332.20	332.00	0.20	30.41	31.43	-1.02
23	395.70	398.05	-2.35	41.2	40.25	0.95
24	394.00	395.92	-1.92	40.88	41.84	-0.96
25	372.70	366.91	5.79	36.91	35.95	0.96
26	399.20	392.68	6.52	42.1	39.21	2.89
27	423.65	417.81	5.84	40.1	41.70	-1.60
28	445.20	442.20	3.00	42.3	43.22	-0.92
29	468.15	466.06	2.09	41.4	43.57	-2.17
30	488.35	489.58	-1.23	43.3	42.52	0.78

## REFERENCES

- M. Randić, *New J. Chem.* 20,1001, (1996)
- I. Gutman, S. Klavžar, *J. Chem. Inf. Comput. Sci.* 35,1011, (1995)
- M. Randić, *J. Am. Chem. Soc.* 97,6609, (1975)
- D. J. Klein, I. Lukovits, I. Gutman, *J. Chem. Inf. Comput. Sci.* 35,50, (1995)
- F. Shafiei, *Iranian. J. Math. Chem.* 6,15, (2015)
- A. S. Bahjat, S. E. Rita, H. I. Sadigm, A. H. Kawkab, *Am. J. Sci.* 8,773, (2011)
- F. Ashrafi, R. Saadati, A. Behboodi Amlashi, *African. J. Pure. Appl. Chem.* 2,116,(2008)
- F. Shafiei, H. Hosseini, *MATCH Commun. Math. Comput. Chem.* 75,583, (2016)
- V. Sharma, R. Goswami, A.K. Madan, *J. Chem. Inf. Comput. Sci.* 37,273,(1997)
- G.L. Amidon, S.T. Anik, *J. Pharm. Sci.* 65,801,( 1976 )
- V. G. Uryadov, A.I. Kourdioukov, N. Aristova, E. N. Vand Ofitserov, *Butlerov Communications.* 3,67,(2000)
- J. Hao, W. Xinghao, Y. Fen, W. Zunyao, *J. Chemosphere.* 80,665,(2010)
- K. Roy, D. KumarPal, C. Sengupta, *Indian. J. Chem.* 40,129,(2001)
- E. D. Nikitin, A. P. Popov, N.S. Bogatishcheva, *J. Chem. Thermodynamics.* 68,288, (2014)
- G. Espinosa, D. Yaffe, Y. Cohen, A. Arenas, F. Giral, *J. Chem. Inf. Comput. Sci.* 40,859,(2000)
- A.R. Katritzky, V.S. Lobanov, M. Karelson, *J. Chem. Inf. Comput. Sci.* 38,28,(1998)
- Standard Reference Database Number 69; software available at <http://www.nist.gov>.
- M. Randić, *J. Math. Chem.* 7,155, (1991)
- A. T. Balaban, *Chem. Phys. Lett.* 89,399, (1982)
- K. C. Das, B. Zhou, N. Trinajstić, *J. Math. Chem.* 46,1369, (2009)
- Zhou, I. Gutman, *Chem. Phys. Lett.* 394,93, (2004)
- H. Deng, H. Xiao, F. Tang, *MATCH Commun. Math. Comput. Chem.* 63,257,(2010)
- M. Liu, B. Liu, *MATCH Commun. Math. Comput. Chem.* 66,293, (2011)
- Web search engine developed by Chem Axon; software available at <http://WWW.Chemicalize.Org>.