**Journal of Physical and Theoretical Chemistry**
of Islamic Azad University of Iran, 13 (3) 219-226: Fall 2016
(J. Phys. Theor. Chem. IAU Iran)
ISSN 1735-2126

# Application of Graph Theory: Investigation of Relationship Between Boiling Temperatures of Olefins and Topological Indices

Esmat Mohammadinasab

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

Received August 2016; Accepted November 2016

## ABSTRACT

In this study an appropriate computational approach was presented for estimating the boiling temperatures of 41 different types of olefins and their derivatives. Based on the guidelines of this approach, several structural indices related to the organic components were applied using graph theory. Meanwhile, in addition to evaluating the relation between the boiling temperatures of olefins with the structural indices, the property estimation was done with the help of multiple non-linear regression model and other suitable coefficients. For specifying the best structural descriptors out of seven descriptors for determining the considered boiling temperatures, the most appropriate one was specified with the help of multiple non-linear regression model. It was determined that a combined model of Harary and Randic indices is appropriate for determining the boiling temperatures of olefins. The best model to predict $T_{boil}$ of olefins was obtained as follows: $T_{boil}/K = 0.112\ H^{**2} + 2.148 \exp(\chi) + 290.606$.

**Keywords:** Graph theory; Olefin; Boiling temperature; Topological indices

# INTRODUCTION

Olefins, Alkenes or non-saturated aliphatic hydrocarbons have the general form of $C_nH_{2n}$ with one or more double bonds that the carbons participating in the bonds have hybridization $SP^2$[1].The boiling points of olefins increase with the increase of carbon numbers. Due to the existence of Van der Waals forces between the molecules, the boiling temperatures of olefins increase between 20˚C to 30˚C for increasing one carbon atom, except for small olefins. More branches for olefins will cause the decrease of boiling temperature [2]. Boiling point of the material, in addition to the inter-molecular forces, is related to the external pressure.

Generally, organic components do not specify using their material component and their molecular masses, but the other physical properties such as boiling temperatures should be used as well.

Table of physical constants like boiling temperature is so helpful for recognizing the olefins, since with observing the boiling temperature some information about the studied material can be obtained.

The boiling point of a liquid is the temperature at which its vapor pressure is

*Corresponding author: e-mohammadinasab@iau-arak.ac.ir

equal to the pressure of the gas above it, so compared to the melting point, boiling point is more sensitive to changing the air pressure. Hence, boiling point is used as a factor for recognizing the materials. The pressure in which the boiling point has been determined should be specified exactly. The first methods for determining the boiling points were experimental. One of the methods that were used for determining boiling points of olefins was distillation method, though this method can't be used in the cases that the sufficient amount of liquid is not available.

Corresponding to the existing problems for determining the experimental value of this physical constant, and because of the lack of information about boiling temperatures of heavy olefins, the use of non-laboratory methods for determining this property is of prime importance.

In this study, an appropriate computational approach has been presented for estimating the boiling temperature of olefins and their derivatives [3]. The independent variables in the multiple non-linear regression method can be obtained with several methods. Also, graph theory is an appropriate tool for calculating the topological descriptor in the form of independent variables [4].

After plotting the chemical graph of a molecule, it is easy to extract the topological index for that graph [5]. Drawing method and marking graph of the mentioned number is a constant value. In the chemical graph theory, these indices help to predict some of the chemical properties and pharmaceutical material, and these indices presented a considerable results compared to the experimental results. In this study, such topological indices as Hyper-Wiener, Wiener-Polarity, Wiener, Randic, Platt, Balaban, Harary have been used for evaluating this relation [6-17]. (see Table 1).

**Table 1.** Indices, Formula, Description, References of used descriptors

| Index | Formula | Description | Ref | Number |
|---|---|---|---|---|
| Wiener | $W = \frac{1}{2}\sum_{i=}^{n}\sum_{j=1}^{n} D_{ij}$ | $D_{ij}$ is the distance of two vertices i and j in the graph G. | [6] | (1) |
| Hyper-Wiener | $WW(G) = \sum d(U,V)^2 + \sum d(U,V)/2$ | d(U,V) denotes the distance between the vertices U and V in the graph G and the summations run over all (unordered) pairs of vertices of G. | [7] | (2) |
| Randic | $\chi = \sum (\frac{1}{d_i d_j})^{\frac{1}{2}}$ | where $d_i$ and $d_j$ are the degrees of the vertices representing atoms. | [8-11] | (3) |
| Balaban | $J = \frac{m}{\mu+1}\sum_{i=1}^{n}\sum_{j=1}^{n} \left[(D_i)(D_j)\right]^{-0/5}$ | $\mu = m-n+1$ is the the cyclomatic number. | [12-13] | (4) |
| Harary | $\mathrm{H} = \frac{1}{2}\sum_{i=1}^{n} \sum_{j=1}^{n} (D_{ij})^{-2}$ | $D^{-2}$ is the matrix whose elements are the squares of the reciprocal distances. | [14] | (5) |
| Wiener Polarity | $W_P(G) = \left\|\{\{u,v\} \middle\| d_G(u,v) = 3, u,v \in V\}\right.$ | Wp is the number of unordered pairs of vertices {u,v} of G such that $d_G(u,v)=3$ where $d_G$ which is the number of unordered pairs of vertices {u,v}of G. | [15] | (6) |
| Platt number | $F = \sum_{i=1}^{M} D_e$ | F index is the total sum of degree of edges in a graph. | [16-17] | (7) |

## THEORETICAL METHOD

First, the values of experimental normal boiling temperatures of some olefins and their derivatives were taken from reference [18] and were listed in Table 2. Second, the values of Randic ($\chi$), Harary ($H$), Balaban ($J$), Wiener ($W$), Platt ($F$) and HyperWiener ($WW$) and Wiener-Polarity ($Wp$) topological indices were calculated by formula 1-7 using graph theory for 41 different types of olefins' derivatives.

Third, the relationship between experimental normal boiling temperatures of the used olefins and their derivatives with seven different types of topological indices was investigated for 41 various types of olefins derivatives using excel software and the relevant equations were extracted [19-20].

Fourth, the estimation of experimental normal boiling temperatures of the used olefins and their derivatives was performed by using SPSS software version 16 with multiple linear regression method and backward procedure [21]. According to the important determining factors of this method such as correlation coefficient, square correlation coefficient, adjust square correlation coefficient, Fisher statistics, durbin Watson,…. the best topological indices were determined for predicting the boiling temperature of studied molecules.

## RESULTS AND DISCUSSION

The values of experimental normal boiling temperatures of all used olefins were collected in Table. 2.

After calculating the topological indices of studied olefins, the relationship between experimental normal boiling temperatures of olefins with seven the topological indices was investigated using excel software with multiple linear regression method (see equations:8-14).

**Table 2.** Used Olefins, Experimental Data of Boiling Temperature ($T_{boil}$/K) of Studied Olefins

| No | olefin | $T_{boil}$/K | No | olefin | $T_{boil}$/K |
|---|---|---|---|---|---|
| 1 | 3-methyl-1,2-butadiene | 313.95 | 22 | trans-3-hexene | 340.3 |
| 2 | 2,3-pentadiene | 321.35 | 23 | 2,4-dimethyl-1,3-pentadiene | 366.9 |
| 3 | trans-2-pentene | 309.5 | 24 | 2-methyl-1,5-hexadiene | 362 |
| 4 | 1,5-hexadien | 332.55 | 25 | 1,2-heptadiene | 376.9 |
| 5 | 2,3-dimethyl-1,3-butadiene | 343.15 | 26 | 1,5-heptadiene | 366.85 |
| 6 | 3-methyl-1-pentene | 327.37 | 27 | 2,3-dimethyl-1-pentene | 357.5 |
| 7 | trans-3-methyl-2-pentene | 343.5 | 28 | 2,3-dimethyl-2-pentene | 370.55 |
| 8 | trans-4-methyl-2-pentene | 331.7 | 29 | 2,4-dimethyl-1-pentene | 354.73 |
| 9 | 2,3-dimethyl-1-butene | 328.76 | 30 | 2-methyl-1-hexene | 364.65 |
| 10 | 2-ethyl-1-butene | 340.65 | 31 | trans-2-methyl-3-hexene | 359.02 |
| 11 | 2-methyl-1-pentene | 335.26 | 32 | trans-3-methyl-3-hexene | 366.67 |
| 12 | trans-3-methyl-2-pentene | 343.5 | 33 | 2-ethyl-1-pentene | 365.55 |
| 13 | trans-4-methyl-2-pentene | 331.7 | 34 | 2-ethyl-3-methyl-1-butene | 362.05 |
| 14 | 3,3-dimethyl-1-butene | 314.43 | 35 | 3,3-dimethyl-1-pentene | 350.69 |
| 15 | cis-2-hexene | 342.01 | 36 | 3,4-dimethyl-1-pentene | 353.93 |
| 16 | cis-3-hexene | 339.65 | 37 | 3-ethyl-1-pentene | 357.28 |
| 17 | cis-3-methyl-2-pentene | 340.86 | 38 | 3-methyl-1-hexene | 357.09 |
| 18 | cis-4-methyl-2-pentene | 329.45 | 39 | 4,4-dimethyl-1-pentene | 345.35 |
| 19 | trans-2-hexene | 340.24 | 40 | 4-methyl-1-hexene | 359.97 |
| 20 | 3-methyl-2-pentene | 340.65 | 41 | 5-methyl-1-hexene | 358.65 |
| 21 | 2,3,3-trimethyl-1-butene | 350.85 | | | |

The value of dispersion coefficient (R2 < 0.9) in the resulting equations showed that the simple linear regression pattern isn't appropriate for showing the structure-property relation. It was also specified that none of the proposed models have got enough efficiency for predicting the boiling temperatures of olefins. Therefore, the relation between boiling temperatures of olefins with structural descriptors were studied using multiple non-linear regression method through selecting the recycling method and considering the error of 0.05. Some of the mathematical operations such as Powering, Logarithm, Square Root, Exponential and etc. were studied on all of the structural indices and the way that they are related to the boiling temperatures of olefins was investigated. Four models out of all built patterns were selected based on Table. 3.

According to the Table 3, there were 4 models for estimating $T_{boil}$, with sig =0.000, F: 45.399< 58.300<78.128< 120.284, σ: 6.480 > 6.393> 6.367> 6.284, respectively. Therefore, the best model for predicting $T_{boil}$ was obtained using Durbin Watson=2.055, R=0.929, $R^2$ = 0.864, $R^2_{Adjust}$ =0.856, F =120.284, σ =6.28453 K, M-S=4750.669.

**equations: 8-14**

$T_{boil}$=3.1144 (*F*) +311.34 $R^2$= 0.2530 (8)

$T_{boil}$= 52.515 (χ) + 190.01 $R^2$= 0.8639 (9)

$T_{boil}$=17.04 (*J*) + 299.23 $R^2$=0.1051 (10)

$T_{boil}$=8.1682 (*H*) + 263.29 $R^2$=0.7134 (11)

$T_{boil}$=1.4618 (*W*) +288.78 $R^2$=0.8545 (12)

$T_{boil}$=0.5939 (*WW*) + 302.04 $R^2$=0.7958 (13)

$T_{boil}$=9.935 ( *Wp*) +306.06 $R^2$=0.4836 (14)

**Table 3.** Predictors, R, $R^2$, $R^2_{Adjust}$, STD. Error of the estimate, Fisher Coefficient, Mean Square, Significant in estimating of $T_{boil}$

| No | Model | Predictors | *R* | $R^2$ | $R^2_{Adjust}$ | σ | *F* | *M-S* | *Sig* |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 1 | $WP^{**3}$, $H^{**2}$, $WW^{**3}$, $Sz^{**3}$, *exp(X)* | 0.931 | 0.866 | 0.847 | 6.480 | 45.399 | 1906.478 | 0.000 |
| 16 | 2 | $H^{**2}$, $WW^{**3}$, $Sz^{**3}$, *exp(X)* | 0.931 | 0.866 | 0.851 | 6.393 | 58.300 | 2382.709 | 0.000 |
| 17 | 3 | $H^{**2}$, $WW^{**3}$, *exp(X)* | 0.929 | 0.864 | 0.853 | 6.367 | 78.128 | 3167.380 | 0.000 |
| 18 | 4 | $H^{**2}$, *exp(X)* | 0.929 | 0.864 | 0.856 | 6.284 | 120.28 | 4750.669 | 0.000 |

**Models 4**

$T_{boil}$ = 0.006 $W_P^3$+0.190 $H^2$+1.287E-5 $WW^3$+0.000 $Sz^3$+2.309 exp(χ)+285.243 (Model:1)

$T_{boil}$ = 0.204 $H^2$+1.287E-5 $WW^3$+0.000 $Sz^3$+2.382 exp (χ)+283.678 (Model:2)

$T_{boil}$ = 0.115 $H^2$+6.894E-7 $WW^3$+2.075 exp (χ)+291.325 (Model:3)

$T_{boil}$ = 0.112 $H^2$+2.148 exp(χ)+290.606 (Model:4)

The results of this study indicated that between the proposed models for making the structure-property relation, the fourth model had a better validity for predicting the considered property. It was specified that in this model, boiling temperatures of olefins with appropriate coefficients can make an appropriate and sufficient relation with the second power of Harary index and also with the exponential values of Randic index. This proves that the fourth equation has more acceptable statistical coefficients than the proposed equations. One of the important coefficients in the statistical computations is the level of significance. As this level is closer to zero, it shows more significant level between values.

Another important factor is Durbin Watson statistic that has a variation limit of 0-4. The value of 2.055 for the mentioned coefficient in the fourth model shows less correlation between errors. Also, higher values for $R^2$ in this model show a good correlation between boiling temperature with second power of Harary index ($H^{**2}$) and exponential value of Randic index (exp(x)). Furthermore, R being close to $R^2_{adjust}$ shows that the used indices for predicting boiling temperatures are selected correctly. Therefore, the best model to predict $T_{boil}$ of olefins was obtained as follows:

$T_{boil}/K = 0.112\ H^{**2} + 2.148 \exp \chi + 290.606$

Table 4. indicates the values of predicted boiling temperature and residuals of studied olefins.

**Table 4.** The values of predicted $T_{boil}$ and residuals of olefins

| No | $T_{boil(Pred)}$ /K | Residual /K | No | $T_{boil(Pred)}$ /K | Residual /K |
|---|---|---|---|---|---|
| **1** | 316.3605 | -2.41051 | **22** | 338.4797 | 1.820343 |
| **2** | 319.1193 | 2.230692 | **23** | 355.3532 | 11.54681 |
| **3** | 319.1193 | -9.61931 | **24** | 361.8252 | 0.174771 |
| **4** | 338.4797 | -5.92966 | **25** | 369.4838 | 7.416245 |
| **5** | 330.4184 | 12.73156 | **26** | 369.4838 | -2.63376 |
| **6** | 335.4819 | -8.11186 | **27** | 358.3241 | -0.82414 |
| **7** | 335.4819 | 8.018141 | **28** | 358.3241 | 12.22586 |
| **8** | 333.9216 | -2.22163 | **39** | 355.3532 | -0.62319 |
| **9** | 330.4184 | -1.65844 | **30** | 361.8252 | 2.824771 |
| **10** | 335.4819 | 5.168141 | **31** | 361.8252 | -2.80523 |
| **11** | 333.9216 | 1.338366 | **32** | 364.4924 | 2.17761 |
| **12** | 335.4819 | 8.018141 | **33** | 364.4924 | 1.05761 |
| **13** | 333.9216 | -2.22163 | **34** | 358.3241 | 3.725856 |
| **14** | 328.4617 | -14.0317 | **35** | 355.9931 | -5.30311 |
| **15** | 338.4797 | 3.530343 | **36** | 358.3241 | -4.39414 |
| **16** | 338.4797 | 1.170343 | **37** | 367.2318 | -9.95181 |
| **17** | 335.4819 | 5.378141 | **38** | 364.4924 | -7.40239 |
| **18** | 333.9216 | -4.47163 | **49** | 352.699 | -7.34899 |
| **19** | 338.4797 | 1.760343 | **40** | 364.4924 | -4.52239 |
| **20** | 335.4819 | 5.168141 | **41** | 361.8252 | -3.17523 |
| **21** | 348.6714 | 2.178572 | | | |

The residual values were shown at a fairly random pattern (see Fig. 1). Residuals were used to assess the normality assumption.

For evaluating the uniform distribution form of errors, the curves for relation between experimental values and residual values were plotted and it was shown that the errors around x-axis have almost a uniform distribution. This proves the suitability of selected pattern for proposed boiling points of olefins (Fig. 2).

The comparison of these computational results with experimental ones indicated that multiple non-linear regression method enjoys a high precision and accuracy for estimating the boiling temperatures of olefins and their derivatives. Therefore, the histogram and figures did not break the normality assumption.

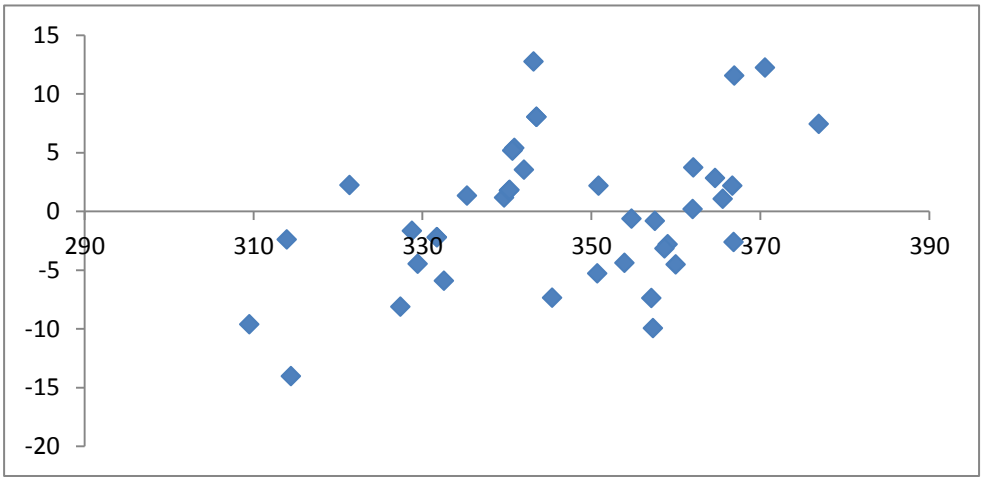


**Fig. 1.** The scatter curve of residual against experimental values of boiling temperatures.

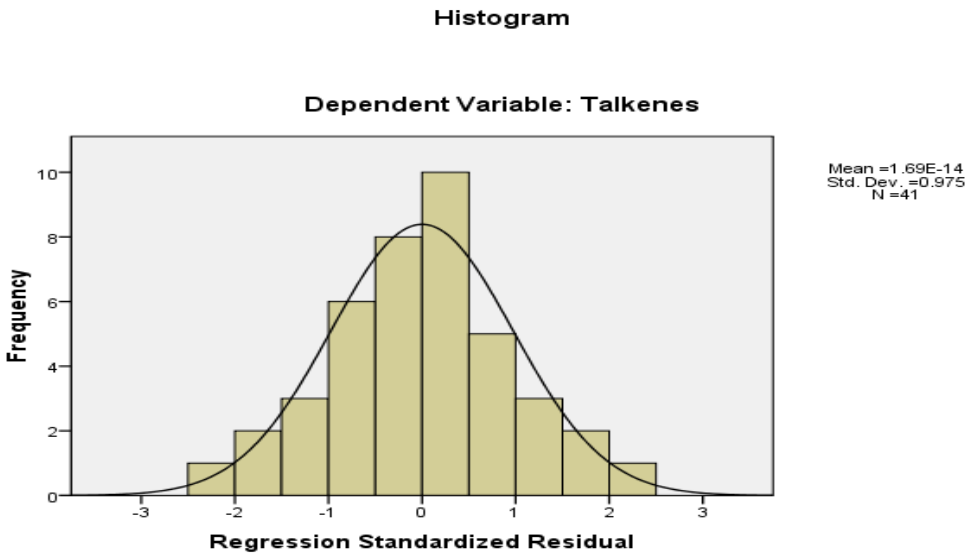


**Fig. 2.** The Histogram of Experimental $T_{boil}$ against Regression Standardized Residual.

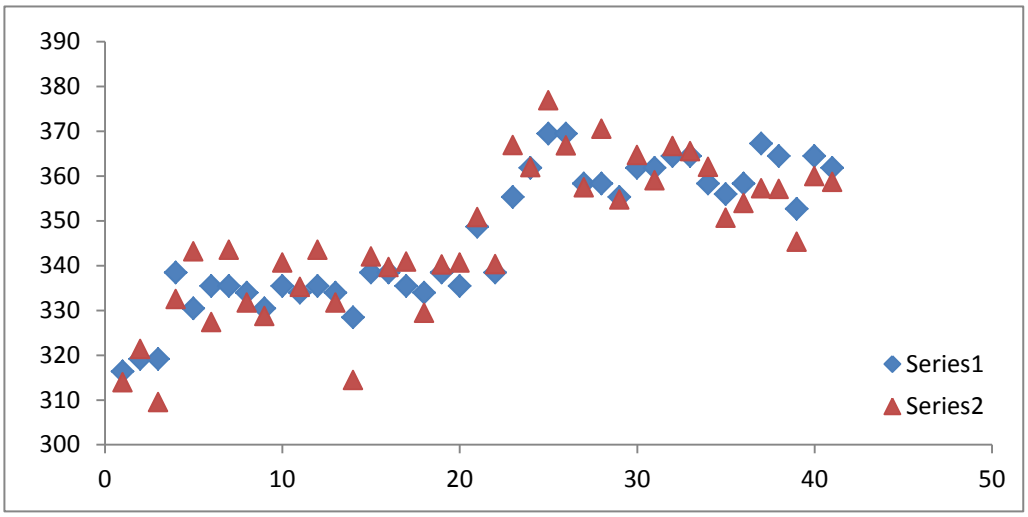


**Fig. 3.** The plot of comparison between the experimental and perdicted $T_{boil}$.

It is also worth mentioning that the comparison between experimental and predicted $T_{boil}$ showed that a good correlation exists between experimental boiling temperatures of the olefins derivatives and the predicted values of non-linear regression model. As it was indicated in the Fig. 3, the predicted $T_{boil}$ values are very close to the experimental $T_{boil}$ values. So, it is observed that the proposed patterns in this model had been selected correctly for determining the boiling temperatures of the studied olefins.

## CONCLUSION

The results of present research indicate that the simple linear regression model with dispersion coefficient (alone) is not sufficient to predict the boiling temperatures of olefins. The non-linear regression model is a statistical means for analyzing the correlation between boiling temperatures as the independent variables and topological indices as dependent variables. However, multiple non-linear regression model benefiting from various descriptors, factors and efficient coefficients can suggest the best algorithm for determining this physical property of chemical molecules.

Also, the research findings revealed that topological indices play a significant role in modeling non-linear regression and determining the boiling temperatures olefines and their derivatives. The comparison of these computations with experimental results indicates that the multiple non-linear regression enjoys a high precision and accuracy for estimation of boiling temperatures of this class of olefines. Finally, the results of the study indicated that Randic and Harary indices have an important role in predicting the boiling temperatures of olefins.

## REFERENCES

[1] L. G. Wade, Organic Chemistry; (6th ed.), Pearson Prentice Hall, 2006.
[2] Arora, Hydrocarbons, Alkanes, Alkens and Alkynes; Discovery Publishing House, 2006.
[3] M. Randic, S. C. Basak, SAR QSAR Environ. Res. 11(1) (2000) 1.
[4] R. J. Wilson, Introduction to *GRAPH THEORY;* (4th ed.), Prentice Hall, 1995.

[5] M. Randic, Chemical graph theory-facts and fiction, Indian J. Chem. 42A (2003) 1207.
[6] H. Wiener, J. Am. Chem. Soc. 69 (1947) 17.
[7] G. Cash, S. Klavzar, Marko Petkovsek, J. Chem. Inf. Comput. Sci. 42 (2002) 571.
[8] X. Li, Y. Shi, MATCH Commun. Math. Comput. Chem. 59 (1) (2008) 127.
[9] M. Randic, MATCH Commun. Math. Comput. Chem. 7 (1979) 5.
[10] B. Liu, I. Gutman, MATCH Commun. Math. Comput. Chem. 58 (1) (2007) 147.
[11] M. Randic, J. Am. Chem. 97(23) (1975) 6609.
A. T. Balaban and T. S. Balaban, Math. Chem. 8 (1991) 383.
[12] T. Balaban, Pure Appl. Chem. 55 (1983) 199.
[13] K. Xu and K. C. Das, Discrete Appl. Math. 159 (15) (2011) 1631.
[14] H. Deng and H. Xiao, J. Serb. Chem. Soc. 75 (10) (2010) 1405.
[15] J. R. Platt, J. Phys. Chem. 56 (1952) 328.
[16] J. R. Platt, J. Chem. Phys. 15 (1947) 419.
[17] Owczarek and K. Blazej, J. Phys. Chem. Ref. Data. 35 (2006) 4.
[18] M. Goodarzi and E. Mohammadinasab, Fuller Nanotub Car N. 21 (2) (2013) 102.
[19] E. Mohammadinasab and M. Goodarzi, J. Fuller Nanotub Car N. 19 (2011) 550.
[20] G.W. Snedecor and W. G. Cochran, Statistical methods; Oxford and IBH Publishing Co, New Delhi, 1967.