

Modeling Ultrasonic Velocity Of 15 Halogen Containing Organic Liquids

Arvind Kumar^{1*}, Vijay K. Agrawal²

^{1*}, ²QSAR Laboratories, Department of Chemistry, APS University, Rewa-486003, MP, India, Email: arvindspatel184@rediffmail.com

*Corresponding Author: - Arvind Kumar

*QSAR Laboratories, Department of Chemistry, APS University, Rewa-486003, MP, India, Email: arvindspatel184@rediffmail.com
DOI: 10.47750/pnr.2023.14.03.188

Abstract

In this study we have tried to model ultrasonic velocity of 15 halogen containing organic liquids. For this purpose, we have used physicochemical and topological parameters. We found that the ultrasonic velocity can be modelled using both the parameters. The best model obtained, after regression analysis, contains surface tension, density and second order-valence connectivity index. The R² value comes out to be 0.9416. The model has been tested for any kind of defect and observed that the proposed model is free from any defect.

Keywords: Ultrasonic velocity, Topological modelling, Organic liquids, physico-chemical parameters, Regression analysis

1. INTRODUCTION

The ultrasonic velocity plays an important role in understanding the ionic interactions of liquids. The chemists study the change in the property of such liquids by the study of changes in excess volume of liquids employing ultrasonic velocity measurements. The thermodynamic properties of pure and mixture organic liquids have been used by scientists [1-2] to study the behaviour and mechanism of action of organic liquids.

In past, we have worked on modelling of drug activities and successfully demonstrated that the physico-chemical parameters as well as topological indices may be used for theoretical estimation of biological activities of drug molecules.

Very recently we have also tried to model the ultrasonic velocity of some 20 organic liquids which were experimentally studied by Parthasarthy and co-workers [3]. We found that excellent results were obtained when the physico-chemical and topological indices are taken together. Such results have been published [4-9] in reputed journals.

To extend this study we tried to model the ultrasonic velocity of halogen derivatives of 15 organic liquids which were studied by Parthasarthy et al [10].

Topological descriptors [11,12] are simple numbers which are derived from the molecular graphs of a molecule. Molecular graphs are constructed from the structure by the atom connections. These graphs are used for the construction of distance matrix and using the mathematical relationships one can obtain the descriptors which are numerical values. It is worthy to mention that the Chemical application of Graph theory is a branch of chemistry which has wide application in mathematical modelling.

2. METHODOLOGY USED

The methodology given by C. Hansch [13] is very useful in modelling any property of organic molecules. Hansch proposed that the properties of molecules can be modelled using physico-chemical parameters and a linear relationship can be established. The mathematical relationship relevant for this is as below

$$Y = \sum M_i X_i + C$$

Where M_i are the coefficients and X_i are parameters (variables). Here Y is dependent variable and X_i are independent variables. In statistical terms this relationship is called multi-parametric correlation. This method is used for establishing correlation between dependent and independent variables.

The following steps are involved in modelling:

1. Drawing of structure and saving mol file.
2. Calculation of Physico-chemical parameters

3. Calculation of Topological descriptors
4. Regression analysis
5. Proposed models
6. Testing of proposed model

The Chemskech software [14] is an effective software which not only helps in drawing of structure of any molecule but also calculates physic-chemical parameters of drawn molecules. It has a built-in module for that. Therefore, we used Chemskech to draw the structure of molecules and calculated the physic-chemical properties. The molecules are saved in the form of mol file which will be used by the Dragon software for calculation of Topological indices [15].

As mentioned earlier, the mol files are used for the calculation of topological descriptor by Dragon software [16]. This software calculates topological, electronic and many other semi-empirical properties of molecules. The parameters (descriptors) are saved and used for the modelling purposes. Thousands of descriptors can be obtained using Dragon but we have to use only few of them for obtaining fruitful correlations.

For obtaining correlations NCSS [17] software has been used. This software also calculates correlation matrix which shows which parameters are suitable for braining suitable models. The different modules in the software also help the user to test the model for any kind of defect. It calculates VIF (valence Inflation Ratio) which is a parameter for testing the defect due to collinearity. LOO (leave one out method) for acceptability of proposed models and in deciding which model is the best by checking the cross validated values of parameters [18,19].

Presentation of Data:

The values of ultrasonic velocity given by Parthasarthy [10] has been presented in Table 1.

Table 1 Compounds used in the present study and their ultrasonic velocity

| Comp. No. | Name/structure of compound | Velocity of sound in meters per second |
|-----------|----------------------------|--|
| 1 | Ethyl bromide | 892 |
| 2 | Butyl bromide | 1016 |
| 3 | Butyl iodide | 959 |
| 4 | Allyl chloride | 1088 |
| 5 | Acetylene dichloride (Cis) | 1025 |
| 6 | AcetyleneTetrachloride | 1155 |
| 7 | Acetylene Tetrabromide | 1007 |
| 8 | Tetrachlorethylene | 1027 |
| 9 | AOrtho Dichlorobenzene | 1246 |
| 10 | Meta Dichlorobenzene | 1232 |
| 11 | Benzoyl chloride | 1318 |
| 12 | Benzyl chloride | 1377 |
| 13 | Chlorobenzene | 1302 |
| 14 | Bromobenzene | 1134 |
| 15 | Ortho-Monochloronapthalene | 1462 |

The physico-chemical parameters (MV, ST, D, Pol) obtained from Chemskech software for present set of compounds are reported in Table 2.

Table 2. Values of calculated physicochemical descriptors for the compounds used in the present study

| S. No. | MW | MR | ST | D | POL. |
|--------|---------|-------|-------|-------|-------|
| 1 | 108.965 | 19.04 | 23.13 | 1.458 | 7.55 |
| 2 | 137.018 | 28.31 | 26.03 | 1.271 | 11.22 |
| 3 | 184.018 | 33.54 | 29.93 | 1.622 | 13.3 |
| 4 | 76.524 | 20.52 | 20.03 | 0.905 | 8.131 |
| 5 | 96.943 | 21.07 | 25.93 | 1.243 | 8.35 |
| 6 | 167.849 | 30.62 | 33.93 | 1.556 | 12.14 |
| 7 | 343.637 | 41.98 | 60.53 | 3.196 | 16.64 |
| 8 | 165.833 | 30.45 | 35.63 | 1.653 | 12.07 |
| 9 | 147.001 | 36.04 | 36.73 | 1.297 | 14.28 |
| 10 | 147.001 | 36.04 | 36.73 | 1.297 | 14.28 |
| 11 | 140.567 | 36.49 | 39.83 | 1.213 | 14.46 |
| 12 | 126.58 | 36.01 | 33.83 | 1.08 | 14.27 |
| 13 | 112.556 | 31.14 | 33.03 | 1.11 | 12.34 |
| 14 | 157.007 | 33.94 | 35.43 | 1.486 | 13.45 |
| 15 | 162.61 | 48.99 | 42.93 | 1.2 | 19.42 |

The independent topological parameters Randic connectivity index [20,21], Kier and Hall Valence connectivity indices [22,23], Wiener index [24], Balaban index [25] have been calculated using Dragon software. The mathematical formulae for the calculation of these descriptors are discussed in the literature [15]. Such values are presented in Table 3.

Table 3 Values of calculated topological descriptors for the compounds used in the present study

| S. NO. | W | JhetZ | Jhetm | ¹ c _c | ⁰ c _c ^r | ¹ c _c ^r | ² c _c ^r | ³ c _c ^r |
|--------|-----|--------|--------|-----------------------------|--|--|--|--|
| 1 | 4 | 2.849 | 2.907 | 1.414 | 3.671 | 2.096 | 1.389 | 0 |
| 2 | 20 | 2.57 | 2.582 | 2.414 | 5.085 | 3.096 | 1.836 | 1.048 |
| 3 | 20 | 2.604 | 2.614 | 2.414 | 5.657 | 3.5 | 2.121 | 1.25 |
| 4 | 10 | 2.898 | 2.915 | 1.914 | 3.125 | 1.618 | 0.752 | 0.327 |
| 5 | 10 | 4.717 | 4.807 | 1.914 | 3.422 | 1.643 | 0.756 | 0.429 |
| 6 | 29 | 5.079 | 5.159 | 2.643 | 5.69 | 2.952 | 2.996 | 1.714 |
| 7 | 29 | 10.271 | 10.772 | 2.643 | 8.856 | 4.178 | 5.821 | 3.857 |
| 8 | 29 | 7.354 | 7.52 | 2.643 | 5.536 | 2.518 | 2.42 | 1.286 |
| 9 | 60 | 3.752 | 3.769 | 3.805 | 5.577 | 2.961 | 2.229 | 1.58 |
| 10 | 61 | 3.656 | 3.671 | 3.788 | 5.577 | 2.955 | 2.313 | 1.257 |
| 11 | 88 | 3.214 | 3.219 | 4.305 | 5.429 | 2.932 | 1.983 | 1.219 |
| 12 | 64 | 3.042 | 3.047 | 3.932 | 5.228 | 3.066 | 1.886 | 1.306 |
| 13 | 42 | 3.362 | 3.371 | 3.394 | 4.521 | 2.478 | 1.732 | 0.985 |
| 14 | 42 | 3.48 | 3.495 | 3.394 | 5.351 | 2.893 | 2.211 | 1.262 |
| 15 | 140 | 3.069 | 3.073 | 5.377 | 6.675 | 3.888 | 2.868 | 2.066 |

The correlation matrix showing inter-correlation among all the parameters are demonstrated in Table 4.

Table 4. Correlation matrix for the inter correlation of the structural descriptors and their correlation with ultrasonic velocity (UV)

| | UV | ² c _c ^r | MW | ⁰ c _c ^r | ¹ c _c ^r | JhetZ | ¹ c _c ^r | Jhetm | ST | D | POL | MR | W | ¹ c _c |
|--|---------|--|--------|--|--|--------|--|--------|-------|---------|--------|--------|--------|-----------------------------|
| UV | 1 | | | | | | | | | | | | | |
| ² c _c ^r | 0.175 | 1 | | | | | | | | | | | | |
| MW | -0.0769 | 0.9062 | 1 | | | | | | | | | | | |
| ⁰ c _c ^r | 0.1519 | 0.9419 | 0.9249 | 1 | | | | | | | | | | |
| ¹ c _c ^r | 0.3571 | 0.9071 | 0.7864 | 0.8686 | 1 | | | | | | | | | |
| JhetZ | 0.0893 | 0.5821 | 0.4057 | 0.3986 | 0.5214 | 1 | | | | | | | | |
| ¹ c _c | 0.1143 | 0.6464 | 0.6881 | 0.7846 | 0.7214 | 0.0429 | 1 | | | | | | | |
| Jhetm | 0.0893 | 0.5821 | 0.4057 | 0.3986 | 0.5214 | 1 | -0.0429 | 1 | | | | | | |
| ST | 0.4629 | 0.8275 | 0.6637 | 0.7961 | 0.7971 | 0.5344 | 0.5916 | 0.5344 | 1 | | | | | |
| D | -0.5952 | 0.613 | 0.7567 | 0.5599 | 0.3789 | 0.4522 | 0.2842 | 0.4522 | 0.288 | 1 | | | | |
| POL | 0.5702 | 0.6631 | 0.5528 | 0.7388 | 0.7328 | 0.2663 | 0.706 | 0.2663 | 0.907 | 0.0519 | 1 | | | |
| MR | 0.5702 | 0.6631 | 0.5528 | 0.7388 | 0.7328 | 0.2663 | 0.706 | 0.2663 | 0.907 | 0.0519 | 1 | 1 | | |
| W | 0.8769 | 0.4762 | 0.277 | 0.4748 | 0.5822 | 0.1653 | 0.4421 | 0.1653 | 0.768 | -0.2554 | 0.8489 | 0.8489 | 1 | |
| ¹ c _c | 0.8805 | 0.4726 | 0.277 | 0.4748 | 0.5966 | 0.1689 | 0.4457 | 0.1689 | 0.768 | -0.2554 | 0.8489 | 0.8489 | 0.9964 | 1 |

The data was subjected to regression analysis [26] and many correlations have been obtained. Such correlations are reported in Table 5.

Table 5. Quality of obtained statistical parameters for different models

| S. No. | Parameters | Ai | B | Se | R ² | R ² Adj | F Ratio | Q= R/Se |
|--------|-----------------------------|---------------------|---------|--------|----------------|--------------------|---------|---------|
| 1 | W | 4.0891(±0.6376) | 35.2431 | 0.0741 | 0.7598 | 0.7414 | 41.129 | 11.7634 |
| 2 | ¹ c | 144.4298(±17.9137) | 57.901 | 0.0617 | 0.8333 | 0.8205 | 65.004 | 14.795 |
| 3 | D | -457.1671(±59.2298) | 59.7592 | 0.0633 | 0.8382 | 0.8113 | 31.094 | 14.4634 |
| | ³ c ^v | 250.1539(±35.4468) | | | | | | |
| 4 | ¹ c | 141.1162(±17.9797) | 73.769 | 0.0611 | 0.8492 | 0.8241 | 33.787 | 15.0822 |
| | JhetZ | -10.1722(±9.0561) | | | | | | |
| 5 | ¹ c | 140.9468(±17.9758) | 72.9966 | 0.061 | 0.8496 | 0.8245 | 33.889 | 15.1105 |
| | Jhetm | -9.7413(±8.5583) | | | | | | |
| 6 | ¹ c | 148.1380(±15.0715) | 60.0445 | 0.0517 | 0.8921 | 0.8741 | 49.615 | 18.2691 |
| | MW | -0.6723(±0.2629) | | | | | | |
| 7 | ¹ c | 163.2956(±16.5876) | 66.0388 | 0.0514 | 0.8934 | 0.8756 | 50.264 | 18.3891 |
| | ⁰ c ^v | -33.0876(±12.7320) | | | | | | |
| 8 | ¹ c | 207.6403(±27.1513) | 65.4774 | 0.0502 | 0.8981 | 0.8811 | 52.855 | 18.8781 |
| | MR | -9.8907(±3.5836) | | | | | | |
| 9 | ¹ c | 207.5677(±27.1224) | 65.4562 | 0.0502 | 0.8981 | 0.8811 | 52.869 | 18.8781 |
| | POL | -24.9352(±9.0319) | | | | | | |

| | | | | | | | | |
|----|-----------------------------|---------------------|----------|--------|--------|--------|--------|---------|
| 10 | ¹ c | 170.6324(±17.1813) | 64.0459 | 0.0498 | 0.8998 | 0.8831 | 53.892 | 19.0477 |
| | ¹ c ^v | -71.5211(±25.3454) | | | | | | |
| 11 | ST | 22.6737(±1.8285) | 43.8067 | 0.0386 | 0.9397 | 0.9296 | 93.489 | 25.1135 |
| | D | -433.5393(±33.2448) | | | | | | |
| 12 | ST | 24.0188(±5.0833) | 47.0695 | 0.0402 | 0.9401 | 0.9238 | 57.582 | 24.1191 |
| | D | -455.1437(±83.2436) | | | | | | |
| | W | -0.2690(±0.9428) | | | | | | |
| 13 | ST | 21.9006(±3.0329) | 60.1178 | 0.0402 | 0.9403 | 0.924 | 57.724 | 24.1217 |
| | D | -457.8124(±81.8561) | | | | | | |
| | MW | 0.3203(±0.9793) | | | | | | |
| 14 | ST | 21.8857(±2.5347) | 51.0315 | 0.04 | 0.9409 | 0.9247 | 58.337 | 24.25 |
| | D | -432.8047(±34.4200) | | | | | | |
| | ¹ c ^v | 12.9363(±27.7072) | | | | | | |
| 15 | ⁰ c ^v | -151.9775(±32.6942) | 105.3067 | 0.0354 | 0.9535 | 0.9409 | 75.261 | 27.584 |
| | ³ c ^v | 186.9193(±49.5151) | | | | | | |
| | ¹ c | 172.0336(±11.6670) | | | | | | |
| 16 | ST | 17.4054(±3.3137) | 76.7596 | 0.0353 | 0.9539 | 0.9413 | 75.859 | 27.6679 |
| | D | -457.4780(±33.0290) | | | | | | |
| | ³ c ^v | 72.1990(±39.2241) | | | | | | |
| 17 | ³ c ^v | 197.5872(±51.3169) | 83.9064 | 0.0352 | 0.954 | 0.9415 | 76.125 | 27.748 |
| | ¹ c | 99.9946(±16.1830) | | | | | | |
| | MW | -3.2605(0.6957) | | | | | | |
| 18 | ST | 18.0515(±2.9935) | 70.8757 | 0.0352 | 0.9541 | 0.9416 | 76.223 | 27.7494 |
| | D | -499.0359(±46.4709) | | | | | | |
| | ² c ^v | 64.7003(±34.8132) | | | | | | |

This table also reports the quality of various statistical parameters. On the basis of such values, one may decide which model is most suitable for modelling the ultrasonic velocity of present set of compounds.

The quality factor, Q, proposed by Pogliani [27, 28] which is ratio of R/Se is also recorded in this table for different correlations. Higher the value of Q the better will be the model.

The value of cross validated parameters has been calculated and they are reported in Table 7. The VIF and other parameters for the best proposed model have been calculated and are presented in Table 8. The estimated values of ultrasonic velocity of present set of compounds have been estimated using best model. Such values are presented in Table 7. A graph has been plotted for the comparison of observed and estimated ultrasonic velocity and the same is shown in Fig. 1.

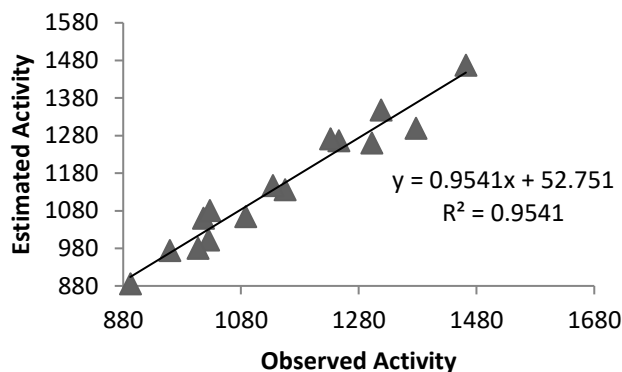


Fig. 1. Correlation between observed and estimated activity using best model

3. RESULTS AND DISCUSSION

We have tried to obtain one-, two-, three- parametric correlations. On the basis of value of statistical parameters, we decide which correlations are of use. A close look of Table indicates that:

In one-parametric correlation with $^1\chi$ only is the best model. It has better potential to model the ultrasonic activity of present setoff compounds. The correlation with Ultrasonic velocity is 0.8805. The second parameter which has capability of correlation is Wiener index. Its value is 0.8769. ST has the minimum capability and D, Pol, MR have moderate potential as their values range from 0.57 to 0.5952. Interestingly the density shows negative correlation (-0.5952).

We have considered all the above information while attempting in higher order correlations. The correlations obtained with their statistical parameters are summarized in Table 5. The results are discussed below.

One-parametric model:

Out of many one-parametric model one with $^1\chi$ gave excellent result as compared to others. The model is as under:

$$UV = 144.4298(\pm 17.9137) ^1\chi + 57.901 \quad \dots (1)$$

N=15, Se = 0.0617, R²=0.8333, Adj R²=0.8205, F-ratio=65.004, Q=14.795

Here and hereafter, N is number of compounds, Se is error of estimation, R² is square value of correlation coefficient, Adj R² is adjusted R², Q is Pogliani's Quality Factor which is ratio of R/Se. Higher the value of Q better will be the proposed model.

Two-parametric model:

We tried two-parametric model by adding parameters to above model. Few statistically better models have been resulted with $^1\chi$ as correlating parameter. The R² value for the model changes from 0.8333 to 0.8998. This improvement in R² is due to addition of first order valence connectivity parameter in to above model. The significant change in adjusted R² clearly indicates that the added parameter has its fair share in the proposed model. The best two-parametric model with $^1\chi$ is reported as under.

$$UV = 170.6324(\pm 17.1813) ^1\chi - 71.5211(\pm 25.3454) ^1\chi' + 64.0459 \quad \dots (2)$$

N=15, Se = 0.0498, R²=0.8998, Adj R²=0.8831, F-ratio=53.892, Q=19.0477

Interestingly one more two-parametric model with better R² value has been observed where two different parameters have been found as correlating the ultrasonic velocity. One is surface tension (ST) and other is density (D) These are physico-chemical parameters whereas above model was with topological parameters. Best two-parametric model with physico-chemical parameters is as under.

$$UV = 22.6737(\pm 1.8285) ST - 433.5393(\pm 33.2448) D + 43.8067 \quad \dots (3)$$

N=15, Se = 0.0386, R²= 0.9397, Adj R²= 0.9296, F-ratio= 93.489, Q =25.1135

In this model the coefficient for Density is negative suggesting that the density has a negative effect towards modelling ultrasonic velocity of present set of compounds while ST has a positive role.

Three-parametric model:

We have tried for three-parametric model. First, we tried with physic-chemical parameters and found that when MW is added to the two parametric model with physic-chemical parameters the R^2 value shows a significant change from 0.9397 to 0.9403. This improvement is justified due to addition of MW or not can be understood on the basis of adjusted R^2 value. In fact, this value has been reduced from 0.9296 to 0.924. Therefore, it may infer that the addition of MW though improved the R^2 value but decrease in adjusted R^2 does not permit the model to be accepted. The value of Pogliani's quality factor Q also shows a dip (24.1217) in the data therefore, the model is also rejected on this count also. The model obtained with physicochemical parameter is as under:

$$UV=21.9006(\pm 3.0329) ST-457.8124(\pm 81.8561) D+0.3203(\pm 0.9793) MW+60.1178 \quad \dots (4)$$

$$N=15, Se =0.0402, R^2= 0.9403, Adj R^2= 0.924 F\text{-ratio}= 57.724, Q =24.1217$$

In the above model the error of estimation incase of MW coefficient come out to be 3 times the value of coefficient hence addition of MW is also rejected on this observation.

When in place of MW we added ${}^2\chi^v$ which is a topological parameter, a three-parametric model with improved statistics has been yielded. The model has shown improved values of R^2 (from 9397 to 0.9541) Adj R^2 (from 9296 to 0.9416) and Q (from 25.1135 to 27.7494). Therefore, we may conclude that the addition of ${}^2\chi^v$ is justified. The model is as given below:

$$UV=18.0515(\pm 2.9935) ST-499.0359(\pm 46.4709) D+64.7003(\pm 34.8132) {}^2\chi^v +70.8757 \quad \dots (5)$$

$$N=15, Se =0.0352, R^2= 0.9541, Adj R^2= 0.9416, F\text{-ratio}= 76.223, Q =27.7494$$

The coefficient of second-order valence connectivity is negative indicating that this parameter has retarding effect towards ultrasonic velocity of present set of compounds.

We have also found another three-parametric model with similar statistical values. In this model out of three correlating parameters one is physicochemical (MW) and two are topological third-order valence connectivity (${}^3\chi^v$) and first order Randic connectivity index (${}^1\chi$).

$$UV=197.5872(\pm 51.3169) {}^3\chi^v+99.9946(\pm 16.1830) {}^1\chi -3.2605(0.6957) MW+83.9064 \quad \dots (6)$$

$$N=15, Se =0.0352, R^2= 0.954, Adj R^2= 0.9415, F\text{-ratio}= 76.125, Q =27.748$$

A comparison of two models (eqns. 5 and 6) reveal that they have almost identical values of statistical parameters. But on the basis of Adj R^2 and Q values we will prefer three-parametric model (Eqn. 5) with ST, D and ${}^2\chi^v$ as the best model for modeling ultrasonic velocity of present set of compounds. The rule of thumb [29,30] says that it is not permitted to use more than 3 parameters for modeling when the number of compounds are 15. Hence, no higher parametric models were tried.

To confirm the above finding we have estimated the ultrasonic velocity of present set of compounds with the best model. The estimated values are reported in Table 6.

Table 6. Cross validated parameter for the proposed models

| S.No. | Parameters | PRESS/SSY | R ² CV | SPRESS | PSE | PRESS | SSY |
|-------|-----------------------------|-----------|-------------------|---------|---------|----------|----------|
| 1 | ¹ c | 0.2 | 0.8 | 70.9316 | 66.0337 | 65406.79 | 327056 |
| 2 | ¹ c | 0.1113 | 0.8887 | 57.2401 | 51.1971 | 39317 | 353146.3 |
| | ¹ c ^v | | | | | | |
| 3 | ST | 0.0642 | 0.9358 | 44.4116 | 39.7229 | 23668.68 | 368794.7 |
| | D | | | | | | |
| 4 | ST | 0.0635 | 0.9365 | 46.1624 | 39.5311 | 23440.67 | 369022.7 |
| | D | | | | | | |
| | MW | | | | | | |
| 5 | ST | 0.0483 | 0.9517 | 40.5588 | 34.7325 | 18095.19 | 374368.2 |
| | D | | | | | | |
| | ¹ c ^v | | | | | | |
| 6 | ¹ c ^v | 0.0482 | 0.9518 | 40.4911 | 34.6745 | 18034.84 | 374428.5 |
| | ¹ c | | | | | | |
| | MW | | | | | | |
| 7 | ST | 0.0481 | 0.9519 | 40.4662 | 34.6532 | 18012.68 | 374450.7 |
| | D | | | | | | |
| | ² c ^v | | | | | | |

The good agreement in the observed and estimated values suggest that the model (Eqn. 5) is best for modelling the ultrasonic activity of the present set of compounds.

Further confirmation is obtained by plotting a graph between observed and estimated values. The trend line regression equation of this graph shows that more than 95% data is explained by the model hence may be accepted. The predicting power of the model comes out to be 0.954.

Test of the model:

The models proposed have been tested for any defect. The VIF (variance inflation factor) [32] is very effective method for checking the collinearity defect. If the value of VIF is more than 10 for any parameter involve in the model the model suffers from the defect of collinearity.

In this study for proposed models the value of VIF is less than 10. Therefore, we come to the conclusion that the proposed models are free from the defect of collinearity.

In Table 8, we have reported all the values calculated for this factor for the best model.

On the basis of λ

The Eigen values calculated for the models are also used to decide whether the model is acceptable or not. If the value of λ is greater than 5 then model suffers from collinearity otherwise it is free from this defect. In all the proposed models the λ values have been found to be less than 5 (Table 7).

Table 7. Observed and estimated values for the present set of compounds using best model

| S.No. | Observed UV | Estimated UV | Residual |
|-------|-------------|--------------|----------|
| 1 | 892 | 885.484 | 6.516 |
| 2 | 1016 | 1060.074 | -44.074 |
| 3 | 959 | 973.753 | -14.753 |
| 4 | 1088 | 1064.277 | 23.723 |
| 5 | 1025 | 1002.365 | 22.635 |
| 6 | 1155 | 1135.508 | 19.492 |
| 7 | 1007 | 980.037 | 26.963 |
| 8 | 1027 | 1080.522 | -53.522 |
| 9 | 1246 | 1265.677 | -19.677 |
| 10 | 1232 | 1271.112 | -39.112 |
| 11 | 1318 | 1347.64 | -29.64 |
| 12 | 1377 | 1299.426 | 77.574 |
| 13 | 1302 | 1260.05 | 41.95 |
| 14 | 1134 | 1146.728 | -12.728 |
| 15 | 1462 | 1467.347 | -5.347 |

On the basis of condition number (k)

The condition number (k) is another parameter to test the presence of collinearity. If k is more than 100 then the collinearity exists in the model, however, the values reported in Table shows that k is less than 100 for all the proposed models (Table 8).

Likewise, the tolerance value (T) equal to 1 or less indicates absence of collinearity.

Table 8. Ridge regression parameters for the best obtained models

| Model No. | Parameters Used | VIF | T | λ_i | K |
|-----------|-----------------|--------|--------|-------------|-------|
| 31 | ST | 7.0913 | 0.141 | 2.6922 | 1 |
| | D | 5.1697 | 0.1934 | 0.2635 | 10.22 |
| | R^2_{CV} | 14.499 | 0.069 | 0.0443 | 60.83 |

The Ridge trace and VIF trace also show that the proposed best model (Table 8) is free from the defect of collinearity. Hence, we may safely say that the proposed best model (Eqn.) is free from any defect of collinearity and is most suitable for modelling the ultrasonic velocity of present set of compounds.

Ridge trace and VIF graphs have been depicted in Figs. 2 and 3 respectively.

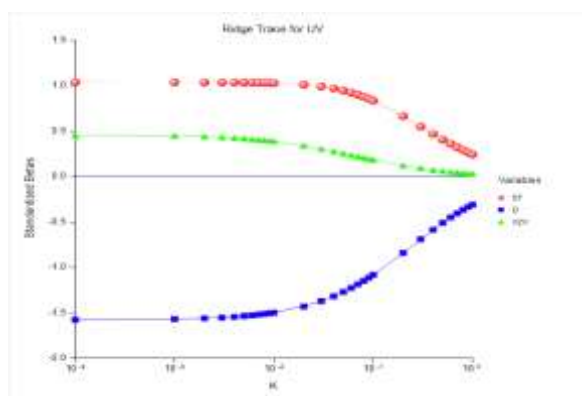


Fig. 2. Ridge plot for UV

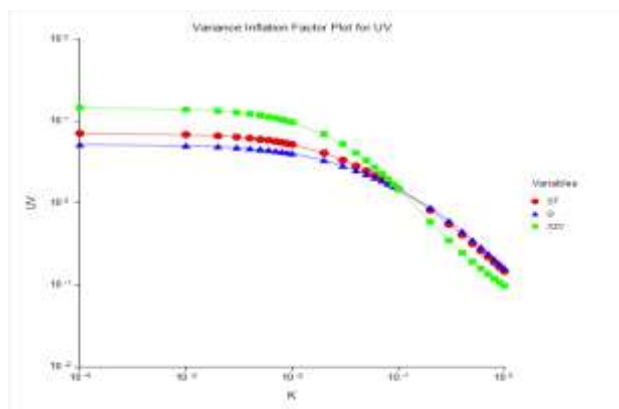


Fig. 3. VIF plot for UV

On the basis of Cross validated parameters

In Cross-validation [Table 7] we obtain the values of PRESS, SSY, S_{PRESS} , R^2_{CV} and PSE. These parameters help us in deciding the predictive power of the proposed model. The PRESS value is a good estimate of the real prediction error of the model. If it is smaller than SSY the model predicts better than chance and the proposed model can be considered “statistically” significant.

It is also important to mention that to be a reasonable QSPR model, the value of PRESS/SSY should be smaller than 0.4. However, if the value of this ratio is smaller than 0.1 the model is accepted as an excellent model. If the PRESS value is transformed in a dimension less quantity by relating it to the initial sum of squares, we obtain R^2_{CV} i.e., the complement to the traces of unexplained variance over the total variance. Thus, PRESS and R^2_{CV} have good properties and are used for deciding whether the proposed model is acceptable or not.

A perusal of Table 7 shows that on the above counts the proposed model (Eqn.) is most appropriate for modeling the ultrasonic velocity of present set of compounds. The R^2_{CV} value comes out to be 0.9518. Also PRESS/SSY for the model comes out to be 0.0481, PSE 34.6532 which is less than SSY value (Table 6).

4. CONCLUSION

On the basis of above discussion, we may conclude that

1. The ultrasonic velocity can be modeled using physic-chemical and topological parameters.
2. Connectivity indices related to connectivity of atoms play dominant role in modeling the ultrasonic velocity of present set of compounds.
3. Best results are obtained when physic-chemical and topological indices are taken together.
4. The coefficient of ST in the model has positive coefficient suggesting that surface tension has a positive effect towards ultrasonic velocity.
5. The density has negative coefficient in the proposed model suggesting that it has retarding role towards ultrasonic velocity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

FUNDINGS

There is no fundings for this research work.

ACKNOWLEDGEMENTS

We would like to express our special thanks of gratitude to Prof J D Pandey who gave us the golden opportunity to do this wonderful research, which also helped us in doing a lot of Research and we came to know about so many new things. We are thankful to them.

REFERENCES

1. Pandey JD, Chaturvedi, B.R. Pant, N., *Acoustics Letters (UK)* 1981; 4:128.
2. Pandey JD, Dey R, *Ind. J. Phys.* 2000, vol. 74(A), no. 3, p. 319.
3. Parthasarathy S, Determination of ultrasonic velocity in 52 organic liquids. *Proc. Ind. Acad. Sci.* 1935;2(5):497. <https://doi.org/10.1007/BF03046893>
4. Agrawal VK, Bano S, Khadikar PV, *Bioorg. Med. Chem.* 2003;11(18):4039. [https://doi.org/10.1016/S0968-0896\(03\)00370-5](https://doi.org/10.1016/S0968-0896(03)00370-5)
5. Singh J, Dubey VK, Agrawal VK, Khadikar PV, *Oxid. Commun.* 2008;1:27.
6. Agrawal VK, Khadikar PV, *Revue Roumaniae de Chimie* 2003;48(10):795.
7. Agrawal VK, Sachan S, Chaturvedi S, Khadikar PV, *Proceedings of the Nat. Acad. Sci. India* 2003;73b:283.
8. Agrawal VK, Singh J, Khadikar PV, *Bioorg. Med. Chem.* 2002;10(12):3981. [https://doi.org/10.1016/S0968-0896\(02\)00299-7](https://doi.org/10.1016/S0968-0896(02)00299-7)
9. Agrawal VK, Dubey VK, Shaik B, Singh J, Singh K, Khadikar PV, *J. Ind. Chem. Soc.* 2009;86:337.
10. Parthasarathy S, *Proc. Ind. Acad. Sci.* 1936;3:519. <https://www.ias.ac.in/article/fulltext/seca/003/06/0519-0522>
11. *Handbook of Graph Theory* (pp.1538-1558), Publisher: CRC Press, Editors: Jonathan L. Gross, Jay Yellen, Ping Zhang.
12. Rouvray DH, *The role of the topological distance matrix in chemistry. Mathematical and computational concepts in chemistry*, (Dubrovnik, 1985) New York, Ellis Horwood; 1986:295.
13. Hansch C, (Ed.) (1991) *Comprehensive Medicinal Chemistry*, Pergamon Press, New York, vol. 5.
14. Chem Sketch: www.acdlabs.com
15. Todeschini R, Consonni V, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim (GER) (2000).
16. DRAGON software for calculation of topological indices, www.disat.unimib.it
17. NCSS software: www.ncss.com
18. Wold S, Eriksson L, *Statistical Validation of QSAR Results, Chemometrics. Methods in Molecular Design*. In Edited by van de Waterbeemd) H. Weinheim New York: VCH; 1995:309.
19. Stuart S, Brenhard G, *Inhibition of oral bacteria by phenolic compounds. Part 1. QSAR analysis using molecular connectivity*, *Quant. Struct. Act. Relat.*, 1998;17(04):327. [https://doi.org/10.1002/\(SICI\)1521-3838\(199808\)17:04%3C327::AID-QSAR327%3E3.0.CO;2-O](https://doi.org/10.1002/(SICI)1521-3838(199808)17:04%3C327::AID-QSAR327%3E3.0.CO;2-O)
20. Randic M. *J. Mol. Graf. Model.* 2001;20(1):19. [https://doi.org/10.1016/S1093-3263\(01\)00098-5](https://doi.org/10.1016/S1093-3263(01)00098-5)
21. Randic M, *J. Chem. Inf. Comput. Sci.* 1997; 37(6):1063. <https://doi.org/10.1021/ci9702407>
22. Kier LB, Hall LH, *Molecular Connectivity in Structure-Activity Analysis*. J. Wiley & Sons, New York, (1986).
23. Kier LB, *J. Pharm. Sci.* 1980;69(7):807. <https://doi.org/10.1002/jps.2600690717>
24. Wiener H, *J. Am. Chem. Soc.* 1947; 69:17-20. <https://doi.org/10.1021/ja01193a005> ; Balaban AT, *Chem. Phys. Lett.* 1982; 89(5):399. [https://doi.org/10.1016/0009-2614\(82\)80009-2](https://doi.org/10.1016/0009-2614(82)80009-2)
25. Chatterjee S, Hadi AS, Price B, *Regression Analysis by Examples*, 3rd Ed. Wiley: New York, (2000).
26. Pogliani L, *Amino Acids* 1994;6(2):141. <https://doi.org/10.1007/BF00805842>
27. Pogliani L, *J. Phys. Chem.* 1996;100(4):18065. <https://doi.org/10.1021/jp961434c>
28. Tute MS, *History and Objectives of Quantitative Drug Design in Advances in Drug Research* Harper NJ, Simmond AB, (Eds) Academic Press: Lond. 1971;6:1-17.