A Comparative QSPR Study of Alkanes with the Help of Computational Chemistry

Srivastava Hemant Kumar

Department of Medicinal Chemistry and Natural Products, School of Pharmacy, The Hebrew University of Jerusalem, Israel
E-mail: hemantkrsri@rediffmail.com
Received August 1, 2008, Accepted November 5, 2008

The development of a variety of methods like AM1, PM3, PM5 and DFT now allows the calculation of atomic and molecular properties with high precision as well as the treatment of large molecules with predictive power. In this paper, these methods have been used to calculate a number of quantum chemical descriptors (like Klopman atomic softness in terms of $E_n^{\ddagger}$ and $E_m^{\ddagger}$, chemical hardness, global softness, electronegativity, chemical potential, electrophilicity index, heat of formation, total energy etc.) for 75 alkanes to predict their boiling point values. The 3D modeling, geometry optimization and semiempirical & DFT calculations of all the alkanes have been made with the help of CAChe software. The calculated quantum chemical descriptors have been correlated with observed boiling point by using multiple linear regression (MLR) analysis. The predicted values of boiling point are very close to the observed values. The values of correlation coefficient (r^2) and cross validation coefficient (r_{cv}^2) also indicates the generated QSPR models are valuable and the comparison of all the methods indicate that the DFT method is most reliable while the addition of Klopman atomic softness $E_n^{\ddagger}$ in DFT method improves the result and provides best correlation.

Key Word : DFT, Semiempirical methods, QSPR, Boiling point, Klopman atomic softness

Introduction

In our previous communications[1-6] we have employed semiempirical techniques to develop QSAR/QSPR models in which the quantum chemical descriptors have been successfully correlated with observed biological activity. Density functional theory (DFT)[7-8] founded within the two basic theorems provided by Hohenberg and Khon in 1960s[9]. The performance of DFT method in description of structural, energetic and magnetic molecular properties has been reviewed quite substantially in recent time. DFT methods are in general capable of generating a variety of isolated molecular properties such as ionization energies,[10-11] dipole moment,[12-13] electrostatic potential,[14-15] electron affinities,[16-17] electronegativities[16-17] and chemical hardness[16-17] *etc.* quite accurately.

Thanikaivelan *et al.*[18] presented a QSPR model for the prediction of boiling points of alkanes using ionization potential as quantum chemical descriptor (r^2 0.968, SE 8.7 degree). Grigoras[19] QSPR model (r^2 0.941, SE 14.1 degree) for the prediction of boiling point employs MSI (molecular surface interactions) descriptors calculated from atomic surface areas and EHT net atomic charges. Murugan *et al.*[20] also found similar results like Grogros[19] by independent correlations of boiling points, melting points and flash points of substituted pyridines (r^2 0.943 and SE 14.5 degree). Balban *et al.*[21] presented the molecular descriptors based QSPR study of ketones and aldehydes (r^2 0.970, SE 6.49 degree), esters (r^2 0.987, SE 4.0 degree) and the combination of these two series of compounds (r^2 0.964 and SE 6.93 degree). Katrizki *et al.*[22] employed the CODESSA software package[23-24] to derive QSPR models for the boiling point of 298 diverse organic compounds. Their best 2-parameter regression (r^2 0.954, SE 16.15 degree) was obtained with the SQRC (Grav Ind all bonds) and the HA dependent HDCA-2/TMSA descriptors.[25] Katritzky *et al.*[22] extended their QSPR approach[26] to a set of 584 diverse organic compounds representative of all major classes of organic compounds containing C, H, O, N, S, F, Cl, Br, and I. The best correlations were obtained by a 6- and an 8-parameter model, (r^2 0.946, SE 18.9 degree). Cocchi *et al.*[27] presented the QSPR model for the boiling point prediction by using GOLPE[28] procedure (r^2 0.931, SE 17.27 degree). GOLPE procedure is based on statistical design and is aimed to improve the predictive ability of the models. Although numerous attempts have been made to correlate physical properties of organic compounds (particularly boiling point) with structural parameters,[18-35] there are very few papers employing quantum chemical descriptors for this purpose.[20,55] In present work we have collected 75 alkanes[36] for the prediction of their boiling point values with the help of quantum chemical descriptors. The comparative QSPR models have been made with the help of AM1,[37] PM3,[38] PM5[39] and DFT[7-8] methods. As we expected, the DFT method based model provides better result than the semiempirical method based models. We also have calculated Klopman atomic softness[40] values in terms of $E_n^{\ddagger}$ at all the atoms of alkanes. The addition of highest $E_n^{\ddagger}$ values of the atom in DFT model surprisingly improve the result of the model (especially in the values of standard error) and provide best correlation. These results are better than the previously known quantum chemical descriptor based study of boiling point.[20,55]

Theory

In DFT the electronegativity commonly known by chemist is defined as negative of partial derivative of energy E of an atomic or molecular system with respect to the number of electron N for a constant external potential $v(r)$[17]

$$\mu = -\chi = -(\partial E/\partial N)_{v(r)} \quad (1)$$

In accordance with the earlier work of Iczkowski and Margrave,[41] it should be remarked that when assuming a quadratic relationship between E and N and in a finite difference approximation equation-1 may be rewritten as

$$\chi = -\mu = -(I+A)/2 \quad (2)$$

where I and A are the vertical ionization energy and electron affinity respectively, there by recovering the electronegativity definition of Mulliken.[42] More over theoretical justification was provided for Sanderson's principle of electronegativity equalization which state that when two or more atoms come together to form molecule, their electronegativities become adjusted to the same intermediate value.[43-45] The absolute hardness h is defined as,[46]

$$\eta = 1/2(\delta\mu/\delta N)v(r) = 1/2(\delta^2 E/\delta N^2)v(r) \quad (3)$$

Where E is the total energy, N the number of electrons of the chemical species and n(r) the external potential. The operational definition of absolute hardness and electro negativity is as

$$\eta = 1/2(I-A) \quad (4)$$

According to the Koopman's theorem, the I is simply the eigen value of HOMO with change of sign and A is the eigen value of LUMO with change of sign,[17] hence the equation-2 and equation-4 may be written as

$$\eta = 1/2(\varepsilon_{LUMO} - \varepsilon_{HOMO}) \quad (5)$$

$$\chi = -\mu = -1/2(\varepsilon_{LUMO} - \varepsilon_{HOMO}) \quad (6)$$

We have also taken a general but important property of a molecular system the molecular weight as a descriptor. In the matter of QSPR of chemical system the total energy also plays important role. Total energy of a molecular system is the sum of the total electronic energy, E_{ee} and the energy of internuclear repulsion, E_{nr}. The total electronic energy of the system is given by,[47-48]

$$E = 1/2\ P\ (H + F), \quad (7)$$

where P is density matrix and H is one-electron matrix.

The softness of an atom in a molecule was described by Klopman[40] and modified by Singh *et al*.[2,49] The Klopman equation is as follows.

$$E_m^{\ddagger} = IP_m - a^2(IP_m - EA_m) - [\chi_r(C_r^m)^2 / R_r](1 - 1/\in) \times [q_r + 2b^2\chi_r(C_r^m)^2] \quad (8)$$

$$E_m^{\ddagger} = IP_n - b^2(IP_n - EA_n) - [\chi_s(C_s^n)^2 / R_s](1 - 1/\in) \times [q_s + 2b^2\chi_s(C_s^n)^2] \quad (9)$$

Where

$E_n^{\ddagger}$ = Softness of a Lewis acid

$E_m^{\ddagger}$ = Softness of a Lewis base

IP = Ionization potential of an atom in a molecule

EA = Electron affinity of an atom in a molecule

$\in$ = Dielectric constant of the medium in which reaction is carried out

R and q = Radius and charge of atom s & r

C = Electron density

χ_r = q - (q-1) $\sqrt{}$ k and k = 0.75

a & b = Variational parameter defined as $a^2+b^2=1$

The ionization potential (IP) of an atom in a molecule, electron affinity (EA) of an atom in a molecule, charge (q) of an atom in a molecule and electron density (C) of an atom in a molecule are essential requirements for the solution of Klopman equations. The method for the calculation of ionization potential of an atom in a molecule (IP) has been described by Dewar and Morita.[50] Method for the calculation of electron affinity of an atom in a molecule (EA) has been described by us earlier.[2] The charge and electron density of an atom in a molecule can be obtained by the quantum chemical calculations. The softness values represented by $E_n^{\ddagger}$ describe the electrophilic character of the molecule, whereas the softness values represented by $E_m^{\ddagger}$ describe the nucleophilic character of the compound.[51] Both $E_n^{\ddagger}$ and $E_m^{\ddagger}$ have been evaluated at all the atoms of each alkanes. The most reactive site as an acid will be that site which has the highest value of $E_n^{\ddagger}$ and the most reactive site as a base will be that site which has the highest value of $E_m^{\ddagger}$. We have used highest $E_n^{\ddagger}$ values as a descriptor because it provides better correlation. Charge and electron density have been calculated at the atom for which the $E_n^{\ddagger}$ value is highest. Dielectric constant of water[52] is used in all the Klopman atomic softness calculations.

Materials and Method

75 alkanes have been used as study material reported under Table-1 along with their observed boiling point values.[36] For QSPR prediction, the 3D modeling and geometry optimization of all the alkanes have been carried out with the help of CAChe Pro software by using semiempirical (AM1, PM3 and PM5) and DFT method. The DFT calculations have been made by using the B88-PW91 functional with DZVP basis set. The values of various descriptors like molecular weight (Mw), HOMO energy (eV) (ε_{HOMO}), LUMO energy (eV) (ε_{LUMO}), ionization potential (eV) (I), electron affinity (eV) (A), chemical hardness (η), global softness (S), electronegativity (χ), chemical potential (μ), electrophilicity index (ω), total energy (Hartree) (T_E), Klopman atomic softness (in terms of $E_n^{\ddagger}$ and $E_m^{\ddagger}$) have been obtained from the quantum chemical calculations or calculated by solving the equations given in theory.

Multiple Linear Regression Analysis (MLR). The Project Leader program associated with CAChe Pro of Fujitsu and DATA Fit program, have been used for multiple linear regression (MLR) analysis. The Quantum chemical descriptors used as independent variables and the boiling point values as the dependent variable. In the statistical analyses, the systematic search was performed to determine the significant descriptors. In order to explore the reliability of the proposed model we have used the cross-validation method. Prediction error sum of squares (PRESS) is a standard index

Table 1. The name of alkanes used as study material

No.	Name	No.	Name
1	hexane	39	2,4-dimethylhexane
2	2-methylpentane	40	2,2-dimethylhexane
3	2,2-dimethylbutane	41	3-ethyl-2-methylpentane
4	isopropylcyclopropane	42	2,2,4-trimethylpentane
5	1-ethyl-2-methylcyclopropane	43	3-ethyl-3-methylpentane
6	1,1,2-trimethylcyclopropane	44	pentacyclopropane
7	ethylcyclobutane	45	(1-methylbutyl)cyclopropane
8	1,1-dimethylcyclobutane	46	1-butyl-2-methylcyclopropane
9	methylcyclopentane	47	(1,2-dimethylpropyl)cyclopropane
10	1,1'-bi(cyclopropyl)	48	1-isobytyl-2-methycyclopropane
11	bicyclo[2.2.0]hexane	49	2-ethyl-1,1,2-trimethylcyclopropane
12	I-methylbicyclo[2.1.0]pentane	50	1,1,2,2,3-pentamethylcyclopropane
13	heptane	51	sec-butylcyclobutane
14	2-methylhexane	52	1,2-diethylcyclobutane
15	2,4-dimethylpentane	53	propylcyclopentane
16	2,3-dimethylpentane	54	isopropylcyclopentane
17	2,2,3-trimethylbutane	55	1,2,4-trimethylcyclopentane
18	butylcyclopropane	56	1-ethyl-1-methylcyclopentane
19	1,2-diethylcyclopropane	57	1,1,2-trimethylcyclopentane
20	1-methyl-1-propylcyclopropane	58	1-ethyl-cyclohexane
21	1,1-diethylcyclopropane	59	1,2-dimethylcyclohexane
22	1-isopropyl-1-methylcyclopropane	60	1,1-dimethylcyclohexane
23	1-ethyl-1,2-dimethylcyclopropane	61	(2-cyclopropylethyl)cyclopropane
24	1,1,2,2-tetramethylcyclopropane	62	bicyclo[3.3.0]octane
25	isopropylcyclobutane	63	bicyclo[5.1.0]octane
26	1-ethyl-3-methylcyclobutane	64	2-methylbicyclo[2.2.1]heptane
27	1,3-dimethylcyclopentane	65	2-methylbicyclo[3.2.0]heptane
28	1,2-dimethylcyclopentane	66	1-methylbicyclo[2.2.1]heptane
29	cycloheptane	67	1-methylbicyclo[4.1.0]heptane
30	(cyclopropylmethyl)cyclopropane	68	3,3-dimethylbicyclo[3.1.0]hexane
31	bicyclo[3.2.0]heptane	69	2,2,4,4-tetramethylbicycclo[1.1.0]butane
32	2-methylbicyclo[3.1.0]hexane	70	1,2,2,3-tetramethylbicycclo[1.1.0] butane
33	1,2-diethyl-3-methylcyclopropane	71	tricyclo[5.1.0.0 3,5]octane
34	1-methylbicyclo[3.1.0]hexane	72	tricyclo[3.2.1.0 2,4]octane
35	tricyclo[4.1.0.0 2,4]heptane	73	3-methyltetracyclo[2.2.1.0]heptane
36	2-methylpentane	74	1-methylcyclohexane
37	3-methylpentane	75	2,2,3-trimethylpentane
38	3-ethylhexane		

to measure the accuracy of a modeling method based on the cross-validation technique. The $r_{cv}2$ can easily be calculated by using equation-10 based on the PRESS and SSY (Sum of squares of deviations of the experimental values from their mean).

$$r_{cv}^2 = 1 - \frac{PRESS}{SSY} = 1 - \frac{\sum_{i=1}^{n}(y_{\exp} - y_{pred})^2}{\sum_{i=1}^{n}(y_{\exp} - \bar{y})^2} \qquad (10)$$

Various regression equations have been developed by using selected quantum chemical descriptors. The best fitted regression equations have been used for the calculation of boiling points (BP_{Pred}). In Models 1-4, the use of three descriptors (chemical hardness, total energy and molecular weight) as independent variables gave nice correlation. In Model-5, the use of E_n^{+} values in combination with above three descriptors as independent variables provide best result.

Results and Discussion

Previously a number of descriptors have been identified and they are capable to correlate successfully the properties with structure of a chemical system.[1-4,53-54] In this paper we are using chemical hardness (η), total energy (T_E in Hartree), molecular weight (M_W) and Klopman atomic softness (E_n^{+}) as descriptor for the prediction of boiling point of alkanes. The importance of hardness and total energy for the prediction was described earlier by Karlsen *et al*.[55] Generally, the boiling point increases as the length of the carbon atom chain increase, so the molecular weight can be an important

Table 2. The AM1 based global descriptors and observed & predicted boiling points

No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual	No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual
1	86.177	-44.281	7.411	68.7	56.66	12.04	39	114.230	-58.572	7.302	109.4	113.49	-4.09
2	86.177	-44.282	7.409	60.3	56.62	3.68	40	114.230	-58.583	7.348	106.8	113.54	-6.74
3	86.177	-44.286	7.468	49.7	57.03	-7.33	41	114.230	-58.566	7.250	115.6	113.19	2.41
4	84.161	-42.761	6.857	58.3	64.39	-6.09	42	114.230	-58.576	7.326	99.2	113.55	-14.35
5	84.161	-42.765	6.719	63.0	62.98	0.02	43	114.230	-58.565	7.239	118.2	113.13	5.07
6	84.161	-42.662	6.541	52.6	64.88	-12.28	44	112.214	-57.060	6.873	128.0	122.11	5.90
7	84.161	-42.830	7.259	70.7	65.76	4.94	45	112.214	-57.055	6.829	117.7	121.86	-4.16
8	84.161	-42.835	7.201	53.6	65.05	-11.45	46	112.214	-57.063	6.697	124.0	120.33	3.67
9	84.161	-42.871	7.289	71.8	64.63	7.17	47	112.214	-57.060	6.894	115.5	122.29	-6.79
10	82.145	-41.246	6.503	76.0	73.81	2.19	48	112.214	-57.062	6.694	110.0	120.36	-10.36
11	82.145	-41.195	6.041	83.0	71.27	11.73	49	112.214	-57.067	6.417	104.5	117.59	-13.09
12	82.145	-41.286	6.258	60.5	70.16	-9.66	50	112.214	-57.067	6.353	100.5	116.98	-16.48
13	100.203	-51.429	7.377	98.5	85.17	13.33	51	112.214	-57.126	7.205	123.0	122.92	0.08
14	100.203	-51.430	7.361	90.0	84.99	5.01	52	112.214	-57.126	7.122	119.0	122.14	-3.14
15	100.203	-51.430	7.359	80.5	84.96	-4.46	53	112.214	-57.166	7.260	131.0	122.04	8.96
16	100.203	-51.426	7.362	89.8	85.14	4.66	54	112.214	-57.168	7.217	126.4	121.58	4.82
17	100.203	-51.431	7.310	80.9	84.47	-3.57	55	112.214	-57.168	7.221	115.0	121.60	-6.60
18	98.188	-49.911	6.878	98.0	93.34	4.67	56	112.214	-57.166	7.236	121.5	121.82	-0.32
19	98.188	-49.912	6.745	90.0	92.05	-2.05	57	112.214	-57.168	7.211	114.0	121.51	-7.51
20	98.188	-49.916	6.676	84.9	91.30	-6.40	58	112.214	-57.182	7.169	131.8	120.64	11.16
21	98.188	-49.908	6.639	88.6	91.21	-2.61	59	112.214	-57.185	7.196	126.6	120.77	5.83
22	98.188	-49.913	6.661	81.5	91.24	-9.74	60	112.214	-57.183	7.162	119.5	120.53	-1.03
23	98.188	-49.917	6.551	85.2	90.09	-4.89	61	110.199	-55.542	6.755	129.0	133.84	-4.84
24	98.188	-49.920	6.453	78.0	89.06	-11.06	62	110.199	-55.725	7.093	137.0	130.67	6.33
25	98.188	-49.981	7.230	92.7	94.24	-1.54	63	110.199	-55.654	6.694	141.0	129.38	11.62
26	98.188	-49.980	7.150	89.5	93.52	-4.02	64	110.199	-55.652	5.566	125.0	118.87	6.13
27	98.188	-50.020	7.253	91.3	93.07	-1.77	65	110.199	-55.620	6.628	130.5	129.95	0.56
28	98.188	-50.019	7.266	95.6	93.23	2.37	66	110.199	-55.739	4.404	117.0	105.00	12.00
29	98.188	-50.019	7.258	118.4	93.17	25.23	67	110.199	-55.660	6.545	125.0	127.76	-2.76
30	96.172	-48.393	6.760	102.0	105.09	-3.09	68[a]	110.199	-55.651	6.673	115.0	129.29	-14.29
31	96.172	-48.473	6.650	110.5	101.28	9.22	69[a]	110.199	-55.470	5.997	104.0	129.22	-25.22
32	96.172	-48.502	6.658	100.0	100.37	-0.37	70[a]	110.199	-55.476	5.625	105.0	125.53	-20.53
33	96.172	-48.505	6.585	103.0	99.55	3.45	71	108.183	-54.139	6.535	142.0	140.64	1.36
34	96.172	-48.506	6.482	92.0	98.57	-6.57	72	108.183	-53.849	4.273	136.0	129.48	6.52
35	94.156	-46.984	6.525	105.0	111.96	-6.96	73	110.199	-55.732	6.469	120.5	124.57	-4.07
36	114.230	-58.578	7.334	117.6	113.55	4.05	74	98.188	-50.036	7.216	101.0	92.17	8.83
37	114.230	-58.575	7.314	118.9	113.47	5.44	75	114.230	-58.577	7.303	109.8	113.32	-3.52
38	114.230	-58.567	7.254	118.5	113.18	5.32							

M_W is the molecular weight, E_T is the total energy, η is the chemical hardness, BP_{Obs} is the observed boiling point in centigrade and BP_{Pred} is the predicted boiling point with the help of the AM1 method. [a]data points are not included in deriving regression equation. Residual is the difference between observed and predicted boiling point.

descriptor for the prediction of boiling point. It is known that the charge and electron density are important in many physico-chemical properties of the compound.[55] Klopman atomic softness strongly depends on the charge and electron density of atom and has been proven a useful quantum chemical descriptor.[3-4]

The values of all these descriptors for all the alkanes have been calculated with the help of AM1, PM3, PM5 and DFT methods by using CAChe Pro software. To make QSPR model based on AM1 method we have generated various equations by employing all the independent variables and the best-fitted equation of this class is equation-11. The predicted boiling point (BP_{Pred}) from equation-11 is reported under Table 2. The statistical quality of the equation is good as is clear from its correlation coefficient r^2 value, which is 0.891, and the cross validation coefficient r_{cv}^2 value, which is 0.885. In this model compounds 68, 69 and 70 have been identified as outlier and the descriptor values of these compounds are not included in deriving regression equation. The Plot of experimental *vs.* predicted boiling point values for this model is given in Figure 1.

$$B.P.\ AM1 = 19.69\ M_W + 34.6063\ E_T + 9.36229\ \eta - 177.126$$
$$r_{cv}^2 = 0.885\ \ r^2 = 0.891,\ SE = 7.9,\ \text{Degree of freedom} = 0.88 \quad (11)$$

The second QSPR model has been made with the help of PM3 based results. In this model we have generated various equations by employing several quantum chemical descriptors and the only best fitted equation is reported here. The

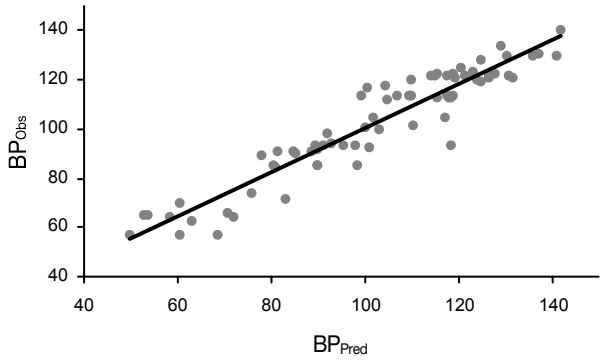

Figure 1. The trend of observed and predicted boiling points using AM1 based model

best fitted equation for the prediction of boiling point (BP_{Pred}) is equation-12 and the predicted values are reported under Table 3. The statistical quality for this model is better than AM1 model. The correlation coefficient r^2 value is 0.910, and the cross validation coefficient r_{cv}^2 value is 0.904. In this model the outlier compounds are 11, 29 and 72. The descriptor values of these compounds are not included in deriving regression equation. The Plot of experimental *vs*. predicted boiling point values for this model is given in Figure 2.

$$B.P.\ PM3 = 19.3459\ M_W + 33.9725\ E_T + 23.6533\ \eta - 273.751$$
$$r_{cv}^2 = 0.904\ r^2 = 0.910,\ SE = 7.5,\ \text{Degree of freedom} = 0.89 \quad (12)$$

Table 3. The PM3 based global descriptors and observed & predicted boiling points

No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual	No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual
1	86.177	-44.397	7.3205	68.7	58.30	10.40	39	114.230	-58.718	7.3040	109.4	114.10	-4.70
2	86.177	-44.402	7.3350	60.3	58.48	1.82	40	114.230	-58.732	7.2290	106.8	111.90	-5.10
3	86.177	-44.413	7.4955	49.7	61.90	-12.20	41	114.230	-58.709	7.0470	115.6	108.30	7.30
4	84.161	-42.846	6.9980	58.3	64.35	-6.05	42	114.230	-58.732	7.2870	99.2	113.20	-14.00
5	84.161	-42.855	6.8850	63.0	61.38	1.62	43	114.230	-58.710	7.1405	118.2	110.50	7.70
6	84.161	-42.663	6.4110	52.6	56.68	-4.08	44	112.214	-57.164	6.9985	128.0	120.70	7.30
7	84.161	-42.913	7.2165	70.7	67.24	3.46	45	112.214	-57.161	6.9385	117.7	119.40	-1.70
8	84.161	-42.931	7.2790	53.6	68.11	-14.51	46	112.214	-57.178	6.8475	124.0	116.60	7.40
9	84.161	-42.960	7.2605	71.8	66.70	5.10	47	112.214	-57.170	7.0420	115.5	121.50	-6.00
10	82.145	-41.291	6.6415	76.0	69.77	6.23	48	112.214	-57.180	6.8315	110.0	116.20	-6.20
11a	82.145	-41.250	6.0965	83.0	58.25	24.75	49	112.214	-57.207	6.5700	104.5	109.10	-4.60
12	82.145	-41.362	6.4390	60.5	62.57	-2.07	50	112.214	-57.215	6.5060	100.5	107.30	-6.80
13	100.203	-51.557	7.2865	98.5	85.62	12.88	51	112.214	-57.234	7.1665	123.0	122.30	0.70
14	100.203	-51.562	7.2670	90.0	84.99	5.01	52	112.214	-57.241	7.1545	119.0	121.80	-2.80
15	100.203	-51.567	7.3230	80.5	86.15	-5.65	53	112.214	-57.277	7.1970	131.0	121.50	9.50
16	100.203	-51.557	7.3830	89.8	87.88	1.92	54	112.214	-57.280	7.1865	126.4	121.20	5.20
17	100.203	-51.570	7.3630	80.9	86.96	-6.06	55	112.214	-57.292	7.2615	115.0	122.50	-7.50
18	98.188	-50.005	7.0025	98.0	92.63	5.37	56	112.214	-57.284	7.2390	121.5	122.30	-0.80
19	98.188	-50.015	6.8515	90.0	88.69	1.31	57	112.214	-57.292	7.2200	114.0	121.60	-7.60
20	98.188	-50.022	6.8290	84.9	87.93	-3.03	58	112.214	-57.287	7.2290	131.8	121.90	9.90
21	98.188	-50.010	6.7490	88.6	86.44	2.16	59	112.214	-57.298	7.3165	126.6	123.60	3.00
22	98.188	-50.022	6.8185	81.5	87.69	-6.19	60	112.214	-57.293	7.2430	119.5	122.10	-2.60
23	98.188	-50.031	6.7045	85.2	84.67	0.53	61	110.199	-55.612	6.8725	129.0	131.40	-2.40
24	98.188	-50.044	6.6070	78.0	81.93	-3.93	62	110.199	-55.812	7.0845	137.0	129.60	7.40
25	98.188	-50.078	7.1900	92.7	94.58	-1.88	63	110.199	-55.723	6.8300	141.0	126.60	14.40
26	98.188	-50.083	7.1930	89.5	94.48	-4.98	64	110.199	-55.820	6.8865	125.0	124.70	0.30
27	98.188	-50.127	7.2505	91.3	94.35	-3.05	65	110.199	-55.711	6.7545	130.5	125.30	5.20
29[a]	98.188	-50.106	7.2585	95.6	95.25	0.35	66	110.199	-55.835	7.1915	117.0	131.40	-14.40
28	98.188	-50.125	7.2385	118.4	94.12	24.28	67	110.199	-55.751	6.7035	125.0	122.70	2.30
30	96.172	-48.451	6.8735	102.0	103.30	-1.30	68	110.199	-55.744	6.8325	115.0	126.00	-11.00
31	96.172	-48.546	6.7760	110.5	97.82	12.68	69	110.199	-55.595	6.1140	104.0	114.10	-10.10
32	96.172	-48.576	6.7825	100.0	96.96	3.04	70	110.199	-55.627	5.8160	105.0	105.90	-0.90
33	96.172	-48.583	6.6985	103.0	94.75	8.25	71	108.183	-54.195	6.6310	142.0	134.90	7.10
34	96.172	-48.591	6.6295	92.0	92.83	-0.83	72[a]	108.183	-54.248	5.4455	136.0	105.00	31.00
35	94.156	-47.026	6.6290	105.0	107.00	-2.00	73	110.199	-55.817	7.2325	120.5	133.00	-12.50
36	114.230	-58.721	7.2330	117.6	112.30	5.30	74	98.188	-50.131	7.3270	101.0	96.00	5.00
37	114.230	-58.717	7.2305	118.9	112.40	6.50	75	114.230	-58.726	7.3000	109.8	113.80	-4.00
38	114.230	-58.710	7.1150	118.5	109.90	8.60							

M_W is the molecular weight, E_T is the total energy, η is the chemical hardness, BP_{Obs} is the observed boiling point in centigrade and BP_{Pred} is the predicted boiling point with the help of the PM3 method. [a]data points are not included in deriving regression equation. Residual is the difference between observed and predicted boiling point.

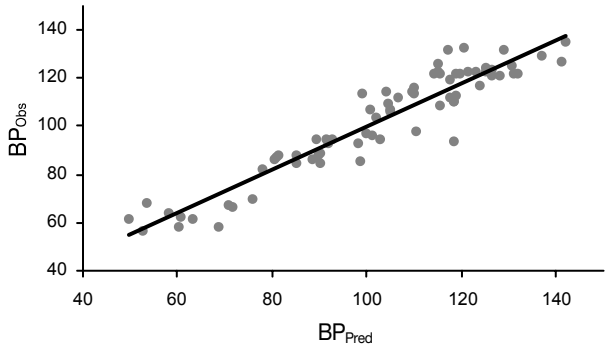

Figure 2. The trend of observed and predicted boiling points using PM3 based model

The third QSPR model has been made with the help of PM5 based results. The various equations for boiling point prediction have been generated by employing different descriptors based on PM5 calculations. The best-fitted equation for this class is equation-12. The predicted boiling points-from equation-13 are given under Table 4. The results are reliable as is clear from correlation coefficient r^2, which is 0.911, and the cross validation coefficient r_{cv}^2, which is 0.905. In this model also three compounds 11, 13 and 29 are outliers and the descriptor values of these compounds are not included in deriving regression equation. The statistical quality of this method is similar like PM3 method. The Plot of experimental *vs.* predicted boiling point values for this model is given in Figure 3.

Table 4. The PM5 based global descriptors and observed & predicted boiling points

No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual	No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual
1	86.177	-44.353	6.8280	68.70	56.32	12.38	39	114.230	-58.659	6.8320	109.40	113.40	-4.00
2	86.177	-44.356	6.8920	60.30	58.04	2.26	40	114.230	-58.670	6.8060	106.80	112.30	-5.50
3	86.177	-44.362	7.0270	49.70	61.63	-11.93	41	114.230	-58.654	6.7605	115.60	111.60	4.00
4	84.161	-42.810	6.6900	58.30	65.46	-7.16	42	114.230	-58.666	6.8585	99.20	113.90	-14.70
5	84.161	-42.819	6.5510	63.00	61.24	1.76	43	114.230	-58.659	6.7460	118.20	111.00	7.20
6	84.161	-42.664	6.2110	52.60	56.78	-4.18	44	112.214	-57.114	6.6285	128.00	120.80	7.20
7	84.161	-42.890	6.8085	70.70	66.14	4.56	45	112.214	-57.117	6.6040	117.70	120.00	-2.30
8	84.161	-42.899	6.8915	53.60	68.17	-14.57	46	112.214	-57.128	6.4895	124.00	116.40	7.60
9	84.161	-42.925	6.8590	71.80	66.41	5.39	47	112.214	-57.118	6.6810	115.50	122.10	-6.60
10	82.145	-41.254	6.3640	76.00	69.70	6.30	48	112.214	-57.130	6.4945	110.00	116.50	-6.50
11[a]	82.145	-41.256	5.8535	83.00	55.25	27.75	49	112.214	-57.154	6.2545	104.50	108.90	-4.40
12	82.145	-41.333	6.2005	60.50	62.47	-1.97	50	112.214	-57.167	6.2270	100.50	107.70	-7.20
13[a]	100.203	-51.507	6.7865	98.50	83.61	14.89	51	112.214	-57.200	6.7345	123.00	120.90	2.10
14	100.203	-51.510	6.8225	90.00	84.54	5.46	52	112.214	-57.205	6.7125	119.00	120.10	-1.10
15	100.203	-51.512	6.8810	80.50	86.10	-5.60	53	112.214	-57.233	6.7815	131.00	121.10	9.90
16	100.203	-51.511	6.8830	89.80	86.20	3.60	54	112.214	-57.235	6.7885	126.40	121.30	5.10
17	100.203	-51.514	6.8625	80.90	85.51	-4.61	55	112.214	-57.238	6.8055	115.00	121.60	-6.60
18	98.188	-49.960	6.6460	98.00	92.81	5.19	56	112.214	-57.238	6.7890	121.50	121.20	0.30
19	98.188	-49.973	6.5550	90.00	89.81	0.19	57	112.214	-57.241	6.7860	114.00	121.00	-7.00
20	98.188	-49.975	6.5080	84.90	88.42	-3.52	58	112.214	-57.247	6.7475	131.80	119.70	12.10
21	98.188	-49.972	6.4805	88.60	87.75	0.85	59	112.214	-57.253	6.7720	126.60	120.20	6.40
22	98.188	-49.975	6.5340	81.50	89.14	-7.64	60	112.214	-57.248	6.7305	119.50	119.20	0.30
23	98.188	-49.988	6.3860	85.20	84.55	0.65	61	110.199	-55.567	6.5080	129.00	130.50	-1.50
24	98.188	-50.000	6.3165	78.00	82.20	-4.20	62	110.199	-55.771	6.7065	137.00	129.40	7.60
25	98.188	-50.047	6.7620	92.70	93.18	-0.48	63	110.199	-55.702	6.5540	141.00	127.30	13.70
26	98.188	-50.051	6.7650	89.50	93.13	-3.63	64	110.199	-55.772	6.7575	125.00	130.70	-5.70
27	98.188	-50.082	6.8145	91.30	93.51	-2.21	65	110.199	-55.687	6.3685	130.50	122.60	7.90
28	98.188	-50.082	6.8285	95.60	93.91	1.69	66	110.199	-55.783	6.7595	117.00	130.40	-13.40
29[a]	98.188	-50.083	6.8380	118.40	94.14	24.26	67	110.199	-55.714	6.4110	125.00	122.90	2.10
30	96.172	-48.412	6.5305	102.00	102.70	-0.70	68	110.199	-55.689	6.5605	115.00	128.00	-13.00
31	96.172	-48.527	6.4030	110.50	95.32	15.18	69	110.199	-55.538	5.9395	104.00	115.50	-11.50
32	96.172	-48.534	6.5255	100.00	98.55	1.45	70	110.199	-55.580	5.5905	105.00	104.20	0.80
33	96.172	-48.544	6.4395	103.00	95.79	7.21	71	108.183	-54.171	6.3855	142.00	135.20	6.80
34	96.172	-48.550	6.3660	92.00	93.52	-1.52	72	108.183	-54.204	6.4970	136.00	137.20	-1.20
35	94.156	-46.985	6.3545	105.00	106.90	-1.90	73	110.199	-55.774	6.7835	120.50	131.40	-10.90
36	114.230	-58.664	6.7780	117.60	111.70	5.90	74	98.188	-50.093	6.8020	101.00	92.77	8.23
37	114.230	-58.663	6.7825	118.90	111.90	7.00	75	114.230	-58.666	6.8455	109.80	113.60	-3.80
38	114.230	-58.655	6.7705	118.50	111.80	6.70							

M_W is the molecular weight, E_T is the total energy, h is the chemical hardness, BP_{Obs} is the observed boiling point in centigrade and BP_{Pred} is the predicted boiling point with the help of the PM5 method. [a]data points are not included in deriving regression equation. Residual is the difference between observed and predicted boiling point.

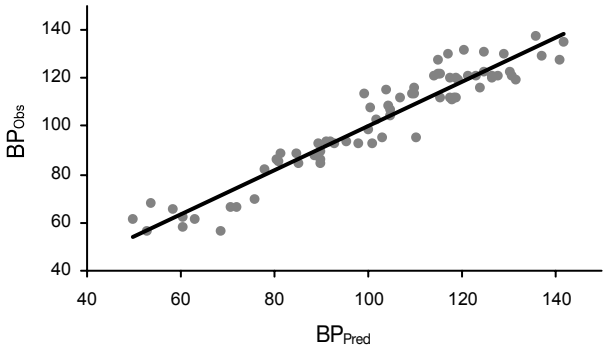

Figure 3. The trend of observed and predicted boiling points using PM5 based model

$$B.P.\ PM5 = 18.9745\ M_W + 33.2244\ E_T + 28.1835\ \eta - 297.672$$
$$r_{cv}^2 = 0.905,\ r^2 = 0.911,\ SE = 7.09,\ \text{Degree of freedom} = 0.90 \quad (13)$$

The fourth QSPR model has been developed on the basis of descriptor values derived from DFT calculation using the B88-PW91 method with the DZVP basis set. In this model we have generated various equations by employing descriptors as independent variables and the best fitted equation is equation-14. The predicted boiling point (BP_{Pred}) from equation-14 is reported in Table-5. The correlation coefficient r^2 is 0.941, and the cross validation coefficient r_{cv}^2 is 0.938. Here compounds 11, 29 and 63 are outliers and the descriptor values of these compounds are not included in deriving

Table 5. The DFT based global descriptors and observed & predicted boiling points

No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual	No.	M_W	E_T	η	BP_{Obs}	BP_{Pred}	Residual
1	86.177	-237.076	4.6055	68.70	64.89	3.81	39	114.23	-315.692	4.3280	109.40	113.22	-3.82
2	86.177	-237.074	4.5170	60.30	62.22	-1.92	40	114.23	-315.697	4.2850	106.80	111.94	-5.14
3	86.177	-237.073	4.3380	49.70	56.84	-7.14	41	114.23	-315.688	4.3570	115.60	114.07	1.53
4	84.161	-235.832	4.0740	58.30	60.64	-2.34	42	114.23	-315.693	4.2145	99.20	109.81	-10.61
5	84.161	-235.835	4.0725	63.00	60.61	2.39	43	114.23	-315.686	4.1730	118.20	108.54	9.66
6	84.161	-235.838	3.9240	52.60	56.16	-3.56	44	112.214	-314.458	4.1645	128.00	120.06	7.94
7	84.161	-235.836	4.2505	70.70	65.96	4.74	45	112.214	-314.453	4.0275	117.70	115.92	1.78
8	84.161	-235.838	4.1530	53.60	63.04	-9.44	46	112.214	-314.461	4.0760	124.00	117.41	6.59
9	84.161	-235.865	4.4700	71.80	72.66	-0.86	47	112.214	-314.454	4.1225	115.50	118.78	-3.28
10	82.145	-234.593	3.8230	76.00	64.84	11.16	48	112.214	-314.458	3.9145	110.00	112.55	-2.55
11[a]	82.145	-234.535	3.5455	83.00	56.29	26.71	49	112.214	-314.458	3.7210	104.50	106.73	-2.23
12	82.145	-234.599	3.8545	60.50	65.81	-5.31	50	112.214	-314.459	3.6580	100.50	104.84	-4.34
13	100.203	-276.387	4.5380	98.50	91.20	7.30	51	112.214	-314.458	4.1740	123.00	120.34	2.66
14	100.203	-276.386	4.4550	90.00	88.71	1.29	52	112.214	-314.462	4.2120	119.00	121.50	-2.50
15	100.203	-276.385	4.3675	80.50	86.07	-5.57	53	112.214	-314.488	4.3885	131.00	126.90	4.10
16	100.203	-276.382	4.3040	89.80	84.16	5.64	54	112.214	-314.487	4.3735	126.40	126.44	-0.04
17	100.203	-276.381	4.2320	80.90	81.99	-1.09	55	112.214	-314.491	4.3535	115.00	125.86	-10.86
18	98.188	-275.146	4.1900	98.00	92.48	5.52	56	112.214	-314.485	4.1925	121.50	121.00	0.50
19	98.188	-275.146	4.0770	90.00	89.08	0.92	57	112.214	-314.487	4.1260	114.00	119.01	-5.01
20	98.188	-275.146	3.9665	84.90	85.76	-0.86	58	112.214	-314.496	4.2635	131.80	123.17	8.63
21	98.188	-275.142	3.9460	88.60	85.13	3.47	59	112.214	-314.494	4.1130	126.60	118.64	7.96
22	98.188	-275.142	3.8965	81.50	83.65	-2.15	60	112.214	-314.496	4.2835	119.50	123.77	-4.27
23	98.188	-275.148	3.8450	85.20	82.12	3.08	61	110.199	-313.217	4.1015	129.00	129.90	-0.90
24	98.188	-275.149	3.8575	78.00	82.50	-4.50	62	110.199	-313.270	4.2935	137.00	135.85	1.15
25	98.188	-275.148	4.1795	92.70	92.17	0.53	63[a]	110.199	-313.246	3.8520	141.00	122.50	18.50
26	98.188	-275.150	4.1150	89.50	90.24	-0.74	64	110.199	-313.273	4.0800	125.00	129.45	-4.45
27	98.188	-275.178	4.3790	91.30	98.27	-6.97	65	110.199	-313.223	3.9635	130.50	125.77	4.73
28	98.188	-275.178	4.3660	95.60	97.88	-2.28	66	110.199	-313.278	4.0610	117.00	128.90	-11.90
29[a]	98.188	-275.173	4.3060	118.40	96.06	22.34	67	110.199	-313.250	3.7600	125.00	119.75	5.25
30	96.172	-273.904	4.0795	102.00	100.89	1.11	68	110.199	-313.248	3.8835	115.00	123.46	-8.46
31	96.172	-273.911	4.0190	110.50	99.09	11.41	69	110.199	-313.196	3.2065	104.00	102.93	1.07
32	96.172	-273.941	3.9860	100.00	98.21	1.79	70	110.199	-313.212	3.4045	105.00	108.94	-3.94
33	96.172	-273.943	4.0245	103.00	99.37	3.63	71	108.183	-312.010	3.6715	142.00	128.83	13.17
34	96.172	-273.944	3.8825	92.00	95.11	-3.11	72	108.183	-312.032	3.9500	136.00	137.28	-1.28
35	94.156	-272.703	4.0490	105.00	111.84	-6.84	73	110.199	-313.274	4.0925	120.50	129.83	-9.33
36	114.23	-315.698	4.4135	117.60	115.81	1.79	74	98.188	-275.185	4.2925	101.00	95.70	5.30
37	114.23	-315.696	4.3810	118.90	114.82	4.08	75	114.23	-315.690	4.2075	109.80	109.59	0.21
38	114.23	-315.692	4.3820	118.50	114.84	3.66							

M_W is the molecular weight, E_T is the total energy, η is the chemical hardness, BP_{Obs} is the observed boiling point in centigrade and BP_{Pred} is the predicted boiling point with the help of the DFT method. [a]data points are not included in deriving regression equation. Residual is the difference between observed and predicted boiling point.

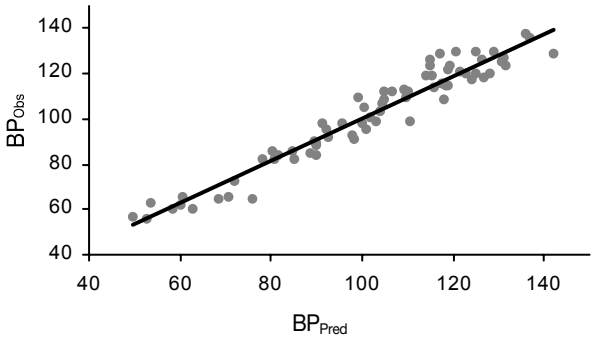

Figure 4. The trend of observed and predicted boiling points using DFT based model

regression equation. The Plot of experimental *vs.* predicted boiling point values for this model is given in Figure 4.

$$\text{B. P. DFT} = -8.02625\ M_W - 3.58485\ E_T + 30.045\ \eta - 231.69$$
$$r_{cv}^2 = 0.938,\ r^2 = 0.941,\ SE = 5.6,\ \text{Degree of freedom} = 0.93 \quad (14)$$

The above discussion indicates that the DFT method is the more reliable than semiempirical methods for the prediction of boiling point of alkanes which is in accordance with our expectations. Although the statistical quality of DFT method is very good but the absolute values of the predicted boiling point is a bit different. To improve the absolute values of the boiling point as well as the statistical quality we have calcu-

Table 6. The DFT based local & global descriptors and observed & predicted boiling points

No.	Site	$E_n^{\ddagger}$	Mw	E_T	η	BP_{Obs}	BP_{Pred}	Residual
1	6	160.8562	86.177	-237.076	4.6055	68.70	68.24	0.46
2	1	174.0529	86.177	-237.074	4.5170	60.30	60.40	-0.10
3	6	177.2726	86.177	-237.073	4.3380	49.70	54.40	-4.70
4	2	168.8818	84.161	-235.832	4.0740	58.30	62.05	-3.75
5	4	164.9671	84.161	-235.835	4.0725	63.00	63.67	-0.67
6	3	178.1847	84.161	-235.838	3.9240	52.60	54.28	-1.68
7	6	159.8139	84.161	-235.836	4.2505	70.70	70.45	0.25
8	5	177.3055	84.161	-235.838	4.1530	53.60	60.58	-6.98
9	6	181.4787	84.161	-235.865	4.4700	71.80	67.12	4.68
10	5	157.5123	82.145	-234.593	3.8230	76.00	71.30	4.70
11[a]	3	144.8478	82.145	-234.535	3.5455	83.00	69.24	13.77
12	6	180.2796	82.145	-234.599	3.8545	60.50	62.56	-2.06
13	7	162.8379	100.203	-276.387	4.5380	98.50	94.21	4.29
14	1	172.9725	100.203	-276.386	4.4550	90.00	87.79	2.21
15	5	179.4718	100.203	-276.385	4.3675	80.50	82.79	-2.29
16	6	173.0481	100.203	-276.382	4.3040	89.80	83.84	5.96
17	5	175.7504	100.203	-276.381	4.2320	80.90	80.83	0.07
18	1	164.0661	98.188	-275.146	4.1900	98.00	95.63	2.37
19	4	152.6404	98.188	-275.146	4.0770	90.00	97.51	-7.51
20	7	175.7985	98.188	-275.146	3.9665	84.90	84.91	-0.01
21	2	165.0029	98.188	-275.142	3.9460	88.60	88.90	-0.30
22	5	181.0449	98.188	-275.142	3.8965	81.50	80.87	0.63
23	2	172.1470	98.188	-275.148	3.8450	85.20	83.30	1.90
24	5	174.8789	98.188	-275.149	3.8575	78.00	82.48	-4.48
25	7	175.2970	98.188	-275.148	4.1795	92.70	90.64	2.06
26	7	172.4479	98.188	-275.150	4.1150	89.50	90.17	-0.67
27	6	175.7110	98.188	-275.178	4.3790	91.30	95.74	-4.44
28	6	180.2740	98.188	-275.178	4.3660	95.60	93.48	2.12
29	4	135.1205	98.188	-275.173	4.3060	118.40	110.90	7.51
30	6	162.8398	96.172	-273.904	4.0795	102.00	104.24	-2.24
31	5	157.9727	96.172	-273.911	4.0190	110.50	104.74	5.76
32	6	175.0174	96.172	-273.941	3.9860	100.00	96.82	3.18
33	7	162.8846	96.172	-273.943	4.0245	103.00	102.93	0.07
34	6	177.8814	96.172	-273.944	3.8825	92.00	92.95	-0.95
35	6	177.0343	94.156	-272.703	4.0490	105.00	108.57	-3.57
36	8	172.6029	114.230	-315.698	4.4135	117.60	115.41	2.19
37	1	165.9926	114.230	-315.696	4.3810	118.90	117.34	1.56
38	8	161.3686	114.230	-315.692	4.3820	118.50	119.30	-0.80

$E_n^{\ddagger}$ is the highest Klopman atomic softness value, M_W is the molecular weight, E_T is the total energy, η is the chemical hardness, BP_{Obs} is the observed boiling point in centigrade and BP_{Pred} is the predicted boiling point from the DFT_{Local} (DFT_{Local} is the addition of $E_n^{\ddagger}$ values in DFT model) method. [a]Data points are not included in deriving regression equation. Residual is the difference between observed and predicted boiling point.

Table 6. Continued

No.	Site	$E_n^{\ddagger}$	Mw	E_T	η	BP_{Obs}	BP_{Pred}	Residual
39	3	173.6211	114.230	-315.692	4.3280	109.40	112.75	-3.35
40	7	174.3305	114.230	-315.697	4.2850	106.80	111.36	-4.56
41	6	170.7109	114.230	-315.688	4.3570	115.60	114.71	0.89
42	1	178.5331	114.230	-315.693	4.2145	99.20	107.75	-8.55
43	6	173.7142	114.230	-315.686	4.1730	118.20	108.68	9.52
44	1	165.7014	112.214	-314.458	4.1645	128.00	122.83	5.17
45	7	163.3517	112.214	-314.453	4.0275	117.70	120.25	-2.55
46	2	175.3007	112.214	-314.461	4.0760	124.00	116.51	7.49
47	2	175.3899	112.214	-314.454	4.1225	115.50	117.65	-2.15
48	7	175.1585	112.214	-314.458	3.9145	110.00	112.38	-2.38
49	8	184.3572	112.214	-314.458	3.7210	104.50	103.50	1.00
50	6	174.5292	112.214	-314.459	3.6580	100.50	106.00	-5.50
51	8	167.4685	112.214	-314.458	4.1740	123.00	122.33	0.67
52	6	155.4407	112.214	-314.462	4.2120	119.00	128.39	-9.38
53	6	156.5930	112.214	-314.488	4.3885	131.00	132.56	-1.56
54	7	171.5713	112.214	-314.487	4.3735	126.40	125.87	0.53
55	7	175.7776	112.214	-314.491	4.3535	115.00	123.60	-8.60
56	6	175.7491	112.214	-314.485	4.1925	121.50	119.42	2.08
57	6	190.1808	112.214	-314.487	4.1260	114.00	111.64	2.36
58	8	155.3382	112.214	-314.496	4.2635	131.80	129.88	1.92
59	7	180.6932	112.214	-314.494	4.1130	126.60	115.31	11.29
60	8	176.4311	112.214	-314.496	4.2835	119.50	121.53	-2.03
61	8	164.1046	110.199	-313.217	4.1015	129.00	132.82	-3.82
62	8	164.4965	110.199	-313.270	4.2935	137.00	137.81	-0.81
63[a]	8	190.0903	110.199	-313.246	3.8520	141.00	115.53	25.47
64	7	179.6043	110.199	-313.273	4.0800	125.00	125.93	-0.93
65	7	153.7925	110.199	-313.223	3.9635	130.50	133.60	-3.10
66	7	192.7253	110.199	-313.278	4.0610	117.00	119.94	-2.94
67[a]	8	193.0943	110.199	-313.250	3.7600	125.00	111.90	13.11
68	7	171.7252	110.199	-313.248	3.8835	115.00	124.07	-9.07
69	5	187.9669	110.199	-313.196	3.2065	104.00	99.53	4.47
70	7	184.6418	110.199	-313.212	3.4045	105.00	106.11	-1.11
71	8	164.8661	108.183	-312.010	3.6715	142.00	132.43	9.57
72	7	187.8743	108.183	-312.032	3.9500	136.00	130.05	5.95
73	8	174.7390	110.199	-313.274	4.0925	120.50	128.31	-7.81
74	7	176.9594	98.188	-275.185	4.2925	101.00	92.99	8.01
75	1	183.8653	114.230	-315.690	4.2075	109.80	105.32	4.49

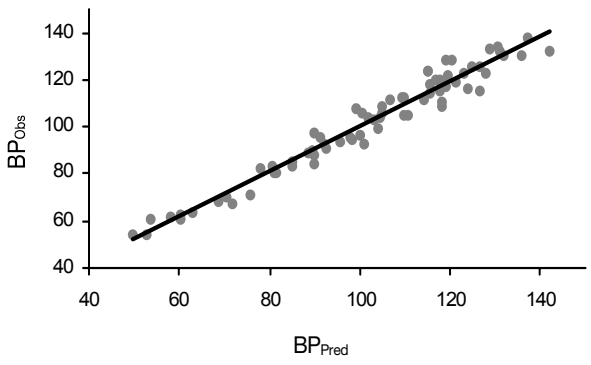

Figure 5. The trend of observed and predicted boiling points using DFT_{Local} based model

lated Klopman atomic softness values at all the atoms of all the alkanes and selected the highest $E_n^{\ddagger}$ value. The MLR analysis with this highest $E_n^{\ddagger}$ values in combination with the descriptors of DFT model (molecular weight, total energy and chemical hardness) provides most valuable QSPR model for the prediction of boiling point of alkanes. We called this model DFT_{Local} model. The regression equation for this model is equation-15 and the predicted values are given in Table-6. In this model the correlation coefficient r^2 is 0.959 and the cross validation coefficient r_{cv}^2 is 0.954. Here compounds 11, 63 and 67 are outliers and the descriptor values of these compounds are not included in deriving regression equation. The Plot of experimental *vs.* predicted boiling point values for this model is given in Figure 5.

$$\text{B. P. DFT}_{\text{Local}} = -0.420475\ E_n^{\ddagger} - 7.53796\ M_W - 3.41562\ E_T + 25.8986\ \eta - 143.56$$
$$r_{cv}^2 = 0.954,\ r^2 = 0.959,\ \text{SE} = 4.7,\ \text{Degree of freedom} = 0.96 \quad (15)$$

The correlation coefficient (r^2), cross validation coefficient (r_{cv}^2) and standard error (SE) values for all the models are as follows.

Method	r^2	r_{cv}^2	SE
AM1	0.891	0.885	7.9
PM3	0.910	0.904	7.5
PM5	0.911	0.905	7.1
DFT	0.941	0.938	5.6
$\text{DFT}_{\text{Local}}$	0.959	0.954	4.7

Conclusion

In this contribution the calculation of semiempirical and DFT based descriptors is carried out and the regression models have been generated for the determination of boiling point of a series of alkanes. The comparison of all the models indicates that the DFT model is more reliable than others and has high predictive power. The addition of Klopman atomic softness in DFT model improves the result of the model and makes it more reliable and important for the prediction of the boiling point of alkanes.

Reference

1. Singh, P. P.; Srivastava, H. K.; Pasha, F. A. *Bioorg. Med. Chem.* **2004**, *12*, 171.
2. Singh, P. P.; Pasha, F. A.; Srivastava, H. K. *QSAR & Comb. Sci.* **2003**, *22*, 843.
3. Srivastava, H. K.; Pasha, F. A.; Singh, P. P. *Int. J. Quantum Chem*. **2005**, *103*, 237**.**
4. Pasha, F. A.; Srivastava, H. K.; Singh, P. P. *Mol. Divers*. **2005**, *9*, 215.
5. Pasha, F. A.; Srivastava, H. K.; Singh, P. P. *Int. J. Quantum. Chem.* **2005**, *104*, 87.
6. Pasha, F. A.; Srivastava, H. K.; Singh, P. P. *Bioorg. Med. Chem.* **2005**, *13*, 6823.
7. Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
8. Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.
9. Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864.
10. Chattaraj, P. K.; Cedillo, A.; Parr, R. G. *J. Phys. Chem.* **1991**, *103*, 7645.
11. Ayers, P. W.; Parr, R. G. *J. Am. Chem. Soc.* **2000**, *122*, 2010.
12. De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Leff.* **1996**, *250*, 393.
13. Geerlings, P.; De Proft, F.; Martin, J. M. L. In *Theoretical and Computational Chemistry*; Seminario, J., Ed.; Elsveir: Amsterdam, 1996; Vol-4 (Recent Developments in Density Functional Theory), p 773.
14. De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *256*, 400.
15. De Proft, F.; Geerlings, P. *J. Chem. Phys.* **1997**, *106*, 3270.
16. Geerlings, P.; De Proft, F.; Langenaeker, W. *Adv. Quantum Chem.* **1999**, *33*, 303.
17. Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801.
18. Thanikaivelan, P.; Subramanian, V.; Rao, R. J.; Nair, B. U. *Chem. Phys. Lett.* **2000**, *323*, 59.
19. Grigoras, S. *J. Comput. Chem.* **1990**, *4*, 493.
20. Murugan, R.; Grendze, M. P.; Toomey, J. E., Jr.; Katritzky, A. R.; Karelson, M.; Lobanov, V. S.; Rachwal, P. *CHEMTECH* **1994**, *24*, 17.
21. Balaban, A. T.; Denise, M.; Basak, S. C. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 758.
22. Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karleson, M. *J. Phys. Chem.* **1996**, *100*, 10400.
23. Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA Training Manual* 1995.
24. Katritzky, A. R.; Lobanov, W. S.; Karelson, M. *Chem. Soc. Rev.* **1995**, 279.
25. Stanton, D. T.; Egolf, L. M.; Jurs, P. C.; Hicks, M. G. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306.
26. Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 28.
27. Cocchi, M.; De Benedetti, P. G.; Seeber, R.; Tassi, L.; Ulrici, A. *J. Chem. Inf. Comput. Sci.* **1999**, *39,* 1190.
28. Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. *Quant. Struct.-Act. Relat.* **1993**, *12*, 9.
29. Needham, D. E.; Wei, I.-C.; Seybold, P. G. *J. Am. Chem. Soc.* **1988**, *110*, 4186.
30. Stanton, D. T.; Egolf, L. M.; Jurs, P. C.; Hicks, M. G. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306.
31. Balaban, A. T.; Joshi, H.; Kier, L. B.; Hall, L. H. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 233.
32. Balaban, A. T.; Kier, L. B.; Joshi, H. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 237.
33. Stanton, D. T.; Jurs, P. C.; Hicks, M. G. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 301.
34. Egolf, L. M.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 616.
35. Egolf, L. M.; Wessel, M. D.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 947.
36. Rucker, G.; Rucker, C. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 788.
37. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
38. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 209.
39. Stewart, J. J. P. *MOPAC2002* Fujitsu Limited: Tokyo, Japan, 2002.
40. Klopman, G. *J. Am. Chem. Soc.* **1968**, *90*, 223.
41. Iczkowski, R. P.; Margrave, J. L. *J. Am. Chem. Soc.* **1961**, *83*, 3547.
42. Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782.
43. Sanderson, R. T. *Science* **1955**, *121*, 207.
44. Sanderson, R. T. *Chemical Bonds and Bond Energy*; Academic: New York, 1976.
45. Sanderson, R. T. *Polar Covalence*; Academic Press: New York, 1983.
46. Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.
47. Clare, B. W. *Aust. J. Chem.* **1995**, *48*, 1385.
48. Parr, R. G.; Szentpaly, L. V.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922.
49. Singh, P. P.; Srivastava, S. K.; Srivastava, A. K. *J. Inor. Nucl. Chem.* **1980**, *42*, 521.
50. Dewar, M. J. S.; Morita, T. F. *J. Am. Chem. Soc.* **1969**, *91*, 796.
51. Singh, P. P. *Coord. Chem. Rev.* **1980**, *32*, 33.
52. Daniels, D. J. *Surface-penetrating radar-IEE Radar, Sonar, Navigation and Avionics*, Series 6; The Institute of Electrical: London, Engineers, pp320, **1996**, p32.
53. Parthasarathi, R.; Subramanian, V.; Roy, D. R.; Chattaraj, P. K. *Bioorg. Med. Chem.* **2004**, *12*, 5533.
54. Hall, L. H.; Vaughn, T. A. *Med. Chem. Res*. **1997**, *7*, 407.
55. Karelson, M.; Lobanov, V. S. *Chem. Rev.* **1996**, *96*, 1027.