J. Chem. Eng. Data **2004,** *49,* 858−863

Melting-Point Models of Alkanes

Kimberly Jordan Burch

Department of Mathematical Sciences, Montclair State University, Montclair, New Jersey 07043

Earl Glen Whitehead, Jr.

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Multivariable models were developed to predict the melting points of alkanes having 10−20 carbon atoms and only one methyl group. These types of alkanes are of special interest to petroleum engineers manufacturing synthetic diesel fuel. Models were based on the number of carbon atoms, the Wiener Path numbers, the mean Wiener index, and the methyl locant index. Nonlinear regression techniques were used which employed web-based NEOS solvers to determine the coefficients and exponents of the indices in the models. The models were used to predict the melting points of alkanes having 10−20 carbon atoms for which no experimental melting-point data existed.

1. Introduction

Boiling-point models were previously constructed for alkanes having 1−12 carbon atoms.[1] The boiling points of alkanes are determined by the forces of attraction between the molecules in the liquid form. Many more factors are involved in determining the melting point of alkanes since the molecules are solids and have a rigid three-dimensional structure. Hence, the melting points of alkanes are more difficult to model. This work models the melting points of a select group of alkanes having a single methyl group. This group was proposed by Dr. John Tierney and Dr. Irving Wender, professors of chemical and petroleum engineering at the University of Pittsburgh. The alkanes having between 10 and 20 carbon atoms were of special interest because they constitute most of the molecules in jet and diesel fuels. Normal alkanes have high cetane numbers and high melting points. Methyl-substituted alkanes have lower, but still adequate, cetane numbers and lower melting points; these alkanes are composed of a straight chain of carbon atoms with one and only one carbon atom (methyl group) attached by a single bond to one of the carbon atoms in the straight chain. Examples of the carbon trees of two such alkanes are shown in Figure 1.

Fuels containing methyl-substituted alkanes have lower pour points than fuels based on normal alkanes. Such fuel can therefore remain fluid at colder temperatures. This is especially important for jet fuels since they are subjected to low temperatures at high altitudes.

Most diesel fuel in the United States is made from petroleum and may have undesirable amounts of sulfur and aromatics. Fuels made via the Fischer−Tropsch process offer an attractive alternative to petroleum-based fuels. They can be made from any carbon-based material such as natural gas, coal, or biomass and have no sulfur or aromatics. The present Fischer−Tropsch manufacturing practice produces mostly normal alkanes, and these must be isomerized for use as fuels.[2]

The models presented in this paper are nonlinear models in that the indices are raised to powers different than 1. Nonlinear models were used because natural laws are often nonlinear. Light intensity, gravitational force, electric fields, and sound intensity all satisfy inverse square laws. In searching for nonlinear models, a linear model would be found if it optimized the objective function.

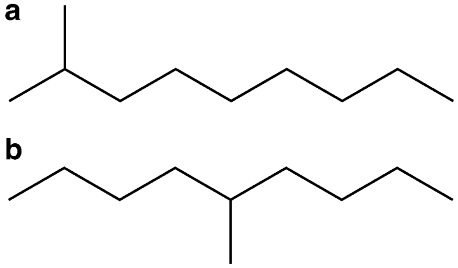

Figure 1. a, 2-Methylnonane; b, 5-Methylnonane.

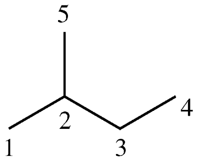

Figure 2. 2-Methylbutane.

2. Indices and Error Vector

This section describes the indices and the error vector used in the models in the following sections. The notation used for the indices is in agreement with the notation used by Todeschini and Consonni.[3]

Wiener Index (W) is the sum of the distances between all pairs of vertices in a graph. It can also be defined as $W = {}^1/_2 \sum_i \sum_j d_{ij}$, where d_{ij} is the ijth entry in the distance matrix, D(G).[4]

Figure 2 is the graph of 2-methylbutane. The distance matrix for 2-methylbutane is the following matrix

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{pmatrix}$$

10.1021/je034185b CCC: $27.50

Table 1. Melting-Point Data

alkane	T_{fus}/K	source	data pts	avg	std dev
decane	243.7	Beilstein	13[a]	243.4	0.7
2-methylnonane	198.6	Beilstein	4	198.6	0.2
3-methylnonane	186.6	NIST	1	NA	NA
4-methylnonane	174.5	Beilstein	1	NA	NA
5-methylnonane	186.2	Beilstein	3	186.2	0.6
undecane	247.4	Beilstein	7	247.4	0.4
2-methyldecane	224.0	Beilstein	2	224.0	0.5
3-methyldecane	189.9	Beilstein	2	189.9	5.3
4-methyldecane	180.3	Beilstein	1	NA	NA
5-methyldecane	183.2	Beilstein	1	NA	NA
dodecane	263.6	Beilstein	15[a]	263.4	0.6
2-methylundecane	227.4	Beilstein	2	227.4	0.3
3-methylundecane	216.9	Beilstein	2	216.8	2.3
4-methylundecane	205.3	Beilstein	2	205.3	1.6
5-methylundecane	204.3	Beilstein	2	204.3	1.6
6-methylundecane	209.3	Beilstein	1	NA	NA
tridecane	267.6	Beilstein	9	267.6	0.3
2-methyldodecane	246.4	Beilstein	2	246.4	1.1
3-methyldodecane	213.0	Beilstein	2	213.0	1.0
4-methyldodecane	222.0	Beilstein	2	222.0	2.3
5-methyldodecane	203.9	Beilstein	2	203.9	0.4
6-methyldodecane	215.0	Beilstein	1	NA	NA
tetradecane	279.0	Beilstein	13	278.9	0.2
2-methyltridecane	247.0	Beilstein	2	247.0	0.4
3-methyltridecane	232.4	Beilstein	2	232.4	4.6
4-methyltridecane	228.5	Beilstein	2	228.5	1.6
5-methyltridecane	222.1	Beilstein	2	222.1	3.6
6-methyltridecane	222.6	Beilstein	1	NA	NA
7-methyltridecane	234.5	Beilstein	2	234.5	2.1
pentadecane	282.9	Beilstein	11[a]	282.9	0.2
2-methyltetradecane	264.6	Beilstein	2[a]	264.6	0.4
3-methyltetradecane	237.2	Beilstein	1	NA	NA
4-methyltetradecane	241.5	Beilstein	2	241.5	1.7
5-methyltetradecane	226.5	Beilstein	2	226.5	3.7
6-methyltetradecane	237.8	Beilstein	1	NA	NA
7-methyltetradecane	225.2	Beilstein	1	NA	NA
hexadecane	291.7	Beilstein	30[b]	291.4	0.9
2-methylpentadecane	263.3	Beilstein	3	263.3	0.9
3-methylpentadecane	251.8	Beilstein	4	251.8	2.7
4-methylpentadecane	250.9	Beilstein	3	250.9	1.8
5-methylpentadecane	242.1	Beilstein	3	242.1	2.3
6-methylpentadecane	241.8	Beilstein	2	241.8	1.1
7-methylpentadecane	243.1	Beilstein	2	243.1	1.0
8-methylpentadecane	245.8	Beilstein	2[a]	245.7	2.7
heptadecane	295.2	Beilstein	18[a]	294.9	0.9
2-methylhexadecane	278.1	Beilstein	1	NA	NA
3-methylhexadecane	258.5	Beilstein	1	NA	NA
4-methylhexadecane	258.4	Beilstein	1	NA	NA
5-methylhexadecane	241.3	Beilstein	1	NA	NA
octadecane	301.7	Beilstein	32[a]	301.3	0.8
2-methylheptadecane	278.5	Beilstein	2	278.5	0.5
3-methylheptadecane	267.0	Beilstein	1	NA	NA
4-methylheptadecane	264.9	Beilstein	1	NA	NA
5-methylheptadecane	253.3	Beilstein	1	NA	NA
9-methylheptadecane	265.0	NIST	1	NA	NA
nonadecane	305.1	Beilstein	23	304.9	0.7
2-methyloctadecane	284.4	Beilstein	2	285.2	1.2
3-methyloctadecane	274.5	Beilstein	2	274.5	1.2
4-methyloctadecane	270.4	Beilstein	2	270.4	2.5
5-methyloctadecane	257.4	Beilstein	2	258.5	1.6
6-methyloctadecane	269.2	Beilstein	1	NA	NA
7-methyloctadecane	256.9	Beilstein	1	NA	NA
8-methyloctadecane	263.2	Beilstein	1	NA	NA
9-methyloctadecane	256.7	Beilstein	1	NA	NA
eicosane	311.0	Beilstein	40	309.8	1
2-methylnonadecane	291.9	Beilstein	3	291.9	1.1
3-methylnonadecane	280.8	Beilstein	1	NA	NA
4-methylnonadecane	271.9	Beilstein	1	NA	NA
5-methylnonadecane	266.2	Beilstein	1	NA	NA

[a] One additional bad data point omitted in average and standard deviation calculations. [b] Two additional bad data points omitted in average and standard deviation calculations.

Therefore

$$W = \frac{8(1) + 8(2) + 4(3)}{2} = 18$$

Mean Wiener Index ($\bar{W}$) is the average of the distances between all pairs of vertices in a graph. For a graph with n vertices

$$\bar{W} = \frac{W}{\binom{n}{2}}$$

For 2-methylbutane, $\bar{W} = 18/10 = 1.8$.

Wiener Path Numbers (^1P, ^2P, ^3P, ^4P) are the numbers ^iP, where ^iP is the number of pairs of vertices in the graph separated by i edges (chemical bonds). Again using the distance matrix, ^iP can be obtained by counting the number of times i appears in the upper triangular part of the matrix. For example, 2-methylbutane has

$$^1P = 4 \qquad ^2P = 4 \qquad ^3P = 2 \qquad ^4P = 0$$

Methyl Locant (loc) is defined to be the number of the carbon to which the methyl branch is attached to the straight chain whose carbons are numbered 1 to $n - 1$, consistent with IUPAC nomenclature. The methyl locant number is defined as 0 for the normal alkanes.

For example, 2-methylbutane has loc = 2.

Error Vector (R^2, number of data points, standard deviation) where R^2 is the coefficient of determination,[5] the number of data points is the number of melting-point data in the model, and standard deviation, denoted σ, is the square root of the sum of the squares divided by the number of data points minus the number of indices used.

3. Regression Techniques

The objective functions for the models were written in AMPL.[6] AMPL is a modeling language for mathematical programming. Problems written in AMPL can be solved in two ways. One can solve problems locally using packages included with the AMPL software such as MINOS, or one can solve problems over the Internet using NEOS solvers at Argonne National Laboratory.[7−9]

The models in this work were optimized using FilterSQP, a NEOS solver that is a sequential quadratic programming (SQP) algorithm.[10] FilterSQP solves nonlinear programming problems with a "filter" to promote global convergence. This solver uses second derivative information from the objective function and constraints in a meaningful way. As a result, FILTER is less prone to converge to shallow local minima.

4. Models

Experimental melting-point data is available for 69 of the 80 alkanes having between 10 and 20 carbon atoms and a single methyl group.[11,12] These data are presented in Table 1. All temperatures are given in Kelvin (K). The predictive abilities of the following models were tested on a group of data that was withheld from the existing data set. A random number generator was used to select this predictive set for each of the models presented. Finally, the models can be used to predict the melting points of the 11 alkanes for which no experimental data exists.

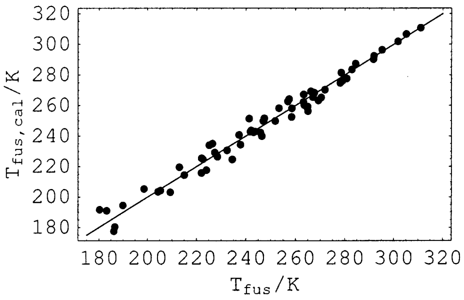

Figure 3. Experimental MP vs Model MP for Model 4.1.

Table 2. Predictive Ability of Model 4.1

alkane	T_{fus}/K	$T_{\mathrm{fus,cal}}$/K	ΔT/K
4-methylnonane	174.5	178.0	3.5
3-methylundecane	216.9	207.6	9.3
5-methyldodecane	203.9	214.9	11.1
4-methyltetradecane	241.5	236.1	5.4
4-methylpentadecane	250.9	244.7	6.2
3-methyloctadecane	274.5	272.1	2.4
9-methyloctadecane	256.7	262.1	5.4

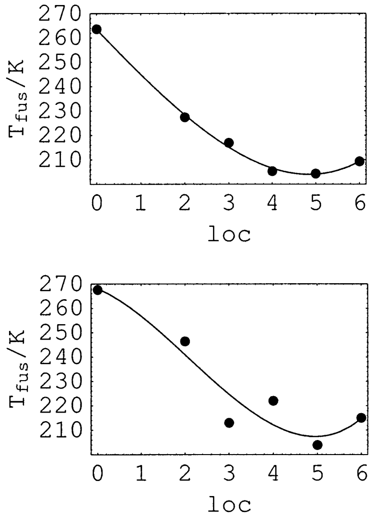

Figure 4. a, MP of the single-branch dodecanes; b, MP of the single-branch tridecanes.

Model 4.1 is the fit to 62 of the 69 alkanes with the remaining 7 withheld as predictive data. This model uses the Wiener ${}^{i}P$ numbers ($1 \leq i \leq 4$) and the mean Wiener index.

Model 4.1.

$$f_{4.1}({}^1P,{}^2P,{}^3P,{}^4P,\overline{W}) = 1625 + 2971({}^1P)^{0.1428} - 4460({}^2P)^{0.05842} - 236.7({}^3P)^{0.3554} - 0.01762\,({}^4P)^{2.727} + 16.11\overline{W}^{1.047}$$

The error vector is (0.97659, 62, 5.1).

Figure 3 shows a plot of the experimental melting points of the 62 alkanes vs the model melting points determined by Model 4.1. The straight line represents an exact prediction.

Table 2 illustrates the predictive ability of Model 4.1 for the given set of seven withheld alkanes. The high errors for some of the alkanes prompted a more careful analysis of the data. After plotting the data, it was observed that the melting points of the alkanes with an odd number of alkanes oscillate as the methyl locant varies. This is illustrated in Figure 4 with the graphs of the melting points of the single branch dodecanes and tridecanes vs their methyl locant indices. The curves in Figure 4 are cubic functions fit to the data. This phenonmenon is a well-known fact, first observed in the late nineteenth century.[13]

Table 3. Predictive Ability of Model 4.2

alkane	T_{fus}/K	$T_{\mathrm{fus,cal}}$/K	ΔT/K
dodecane	263.6	261.9	1.7
2-methylundecane	227.4	225.3	2.0
4-methylundecane	205.3	206.1	0.8
2-methyltridecane	247.0	246.0	0.9
4-methylpentadecane	250.9	244.8	6.1
eicosane	311.0	315.1	4.0

Table 4. Predictive Ability of Model 4.3

alkane	T_{fus}/K	$T_{\mathrm{fus,\,cal}}$/K	ΔT/K
3-methyldodecane	213.0	217.9	4.9
4-methyltetradecane	241.5	238.3	3.1
4-methylhexadecane	258.4	256.1	2.2
nonadecane	305.1	306.9	1.8
2-methyloctadecane	284.4	291.9	7.5

On the basis of these findings, the data set was split into alkanes having an even number of carbons and alkanes having an odd number of carbons. The predictive data for each set was again chosen using a random number generator. The first model discussed is the model based on the alkanes having an even number of carbon atoms.

Model 4.2 is the fit to 31 of the 37 alkanes with the remaining 6 withheld as predictive data. These alkanes have an even number of carbon alkanes. This model uses the Wiener ${}^{i}P$ numbers ($1 \leq i \leq 4$) and the mean Wiener index.

Model 4.2.

$$f_{4.2}({}^1P,{}^2P,{}^3P,{}^4P,\overline{W}) = -552 + 1731({}^1P)^{0.2345} - 1090.5({}^2P)^{0.2120} - 4.866({}^3P)^{1.245} - 254.0({}^4P)^{0.1865} + 0.006393\overline{W}^{4.092}$$

The error vector is (0.99950, 31, 3.8).

The standard deviation of Model 4.2 improved as well as its predictive ability as seen in Table 3. Four of the six predicted melting points are within an error of 2 K or less.

The melting points of the alkanes having an odd number of carbon atoms oscillate as the methyl locant changes, as seen in Figure 4. For this reason, a “wiggle” factor was included in the model for the odd carbon alkanes. Model 4.3 is the fit to 27 of the 32 alkanes with the remaining five withheld as predictive data. This model uses the Wiener ${}^{i}P$ numbers ($1 \leq i \leq 4$), the mean Wiener index, the number of carbon atoms (n), and the methyl locant index.

Model 4.3.

$$f_{4.3}({}^1P,{}^2P,{}^3P,{}^4P,\overline{W},n,\mathrm{loc}) = -6695 + 2159({}^1P)^{0.16275} - 240.5({}^2P)^{0.3518} + 1076({}^3P)^{-0.4243} - 2.537({}^4P)^{1.407} + 3750\,\overline{W}^{0.03078} + 1.203n^{0.5095}\cos(\mathrm{loc}\,\pi)$$

The error vector is (0.99241, 27, 3.0).

The predictive ability of Model 4.3 can be seen in Table 4. The five alkanes used in the predictive set were again chosen by a random number generator.

Models 4.1−4.3 could alternatively be expressed in only two variables, the number of carbon atoms and the methyl locant index. This was suggested at the 2003 SIAM Annual meeting by Alan Sokal of New York University and by Robert Nachbar of Merck Research Laboratories.

We begin by defining the Wiener path numbers and the Wiener index in terms of the number of carbons and the methyl locant index only. Recall that the alkanes studied here have between 10 and 20 carbon atoms and may have one and only one methyl branch

$$^{1}P(n,\mathrm{loc}) = n - 1$$

$$^{2}P(n,\mathrm{loc}) = \begin{cases} n - 2 & \text{if loc} = 0 \\ n - 1 & \text{otherwise} \end{cases}$$

$$^{3}P(n,\mathrm{loc}) = \begin{cases} n - 3 & \text{if loc} \le 2 \\ n - 2 & \text{otherwise} \end{cases}$$

$$^{4}P(n,\mathrm{loc}) = \begin{cases} n - 4 & \text{if loc} \le 3 \\ n - 3 & \text{otherwise} \end{cases}$$

$$W(n,\mathrm{loc}) = \begin{cases} \dfrac{(n-1)n(n+\mathrm{l})}{6} & \text{if loc} = 0 \\ \dfrac{(n-1)n(n+\mathrm{l})}{6} - ((\mathrm{loc}-1)n - \mathrm{loc}^2 + 1) & \text{otherwise} \end{cases}$$

It is well known that the Wiener index of the normal alkanes (loc = 0) is a binomial coefficient.[4] The following gives a proof of the Wiener index when the methyl locant is not zero.

All of the alkanes considered here will have n carbon atoms. First, we consider the difference between the Wiener index for a normal alkane and the Wiener index of an alkane with loc = 2. We will denote this $W_{\mathrm{normal}} - W_{2-\mathrm{methyl}}$. It is well known that $W_{\mathrm{normal}} = 1 + 2 + 3 + ... + n - 1$. We work out that $W_{2-\mathrm{methyl}} = 1 + 2 + 2 + 3 + 4 + ... + n - 2$.

Hence we have that

$$\begin{aligned} W_{\mathrm{normal}} - W_{2-\mathrm{methyl}} &= 1 + 2 + 3 + ... + n - 1 - \\ &\quad (1 + 2 + 2 + 3 + 4 + ... + n - 2) \\ &= (n - 1) - 2 = n - 3. \end{aligned}$$

Similarly

$$\begin{aligned} W_{2-\mathrm{methyl}} - W_{3-\mathrm{methyl}} &= 1 + 2 + 2 + 3 + 4 + ... + \\ &\quad n - 2 - (1 + 2 + 2 + 3 + 3 + \\ &\quad 4 + 5 + ... + n - 3) \\ &= (n - 2) - 3 = n - 5 \end{aligned}$$

In general

$$\begin{aligned} W_{(j-1)-\mathrm{methyl}} - W_{j-\mathrm{methyl}} &= (1 + 2 + 2 + 3 + 3 + ... + \\ &\quad j - 1 + j - 1 + j + j + 1 + \\ &\quad j + 2 + ... + n - j + 1) - \\ &\quad (1 + 2 + 2 + 3 + 3 + ... + j + \\ &\quad j + j + 1 + j + 2 + ... + n - j) \\ &= (n - j + 1) - j = n - (2j - 1) \end{aligned}$$

By use of these results, we can calculate the difference between the Wiener index of a normal alkane and the Wiener index of a j-methyl alkane where both have n carbon atoms

$$\begin{aligned} W_{\mathrm{normal}} - W_{j-\mathrm{methyl}} &= (n - 3) + (n - 5) + ... + \\ &\quad n - (2j - 1) \\ &= (n - 1) + (n - 3) + (n - 5) + ... + \\ &\quad n - (2j - 1) - (n - 1) \\ &= jn - j^2 - (n - 1) \end{aligned}$$

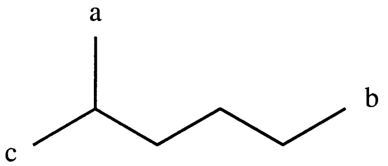

Figure 5. Defining x_1, x_2, and x_3.

This allows us to write that $W_{j-\mathrm{methyl}} = W_{\mathrm{normal}} - (jn - j^2 - (n - 1))$. This is the formula given in the definition of $W(n,\mathrm{loc})$ with j replaced by loc. Recall that the mean Wiener index is defined as $\overline{W} = W/\binom{n}{2}$.

With these definitions, Model 4.1 of all the alkanes can be written as follows.

Model 4.4.

$$f_{4.4}(n,\mathrm{loc}) = 1625 + 2971(^{1}P(n,\mathrm{loc}))^{0.1428} - 4460(^{2}P(n,\mathrm{loc}))^{0.05842} - 236.7(^{3}P(n,\mathrm{loc}))^{0.3554} - 0.01762(^{4}P(n,\mathrm{loc}))^{2.727} + 16.\ 11(\overline{W}(n,\mathrm{loc}))^{1.047}$$

Model 4.2 of the alkanes having an even number of carbons can be written as follows.

Model 4.5.

$$f_{4.5}(n,\mathrm{loc}) = -552 + 1731(^{1}P(n,\mathrm{loc}))^{0.2345} - 1090.5(^{2}P(n,\mathrm{loc}))^{0.2120} - 4.866(^{3}P(n,\mathrm{loc}))^{1.245} - 254.0(^{4}P(n,\mathrm{loc}))^{0.1865} + 0.006393(\overline{W}(n,\mathrm{loc}))^{4.092}$$

Model 4.3 of the alkanes having an odd number of carbons can be written as follows.

Model 4.6.

$$f_{4.6}(n,\mathrm{loc}) = -6695 + 2159(^{1}P(n,\mathrm{loc}))^{0.16275} - 240.5(^{2}P(n,\mathrm{loc}))^{0.3518} + 1076(^{3}P(n,\mathrm{loc}))^{-0.4243} - 2.537(^{4}P(n,\mathrm{loc}))^{1.407} + 3750(\overline{W}(n,\mathrm{loc}))^{0.03078} + 1.203n^{0.5095}\cos(\mathrm{loc}\ \pi)$$

Seifer and Smolenskii modeled a similar set of alkanes which all had a basic “y-shaped” carbon tree. Their work modeled the boiling points of 35 alkanes including alkanes having loc = 2, 3, and 4, as well as ethyl and propyl attachments.[14] Seifer and Smolenskii used three parameters in their models, x_1, x_2, and x_3. By use of Figure 5, we define the parameters as follows: $x_1 = \mathrm{dist}(a,b)$, $x_2 = \mathrm{dist}(a,c)$, and $x_3 = \mathrm{dist}(b,c)$.

The equation they used to fit the boiling points of the alkanes was of the following form

$$y = a_0 + a_1(x_1 + x_2 + x_3) + a_2(x_1^{\,2} + x_2^{\,2} + x_3^{\,2}) + a_3(x_1x_2 + x_1x_3 + x_2x_3) + a_4(x_1^{\,3} + x_2^{\,3} + x_3^{\,3}) + a_5(x_1^{\,2}x_2 + x_1^{\,2}x_3 + x_1x_2^{\,2} + x_1x_3^{\,2} + x_2^{\,2}x_3 + x_2x_3^{\,2}) + a_6x_1x_2x_3 \quad (1)$$

The question arose if these techniques could be applied to modeling the melting points of the alkanes studied in this work. The data set examined in this work has 15 alkanes in common with the data Seifer and Smolenskii used in their study. By use of eq 1, the following model was formed which uses the same 62 alkanes that were used in Model 4.1.

Table 5. Predictive Ability of Model 4.7

alkane	T_{fus}/K	$T_{fus,cal}$/K	ΔT/K
4-methylnonane	174.5	185.1	10.6
3-methylundecane	216.9	209.4	7.4
5-methyldodecane	203.9	205.3	1.5
4-methyltetradecane	241.5	231.5	9.9
4-methylpentadecane	250.9	240.7	9.5
3-methyloctadecane	274.5	265.0	9.5
9-methyloctadecane	256.7	238.1	18.6

Table 6. Eleven Predicted Melting Points

alkane	$T_{fus,cal}$/K
6-methylhexadecane	253.9
7-methylhexadecane	243.1
8-methylhexadecane	253.0
6-methylheptadecane	258.8
7-methylheptadecane	258.6
8-methylheptadecanee	258.4
6-methylnonadecane	271.4
7-methylnonadecane	271.1
8-methylnonadecane	270.8
9-methylnonadecane	270.7
10-methylnonadecane	270.7

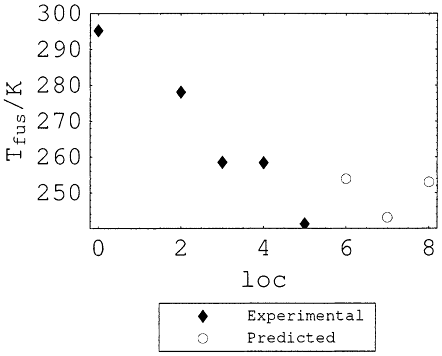

Figure 6. Heptadecanes.

Model 4.7.

$$f_{4.7}(n,\text{loc}) = (x_1, x_2, x_3) = 248.33 - 11.595(x_1 + x_2 + x_3) + 1.629(x_1^2 + x_2^2 + x_3^2) - 0.010(x_1x_2 + x_1x_3 + x_2x_3) - 0.045(x_1^3 + x_2^3 + x_3^3) - 0.006(x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2) + 0.014(x_1x_2x_3).$$

The error vector is (0.954935, 62, 6.9).

The standard deviation of Model 4.7 is nearly 2 K higher than that of Model 4.1. The predictive ability is extremely poor, as seen in Table 5.

5. Conclusions

Splitting the data set of Model 4.1 improved both the fit of the melting-point models and their predictive abilities. The worst predictive error for Model 4.1 was 11.1 K. The models that split the data into even and odd carbon atom sets, Model 4.2 and Model 4.3 respectively, had 6.1 K and 7.5 K as their worst predictive errors.

Table 6 gives the predicted values for the melting points of the eleven alkanes missing from the 80 alkanes having between 10 and 20 carbon atoms and a single methyl group. Model 4.2 was used if the alkane had an even number of carbons, and Model 4.3 was used if the alkane had an odd number of carbons.

Figures 6 and 7 give the plots of these predicted melting points vs their methyl locant indices. Existing experimental data for each of the plots is also included. Only the heptadecanes, octadecanes, and eicosanes are given because experimental data is missing in these sets.

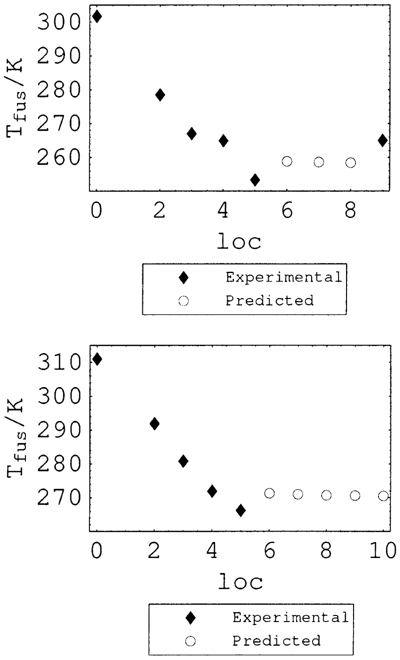

Figure 7. a, Octadecanes; b, Eicosanes.

Plans for future research include using larger data sets of alkanes. Also, we would like to use unbiased cross validation[15] to select the predictive sets for the models. This would allow us to test the stability of our models.

Acknowledgment

We wish to thank Dr. John Tierney and Dr. Irving Wender of the school of engineering, University of Pittsburgh for their help on this project. We also wish to thank Dr. Mihai Anitescu of the mathematics department, University of Pittsburgh for his help with the regression techniques used in this work.

Literature Cited

(1) Burch, K. J.; Wakefield, D. K.; Whitehead, E. G., Jr. Boiling Point Models of Alkanes. *MATCH−Commun. Math. Comput. Chem.* **2003**, *47*, 25−52.
(2) Tierney, J.; Wender, I.; Zhang, S.; Zhang, Y. Anion-Modified Zirconia: Effect of Metal Promotion and Hydrogen Reduction on Hydroisomerization of *n*-Hexadecane and Fischer−Tropsch Waxes. *Fuel Process. Technol.* **2001**, *69*, 59−71.
(3) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim, Germany, 2000.
(4) Rouvray, D. H. The Role of the Topological Distance Matrix in Chemistry. In *Mathematics and Computational Concepts in Chemistry*; Trinajstić, N., Ed.; Ellis Horwood: Chichester, England, 1986; pp 295−306.
(5) Devore, J. L.; Peck, R. *Statistics: The Exploration and Analysis of Data*, 3rd ed.; Duxbury Press: Belmont, CA, 1997.
(6) Fourer, R.; Gay, D. M.; Kernighan, B. W. *AMPL: A Modeling Language for Mathematical Programming*; Boyd and Fraser Publishing Co.: Danvers, MA, 1993.
(7) Czyzyk, J.; Mesnier, M.; Moré, J. The NEOS Server. *IEEE J. Comput. Sci. Eng.* **1998**, *5*, 68−75.
(8) Gropp, W.; Moré, J. Optimization Environments and the NEOS Server. *Approximation Theory and Optimization*; Cambridge University Press: New York, 1997; pp 167−182.
(9) Dolan, E. The NEOS Server 4.0 Administrative Guide. Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, 2001.
(10) Fletcher, R.; Leyffer, S. Nonlinear Programming without a Penalty Function. *Math. Program., Ser. A* **2002**, *91*, 239−269.
(11) Beilstein Crossfire database, accessed March 2004.
(12) Thermodynamics Research Center, NIST Boulder Laboratories, M. Frenkel, Director, “Thermodynamics Source Database” in NIST Chemistry WebBook, NIST Standard Reference Database

Number 69, Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology, Gaithersburg MD, 20899; http://webbook.nist.gov, accessed March 2004.

(13) Baeyer, A. *Ber. Chem. Ges.* **1877**, *10*, 1286.

(14) Seifer, A. L.; Smolenskii, E. A. A Formal Method for Correlating the Properties of Alkanes: I. Normal Boiling Points of Monoalkylalkanes. *Russ. J. Phys. Chem.* **1963**, *37*, 1438−1441.

(15) Mosteller, F.; Tukey, J. W. *Data Analysis and Regression*; Addison-Wesley Publishing Co.: Philippines, 1977.

Received for review September 21, 2003. Accepted March 19, 2004.

JE034185B