

Molecular structure–property relationships for alkenes[☆]

Steven D. Nelson, Paul G. Seybold*

Department of Chemistry, Wright State University, Dayton, OH 45435, USA

Received 28 November 2000; received in revised form 7 February 2001; accepted 7 February 2001

Abstract

Structure–property relationships were obtained for 11 physical and chemical properties (boiling points (bp), melting points (mp), molar refractions (MR), molar volumes (MV), heats of combustion (HCKJ), molar heats of vaporization (HVMOL), flashpoints (FLASHK), second virial coefficients (VIRC2), critical temperatures (T_c), critical pressures (P_c), and viscosities (VISC)) for a data set consisting of 162 C4–C9 monoalkenes. Both molecular connectivity indices and ad hoc descriptors were tested as structural descriptors, and both produced high-quality regression equations for most of the properties. As was observed in an earlier study of alkanes [J. Am. Chem. Soc. 110 (1988) 4186], mp were not well described by either descriptor set. For most properties, the mass/bulk of the molecule was found to be the most important structural feature determining the property, suggesting that dispersion forces play a dominant role in determining those properties influenced by intermolecular interactions. The amount of branching in the molecule and the nature of the double bond environment were also found to be influential features. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Structure–property relationships; Alkenes; Principal component analysis (PCA)

1. Introduction

One of the most fundamental ideas of chemistry is that the physical and chemical properties of a substance are determined, somehow, by its molecular structure — the term ‘structure’ taken here in its broadest sense to include both geometric and electronic aspects. Historically, the difficulty associated with this simple proposition has centered on how to implement it in practice, since finding appropriate mathematical terms to describe ‘molecular structure’ has not been easy. Such mathematical terms (‘descriptors’) are necessary if quantitative relationships are to be constructed between the structure of a compound and its properties. In principle, a full quantum mechanical treatment for a bulk property such as the boiling point (bp) should yield a solution to this conundrum, but in practice a full-blown, accurate quantum mechanical treatment of a sizable collection of molecules is presently out of the question. One turns, therefore, to simpler approaches. One such approach is to attempt to identify suitable quantum chemical descriptors obtained from studies of isolated molecules. This approach, although quite successful in some cases, has in many other cases produced only mixed results. A simpler, alternative approach is to employ

topological descriptors, which fortunately often prove quite adequate to the task and, moreover, highly informative.

Several benefits can be derived from structure–property relationship studies. For example, unmeasured properties of related compounds can be estimated using the equations derived from a structure–property study. On a more fundamental level, a clearer understanding of the roles that specific structural features play in determining properties can be drawn from the equations. And once such an understanding is achieved, this information can be used to design hypothetical structures that might have desirable property values. Finally, the structure–property equations can serve as a useful check on the accuracy of property values already reported in the literature, some of which may be measured incorrectly or misreported.

In an earlier report [1] we applied topological descriptors in a study of eight physical properties of a set of normal and branched alkanes. In that study, it was found that of the topological descriptors examined, molecular connectivity indices [2,3] and ad hoc descriptors [1,4], were especially successful in yielding high-quality structure–property relationships. Good regression equations were obtained for seven of the physical properties of the alkanes (the melting points (mp), traditionally a subtle and difficult property to represent, were an exception). In this report, we employ these same descriptor types as structural measures for a study of the physical and chemical properties of a set of monoalkenes, where a new structural feature, the double bond, is introduced. Only

[☆] Taken in part from the Honors Thesis of SDN at Wright State University.

* Corresponding author. Tel.: +1-937-775-2407; fax: +1-937-775-2717.
E-mail address: paul.seybold@wright.edu (P.G. Seybold).

a few previous QSPR studies have been devoted to the properties of this class of compounds [5–8], and these have generally been limited to single properties.

2. Methods

The properties examined in this study were: bp, molar refractions (MR), molar volumes (MV) at 20°C, heats of combustion (HCKJ), mp, molar heats of vaporization (HVMOL) at 25°C, flashpoints (FLASHK), second virial coefficients (VIRC2) at 25°C, critical temperatures (T_c), critical pressures (P_c), and viscosities (VISC) at 20°C. The property values were taken from reference sources [9–14]. MV were calculated as M_w/d , where M_w is the molecular weight, and d is the density (g/cm^3) at 20°C. MR were calculated using the Lorentz–Lorenz expression

$$\text{MR} = \frac{(n_0^2 - 1)/(n_0^2 + 2)}{M_w/d}$$

where n_0 is the index of refraction.

The compounds examined and their shorthand abbreviations are listed in Appendix A. The property values for each compound are shown in Appendix B. The parameter values for each compound are shown in Appendix C.

Two types of parameters were used in this study: ad hoc descriptors [1,4] and molecular connectivity indices [2,3]. The ad hoc descriptors used were as follows: the number of carbon atoms (NC) represents the mass or bulk of the molecule. The square of the number of carbons (NCSC) and the square root (NCSR) account for non-linear aspects of the bulk dependence. The number of terminal methyl groups (TM) is a measure of the amount of branching in the molecule. The number of paths of length three carbon–carbon single bonds (P3S) is a steric index. The descriptor exterior double bonds (DBE) indicates whether the double bond is exterior or interior. The number of carbon atoms bonded to the double bond carbons (NCDB) measures the amount of crowding in the double bond environment. Several additional ad hoc indices were examined, but did not significantly improve the regression results.

Molecular connectivity indices were originally developed by Randić [15], Kier and Hall [2,3]. The form of these descriptors is

$${}^m\chi_t^v$$

where m is the order of the substructure, indicating the number of bonds, and t is the substructure type indicator. The substructure types included were path (p), cluster (c), path/cluster (pc), double bond (DB), and total (t) [1]. Only the valence (v) type indices, i.e. those which take explicit account of heteroatoms and multiple bonds, have been used in this study. The construction of these indices is described in detail in the alkane study [1] and elsewhere [2,3]. The molecular connectivity indices, used in this study, were calculated using Hall's MOLCONN2 software package [16]. Here, these indices are identified as

$$\text{XV}(t)(m)(f)$$

where f is the functional form (I = inverse, SQ = square, SR = square root). For example, XV1SR is the square root of the first order valence index, and XVPC4 is the fourth order pc index. Also included were the numbers of three-bond clusters (NXC3) and paths length of six bonds (NXP6). Altogether, 22 connectivity index terms were included in the initial screening.

Regression equations and other statistical measures were obtained using options in the SAS software package [17] on the Wright State University IBM 8083E computer. The final equations were selected on the basis of their standard errors (S.E.) and F -statistics. Principal component analysis (PCA) was used to determine the inherent dimensionality of the groups of properties. Orthogonal and oblique rotations did not significantly improve the results.

3. Results

3.1. Correlation analysis

The correlations among the properties examined are shown in Table 1. As can be seen, most of the properties

Table 1
Correlations among the properties examined

	Bp	MR	MV	HCKJ	HVMOL	FLASHK	VIRC2	T_c	P_c	VISC20	Mp
Bp	1.000										
MR	0.967	1.000									
MV	0.946	0.992	1.000								
HCKJ	0.970	0.996	0.992	1.000							
HVMOL	0.993	0.921	0.903	0.938	1.000						
FLASHK	0.933	0.905	0.844	0.897	0.878	1.000					
VIRC2	−0.960	−0.981	−0.969	−0.976	−0.974	−0.975	1.000				
T_c	0.996	0.965	0.979	0.989	0.998	0.997	−0.942	1.000			
P_c	−0.950	−0.940	−0.958	−0.958	−0.920	−0.825	0.903	−0.954	1.000		
VISC20	0.659	0.824	0.804	0.823	0.602	0.431	−0.981	0.925	−1.000	1.000	
Mp	0.664	0.640	0.614	0.525	0.482	0.518	−0.543	0.607	−0.467	0.844	1.000

Table 2

Multiple regression equation for the properties using ad hoc descriptors

Bp (°C)	$-275.58(\pm 3.51) + 139.31(\pm 1.40) \times \text{NCSR} + 7.49(\pm 0.36) \times \text{NCDB} - 7.64(\pm 0.26) \times \text{TM} - 5.93(\pm 0.62) \times \text{DBE} + 1.74(\pm 0.19) \times \text{P3S}$; $n = 162$, $r^2 = 0.9941$, S.E. = 2.28, $F = 5252$
MR (cm ³ /mol)	$20.65(\pm 2.20) + 7.44(\pm 0.35) \times \text{NC} - 14.53(\pm 1.76) \times \text{NCSR} - 0.0862(\pm 0.0130) \times \text{P3S}$; $n = 156$, $r^2 = 0.9976$, S.E. = 0.2224, $F = 21114$
MV (cm ³ /mol)	$26.27(\pm 0.66) + 17.33(\pm 0.12) \times \text{NC} - 1.99(\pm 0.11) \times \text{NCDB} - 1.27(\pm 0.07) \times \text{P3T} + 0.87(\pm 0.11) \times \text{TM}$; $n = 156$, $r^2 = 0.9971$, S.E. = 0.8547, $F = 13105$
HCKJ (kJ/mol)	$84.57(\pm 2.89) + 659.83(\pm 0.45) \times \text{NC} - 8.83(\pm 0.53) \times \text{TM}$; $n = 65$, $r^2 = 1.0000$, S.E. = 3.37, $F = 1000000$
Mp (°C)	$-169.58(\pm 7.06) + 0.9947(\pm 0.1508) \times \text{NCSO}$; $n = 57$, $r^2 = 0.4416$, S.E. = 17.43, $F = 44$
HVMOL (J/mol)	$-27094(\pm 1150) + 24224(\pm 490) \times \text{NCSR} + 1984(\pm 192) \times \text{NCDB} - 1972(\pm 134) \times \text{TM} - 1047(\pm 237) \times \text{DBE}$; $n = 34$, $r^2 = 0.9896$, S.E. = 503, $F = 690$
FLASHK (K)	$22.08(\pm 27.94) + 92.78(\pm 11.25) \times \text{NCSR}$; $n = 18$, $r^2 = 0.8096$, S.E. = 5.81, $F = 68$
VIRC2 (cm ³ /mol)	$3630(\pm 597) + 111.5(\pm 6.3) \times \text{NCSO} + 2427(\pm 343) \times \text{NCSR} - 68.5(\pm 16.7) \times \text{NCDB}$; $n = 14$, $r^2 = 0.9978$, S.E. = 50.8, $F = 1478$
T _c (°C)	$-63.73(\pm 11.54) + 64.51(\pm 3.90) \times \text{NC} \cdot 2.30(\pm 0.31) \times \text{NCSO} - 10.42(\pm 1.50) \times \text{DBE}$; $n = 13$, $r^2 = 0.9981$, S.E. = 2.45, $F = 1607$
P _c (MPa)	$7.585(\pm 0.319) - 1.789(\pm 0.139) \times \text{NCSR}$; $n = 12$, $r^2 = 0.9432$, S.E. = 0.128, $F = 166$
VISC20 (cP)	$0.08661(\pm 0.03813) + 0.00495(\pm 0.00106) \times \text{NCSO} + 0.00884(\pm 0.00167) \times \text{TMSQ}$; $n = 6$, $r^2 = 0.9725$, S.E. = 0.01843, $F = 53$

are highly correlated with one another, with the exception of mp, which is poorly correlated with the other properties. VISC is another possible exception. The remaining nine properties all have correlation coefficients greater than 0.82, and the subset of bp, MR, MV, and HV all have correlations greater than 0.90.

3.2. Multiple regression analysis

Table 2 shows the most successful ad hoc descriptor models with their coefficients of determination (r^2), S.E., and F -values, and Table 3 shows these measures for the most successful molecular connectivity index models. Most of the

Table 3

Multiple regression equation for the properties using connectivity indices

Bp (°C)	$15.04(\pm 16.81) + 74.70(\pm 6.08) \times \text{XV1} + 307.40(\pm 15.99) \times \text{XVDB} - 81.79(\pm 5.36) \times \text{XV0} + 28.49(\pm 1.69) \times \text{XVP3} + 51.20(\pm 3.29) \times \text{XVZ} + 157.89(\pm 16.49) \times \text{XV1SR}$; $n = 162$, $r^2 = 0.9945$, S.E. = 2.20, $F = 46.80$
MR (cm ³ /mol)	$1.91(\pm 0.57) + 8.87(\pm 0.14) \times \text{XV1} + 2.19(\pm 0.04) \times \text{XV2} + 14.00(\pm 1.26) \times \text{XVT}$; $n = 156$, $r^2 = 0.9973$, S.E. = 0.2374, $F = 18527$
MV (cm ³ /mol)	$7.79(\pm 0.89) + 14.19(\pm 0.36) \times \text{XV1} + 14.21(\pm 0.24) \times \text{XV0} + 59.32(\pm 1.58) \times \text{XVDB} + 0.14(\pm 0.26) \times \text{XVP3}$; $n = 156$, $r^2 = 0.9973$, S.E. = 0.8340, $F = 13766$
HCKJ (kJ/mol)	$807.59(\pm 10.34) + 1038.90(\pm 3.81) \times \text{XV1} + 308.07(12.51) \times \text{XV2} + 103.80(\pm 4.82) \times \text{XVP3} + 35.22(\pm 2.68) \times \text{XVDB1}$; $n = 65$, $r^2 = 0.9998$, S.E. = 6.44, $F = 90708$
Mp (°C)	$-414.50(\pm 50.25) + 77.34(\pm 12.22) \times \text{XV1} + 491.59(\pm 107.89) \times \text{XVT}$; $n = 57$, $r^2 = 0.5510$, S.E. = 15.77, $F = 33$
HVMOL (J/mol)	$-19604(\pm 11.63) + 27409(\pm 852) \times \text{XV1SR} + 2644(\pm 246) \times \text{XVDB1} + 1950(\pm 282) \times \text{NXP6}$; $n = 34$, $r^2 = 0.9884$, S.E. = 523, $F = 852$
FLASHK (K)	$-113.57(\pm 59.27) + 208.28(\pm 30.67) \times \text{XV1SR} + 21291(\pm 64.76) \times \text{XVT}$; $n = 18$, $r^2 = 0.8982$, S.E. = 44.5, $F = 64$
VIRC2 (cm ³ /mol)	$1310.1(\pm 145.7) + 405.7(\pm 11.5) \times \text{XV1SQ} + 3421(\pm 402) \times \text{XVT} - 336(\pm 46) \times \text{XVPC4}$; $n = 14$, $r^2 = 0.9982$, S.E. = 44.5, $F = 1901$
T _c (°C)	$-440.18(\pm 56.79) + 416.76(\pm 35.72) \times \text{XV1SR} + 308.33(\pm 50.29) \times \text{XVT} + 5.65(\pm 1.22) \times \text{XVT12}$; $n = 13$, $r^2 = 0.9974$, S.E. = 2.90, $F = 1151$
P _c (MPa)	$7.592(\pm 0.336) + 1.051(\pm 0.100) \times \text{XV0} + 0.141(\pm 0.025) \times \text{XV1SQ} + 4.135(\pm 0.795) \times \text{XVDB2}$; $n = 12$, $r^2 = 0.9878$, S.E. = 0.066, $F = 215$
VISC20 (cP)	$0.2466(\pm 0.0053) + 0.0454(\pm 0.0023) \times \text{NXC3} + 0.1034(\pm 0.0108) \times \text{NXP6}$; $n = 6$, $r^2 = 0.9028$, S.E. = 0.0094, $F = 208$

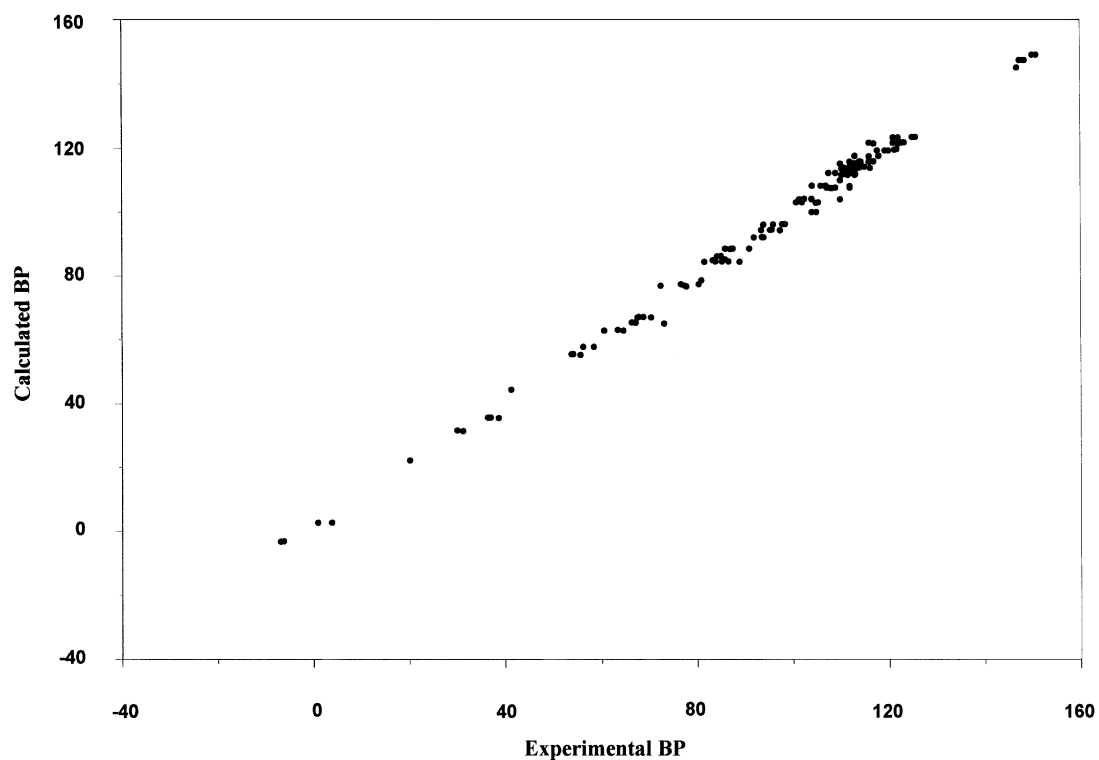


Fig. 1. Plot of calculated (ad hoc) boiling points vs. experimental boiling points.

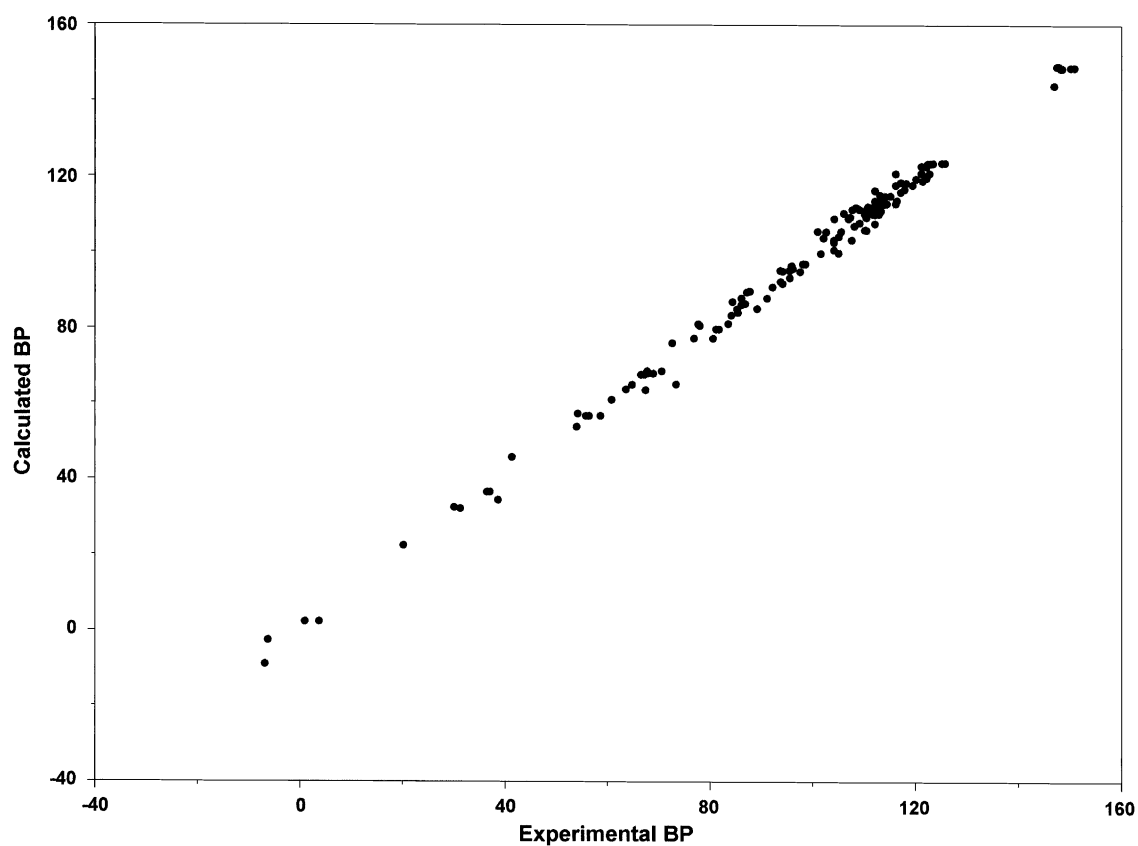


Fig. 2. Plot of calculated (molecular connectivity) boiling points vs. experimental boiling points.

Table 4

Results of the principal components analysis for the properties: eigenvalues with cumulative fractional variance reproduced

Factor	Three properties (156 compounds)	Four properties (57 compounds)	Nine properties (11 compounds)
1	2.937(0.979)	3.446(0.862)	7.632(0.848)
2	0.058(0.998)	0.501(0.987)	1.079(0.968)
3	0.005(1.000)	0.047(0.999)	0.148(0.984)
4		0.005(1.000)	0.075(0.993)
5			0.055(0.999)

properties are reasonably well modeled by these two types of descriptors (Figs. 1 and 2), with the exception of the mp, which were not well modeled by either descriptor set. The FLASHK are less well modeled than most other properties, but this is not entirely unexpected since this property is difficult to measure and some reported values may be relatively inaccurate.

3.3. Factor analysis

Three factor analysis (PCA) studies were performed, focusing on differing sets of properties. These studies ranged from a ‘compound intensive’ study of just three properties (bp, MR, and MV), for which a large number of experimental values (156) were available, to a ‘property intensive’ study including nine properties, for which complete, matching data were available for only 11 compounds.

The results from the first factor analysis, using only bp, MR and MV, are shown in Table 4. As can be seen above, the

three properties are highly correlated (Table 1) and a single factor dominates, accounting for 97.9% of the variation in these properties. As seen in Table 5, all three properties load strongly on the first factor. The data set is quite large ($n = 156$), and both low and high M_w compounds are well represented.

Addition of mp to the analysis considerably reduced the size of the data set, to 57 compounds, but the set still retained a good sampling of both high and low M_w compounds and remained large enough for reliable statistical calculations. The results are shown in Tables 4 and 5. As was seen in Table 1, the mp were not highly correlated with the other properties. The property set now appears to be represented by two factors (Table 4). The first factor accounts for 86.2% of the variance. Including the second factor, related to mp, brings this total to 98.7%. Reminiscent of the alkane study [1], bp, MR, and MV load heavily on the first factor, whereas mp loads strongly on the second factor.

Addition of the heat of combustion, the VIRC2, and the critical properties to the analysis markedly decreased the common compound population, and the population became strongly skewed toward the lower M_w monoalkenes. The molar heat of vaporization was excluded to increase the number of compounds from 8 to 11. The results are shown in Tables 4 and 5. This analysis is included only for the sake

Table 5

PCA factor loadings for the physical properties

Property	Factor		
	1	2	3
Three properties			
Bp	0.981	0.191	0.014
MR	0.997	−0.054	−0.055
MV	0.990	−0.135	0.042
Four properties			
Bp	0.976	−0.126	−0.177
MR	0.982	−0.174	0.056
MV	0.972	−0.204	0.109
Mp	0.766	0.643	0.015
Nine properties			
Bp	0.968	0.163	0.173
MR	0.966	−0.138	−0.110
MV	0.982	−0.093	−0.096
Mp	−0.062	0.993	−0.100
HCKJ	0.998	−0.004	0.017
VIRC2	−0.976	−0.143	−0.067
T_c	0.971	0.106	0.187
P_c	−0.954	0.078	0.212
VISC20	0.996	−0.037	−0.036

Table 6

Modeling of factor scores using structural descriptors

Parameters	r^2	S.E.	F
Factor 1			
NC	0.9947	0.0726	29223
NC, TMSQ	0.9966	0.0585	22553
NC, TMSQ, DBE	0.9986	0.0382	35347
XV1SR	0.9684	0.1783	4720
XV1SR, XV0	0.9936	0.0803	11929
XV1SR, XV0, NXP6	0.9960	0.0638	12630
Factor 2			
NCDB	0.0757	0.9701	5
NCDB, NCSR	0.1236	0.9533	4
NCDB, NCSR, P3S	0.2225	0.9063	5
XVTSQ	0.1564	0.9268	10
XVTSQ, XV1SR	0.3121	0.8446	12
XVTSQ, XV1SR, XV1	0.3610	0.8217	10

of completeness, since its statistical significance is questionable. However, it is interesting to note that the first factor continued to dominate for most properties, accounting for 83.7% of the variance. Addition of the second factor brought this up to 97.0%, and the addition of a third factor brought this up to 99.8%.

In order to clarify, what the abstract factors represent, the factor scores were modeled using the ad hoc and connectivity descriptors. Factor 1 from the first analysis ($n = 156$) was employed to obtain a model for this factor, and Factor 2 was taken from the second analysis ($n = 57$) for the same purpose. The results are shown in Table 6. Factor 1 is clearly a bulk factor, depending on the leading descriptors, whereas Factor 2 was not well modeled by either of the descriptor sets.

4. Discussion

The regression equations presented in Tables 2 and 3 are generally of high-quality for properties other than the mp. Therefore, property values estimated on the basis of these equations, with the exception of mp, should be sufficiently accurate for many practical purposes.

The ad hoc descriptor equations in Table 2 show the relative influences of molecular mass/bulk (NC, NCSR and NCSQ), branching (TM), steric factors (P3S), and the double bond environment (DBE, NCDB) in determining the properties studied. As can be seen from the Table, the molecular mass/bulk clearly exerts the dominant influence for properties other than mp, suggesting that dispersion forces play a dominant role for those properties which depend on intermolecular forces. A similar conclusion was reached in the earlier alkane study [1,4]. This is a reasonable conclusion in the present case for bp, HVMOL, VIRC2, T_c , P_c , and VISC20. For MV the ‘mass/bulk’ dependence can be attributed directly to the larger volume of compounds with higher NC, as later modified by small corrections for branching and steric influences. Likewise, MR depends largely on the higher number of electrons in the larger compounds. For the two strictly “chemical” properties, the HCKJ and the flashpoints (FLASHK), the dependence on the mass/bulk dimension is more accurately attributed to the larger number of reacting bonds in the larger, higher NC compounds.

Branching, steric factors, and the double bond environment exert smaller influences on the properties, as demonstrated by the coefficients in the ad hoc regression equations. Molecular branching, represented by the number of TM sequesters interior parts of these compounds and reduces the extent of contact between neighboring molecules. The latter effect is reflected in the positive influence of TM on the MV. Because dispersion forces are strongly dependent on distance — the interaction energies fall as $1/r^6$, where r is the separation — a decrease in the amount of close contact decreases the cohesive forces

experienced by the compounds. Therefore, bp and HVKJ decrease as TM increases. Steric crowding, included here by the parameter P3S, leads to a small reduction in the MV, as reflected in the relatively small contributions to MV and MR from P3S. Whether the double bond is exterior or interior and the NSDB influence the accessibility of the double bond to its environment. When the double bond is exterior the bp, for example, is reduced on average by 6°C. Thus, DBE appear to exert a negative influence on the intermolecular forces. The NCDB exerts a positive influence on the intermolecular forces, possibly by effectively screening the effects of the double bond. Because of the above influences, those properties that depend on the strength of intermolecular forces, such as bp, HV, and T_c , show positive dependences on NC, NCSQ, NCSR, P3S, and NCDB, and negative dependences on TM and DBE.

The failure of both parameter sets to model the mp is not surprising, since this property was also not well modeled by these same topological parameters in our earlier study of the alkanes [1,4]. This illustrates the greater subtlety of the melting transition as compared to the boiling and critical transitions. The latter transitions involve a direct dependence on the operative intermolecular forces, and so directly reflect the strengths of these forces. The melting transition, in contrast, maintains a condensed phase and involves a partial disruption of intermolecular orientations. Melting, thus, depends on geometric and other factors that are not well addressed by the present topological parameters. This dependence on shape and entropic factors, as opposed to a simple intermolecular force dependence, is reflected in the mp strong loading on the second factor, rather than the first (mass/bulk related) factor, in the factor analysis. Dearden has recently given a comprehensive review of mp predictions [18].

Factor analysis for the monoalkenes shows that a single factor dominates for most properties. In the worst case examined (Analysis 3), Factor 1 still accounted for nearly 84% of the variance. Addition of a second factor in this case raised the variance accounted to 97%. Factor 1 is clearly related to the mass/bulk of the molecules. As can be seen, in Table 6 both the ad hoc and connectivity index mass/bulk dependent descriptors give a good account of this factor. Factor 2, however, is a different story. From the results in Table 5, it is obvious that Factor 2 is related to the mp, but this factor was not well modeled by either the ad hoc or connectivity index descriptors employed in this study. Because of this, it is difficult to determine exactly what structural features are crucial for this dimension.

Acknowledgements

We thank Stephen Peterangelo for checking some of the regressions and preparing the figures.

Appendix A. Compounds

Observation no.	Name	Shorthand name
1	1-Butene	1N4
2	<i>Cis</i> -2-butene	2C4
3	<i>Trans</i> -2-butene	2T4
4	2-Methyl propene	2M1N3
5	1-Pentene	1N5
6	<i>Cis</i> -2-pentene	2C5
7	<i>Trans</i> -2-pentene	2T5
8	2-Methyl-1-butene	2M1N4
9	3-Methyl-1-butene	3M1N4
10	2-Methyl-2-butene	2M2N4
11	1-Hexene	1N6
12	<i>Cis</i> -2-hexene	2C6
13	<i>Trans</i> -2-hexene	2T6
14	<i>Cis</i> -3-hexene	3C6
15	<i>Trans</i> -3-hexene	3T6
16	2-Methyl-1-pentene	2M1N5
17	3-Methyl-1-pentene	3M1N5
18	4-Methyl-1-pentene	4M1N5
19	2-Methyl-2-pentene	2M2N5
20	3-Methyl- <i>cis</i> -2-pentene	3M2C5
21	3-Methyl- <i>trans</i> -2-pentene	3M2T5
22	4-Methyl- <i>cis</i> -2-pentene	4M2C5
23	4-Methyl- <i>trans</i> -2-pentene	4M2T5
24	2-Ethyl-1-butene	2E1N4
25	2,3-Dimethyl-1-butene	23M1N4
26	3,3-Dimethyl-1-butene	33M1N4
27	2,3-Dimethyl-2-butene	23M2N4
28	1-Heptene	1N7
29	<i>Cis</i> -2-heptene	2C7
30	<i>Trans</i> -2-heptene	2T7
31	<i>Cis</i> -3-heptene	3C7
32	<i>Trans</i> -3-heptene	3T7
33	2-Methyl-1-hexene	2M1N6
34	3-Methyl-1-hexene	3M1N6
35	4-Methyl-1-hexene	4M1N6
36	5-Methyl-1-hexene	5M1N6
37	2-Methyl-2-hexene	2M2N6
38	3-Methyl- <i>cis</i> -2-hexene	3M2C6
39	3-Methyl- <i>trans</i> -2-hexene	3M2T6
40	4-Methyl- <i>cis</i> -2-hexene	4M2C6
41	4-Methyl- <i>trans</i> -2-hexene	4M2T6
42	5-Methyl- <i>cis</i> -2-hexene	5M2C6
43	5-Methyl- <i>trans</i> -2-hexene	5M2T6
44	2-Methyl- <i>cis</i> -3-hexene	2M3C6
45	2-Methyl- <i>trans</i> -3-hexene	2M3T6
46	3-Methyl- <i>cis</i> -3-hexene	3M3C6
47	3-Methyl- <i>trans</i> -3-hexene	3M3T6
48	2-Ethyl-1-pentene	2E1N5
49	3-Ethyl-1-pentene	3E1N5
50	2,3-Dimethyl-1-pentene	23M1N5
51	2,4-Dimethyl-1-pentene	24M1N5
52	3,3-Dimethyl-1-pentene	33M1N5
53	3,4-Dimethyl-1-pentene	34M1N5
54	4,4-Dimethyl-1-pentene	44M1N5

Appendix A (Continued)

Observation no.	Name	Shorthand name
55	3-Ethyl-2-pentene	3E2N5
56	2,3-Dimethyl-2-pentene	23M2N5
57	2,4-Dimethyl-2-pentene	24M2N5
58	3,4-Dimethyl- <i>cis</i> -2-pentene	34M2C5
59	3,4-Dimethyl- <i>trans</i> -2-pentene	34M2T5
60	4,4-Dimethyl- <i>cis</i> -2-pentene	44M2C5
61	4, 4-Dimethyl- <i>trans</i> -2-pentene	44M2T5
62	2-Ethyl-3-methyl-1-butene	2E3M1N4
63	2,3,3-Trimethyl-1-butene	233M1N4
64	1-Octene	1N8
65	<i>Cis</i> -2-octene	2C8
66	<i>Trans</i> -2-octene	2T8
67	<i>Cis</i> -3-octene	3C8
68	<i>Trans</i> -3-octene	3T8
69	<i>Cis</i> -4-octene	4C8
70	<i>Trans</i> -4-octene	4T8
71	2-Methyl-1-heptene	2M1N7
72	3-Methyl-1-heptene	3M1N7
73	4-Methyl-1-heptene	4M1N7
74	5-Methyl-1-heptene	5M1N7
75	6-Methyl-1-heptene	6M1N7
76	2-Methyl-2-heptene	2M2N7
77	3-Methyl- <i>cis</i> -2-heptene	3M2C7
78	3-Methyl- <i>trans</i> -2-heptene	3M2T7
79	4-Methyl- <i>cis</i> -2-heptene	4M2C7
80	4-Methyl- <i>trans</i> -2-heptene	4M2T7
81	5-Methyl- <i>cis</i> -2-heptene	5M2C7
82	5-Methyl- <i>trans</i> -2-heptene	5M2T7
83	6-Methyl- <i>cis</i> -2-heptene	6M2C7
84	6-Methyl- <i>trans</i> -2-heptene	6M2T7
85	2-Methyl- <i>cis</i> -3-heptene	2M3C7
86	2-Methyl- <i>trans</i> -3-heptene	2M3T7
87	3-Methyl- <i>cis</i> -3-heptene	3M3C7
88	3-Methyl- <i>trans</i> -3-heptene	3M3T7
89	4-Methyl- <i>cis</i> -3-heptene	4M3C7
90	4-Methyl- <i>trans</i> -3-heptene	4M3T7
91	5-Methyl- <i>cis</i> -3-heptene	5M3C7
92	5-Methyl- <i>trans</i> -3-heptene	5M3T7
93	6-Methyl- <i>cis</i> -3-heptene	6M3C7
94	6-Methyl- <i>trans</i> -3-heptene	6M3T7
95	2-Ethyl-1-hexene	2E1N6
96	3-Ethyl-1-hexene	3E1N6
97	4-Ethyl-1-hexene	4E1N6
98	2,3-Dimethyl-1-hexene	23M1N6
99	2,4-Dimethyl-1-hexene	24M1N6
100	2,5-Dimethyl-1-hexene	25M1N6
101	3,3-Dimethyl-1-hexene	33M1N6
102	3,4-Dimethyl-1-hexene	34M1N6
103	3,5-Dimethyl-1-hexene	35M1N6
104	4,4-Dimethyl-1-hexene	44M1N6
105	4,5-Dimethyl-1-hexene	45M1N6
106	5,5-Dimethyl-1-hexene	55M1N6
107	3-Ethyl- <i>cis</i> -2-hexene	3E2C6
108	3-Ethyl- <i>trans</i> -2-hexene	3E2T6

Appendix A (Continued)

Observation no.	Name	Shorthand name
109	4-Ethyl- <i>cis</i> -2-hexene	4E2C6
110	4-Ethyl- <i>trans</i> -2-hexene	4E2T6
111	2,3-Dimethyl-2-hexene	23M2N6
112	2,4-Dimethyl-2-hexene	24M2N6
113	2,5-Dimethyl-2-hexene	25M2N6
114	3,4-Dimethyl- <i>cis</i> -2-hexene	34M2C6
115	3,4-Dimethyl- <i>trans</i> -2-hexene	34M2T6
116	3,5-Dimethyl- <i>cis</i> -2-hexene	35M2C6
117	3,5-Dimethyl- <i>trans</i> -2-hexene	35M2T6
118	4,4-Dimethyl- <i>cis</i> -2-hexene	44M2C6
119	4,4-Dimethyl- <i>trans</i> -2-hexene	44M2T6
120	4,5-Dimethyl- <i>cis</i> -2-hexene	45M2C6
121	4,5-Dimethyl- <i>trans</i> -2-hexene	45M2T6
122	5,5-Dimethyl- <i>cis</i> -2-hexene	55M2C6
123	5,5-Dimethyl- <i>trans</i> -2-hexene	55M2T6
124	3-Ethyl-3-hexene	3E3N6
125	2,2-Dimethyl- <i>cis</i> -3-hexene	22M3C6
126	2,2-Dimethyl- <i>trans</i> -3-hexene	22M3T6
127	2,3-Dimethyl- <i>cis</i> -3-hexene	23M3C6
128	2,3-Dimethyl- <i>trans</i> -3-hexene	23M3T6
129	2,4-Dimethyl- <i>cis</i> -3-hexene	24M3C6
130	2,4-Dimethyl- <i>trans</i> -3-hexene	24M3T6
131	2,5-Dimethyl- <i>cis</i> -3-hexene	25M3C6
132	2,5-Dimethyl- <i>trans</i> -3-hexene	25M3T6
133	3,4-Dimethyl- <i>cis</i> -3-hexene	34M3C6
134	3,4-Dimethyl- <i>trans</i> -3-hexene	34M3T6
135	2- <i>n</i> -Propyl-1-pentene	2NP1N5
136	2-Isopropyl-1-pentene	2IP1N5
137	2-Ethyl-3-methyl-1-pentene	2E3M1N5
138	2-Ethyl-4-methyl-1-pentene	2E4M1N5
139	3-Ethyl-2-methyl-1-pentene	3E2M1N5
140	3-Ethyl-3-methyl-1-pentene	3E3M1N5
141	3-Ethyl-4-methyl-1-pentene	3E4M1N5
142	2,3,3-Trimethyl-1-pentene	233M1N5
143	2,3,4-trimethyl-1-pentene	234M1N5
144	2,4,4-Trimethyl-1-pentene	244M1N5
145	3,3,4-Trimethyl-1-pentene	334M1N5
146	3,4,4-Trimethyl-1-pentene	344M1N5
147	3-Ethyl-2-methyl-2-pentene	3E2M2N5
148	3-Ethyl-4-methyl- <i>cis</i> -2-pentene	3E4M2C5
149	3-Ethyl-4-methyl- <i>trans</i> -2-pentene	3E4M2T5
150	2,3,4-Trimethyl-2-pentene	234M2N5
151	2,4,4-Trimethyl-2-pentene	244M2N5
152	3,4,4-Trimethyl- <i>cis</i> -2-pentene	344M2C5
153	3,4,4-Trimethyl- <i>trans</i> -2-pentene	344M2T5
154	2-Isopropyl-3-methyl-1-butene	2IP3M1N4
155	2-Ethyl-3,3-dimethyl-1-butene	2E33M1N4
156	1-Nonene	1N9
157	<i>Cis</i> -2-nonene	2C9
158	<i>Trans</i> -2-nonene	2T9
159	<i>Cis</i> -3-nonene	3C9
160	<i>Trans</i> -3-nonene	3T9
161	<i>Cis</i> -4-nonene	4C9
162	<i>Trans</i> -4-nonene	4T9

Appendix B. Property values

Observation no.	Shorthand name	bp	mp	MR	MV	HCKJ	HVMOL	FLASHK	VIRC2	T_c	P_c	VISC20
1	1N4	−6.260	−185.35	22.6628	94.277	2716.8	20204	−	−649.7	146.450	4.023	−
2	2C4	3.720	−138.91	20.5879	90.301	2710.0	21963	−	−712.9	162.430	4.205	−
3	2T4	0.880	−105.55	20.7281	92.857	2706.4	21483	−	−700.1	155.480	4.104	−
4	2M1N3	−6.9000	−140.35	22.6311	94.419	2700.2	20101	−	−647.7	144.760	4.000	−
5	1N5	29.968	−165.22	24.8543	109.43	3375.4	25501	−	−1092.7	191.630	3.526	0.2390
6	2C5	36.942	−151.39	24.9524	106.971	3370.0	26885	−	−1150.3	201.790	3.695	−
7	2T5	36.353	−140.24	25.0200	108.192	3365.4	−	228.15	−1140.0	201.850	3.648	−
8	2M1N4	31.163	−137.56	24.8474	107.826	3361.6	−	−	−1111.0	191.850	3.505	−
9	3M1N4	20.061	−168.49	24.9416	111.814	3368.9	−	−	−973.9	191.850	3.435	−
10	2M2N4	38.568	−133.77	24.9521	105.889	3355.7	27090	228.15	−1264.0	197.240	3.380	−
11	1N6	63.485	−139.82	29.4915	125.014	4034.1	30587	247.15	−1729.0	230.800	3.140	0.2600
12	2C6	68.840	−141.13	29.5309	122.462	4023.8	33744	−	−	−	−	−
13	2T6	67.870	−132.97	29.6684	124.133	4022.2	32136	253.15	−	−	−	−
14	3C6	66.440	−137.82	29.6605	123.777	4028.5	31724	−	−	−	−	−
15	3T6	67.080	−113.43	29.7493	124.287	4021.7	32072	261.15	−	−	−	−
16	2M1N6	60.700	−135.72	29.4759	123.782	4016.8	31042	247.15	−	−	−	−
17	3M1N5	54.140	−153.00	29.4946	126.086	4026.1	29262	245.15	−	−	−	−
18	4M1N5	53.880	−153.63	29.5548	126.798	4024.9	29376	242.15	−	−	−	0.2883
19	2M2N5	67.290	−135.07	29.7378	122.587	4007.3	32088	250.15	−	−	−	−
20	3M2C5	70.450	−138.44	29.5488	122.213	4013.9	31812	−	−	−	−	−
21	4M2C5	67.630	−134.84	29.5498	121.437	4013.0	32536	−	−	−	−	−
22	4M2C5	56.300	−134.43	29.6707	125.760	4018.7	30096	245.15	−	−	−	−
23	4M2T5	58.550	−140.81	29.7533	125.865	4014.7	30550	−	−	−	−	−
24	2E1N4	64.660	−131.53	29.3700	122.040	4020.2	31614	−	−	−	−	0.2885
25	23M1N4	55.670	−157.27	29.4338	124.106	4011.2	29800	255.15	−	−	−	−
26	33M1N4	41.240	−115.20	29.5814	128.856	4015.4	27449	245.15	−	−	−	−
27	23M2N4	73.210	−74.28	29.5944	118.848	4007.3	32476	257.15	−1929.0	−	−	−
28	1N7	93.643	−119.03	34.1347	140.868	4692.6	35484	272.15	−2810.0	264.145	−	0.3500
29	2C7	98.500	−	34.1730	138.852	4686.5	−	−	−	−	−	−
30	2T7	97.950	−109.48	34.2812	140.020	4682.3	−	272.15	−	−	−	−
31	3C7	95.750	−	34.3076	139.701	4686.5	−	−	−	−	−	−
32	3T7	95.670	−136.63	34.4259	140.642	4682.3	−	−	−	−	−	−
33	2M1N6	92.000	−102.84	34.1237	139.681	4678.0	−	267.15	−	−	−	−
34	3M1N6	84.000	−	34.1588	142.005	4688.7	−	267.15	−	−	−	−
35	4M1N6	86.730	−141.45	34.0754	140.561	4688.7	−	−	−	−	−	−
36	5M1N6	85.310	−	34.1444	141.882	4686.0	−	−	−	−	−	−
37	2M2N6	95.410	−130.35	34.3978	138.656	4672.2	−	−	−	−	−	−
38	3M2C6	94.000	−	34.1783	137.183	4674.9	−	−	−	−	−	−

Appendix B (Continued)

Observation no.	Shorthand name	bp	mp	MR	MV	HCKJ	HVMOL	FLASHK	VIRC2	T_c	P_c	VISC20
39	3M2T6	94.000	–	34.1889	137.452	4674.9	–	–	–	–	–	–
40	4N2C6	87.370	–	34.2220	140.360	4682.0	–	–	–	–	–	–
41	4M2T6	87.600	–126.50	34.3471	140.904	4677.8	–	–	–	–	–	–
42	5M2C6	91.000	–	34.2048	139.860	4679.3	–	–	–	–	–	–
43	5M2T6	86.000	–	34.4113	141.759	4675.1	–	–	–	–	–	–
44	2M3C6	86.000	–	34.3721	141.473	4679.3	–	–	–	–	–	–
45	2M3T6	86.000	–	34.5228	142.375	4675.1	–	–	–	–	–	–
46	3M3C6	95.350	–	34.3169	137.728	4674.9	–	–	–	–	–	–
47	3M3T6	93.550	–	34.3448	138.353	4674.9	–	–	–	–	–	–
48	2E1N5	94.000	–	33.9890	138.675	4680.7	–	–	–	–	–	–
49	3E1N5	85.130	–127.40	34.0617	141.066	4691.3	–	–	–	–	–	–
50	23M1N5	84.260	–134.80	33.9975	139.226	4673.8	–	–	–	–	–	–
51	24M1N5	81.640	–123.80	34.1793	141.452	4670.9	33344	–	–	–	–	–
52	33M1N5	77.540	–134.30	34.0085	140.783	4679.4	–	–	–	–	–	–
53	34M1N5	81.000	–	34.0542	140.722	4681.7	–	–	–	–	–	–
54	44M1N5	72.490	–136.60	34.2349	143.859	4674.7	31598	–	–	–	–	–
55	3E2N5	96.010	–	34.1144	136.288	4677.5	–	–	–	–	–	–
56	23M2N5	97.46	–118.30	34.2218	134.921	4667.1	–	–	–	–	–	–
57	24M2N5	83.44	–	34.5314	141.285	4665.1	34418	–	–	–	–	–
58	34M2C5	87.00	–	34.1229	137.606	4667.9	–	–	–	–	–	–
59	34M2T5	87.00	–	34.1452	136.992	4667.9	–	–	–	–	–	–
60	44M2C5	80.42	–135.46	34.2255	140.374	4667.9	32973	–	–	–	–	–
61	44M2T5	76.75	–115.23	34.4054	142.530	4663.9	33187	–	–	–	–	–
62	2E3M1N4	89.00	–	33.9649	138.523	4673.7	34634	–	–	–	–	–
63	233M1N4	77.87	–109.85	33.9878	139.332	4668.2	32485	256.15	–	–	–	0.475
64	1N8	121.26	–101.74	38.7781	156.952	5351.1	41224	–	–3948	–	–	–
65	2C8	125.64	–100.20	38.7944	154.919	–	–	–	–	–	–	–
66	2T8	125.00	–87.70	38.8827	155.866	–	–	–	–	–	–	–
67	3C8	122.90	–	38.8482	155.828	–	–	–	–	–	–	–
68	3T8	123.30	–110.00	39.0883	156.890	–	–	–	–	–	–	–
69	4C8	122.54	–118.70	38.9447	155.585	–	–	–	–	–	–	–
70	4T8	122.25	–93.81	39.0817	157.132	–	–	–	–	–	–	–
71	2M1N7	119.30	–90.00	38.7759	155.736	–	–	–	–	–	–	–
72	3M1N7	111.00	–	38.7649	157.817	–	–	–	–	–	–	–
73	4M1N7	112.80	–	38.7739	156.497	–	–	–	–	–	–	–
74	5M1N7	113.30	–	38.7564	156.628	–	–	–	–	–	–	–
75	6M1N7	113.20	–	38.7944	157.596	–	–	–	–	–	–	–
76	2M2N7	122.60	–	38.9693	154.962	–	–	–	–	–	–	–
77	3M2C7	122.00	–	38.8702	153.920	–	–	–	–	–	–	–

78	3M2T7	122.00	–	38.8702	153.920	–	–	–	–	–	–	–
79	4M2C7	114.00	–	38.8281	156.715	–	–	–	–	–	–	–
80	4M2T7	114.00	–	38.8281	156.715	–	–	–	–	–	–	–
81	5M2C7	118.00	–	38.7819	155.198	–	–	–	–	–	–	–
82	5M2T7	118.00	–	38.7819	155.198	–	–	–	–	–	–	–
83	6M2C7	117.00	–	38.8860	156.279	–	–	–	–	–	–	–
84	6M2T7	117.00	–	38.8860	156.279	–	–	–	–	–	–	–
85	2M3C7	112.00	–	39.1241	158.935	–	–	–	–	–	–	–
86	2M3T7	112.00	–	39.1241	158.935	–	–	–	–	–	–	–
87	3M3C7	121.00	–	38.8421	154.132	–	–	–	–	–	–	–
88	3M3T7	121.00	–	38.8421	154.132	–	–	–	–	–	–	–
89	4M3C7	122.00	–	38.9209	154.770	–	–	–	–	–	–	–
90	4M3T7	122.00	–	38.9209	154.770	–	–	–	–	–	–	–
91	5M3C7	112.00	–	38.9914	157.374	–	–	–	–	–	–	–
92	5M3T7	112.00	–	38.9914	157.374	–	–	–	–	–	–	–
93	6M3C7	115.00	–	38.9914	157.374	–	–	–	–	–	–	–
94	6M3T7	115.00	–	38.9914	157.374	–	–	–	–	–	–	–
95	2E1N6	120.00	–	38.7076	154.344	–	–	–	–	–	–	–
96	3E1N6	110.30	–	38.6317	156.934	–	–	–	–	–	–	–
97	4E1N6	113.00	–	38.4575	154.556	–	–	–	–	–	–	–
98	23M1N6	110.50	–	38.7028	155.542	–	–	–	–	–	–	–
99	24M1N6	111.20	–	38.6952	155.844	–	–	–	–	–	–	–
100	25M1N6	111.60	–	38.8047	156.453	–	–	–	–	–	–	–
101	33M1N6	104.00	–	38.6858	157.154	–	–	–	–	–	–	–
102	34M1N6	112.00	–	38.6461	154.983	–	–	–	–	–	–	–
103	35M1N6	104.00	–	38.7599	158.486	–	–	–	–	–	–	–
104	44M1N6	107.20	–	38.6397	155.888	–	–	–	–	–	–	–
105	45M1N6	109.00	–	38.5155	154.132	–	–	–	–	–	–	–
106	55M1N6	102.50	–	38.7813	158.262	–	–	–	–	–	–	–
107	3E2C6	121.00	–	38.8499	152.250	–	–	–	–	–	–	–
108	3E2T6	121.00	–	38.8499	152.250	–	–	–	–	–	–	–
109	4E2C6	113.00	–	38.5106	154.770	–	–	–	–	–	–	–
110	4E2T6	113.00	–	38.5106	154.770	–	–	–	–	–	–	–
111	23M2N6	121.770	–115.10	38.8737	151.469	–	–	–	–	–	–	–
112	24M2N6	110.600	–	38.6916	155.564	–	–	–	–	–	–	–
113	25M2N6	112.200	–	38.9435	155.844	–	–	–	–	–	–	–
114	34M2C6	116.000	–	38.3677	152.250	–	–	–	–	–	–	–
115	34M2T6	116.000	–	38.3677	152.250	–	–	–	–	–	–	–
116	35M2C6	112.000	–	38.8390	154.770	–	–	–	–	–	–	–
117	35M2T6	112.000	–	38.8390	154.770	–	–	–	–	–	–	–
118	44M2C6	106.000	–	38.7531	155.413	–	–	–	–	–	–	–
119	44M2T6	106.000	–	38.7531	155.413	–	–	–	–	–	–	–
120	45M2C6	110.000	–	38.5928	154.770	–	–	–	–	–	–	–
121	45M2T6	110.000	–	38.5928	154.770	–	–	–	–	–	–	–
122	55M2C6	106.900	–	38.8875	156.518	–	–	–	–	–	–	–

[illegible]

Appendix C. Parameter values

Observation no.	Shorthand name	NC	TM	P3S	P3T	NCDB	DBE	XV0	XV1	XV2	XVP3	XVDB	XVT	XVPC4	NXP6	NXC3
1	1N4	4	1	0	1	1	1	2.9916	1.5236	0.6969	0.2887	0.408248	0.288675	0.0000	0	0
2	2C4	4	2	0	1	2	0	3.1547	1.4880	0.6667	0.3333	0.333333	0.333333	0.0000	0	0
3	2T4	4	2	0	1	2	0	3.1547	1.4880	0.6667	0.3333	0.333333	0.333333	0.0000	0	0
4	2M1N3	4	2	0	0	2	1	3.2071	1.3536	1.2071	0.0000	0.353553	0.353553	0.0000	0	1
5	1N5	5	1	1	2	1	1	3.6987	2.0236	1.0774	0.4928	0.408248	0.204124	0.0000	0	0
6	2C5	5	2	0	2	2	0	3.8618	2.0260	0.9773	0.4714	0.333333	0.235702	0.0000	0	0
7	2T5	5	2	0	2	2	0	3.8618	2.0260	0.9773	0.4714	0.333333	0.235702	0.0000	0	0
8	2M1N4	5	2	1	2	2	1	3.9142	1.9142	1.3107	0.6036	0.353553	0.250000	0.2500	0	1
9	3M1N4	5	2	0	2	1	1	3.8618	1.8963	1.4797	0.4714	0.408248	0.235702	0.2357	0	1
10	2M2N4	5	3	0	2	3	0	4.0774	1.8660	1.3660	0.5774	0.288675	0.288675	0.2887	0	1
11	1N6	6	1	2	3	1	1	4.4058	2.5236	1.4309	0.7618	0.408248	0.144338	0.0000	0	0
12	2C6	6	2	1	3	2	0	4.5689	2.5260	1.3577	0.6910	0.333333	0.166667	0.0000	0	0
13	2T6	6	2	1	3	2	0	4.5689	2.5260	1.3577	0.6910	0.333333	0.166667	0.0000	0	0
14	3C6	6	2	0	3	2	0	4.5689	2.5640	1.2879	0.6381	0.333333	0.166667	0.0000	0	0
15	3T6	6	2	0	3	2	0	4.5689	2.5640	1.2879	0.6381	0.333333	0.166667	0.0000	0	0
16	2M1N5	6	2	2	3	2	1	4.6213	2.4142	1.7071	0.6768	0.353553	0.176777	0.1768	0	1
17	3M1N5	6	2	2	3	1	1	4.5689	2.4343	1.6212	1.0463	0.408248	0.166667	0.4024	0	1
18	4M1N5	6	2	2	3	1	1	4.5689	2.3794	1.9182	0.6381	0.408248	0.166667	0.2357	0	1
19	2M2N5	6	3	0	3	3	0	4.7845	2.4040	1.6897	0.6124	0.288675	0.204124	0.2041	0	1
20	3M2C5	6	3	1	4	3	0	4.7845	2.4267	1.4886	1.0505	0.288675	0.204124	0.4082	0	1
21	3M2T5	6	3	1	4	3	0	4.7845	2.4267	1.4886	1.0505	0.288675	0.204124	0.4082	0	1
22	4M2C5	6	3	0	3	2	0	4.7321	2.3987	1.7698	0.5774	0.333333	0.192450	0.1925	0	1
23	4M2T5	6	3	0	3	2	0	4.7321	2.3987	1.7698	0.5774	0.333333	0.192450	0.1925	0	1
24	2E1N4	6	2	2	4	2	1	4.6213	2.4749	1.4571	1.0000	0.353553	0.176777	0.3536	0	1
25	23M1N4	6	3	2	4	2	1	4.7845	2.2969	2.0011	0.9856	0.353553	0.204124	0.9010	0	2
26	33M1N4	6	3	0	3	1	1	4.7845	2.1969	2.5701	0.6124	0.408248	0.204124	0.4082	0	4
27	23M2N4	6	4	0	4	4	0	5.0000	2.2500	2.0000	1.0000	0.250000	0.250000	1.0000	0	2
28	1N7	7	1	3	4	1	1	5.1129	3.0236	1.7845	1.0118	0.408248	0.102062	0.0000	1	0
29	2C7	7	2	2	4	2	0	5.2760	3.0260	1.7113	0.9600	0.333333	0.117851	0.0000	1	0
30	2T7	7	2	2	4	2	0	5.2760	3.0260	1.7113	0.9600	0.333333	0.117851	0.0000	1	0
31	3C7	7	2	1	4	2	0	5.2760	3.0640	1.6683	0.8577	0.333333	0.117851	0.0000	1	0
32	3T7	7	2	1	4	2	0	5.2760	3.0640	1.6683	0.8577	0.333333	0.117851	0.0000	1	0
33	2M1N6	7	2	3	4	2	1	5.3284	2.9142	2.0607	0.9571	0.353553	0.125000	0.1768	0	1
34	3M1N6	7	2	3	5	1	1	5.2760	2.9343	2.0017	1.1464	0.408248	0.117851	0.3333	0	1
35	4M1N6	7	2	3	5	1	1	5.2760	2.9175	2.0378	1.2660	0.408248	0.117851	0.4553	0	1
36	5M1N6	7	2	3	4	1	1	5.2760	2.8794	2.2599	0.9481	0.408248	0.117851	0.2887	0	1
37	2M2N6	7	3	1	4	3	0	5.4916	2.9040	2.0701	0.8413	0.288675	0.144338	0.2041	0	1
38	3M2C6	7	3	2	5	3	0	5.4916	2.9267	1.8850	1.1371	0.288675	0.144338	0.3485	0	1
39	3M2T6	7	3	2	5	3	0	5.4916	2.9267	1.8850	1.1371	0.288675	0.144338	0.3485	0	1
40	4M2C6	7	3	2	5	2	0	5.4392	2.9367	1.9113	1.1649	0.333333	0.136083	0.3718	0	1

Appendix C (Continued)

Observation no.	Shorthand name	NC	TM	P3S	P3T	NCDB	DBE	XV0	XV1	XV2	XVP3	XVDB	XVT	XVPC4	NXP6	NXC3
41	4M2T6	7	3	2	5	2	0	5.4392	2.9367	1.9113	1.1649	0.333333	0.136083	0.3718	0	1
42	5M2C6	7	3	2	4	2	0	5.4392	2.8819	2.1986	0.8432	0.333333	0.136083	0.2357	0	1
43	5M2T6	7	3	2	4	2	0	5.4392	2.8819	2.1986	0.8432	0.333333	0.136083	0.2357	0	1
44	2M3C6	7	3	0	4	2	0	5.4392	2.9367	2.0804	0.7587	0.333333	0.136083	0.1925	0	1
45	2M3T6	7	3	0	4	2	0	5.4392	2.9367	2.0804	0.7567	0.333333	0.136083	0.1925	0	1
46	3M3C6	7	3	1	5	3	0	5.4916	2.9647	1.8123	1.1103	0.288675	0.144338	0.3485	0	1
47	3M3T6	7	3	1	5	3	0	5.4916	2.9647	1.8123	1.1103	0.288675	0.144338	0.3485	0	1
48	2E1N5	7	2	3	5	2	1	5.3284	2.9749	1.8536	1.1036	0.353553	0.125000	0.3018	0	1
49	3E1N5	7	2	4	6	1	1	5.2760	2.9723	1.8123	1.3821	0.408248	0.117851	0.4512	0	1
50	23M1N5	7	3	4	6	2	1	5.4916	2.8349	2.1556	1.4536	0.353553	0.144338	0.9010	0	2
51	24M1N5	7	3	3	4	2	1	5.4916	2.7701	2.5551	0.7567	0.353553	0.144338	0.3485	0	2
52	33M1N5	7	3	3	5	1	1	5.4916	2.7576	2.5463	1.4638	0.408248	0.144338	0.6969	0	4
53	34M1N5	7	3	4	6	1	1	5.4392	2.8070	2.3388	1.4234	0.408248	0.138083	1.0468	0	2
54	44M1N5	7	3	3	4	1	1	5.4916	2.6700	3.0535	0.7567	0.408248	0.144338	0.4082	0	4
55	3E2N5	7	3	2	6	3	0	5.4916	2.9873	1.6540	1.3165	0.288675	0.144338	0.4330	0	1
56	23M2N5	7	4	1	6	4	0	5.7071	2.8107	2.1339	1.3839	0.250000	0.176777	0.9571	0	2
57	24M2N5	7	4	0	4	3	0	5.6547	2.7767	2.4880	0.6667	0.288675	0.166667	0.3333	0	2
58	34M2C5	7	4	2	6	3	0	5.6547	2.8094	2.1874	1.3660	0.288675	0.166667	0.9553	0	2
59	34M2T5	7	4	2	6	3	0	5.6547	2.8094	2.1874	1.3660	0.288675	0.166667	0.9553	0	2
60	44M2C5	7	4	0	4	2	0	5.6547	2.6994	2.8660	0.6667	0.333333	0.166667	0.3333	0	4
61	44M2T5	7	4	0	4	2	0	5.6547	2.6994	2.8660	0.6667	0.333333	0.166667	0.3333	0	4
62	2E3M1N4	7	3	3	6	2	1	5.4916	2.8576	2.1665	1.2706	0.353553	0.144338	0.8413	0	2
63	233M1N4	7	4	3	6	2	1	5.7071	2.6036	3.0303	1.2803	0.353553	0.176777	1.3839	0	5
64	1N8	8	1	4	5	1	1	5.8200	3.5236	2.1380	1.2618	0.408248	0.072169	0.0000	2	0
65	2C8	8	2	3	5	2	0	5.9831	3.5260	2.0648	1.2100	0.333333	0.083333	0.0000	2	0
66	2T8	8	2	3	5	2	0	5.9831	3.5260	2.0648	1.2100	0.333333	0.083333	0.0000	2	0
67	3C8	8	2	2	5	2	0	5.9831	3.5640	2.0219	1.1267	0.333333	0.083333	0.0000	2	0
68	3T8	8	2	2	5	2	0	5.9831	3.5640	2.0219	1.1267	0.333333	0.083333	0.0000	2	0
69	4C8	8	2	2	5	2	0	5.9831	3.5640	2.0488	1.0774	0.333333	0.083333	0.0000	2	0
70	4T8	8	2	2	5	2	0	5.9831	3.5640	2.0488	1.0774	0.333333	0.083333	0.0000	2	0
71	2M1N7	8	2	4	5	2	1	6.0355	3.4142	2.4142	1.2071	0.353553	0.088388	0.1768	2	1
72	3M1N7	8	2	4	6	1	1	5.9831	3.4343	2.3552	1.4154	0.408248	0.083333	0.3333	1	1
73	4M1N7	8	2	5	6	1	1	5.9831	3.4175	2.4182	1.3505	0.408248	0.083333	0.3708	1	1
74	5M1N7	8	2	5	6	1	1	5.9831	3.4175	2.3794	1.5605	0.408248	0.083333	0.4928	1	1
75	6M1N7	8	2	4	5	1	1	5.9831	3.3794	2.6134	1.1897	0.408248	0.083333	0.2887	2	1
76	2M2N7	8	3	2	5	3	0	6.1987	3.4040	2.4237	1.1103	0.288675	0.102062	0.2041	2	1
77	3M2C7	8	3	3	6	3	0	6.1987	3.4267	2.2386	1.4175	0.288675	0.102062	0.3485	1	1
78	3M2T7	8	3	3	6	3	0	6.1987	3.4267	2.2386	1.4175	0.288675	0.102062	0.3485	1	1
79	4M2C7	8	3	3	6	2	0	6.1463	3.4367	2.2917	1.2650	0.333333	0.096225	0.3027	1	1
80	4M2T7	8	3	3	6	2	0	6.1463	3.4367	2.2917	1.2650	0.333333	0.096225	0.3027	1	1

81	5M2C7	8	3	4	6	2	0	6.1463	3.4199	2.3182	1.4711	0.333333	0.096225	0.4553	1	1
82	5M2T7	8	3	4	6	2	0	6.1463	3.4199	2.3182	1.4711	0.333333	0.096225	0.4553	1	1
83	6M2C7	8	3	3	5	2	0	6.1463	3.3819	2.5402	1.1464	0.333333	0.096225	0.2887	2	1
84	6M2T7	8	3	3	5	2	0	6.1463	3.3819	2.5402	1.1464	0.333333	0.096225	0.2887	2	1
85	2M3C7	8	3	1	5	2	0	6.1463	3.4367	2.4608	0.9763	0.333333	0.096225	0.1925	2	1
86	2M3T7	8	3	1	5	2	0	6.1463	3.4367	2.4608	0.9763	0.333333	0.096225	0.1925	2	1
87	3M3C7	8	3	2	6	3	0	6.1987	3.4647	2.1927	1.3392	0.288675	0.102062	0.3485	1	1
88	3M3T7	8	3	2	6	3	0	6.1987	3.4647	2.1927	1.3392	0.288675	0.102062	0.3485	1	1
89	4M3C7	8	3	2	6	3	0	6.1987	3.4647	2.2087	1.1969	0.288675	0.102062	0.2887	1	1
90	4M3T7	8	3	2	6	3	0	6.1987	3.4647	2.2087	1.1969	0.288675	0.102062	0.2887	1	1
91	5M3C7	8	3	2	6	2	0	6.1463	3.4747	2.2219	1.3443	0.333333	0.096225	0.3718	1	1
92	5M3T7	8	3	2	6	2	0	6.1463	3.4747	2.2219	1.3443	0.333333	0.096225	0.3718	1	1
93	6M3C7	8	3	2	5	2	0	6.1463	3.4199	2.5092	1.0099	0.333333	0.096225	0.2357	2	1
94	6M3T7	8	3	2	5	2	0	6.1463	3.4199	2.5092	1.0099	0.333333	0.096225	0.2357	2	1
95	2E1N6	8	2	4	6	2	1	6.0355	3.4749	2.2071	1.3839	0.353553	0.088388	0.3018	1	1
96	3E1N6	8	2	5	7	1	1	5.9831	3.4723	2.1927	1.5172	0.408248	0.083333	0.4024	0	1
97	4E1N6	8	2	6	7	1	1	5.9831	3.4555	2.2069	1.8547	0.408248	0.083333	0.5261	0	1
98	23M1N6	8	3	5	7	2	1	6.1987	3.3349	2.5361	1.5629	0.353553	0.1020682	0.8413	0	2
99	24M1N6	8	3	5	6	2	1	6.1987	3.3081	2.6747	1.3938	0.353553	0.102062	0.5774	0	2
100	25M1N6	8	3	4	5	2	1	6.1987	3.2701	2.8896	1.1485	0.353553	0.102062	0.4655	0	2
101	33M1N6	8	3	4	7	1	1	6.1987	3.2576	2.9427	1.4469	0.408248	0.102062	0.5774	0	4
102	34M1N6	8	3	6	8	1	1	6.1463	3.3450	2.4804	1.9133	0.408248	0.096225	1.0721	0	2
103	35M1N6	8	3	4	6	1	1	6.1463	3.2901	2.8425	1.2456	0.408248	0.096225	0.5385	0	2
104	44M1N6	8	3	6	7	1	1	6.1987	3.2307	3.0106	1.6540	0.408248	0.102062	0.7887	0	4
105	45M1N6	8	3	6	7	1	1	6.1463	3.2901	2.7457	1.6765	0.408248	0.096225	1.1765	0	2
106	55M1N6	8	3	4	5	1	1	6.1987	3.1701	3.3880	1.0985	0.408248	0.102062	0.5000	0	4
107	3E2C6	8	3	3	7	3	0	6.1987	3.4873	2.0505	1.4335	0.288675	0.102062	0.3907	0	1
108	3E2T6	8	3	3	7	3	0	6.1987	3.4873	2.0505	1.4335	0.288675	0.102082	0.3907	0	1
109	4E2C6	8	3	4	7	2	0	6.1463	3.4747	2.1024	1.5134	0.333333	0.096225	0.4296	0	1
110	4E2T6	8	3	4	7	2	0	6.1463	3.4747	2.1024	1.5134	0.333333	0.096225	0.4296	0	1
111	23M2N6	8	4	1	7	4	0	6.4142	3.3107	2.5303	1.4786	0.250000	0.125000	0.9053	0	2
112	24M2N6	8	4	2	6	3	0	6.3618	3.3147	2.6295	1.2618	0.288675	0.117851	0.5202	0	2
113	25M2N6	8	4	2	5	3	0	6.3618	3.2599	2.9110	0.9975	0.288675	0.117851	0.4398	0	2
114	34M2C6	8	4	4	8	3	0	6.3618	3.3474	2.3420	1.8450	0.288675	0.117851	0.9773	0	2
115	34M2T6	8	4	4	8	3	0	6.3618	3.3474	2.3420	1.8450	0.288675	0.117851	0.9773	0	2
116	35M2T6	8	4	3	6	3	0	6.3618	3.2825	2.7330	1.2230	0.288675	0.117851	0.5261	0	2
117	35M2T6	8	4	3	6	3	0	6.3618	3.2825	2.7330	1.2230	0.288675	0.117851	0.5261	0	2
118	44M2C6	8	4	3	7	2	0	6.3618	3.2600	2.8421	1.5291	0.333333	0.117851	0.6440	0	4
119	44M2T6	8	4	3	7	2	0	6.3618	3.2600	2.8421	1.5291	0.333333	0.117851	0.6440	0	4
120	45M2C6	8	4	4	7	2	0	6.3094	3.3094	2.6289	1.5476	0.333333	0.111111	1.0218	0	2
121	45M2T6	8	4	4	7	2	0	6.3094	3.3094	2.6289	1.5476	0.333333	0.111111	1.0218	0	2

Appendix C (Continued)

Observation no.	Shorthand name	NC	TM	P3S	P3T	NCDB	DBE	XV0	XV1	XV2	XVP3	XVDB	XVT	XVPC4	NXP6	NXC3
122	55M2C6	8	4	3	5	2	0	6.3618	3.1725	3.3338	0.9659	0.333333	0.117851	0.4082	0	4
123	55M2T6	8	4	3	5	2	0	6.3618	3.1725	3.3338	0.9659	0.333333	0.117851	0.4082	0	4
124	3E3N6	8	3	2	7	3	0	6.1987	3.5254	1.9777	1.4010	0.288675	0.102062	0.3907	0	1
125	22M3C6	8	4	0	5	2	0	6.3618	3.2374	3.1766	0.8536	0.333333	0.117851	0.3333	0	4
126	22M3T6	8	4	0	5	2	0	6.3618	3.2374	3.1766	0.8536	0.333333	0.117851	0.3333	0	4
127	23M3C6	8	4	2	7	3	0	6.3618	3.3474	2.5111	1.4368	0.288675	0.117851	0.9065	0	2
128	23M3T6	8	4	2	7	3	0	6.3618	3.3474	2.5111	1.4368	0.288675	0.117851	0.9065	0	2
129	24M3C6	8	4	1	6	3	0	6.3618	3.3374	2.6106	1.1755	0.288675	0.117851	0.4886	0	2
130	24M3T6	8	4	1	6	3	0	6.3618	3.3374	2.6106	1.1755	0.288675	0.117851	0.4886	0	2
131	25M3C6	8	4	0	5	2	0	6.3094	3.3094	2.8729	0.8809	0.333333	0.111111	0.3849	0	2
132	25M3T6	8	4	0	5	2	0	6.3094	3.3094	2.8729	0.8809	0.333333	0.111111	0.3849	0	2
133	34M3C6	8	4	2	8	4	0	6.4142	3.3713	2.2678	1.7892	0.250000	0.125000	0.9571	0	2
134	34M3T6	8	4	2	8	4	0	6.4142	3.3713	2.2678	1.7892	0.250000	0.125000	0.9571	0	2
135	2NP1N5	8	2	4	6	2	1	6.0355	3.4749	2.2500	1.2071	0.353553	0.088388	0.2500	1	1
136	2IP1N5	8	3	4	7	2	1	6.1987	3.3576	2.5629	1.3876	0.353553	0.102062	0.7990	0	2
137	2E3M1N5	8	3	5	8	2	1	6.1987	3.3956	2.3211	1.7634	0.353553	0.102062	0.8835	0	2
138	2E4M1N5	8	3	4	6	2	1	6.1987	3.3307	2.7015	1.1969	0.353553	0.102062	0.4830	0	2
139	3E2M1N5	8	3	6	8	2	1	6.1987	3.3729	2.3598	1.6825	0.353553	0.102062	0.8237	0	2
140	3E3M1N5	8	3	6	9	1	1	6.1987	3.3182	2.5653	2.1082	0.408248	0.102062	0.9433	0	4
141	3E4M1N5	8	3	6	8	1	1	6.1463	3.3450	2.5518	1.6305	0.408248	0.096225	0.9326	0	2
142	233M1N5	8	4	6	9	2	1	6.4142	3.1642	3.0178	2.0392	0.353553	0.125000	1.4357	0	5
143	234M1N5	8	4	6	8	2	1	6.3618	3.2076	2.8790	1.7773	0.353553	0.117851	1.4398	0	3
144	244M1N5	8	4	4	5	2	1	6.4142	3.0607	3.6945	0.8321	0.353553	0.125000	0.4786	0	5
145	334M1N5	8	4	6	9	1	1	6.3618	3.1403	3.1802	2.0141	0.408248	0.117851	1.6464	0	5
146	344M1N5	8	4	6	8	1	1	6.3618	3.1076	3.3904	1.7198	0.408248	0.117851	1.5285	0	5
147	3E2M2N5	8	4	2	8	4	0	6.4142	3.3713	2.3107	1.5607	0.250000	0.125000	0.8536	0	2
148	3E4M2C5	8	4	3	8	3	0	6.3618	3.3701	2.3718	1.5206	0.288675	0.117851	0.8422	0	2
149	3E4M2T5	8	4	3	8	3	0	6.3618	3.3701	2.3718	1.5206	0.288675	0.117851	0.8422	0	2
150	234M2N5	8	5	2	8	4	0	6.5774	3.1934	2.8377	1.6547	0.250000	0.144338	1.4047	0	3
151	244M2N5	8	5	0	5	3	0	6.5774	3.0774	3.5877	0.7217	0.288675	0.144338	0.4330	0	5
152	344M2C5	8	5	3	8	3	0	6.5774	3.1160	3.2217	1.6160	0.288675	0.144338	1.3660	0	5
153	344M2T5	8	5	3	8	3	0	6.5774	3.1160	3.2217	1.6160	0.288675	0.144338	1.3660	0	5
154	2IP3M1N4	8	4	4	8	2	1	6.3618	3.2403	2.8843	1.4832	0.353553	0.117851	1.2130	0	3
155	2E33N1N4	8	4	4	8	2	1	6.4142	3.1642	3.2071	1.4874	0.353553	0.125000	1.2071	0	5
156	1N9	9	1	5	8	1	1	6.5271	4.0238	2.4916	1.5118	0.408248	0.051031	0.0000	3	0
157	2C9	9	2	4	6	2	0	6.6902	4.0260	2.4184	1.4600	0.333333	0.058926	0.0000	3	0
158	2T9	9	2	4	6	2	0	6.6902	4.0260	2.4184	1.4600	0.333333	0.058926	0.0000	3	0
159	3C9	9	2	3	6	2	0	6.6902	4.0640	2.3754	1.3767	0.333333	0.058926	0.0000	3	0
160	3T9	9	2	3	6	2	0	6.6902	4.0640	2.3754	1.3767	0.333333	0.058926	0.0000	3	0
161	4C9	9	2	3	6	2	0	6.6902	4.0640	2.4023	1.3464	0.333333	0.058926	0.0000	3	0
162	4T9	9	2	3	6	2	0	6.6902	4.0640	2.4023	1.3464	0.333333	0.058926	0.0000	3	0

References

- [1] D.E. Needham, I.-C. Wei, P.G. Seybold, *J. Am. Chem. Soc.* 110 (1988) 4186–4194.
- [2] L.B. Kier, L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
- [3] L.B. Kier, L.H. Hall, *Molecular Connectivity in Structure–Activity Analysis*, Wiley, New York, 1986.
- [4] P.G. Seybold, M. May, U.A. Bagal, *J. Chem. Educ.* 64 (1987) 575–581.
- [5] R.H. Rohrbaugh, P.C. Jurs, *Anal. Chem.* 57 (1985) 2770–2773.
- [6] P.J. Hansen, P.C. Jurs, *Anal. Chem.* 59 (1987) 2322–2327.
- [7] S. Liu, R. Zhang, M. Liu, Z.J. Hu, *Chem. Info. Comput. Sci.* 37 (1997) 1146–1151.
- [8] S.P. Verevkin, D. Wandschneider, A. Heintz, *J. Chem. Eng. Data* 45 (2000) 618–625.
- [9] *CRC Handbook of Chemistry and Physics*, 68th Edition, CRC Press, Boca Raton, 1987.
- [10] I.B.D. Smith, R. Srivastava, *Thermodynamic Data for Pure Compounds. Part A. Hydrocarbons and Ketones*, Elsevier, New York, 1986.
- [11] R.R. Dreisbach (Ed.), *Physical Properties of Chemical Compounds. Part II. Advances in Chemistry Series, Vol. 22*, American Chemical Society, Washington, DC, 1959.
- [12] *Selected Values of Properties of Hydrocarbons and Related Compounds*, A&M Research Foundation, Thermodynamic Research Center, College Station, TX, 1985.
- [13] M.P. Doss, *Physical Constants of the Principal Hydrocarbons*, 2nd Edition, The Texas Company, New York, 1939.
- [14] *Aldrich Catalog of Fine Chemicals*, Aldrich Chemical Co., New York, 1981–1982.
- [15] M. Randić, *J. Am. Chem. Soc.* 97 (1975) 6609–6615.
- [16] L.H. Hall, Hall Associates, Quincy, MA.
- [17] SAS Institute, Box 8000, Cary, NC.
- [18] J.C. Dearden, The prediction of melting point, in: M. Charton (Ed.), *Advances in Quantitative Structure Property Relationships*, Vol. 2, JAI Press, New York, 1999, pp. 127–175.