Neural Network Based Quantitative Structural Property Relations (QSPRs) for Predicting Boiling Points of Aliphatic Hydrocarbons

Gabriela Espinosa,[†] Denise Yaffe,[‡] Yoram Cohen,*[,‡] Alex Arenas,[§] and Francesc Giralt[†]

Departament d'Enginyeria Química, ETSEQ and Departament d'Enginyeria Informàtica, ETSE, Universitat Rovira i Virgili, 43006 Tarragona, Catalunya, Spain, and Department of Chemical Engineering, University of California, Los Angeles, Los Angeles, California 90095-1592

Received January 3, 2000

Quantitative structural property relations (QSPRs) for boiling points of aliphatic hydrocarbons were derived using a back-propagation neural network and a modified Fuzzy ARTMAP architecture. With the back-propagation model, the selected molecular descriptors were capable of distinguishing between diastereomers. The QSPRs were obtained from four valance molecular connectivity indices ($^1\chi^v$,$^2\chi^v$,$^3\chi^v$,$^4\chi^v$), a second-order Kappa shape index ($^2\kappa$), dipole moment, and molecular weight. The inclusion of dipole moment proved to be particularly useful for distinguishing between cis and trans isomers. A back-propagation 7−4−1 architecture predicted boiling points for the test, validation, and overall data sets of alkanes with average absolute errors of 0.37% (1.65 K), 0.42% (1.73 K), and 0.37% (1.54 K), respectively. The error for the test and overall data sets decreased to 0.19% (0.81 K) and 0.31% (1.30 K), respectively, using the modified Fuzzy ARTMAP network. A back-propagation alkene model, with a 7−10−1 architecture, yielded predictions with average absolute errors for the test, validation, and overall data sets of 1.96% (6.79 K), 1.83% (6.45 K), and 1.25% (4.42 K), respectively. Fuzzy ARTMAP reduced the errors for the test and overall data sets to 0.19% (0.73 K) and 0.25% (0.95 K), respectively. The back-propagation composite model for aliphatic hydrocarbons, with a 7−9−1 architecture, yielded boiling points with average absolute errors for the test, validation, and overall set of 1.74% (6.09 K), 1.25% (4.68 K), and 1.37% (4.85 K), respectively. The error for the test and overall data sets using the Fuzzy ARTMAP composite model decreased to 0.84% (1.15 K) and 0.35% (1.35 K), respectively. Performance of the QSPRs, developed from a simple set of molecular descriptors, displayed accuracy well within the range of expected experimental errors and of better accuracy than other regression analysis and neural network-based boiling points QSPRs previously reported in the literature.

I. INTRODUCTION

In recent years there has been a growing interest in the application of neural networks in the development of quantitative property structure relations (QSPRs) for the correlation and estimation of physical properties of organic compounds. The premise of QSPRs is that physicochemical properties can be correlated with molecular structural characteristics (geometric and electronic) expressed in terms of appropriate molecular descriptors.[1] Various studies have reported on the use of electronic (i.e., dipole moments, hydrogen bonding parameters), lipophilic (i.e., partition coefficients), and topological (i.e., molecular connectivity indices and other geometric parameters) descriptors as well as other molecular parameters (e.g., molar volume, parachor, and molar refractivity) for correlating structural parameters with physicochemical properties. Examples of some of the more commonly reported topological indices proposed for QSPRs include Randić branching indices,[2] valance molecular connectivity indices,[3,14] Wiener path numbers,[4] Kappa Shape indices,[5] and the electrotopological state indices.[6] A combination of several topological indices or molecular descriptors is typically required to adequately represent molecular structural information for QSPR applications.[7]

QSPRs have been traditionally developed by selecting, a priori, an analytical model (typically linear, polynomial, or log-linear) to quantify the correlation between selected molecular indices and desired physicochemical properties, followed by regression analysis to determine model parameters.[8−19] Although the above approach has proved useful in many applications, it has a number of limitations.[11,17,19−22] The quantitative relationships between structure and physicochemical properties can be complex and highly nonlinear; thus, determining the optimal analytical form of the QSPR model presents a challenge. Moreover, regression analysis becomes complex and less reliable as the number of descriptors increases. Therefore, mapping the relationship between a set of molecular descriptors and multiple physicochemical properties, using a single composite QSPR model, becomes a formidable task using traditional regression methods.

In recent years, a number of investigators have demonstrated that neural network systems can be an effective tool for developing QSPRs. The major advantage of neural networks lies in the fact that QSPRs can be developed without having to a priori specify the analytical form of a particular correlation model. The neural network approach is especially suitable for mapping complex nonlinear rela-

* Corresponding author phone: (310)825-8766; fax: (310)477-3868; e-mail: yoram@ucla.edu.
[†] ETSEQ, Universitat Rovira i Virgili.
[‡] University of California, Los Angeles.
[§] ETSE, Universitat Rovira i Virgili.

10.1021/ci000442u CCC: $19.00
Published on Web 04/14/2000

tionships that may exist between model output (i.e., physicochemical properties) and model input (i.e., molecular descriptors). As in regression analysis, building an accurate QSPR requires sufficient and reliable experimental data.

Back-propagation neural networks are commonly used for predicting physicochemical properties. An error-based learning system is used in back-propagation algorithms, where actual predictions are compared with target values, and the errors are used to change adaptive weights to reduce the errors. Since chemical compounds fall into various structural classes, it may also be feasible to use cognitive classifiers, such as the Fuzzy ARTMAP network, for rapid learning of categories which represent structure and properties in a supervised way. Fuzzy ARTMAP networks can also use predictive disconfirmations to supervise learning of categories that fit the statistics of the input−output environment.[23] This class of neural networks uses a match-based learning, in that it actively searches for recognition categories or hypotheses whose prototype (expectations) provides an acceptable match to input data. The learned prototype represents the cluster of input features that the category deems relevant based upon its past experience. When the search discovers a category that provides an acceptable match, the system locks into an attentive resonance, whereby the input patterns refine that adaptive weight of the category based on any new information that it contains. When a suitable match cannot be found, a new category is automatically initiated. The individual recognition categories of Fuzzy ARTMAP have a similar function as hidden units in back-propagation. Every category group (structure/property) is defined inside an independent module ART,[24] linked by an associative memory that records, according to a vigilance parameter, the connections between structure and property. Two modules are utilized, one for structure (module A) and the second for properties (module B). It should be noted that the original Fuzzy ARTMAP architecture requires modification in order to be used as a predictive system.[25,26]

As the literature reveals, a major challenge in neural network/QSPR development has been to establish a reliable and practical set of molecular descriptors.[8,11,21,22] As a consequence, most recent studies have explored the development of QSPRs for commonly available physicochemical parameters (e.g., boiling point, heat capacity, density, refractive index) for selected organic compound classes for which accurate and rich data sets are available.[8,17,19−22]

The use of boiling point data to test the applicability of various molecular descriptors has been particularly popular given the availability of data for large sets of organic chemical classes. QSPRs for boiling points have been proposed by a number of investigators based on back-propagation neural networks.[8,17,20−22] For example, Gakh et al.[19] developed a composite neural network/QSPR boiling points model, based on 134 noncyclic alkanes ranging from 6 to 10 carbon atoms. Their model, which included five additional physicochemical properties, was based on six Weiner-type structural graph invariants representing the number of pathways for carbon lengths ranging from three to eight atoms and also included the number of carbon atoms. In a later study, Ivanciuc[8] proposed neural network/QSPR boiling point models based on the MolNet neural network model[27] and two topological descriptors: degree, DEG,[28,29] which is based on adjacency matrix, and the reciprocal distance sum, RDS.[30,31] It should be noted that the general application of the above approaches to aliphatic hydrocarbons cannot be reliably established since those earlier models excluded smaller alkanes ranging from 1 to 5 carbon atoms and larger compounds with chain lengths greater than eight carbon atoms. A number of composite models, which also include alkanes, were reported by Egolf et al.[21] and Hall and Story[22] as discussed in the subsequent paragraphs.

A neural network-QSPR, for boiling points of a diverse class of organics, was developed by Egolf et al.[21] using a data set of 298 organic compounds that included about 45 alkanes and 50 alkenes as well as halogens, alcohols, esters, ketones, carboxcyclic acids, and amines. The approach identified the following descriptors: three charged partial surface area structural (CPSA) descriptors,[32] a composite CPSA descriptor, number of oxygens, Wiener number,[4] the summation of all unique paths, or molecular ID index,[33] and charge on carbonyl or cyano carbons. The model, based on 8−3−1 network architecture, predicted boiling points with a mean error of 10.65 K. Although the proposed set of molecular descriptors appeared to distinguish among diastereomers, the use of proprietary techniques and software to derive the required molecular descriptors places a limitation on the general applicability of this derived model.

A unique set of molecular descriptors were applied by Hall and Story[22] for boiling point predictions for the same set of 298 compounds used by Egolf et al.[21] A set of 19 atom-type electrotopological indices were selected to represent the encoding of intrinsic electronic state of atoms, as perturbed by the electronic influence of all other atoms in the molecule, within the context of the topological character of the molecule. The electrotopological indices were effective in representing the general structural characteristics of most compounds in the data set. The model with a 19−5−1 architecture produced a mean absolute error (for 10 runs) of 1.12% (4.57 K) for the test set and an average absolute error of 0.94% (3.93 K) for the combined data set. It should be noted, however, that for diastereomers the indices used in the above model become degenerate and thus fail to differentiate between cis and trans isomers.

In a later study Zhang et al.[20] proposed a neural network/QSPR for boiling points, based on a narrow class of organics, utilizing a data set of 85 single double bond alkenes, ranging in size from 4 to 20 carbon atoms. The set of molecular descriptors included four topological indices (derived from distance matrices) based on the approach of Wiener.[4] An additional isomer index was included to distinguish between cis and trans isomers, by assigning values of 1, −1, or 0 for a trans isomer, cis isomer, and all other compounds (i.e. not diastereomers), respectively. The above approach, however, is questionable since it does not provide a true quantitative descriptor of the isomeric structure.

The success and popularity of neural networks/QSPR models will depend, in part, on the ease by which the molecular descriptors can be determined by the interested user. Thus, in the present study, our primary goal is to investigate the potential applicability of a simple and easily calculable set of molecular indices that includes a descriptor to distinguish between diastereomers. The molecular descriptors include four valance molecular connectivity indices,[4] a second-order kappa shape index,[5] dipole moment, and molecular weight. As in previous studies, the approach was

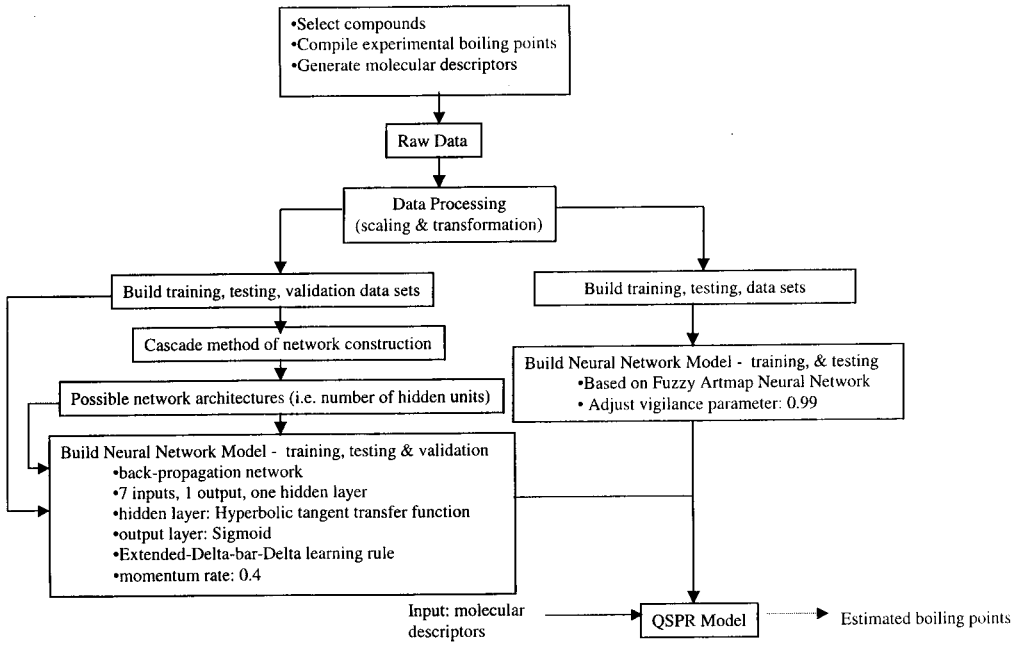

Figure 1. Process flow diagram for developing QSPR/neural networks for predicting boiling points of aliphatic hydrocarbons.

evaluated for boiling points QSPR for alkanes and alkenes for which accurate experimental data are available. Alkynes, which have not been considered in previous studies, are also included in the present work. The lack of boiling point data for a sizable fraction of the alkyne group makes the development of neural network-QSPR models particularly appealing. Our overall goal is to evaluate QSPR models developed using back-propagation neural network and a modified Fuzzy ARTMAP cognitive system[25,26] for aliphatic hydrocarbons with a simple set of molecular descriptors.

II. METHODOLOGY

Overview. QSPRs for the normal boiling points of alkanes, alkenes, alkynes, and for the composite group of compounds were derived, using back-propagation and Fuzzy ARTMAP neural networks, following the methodology shown in Figure 1. The data sets included seven structural descriptors as input parameters and boiling points as a single output parameter. Boiling point data were obtained for 140 alkanes, 144 alkenes, and 43 alkynes composed of both straight and branched chain aliphatic hydrocarbons ranging from one to 10 carbon atoms.[34−36] Tables 1−3 list the boiling points for the complete data set.

Molecular Descriptors. Molecular topology for the aliphatic hydrocarbons was characterized by four valance molecular connectivity indices of orders 1, 2, 3, and 4 ($^1\chi^v$,$^2\chi^v$,$^3\chi^v$,$^4\chi^v$) and the second Kappa shape index,[4,5] $^2\kappa$. Molecular connectivity indices of the path type are topological indices that encode structural information into numerical values or indexes. The molecular structure is expressed topologically by a hydrogen-suppressed graph. The carbons (and heteroatoms) are represented as vertexes, and bonds connecting atoms are represented as edges. Briefly, the connectivity indices $^m\chi^v$ are valance-weighted counts of connected subgraphs. The first-order term $^1\chi^v$ is related to the degree of branching and size of the molecule expressed as the number of non-hydrogen atoms. The second-order term $^2\chi^v$ represents a dissection of the molecular skeleton into "two contiguous bond" fragments. The third-order term $^3\chi^v$ is a weighted count of four atoms (three-bond) fragments representing the potential for rotation around the central bond and is the smallest molecular structure necessary for conformational variability. The $^3\chi^v$ index also reflects the degree of branching at each of the four atoms in the fragment. The fourth-order term $^4\chi^v$ represents path, cluster, path/cluster, and cyclic subgraphs of four edges. Structural information from the $^4\chi^v$ index is useful for compounds with at least five carbon atoms in a chain. To provide for a suitable characterization of the level of branching among isomers, the kappa 2 shape index,[5] $^2\kappa$, was included as an additional molecular descriptor. This index quantifies the structure of compounds within an isomeric series in terms of its relative starlikeness and straight chainlikeness shape.

Combination of the four valance molecular connectivity indices and the $^2\kappa$ shape index provided the ability to distinguish among constitutional isomers. However, geometric isomers of alkenes, known as diastereomers (i.e., cis and trans stereoisomers), which differ only in the way atoms or groups of atoms are oriented in space, can have different physicochemical properties. It appears that, although the dipole moment for the cis and trans configurations is small, the finite difference in dipole moments is sufficient to differentiate between these two isomers. Therefore, the dipole moment was included as an added molecular descriptor. In addition, to increase the uniqueness of the set of descriptors, molecular weight was included as an added parameter. The molecular connectivity indices were determined from molecular structure using available molecular modeling software[37] and following the approach of Kier and Hall.[14] The $^2\kappa$ shape index and dipole moment were determined following Kier[5] and molecular modeling software, respectively.[37] The complete set of descriptors used in the present study are listed in Tables 1−3.

Data Sets and Neural Network Systems. Boiling points and molecular descriptors were divided into three data sets: training, testing, and validating for use with the back-

Table 1. Molecular Descriptors and Experimental Boiling Points for Alkanes[a]

	input parameters							
name	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	$^2\kappa$	MW	dipole moment (Debye)	output reported BP, K
methane	0.00	0.00	0.00	0.00	0.00	16.04	0.0	111.66
ethane	1.00	0.00	0.00	0.00	0.00	30.07	0.0	182.52
propane	1.41	0.71	0.00	0.00	2.00	44.10	0.00748	231.08
butane	1.91	1.00	0.50	0.00	3.00	58.12	0.0	272.65
2-methylpropane	1.73	1.73	0.00	0.00	1.33	58.12	0.01147	261.42
pentane	2.41	1.35	0.71	0.35	4.00	72.15	0.00761	309.24
2-methylbutane	2.27	1.80	0.82	0.00	2.25	72.15	0.01018	301.00
2,2-dimethylpropane	2.00	3.00	0.00	0.00	1.00	72.15	0.0	282.68
hexane	2.91	1.71	0.96	0.50	5.00	86.18	0.0	341.89
3-methylpentane	2.81	1.92	1.39	0.29	3.20	86.18	0.01490	336.43
2-methylpentane	2.77	2.18	0.87	0.58	3.20	86.18	0.01178	333.42
2,2-dimethylbutane	2.56	2.91	1.06	0.00	1.63	86.18	0.00902	322.89
2,3-dimethylbutane	2.64	2.49	1.33	0.00	2.22	86.18	0.0	331.14
heptane	3.41	2.06	1.21	0.68	6.00	100.20	0.00758	371.58
2-methylhexane	3.27	2.54	1.14	0.61	4.17	100.20	0.01031	363.20
3-methylhexane	3.31	2.30	1.48	0.70	4.17	100.20	0.01049	365.00
3-ethylpentane	3.35	2.09	1.73	0.87	4.17	100.20	0.01324	366.63
2,2-dimethylpentane	3.06	3.31	1.00	0.75	2.34	100.20	0.00181	352.34
2,3-dimethylpentane	3.18	2.63	1.78	0.47	3.06	100.20	0.00830	362.93
3,3-dimethylpentane	3.12	2.87	1.91	0.25	2.34	100.20	0.01782	359.21
2,4-dimethylpentane	3.13	3.02	0.94	0.94	3.06	100.20	0.00363	353.65
2,2,3-trimethylbutane	2.94	3.52	1.73	0.00	1.85	100.20	0.01244	354.03
octane	3.91	2.41	1.46	0.85	7.00	114.23	0.0	398.82
2-methylheptane	3.77	2.89	1.39	0.80	5.14	114.23	0.01170	390.80
3-methylheptane	3.81	2.66	1.75	0.76	5.14	114.23	0.01448	392.08
4-methylheptane	3.81	2.68	1.56	1.13	5.14	114.23	0.01182	390.85
3-ethylhexane	3.85	2.47	1.85	1.11	5.14	114.23	0.00880	391.68
2,2-dimethylhexane	3.56	3.66	1.28	0.71	3.11	114.23	0.00906	379.99
2,3-dimethylhexane	3.68	3.01	1.88	0.79	3.94	114.23	0.02145	388.76
2,4-dimethylhexane	3.66	3.14	1.57	0.97	3.94	114.23	0.00462	382.58
2,5-dimethylhexane	3.63	3.37	1.32	0.67	3.94	114.23	0.02017	382.25
3,4-dimethylhexane	3.72	2.77	2.26	0.81	3.94	114.23	0.0	390.88
2,2,3-trimethylpentane	3.48	3.68	2.09	0.61	2.52	114.23	0.01138	382.99
2,2,4-trimethylpentane	3.42	4.16	1.02	1.23	2.52	114.23	0.01047	372.39
2,3,3-trimethylpentane	3.50	3.50	2.47	0.41	2.52	114.23	0.01931	387.91
2,3,4-trimethylpentane	3.55	3.35	2.10	0.77	3.11	114.23	0.01310	386.62
3,3-dimethylhexane	3.62	3.27	1.88	0.85	3.11	114.23	0.01051	385.12
2-methyl-3-ethylpentane	3.72	2.82	1.99	1.23	3.94	114.23	0.01958	388.80
3-methyl-3-ethylpentane	3.68	2.87	2.56	0.75	3.11	114.23	0.02096	391.41
2,2,3,3-tetramethylbutane	3.25	4.50	2.25	0.00	1.75	114.23	0.0	379.62
nonane	4.41	2.77	1.71	1.03	8.00	128.26	0.00741	423.95
2-methyloctane	4.27	3.24	1.64	0.98	6.13	128.26	0.01035	416.41
3-methyloctane	4.31	3.01	2.00	0.95	6.13	128.26	0.01079	417.55
4-methyloctane	4.31	3.04	1.83	1.19	6.13	128.26	0.01067	415.57
3-ethylheptane	4.35	2.82	2.12	1.19	6.13	128.26	0.01350	416.15
4-ethylheptane	4.35	2.85	1.97	1.37	6.13	128.26	0.00981	414.35
2,2-dimethylheptane	4.06	4.02	1.53	0.91	3.92	128.26	0.00180	405.84
3,3-dimethylheptane	4.12	3.62	2.16	0.83	3.92	128.26	0.01799	410.45
4,4-dimethylheptane	4.12	3.66	1.85	1.48	3.92	128.26	0.00301	408.35
2,3-dimethylheptane	4.18	3.36	2.15	0.86	4.84	128.26	0.00845	413.65
2,4-dimethylheptane	4.16	3.52	1.66	1.42	4.84	128.26	0.00325	406.05
3,4-dimethylheptane	4.22	3.15	2.36	1.14	4.84	128.26	0.02312	413.75
2-methyl-4-ethylhexane	4.20	3.31	1.96	1.29	4.84	128.26	0.00977	406.95
2-methyl-3-ethylhexane	4.22	3.20	2.13	1.38	4.84	128.26	0.00747	411.15
3-methyl-3-ethylhexane	4.18	3.27	2.56	1.21	3.92	128.26	0.01463	413.12
3-methyl-4-ethylhexane	4.26	2.96	2.50	1.43	4.84	128.26	0.01783	413.15
2,3,4-trimethylhexane	4.09	3.49	2.59	1.03	3.92	128.26	0.01214	412.25
2,3,5-trimethylhexane	4.04	3.85	1.98	1.02	3.92	128.26	0.00975	404.55
2,2,3-trimethylhexane	3.98	4.06	2.20	0.87	3.24	128.26	0.01320	406.75
2,2,4-trimethylhexane	3.96	4.28	1.66	1.19	3.24	128.26	0.01564	399.69
2,2,5-trimethylhexane	3.92	4.49	1.47	0.72	3.24	128.26	0.01209	397.15
2,3,3-trimethylhexane	4.00	3.89	2.46	0.93	3.24	128.26	0.01386	410.85
3,3,4-trimethylhexane	4.04	3.65	2.86	0.90	3.24	128.26	0.01369	413.65
3,3-diethylpentane	4.24	2.91	3.00	1.50	3.92	128.26	0.02469	419.45
2,3,3,4-tetramethylpentane	3.89	4.13	2.98	0.67	2.72	128.26	0.01450	414.65
2,2,3,3-tetramethylpentane	3.81	4.49	2.91	0.53	2.32	128.26	0.00938	413.35
2,2,4,4-tetramethylpentane	3.71	5.30	1.06	1.59	2.32	128.26	0.01037	395.35
2,2,3,4-tetramethylpentane	3.85	4.40	2.37	1.00	2.72	128.26	0.02222	406.15
2,2-dimethyl-3-ethylpentane	4.02	3.88	2.21	1.51	3.24	128.26	0.00828	406.95
2,3-dimethyl-3-ethylpentane	4.07	3.52	3.01	1.07	3.24	128.26	0.02471	417.85
2,4-dimethyl-3-ethylpentane	4.09	3.56	2.18	1.71	3.92	128.26	0.01142	409.85
decane	4.91	3.12	1.96	1.21	9.00	142.29	0.0	447.25

Table 1 (Continued)

	input parameters							
name	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	$^2\kappa$	MW	dipole moment (Debye)	output reported BP, K
2-methylnonane	4.77	3.60	1.89	1.16	7.11	142.29	0.01153	440.15
3-methylnonane	4.81	3.36	2.25	1.12	7.11	142.29	0.01423	440.95
4-methylnonane	4.81	3.39	2.08	1.38	7.11	142.29	0.01173	438.85
5-methylnonane	4.81	3.39	2.10	1.25	7.11	142.29	0.01474	438.30
3-ethyloctane	4.85	3.18	2.37	1.38	7.11	142.29	0.00874	439.65
4-ethyloctane	4.85	3.21	2.24	1.45	7.11	142.29	0.00917	436.85
4-isopropylheptane	4.72	3.58	2.26	1.55	5.76	142.29	0.02031	432.05
4-propylheptane	4.85	3.23	2.09	1.66	7.11	142.29	0.01194	435.15
2,2-dimethyloctane	4.56	4.37	1.78	1.08	4.76	142.29	0.00904	428.15
2,3-dimethyloctane	4.68	3.72	2.40	1.05	5.76	142.29	0.00085	437.45
2,4-dimethyloctane	4.66	3.88	1.92	1.48	5.76	142.29	0.02118	429.15
2,5-dimethyloctane	4.66	3.87	2.02	1.26	5.76	142.29	0.02065	431.65
2,6-dimethyloctane	4.66	3.84	2.18	1.08	5.76	142.29	0.00460	433.55
2,7-dimethyloctane	4.63	4.07	1.81	1.11	5.76	142.29	0.02017	433.05
3,3-dimethyloctane	4.62	3.97	2.41	1.03	4.76	142.29	0.01029	434.35
3,4-dimethyloctane	4.72	3.51	2.63	1.21	5.76	142.29	0.0	436.58
3,5-dimethyloctane	4.70	3.64	2.28	1.46	5.76	142.29	0.00505	432.55
3,6-dimethyloctane	4.70	3.60	2.55	0.98	5.76	142.29	0.02056	433.95
4,4-dimethyloctane	4.62	4.02	2.13	1.46	4.76	142.29	0.01068	430.65
4,5-dimethyloctane	4.72	3.53	2.46	1.48	5.76	142.29	0.02214	435.25
2-methyl-3-ethylheptane	4.72	3.55	2.40	1.48	5.76	142.29	0.01150	436.15
3-methyl-4-ethylheptane	4.76	3.34	2.63	1.60	5.76	142.29	0.01106	438.15
2-methyl-4-ethylheptane	4.70	3.69	2.08	1.56	5.76	142.29	0.01683	431.15
2-methyl-5-ethylheptane	4.70	3.65	2.31	1.27	5.76	142.29	0.01886	432.85
3-methyl-5-ethylheptane	4.74	3.43	2.59	1.36	5.76	142.29	0.01552	431.35
4-methyl-3-ethylheptane	4.76	3.34	2.60	1.78	5.76	142.29	0.01161	438.15
3-methyl-3-ethylheptane	4.68	3.62	2.84	1.21	4.76	142.29	0.02089	436.95
4-methyl-4-ethylheptane	4.68	3.66	2.56	1.69	4.76	142.29	0.01020	433.95
2,3,4-trimethylheptane	4.59	3.87	2.69	1.38	4.76	142.29	0.01156	434.15
2,4,5-trimethylheptane	4.58	3.99	2.46	1.38	4.76	142.29	0.03244	429.65
2,3,6-trimethylheptane	4.54	4.19	2.34	0.93	4.76	142.29	0.01206	429.15
2,2,3-trimethylheptane	4.48	4.41	2.47	0.94	4.00	142.29	0.01130	430.95
2,2,4-trimethylheptane	4.46	4.66	1.74	1.64	4.00	142.29	0.01148	421.45
2,2,5-trimethylheptane	4.46	4.61	2.08	0.88	4.00	142.29	0.01064	423.95
2,2,6-trimethylheptane	4.42	4.85	1.71	1.04	4.00	142.29	0.01045	421.35
2,3,3-trimethylheptane	4.50	4.25	2.74	0.92	4.00	142.29	0.01947	433.35
3,3,4-trimethylheptane	4.54	4.03	2.97	1.17	4.00	142.29	0.01908	435.05
3,3,5-trimethylheptane	4.52	4.24	2.56	1.23	4.00	142.29	0.02321	428.85
3,4,4-trimethylheptane	4.54	4.05	2.84	1.44	4.00	142.29	0.01116	434.25
2,3,5-trimethylheptane	4.58	3.97	2.61	1.07	4.76	142.29	0.01483	433.85
2,4,6-trimethylheptane	4.52	4.36	1.75	1.71	4.76	142.29	0.01080	420.75
2,5,5-trimethylheptane	4.48	4.45	2.36	0.86	4.00	142.29	0.01910	425.95
3,4,5-trimethylheptane	4.63	3.63	3.08	1.30	4.76	142.29	0.01555	435.65
2,4,4-trimethylheptane	4.48	4.51	1.89	1.88	4.00	142.29	0.01149	424.15
3,3-diethylhexane	4.74	3.31	3.03	1.81	4.76	142.29	0.01908	439.45
2-methyl-3-isopropylhexane	4.59	3.94	2.33	1.77	4.76	142.29	0.01211	438.15
2,2-dimethyl-3-ethylhexane	4.52	4.26	2.35	1.60	4.00	142.29	0.01017	429.25
3,3-dimethyl-4-ethylhexane	4.58	3.86	3.00	1.68	4.00	142.29	0.01175	426.05
2,3-dimethyl-3-ethylhexane	4.57	3.91	3.02	1.45	4.00	142.29	0.01250	439.15
2,3-dimethyl-4-ethylhexane	4.63	3.68	2.84	1.58	4.76	142.29	0.01034	435.15
2,4-dimethyl-4-ethylhexane	4.54	4.12	2.61	1.52	4.00	142.29	0.01599	434.25
3,4-dimethyl-4-ethylhexane	4.60	3.67	3.42	1.44	4.00	142.29	0.02070	435.35
2,2-dimethyl-4-ethylhexane	4.49	4.45	2.06	1.44	4.00	142.29	0.01533	420.15
2,2,3,3-tetramethylhexane	4.31	4.88	2.91	1.00	2.94	142.29	0.00208	433.45
2,2,4,4-tetramethylhexane	4.27	5.26	1.97	1.56	2.94	142.29	0.01959	426.95
2,2,5,5-tetramethylhexane	4.21	5.62	1.63	0.75	2.94	142.29	0.0	410.55
2,3,4,5-tetramethylhexane	4.46	4.21	2.93	1.21	4.00	142.29	0.00232	429.35
2,2,4,5-tetramethylhexane	4.33	4.99	2.07	1.20	3.41	142.29	0.00978	421.05
2,3,3,4-tetramethylhexane	4.43	4.29	3.37	1.10	3.41	142.29	0.00623	437.75
2,2,3,4-tetramethylhexane	4.39	4.54	2.86	1.21	3.41	142.29	0.02459	430.15
2,2,3,5-tetramethylhexane	4.34	4.90	2.30	1.06	3.41	142.29	0.00456	421.56
3,3,4,4-tetramethylhexane	4.37	4.47	3.60	0.96	2.94	142.29	0.0	443.15
2,3,4,4-tetramethylhexane	4.42	4.37	3.14	1.22	3.41	142.29	0.00975	434.75
2,4-dimethyl-3-isopropylpentane	4.46	4.31	2.31	2.31	4.00	142.29	0.01650	430.25
2-methyl-3,3-diethylpentane	4.63	3.58	3.34	1.98	4.00	142.29	0.02505	445.15
2,3,4-trimethyl-3-ethylpentane	4.45	4.17	3.40	1.48	3.41	142.29	0.01775	442.62
2,2,3-trimethyl-3-ethylpentane	4.37	4.52	3.37	1.31	2.94	142.29	0.01324	442.65
2,2,4-trimethyl-3-ethylpentane	4.39	4.62	2.36	2.08	3.41	142.29	0.01796	428.45

[a] Reported data obtained from the following sources: Reid, Prausnitz, and Sherwood, The Property of Gases and Liquids, 1977; Design Institute for Physical Property Data (DIPPR); CRC Press, Inc., Properties of Organic Compounds, − Personal Edition, Version 5.1, 1996. Valance connectivity indices: $^1\chi^v$, $^2\chi^v$, $^3\chi^v$, $^4\chi^v$; second-order kappa index: $^2\kappa$.

Table 2. Molecular Descriptors and Experimental Boiling Points for Alkenes[a]

	input parameters							
name	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	$^2\kappa$	MW	dipole moment (Debye)	output exptl BPs, K
ethylene	0.50	0.00	0.00	0.00	0.00	28.05	0.0	169.40
propene	0.99	0.41	0.00	0.00	2.00	42.08	0.395170	225.75
1,3-butadiene	1.15	0.47	0.17	0.00	3.00	54.09	0.0	268.70
cis-2-butene	1.49	0.67	0.33	0.00	3.00	56.11	0.606490	276.90
trans-2-butene	1.49	0.67	0.33	0.00	3.00	56.11	0.0	274.00
2-methyl-1-propene	1.35	1.21	0.00	0.00	1.33	56.11	0.420820	266.85
trans-1,3-pentadiene	1.65	0.76	0.33	0.14	4.00	68.12	0.407640	315.18
2,3-pentadiene	1.73	0.74	0.33	0.17	4.00	68.12	0.597090	321.40
3-methyl-1,2-butadiene	1.60	1.18	0.35	0.00	2.25	68.12	0.407860	314.00
2-methyl-1,3-butadiene	1.55	1.05	0.35	0.00	2.25	68.12	0.396960	307.20
1-pentene	2.02	1.08	0.49	0.20	4.00	70.13	0.436090	303.12
trans-2-pentene	2.03	0.98	0.47	0.24	4.00	70.13	0.050100	309.50
3-methyl-1-butene	1.90	1.48	0.47	0.00	2.25	70.13	0.447090	293.30
2-methyl-1-butene	1.91	1.31	0.60	0.00	2.25	70.13	0.451750	304.30
cis-1,3,5-hexatriene	1.82	0.86	0.38	0.16	5.00	80.13	0.003850	351.15
1,3-hexadiene	2.19	1.07	0.51	0.23	5.00	82.15	0.444590	346.15
trans,*trans*-2,4-hexadiene	2.15	1.05	0.50	0.22	5.00	82.15	0.0	353.15
1,4-hexadiene	2.14	1.09	0.54	0.23	5.00	82.15	0.358500	266.25
2,3-dimethyl-1,3-butadiene	1.96	1.56	0.73	0.00	2.22	82.15	0.0	341.93
trans-2-hexene	2.53	1.36	0.69	0.33	5.00	84.16	0.049240	314.00
cis-3-hexene	2.56	1.29	0.64	0.33	5.00	84.16	0.588680	339.60
1-hexene	2.52	1.43	0.76	0.35	5.00	84.16	0.439760	336.64
trans-4-methyl-2-pentene	3.44	2.29	1.27	0.88	5.14	84.16	0.060230	331.76
4-methyl-1-pentene	2.38	1.92	0.64	0.33	3.20	84.16	0.429970	327.02
trans-3-methyl-2-pentene	2.43	1.49	1.05	0.08	3.20	84.16	0.381350	343.59
cis-3-methyl-2-pentene	2.43	1.49	1.05	0.08	3.20	84.16	0.360660	340.88
2-methyl-1-pentene	2.41	1.71	0.68	0.43	3.20	84.16	0.465130	335.25
2,3-dimethyl-1-butene	2.30	2.00	0.99	0.00	2.22	84.16	0.433540	328.75
3,3-dimethyl-1-butene	2.20	2.57	0.61	0.00	1.63	84.16	0.420670	314.45
1,3,5-heptatriene	2.32	1.15	0.55	0.25	6.00	94.16	0.407620	396.80
1,5-heptadiene	2.64	1.44	0.77	0.38	6.00	96.17	0.396630	367.15
1,6-heptadiene	2.63	1.51	0.82	0.41	6.00	96.17	0.338520	362.15
2,4-heptadiene	2.69	1.36	0.68	0.33	6.00	96.17	0.063930	381.15
2-methyl-2,4-hexadiene	2.57	1.53	0.88	0.36	4.17	96.17	0.387700	384.65
2,4-dimethyl-1,3-pentadiene	2.43	2.07	0.54	0.49	3.06	96.17	0.703440	366.35
trans-3-heptene	3.06	1.67	0.86	0.45	6.00	98.19	0.007860	368.82
1-heptene	3.02	1.78	1.01	0.54	6.00	98.19	0.437070	366.75
trans-2-heptene	3.03	1.71	0.96	0.49	6.00	98.19	0.055190	371.15
2-methyl-2-hexene	2.90	2.07	0.84	0.43	4.17	98.19	0.309100	368.55
4,4-dimethyl-1-pentene	2.67	3.05	0.76	0.43	2.34	98.19	0.438570	345.65
5-methyl-1-hexene	2.88	2.26	0.95	0.45	4.17	98.19	0.440170	358.45
cis-5-methyl-2-hexene	2.88	2.20	0.84	0.41	4.17	98.19	0.583570	362.65
trans-5-methyl-2-hexene	2.88	2.20	0.84	0.41	4.17	98.19	0.052380	361.25
3-ethyl-1-pentene	2.97	1.81	1.38	0.62	4.17	98.19	0.449350	357.25
cis-3-methyl-3-hexene	2.96	1.81	1.11	0.49	4.17	98.19	0.340140	368.55
3-ethyl-2-pentene	2.99	1.65	1.32	0.66	4.17	98.19	0.404800	369.15
4-methyl-1-hexene	2.92	2.04	1.27	0.45	4.17	98.19	0.439860	359.85
trans-4-methyl-2-hexene	2.94	1.91	1.16	0.46	4.17	98.19	0.045210	360.75
trans-3-methyl-3-hexene	2.96	1.81	1.11	0.49	4.17	98.19	0.328720	366.65
2-methyl-1-hexene	2.91	2.06	0.96	0.48	4.17	98.19	0.462870	365.15
2,4-dimethyl-1-pentene	2.77	2.56	0.76	0.70	3.06	98.19	0.467020	354.75
2,3,3-trimethyl-1-butene	2.60	3.03	1.28	0.00	1.85	98.19	0.427780	351.05
3,3-dimethyl-1-pentene	2.76	2.55	1.46	0.14	2.34	98.19	0.456280	350.65
cis-4,4-dimethyl-2-pentene	2.70	2.87	0.67	0.50	2.34	98.19	0.657170	353.55
2,4-dimethyl-2-pentene	2.78	2.49	0.67	0.67	3.06	98.19	0.365910	356.55
2,3-dimethyl-1-pentene	2.83	2.16	1.45	0.35	3.06	98.19	0.456520	357.45
cis-3,4-dimethyl-2-pentene	2.81	2.19	1.37	0.33	3.06	98.19	0.335810	362.45
trans-3,4-dimethyl-2-pentene	2.81	2.19	1.37	0.33	3.06	98.19	0.399420	365.65
2,4,6-octatriene	2.82	1.44	0.72	0.35	7.00	108.18	0.0	420.65
5-methyl-1,3,6-heptatriene	2.73	1.72	0.92	0.33	5.14	108.18	0.455250	390.15
2,6-octadiene	3.14	1.72	0.97	0.53	7.00	110.20	0.014110	397.65
2,5-dimethyl-2,4-hexadiene	2.91	2.49	0.75	0.33	3.94	110.20	0.0	407.65
3-methyl-1,5-heptadiene	3.05	2.02	1.15	0.45	5.14	110.20	0.414580	373.15
cis-2-octene	3.53	2.06	1.21	0.68	7.00	112.22	0.614440	398.75
trans-2-octene	3.53	2.06	1.21	0.68	7.00	112.22	0.067350	398.15
1-octene	3.52	2.14	1.26	0.72	7.00	112.22	0.440010	394.35
6-methyl-1-heptene	3.38	2.61	1.19	0.67	5.14	112.22	0.438380	386.35
6-methyl-3-heptene	3.42	2.51	1.01	0.54	5.14	112.22	0.011660	388.15
6,6-dimethyl-2-hexene	3.17	3.33	0.97	0.47	3.11	112.22	0.588150	380.05
2,3-dimethyl-2-hexene	3.31	2.53	1.48	0.63	3.94	112.22	0.041900	394.95
2-methyl-1-heptene	3.41	2.41	1.21	0.68	5.14	112.22	0.466000	392.37
cis-3-methyl-2-heptene	3.43	2.24	1.42	0.60	5.14	112.22	0.358650	385.15
3-methyl-3-heptene	3.46	2.19	1.34	0.54	5.14	112.22	0.333060	394.15
4-methyl-1-heptene	3.42	2.42	1.35	0.90	5.14	112.22	0.436990	385.95

Table 2 (Continued)

	input parameters							
name	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	$^2\kappa$	MW	dipole moment (Debye)	output exptl BPs, K
5-methyl-1-heptene	3.42	2.38	1.56	0.61	5.14	112.22	0.436570	386.45
5-methyl-2-heptene	3.42	2.32	1.47	0.54	5.14	112.22	0.056680	391.15
2,3-dimethyl-1-hexene	3.33	2.54	1.56	0.68	3.94	112.22	0.454600	383.65
2,4-dimethyl-1-hexene	3.31	2.67	1.39	0.74	3.94	112.22	0.464020	384.35
1,8-nonadiene	3.63	2.22	1.32	0.75	8.00	124.23	0.338250	415.65
7-methyl-2,4-octadiene	3.55	2.58	1.05	0.54	6.13	124.23	0.061550	422.15
2,4-dimethyl-2,4-heptadiene	3.48	2.58	1.04	0.84	4.84	124.23	0.728320	411.15
2,6-dimethyl-1,3-heptadiene	3.45	2.88	1.01	0.54	4.84	124.23	0.493010	414.15
2,6-dimethyl-2,4-heptadiene	3.44	2.87	0.93	0.47	4.84	124.23	0.367030	413.15
3-ethyl-2-methyl-1,5-hexadiene	3.48	2.48	1.64	0.92	4.84	124.23	0.314350	418.15
1-nonene	4.02	2.49	1.51	0.89	8.00	126.24	0.436950	420.05
trans-3-nonene	4.06	2.38	1.38	0.80	8.00	126.24	0.008880	420.65
trans-4-nonene	4.06	2.40	1.35	0.72	8.00	126.24	0.007700	416.15
7-methyl-3-octene	3.92	2.58	1.31	0.71	6.13	126.24	0.012260	415.15
2-merthyl-4-octene	3.92	2.89	1.23	0.65	6.13	126.24	0.005550	412.15
3-methyl-2-octene	3.93	2.59	1.67	0.80	6.13	126.24	0.361780	418.15
2,3-dimethyl-2-heptene	3.81	2.88	1.76	0.69	4.84	126.24	0.047150	418.35
1,3-decadiene	4.19	2.51	1.50	0.88	9.00	138.25	0.452650	442.15
1-decene	4.52	2.85	1.76	1.07	9.00	140.27	0.439880	443.65
cis-5-decene	4.56	2.76	1.62	0.88	9.00	140.27	0.625140	444.15
2-methyl-1-nonene	4.41	3.12	1.71	1.03	7.11	140.27	0.451020	441.55
4-propyl-3-heptene	4.53	2.77	1.64	1.26	7.11	140.27	0.385980	433.65
2-methyl-3-nonene	4.44	3.17	1.50	0.88	7.11	140.27	0.042890	434.15
propadiene	0.71	0.25	0.00	0.00	2.00	40.06	0.399910	238.70
cis-2-pentene	2.03	0.98	0.47	0.24	4.00	70.13	0.612940	310.09
2-methyl-2-butene	1.87	1.37	0.58	0.00	2.25	70.13	0.342690	311.72
cis,*trans*-2,4-hexadiene	2.15	1.05	0.50	0.22	5.00	82.15	0.626670	356.65
1,5-hexadiene	2.13	1.15	0.57	0.24	5.00	82.15	0.033100	332.60
trans-3-hexene	2.56	1.29	0.64	0.33	5.00	84.16	0.018680	340.30
3-methyl-1-pentene	2.43	1.62	1.05	0.17	3.20	84.16	0.444640	327.31
2,3-dimethyl-2-butene	2.25	2.00	1.00	0.00	2.22	84.16	0.0	346.45
1,2-butadiene	1.22	0.49	0.20	0.00	3.00	54.09	0.0	284.00
2,5-heptadiene	2.64	1.37	0.74	0.35	6.00	96.17	0.107560	381.15
2-ethyl-1-pentene	2.97	1.85	1.10	0.53	4.17	98.19	0.507640	368.15
1,2-pentadiene	1.76	0.82	0.35	0.14	4.00	68.12	0.414410	318.00
trans-4,4-dimethyl-2-pentene	2.70	2.87	0.67	0.50	2.34	98.19	0.074210	349.85
2,5-dimethyl-1,3,5-hexatriene	2.62	2.03	0.65	0.28	3.94	108.18	0.0	419.15
1,7-octadiene	3.13	1.86	1.07	0.58	7.00	110.20	0.033700	388.65
2,5-dimethyl-1,5-hexadiene	2.91	2.41	0.98	0.43	3.94	110.20	0.0	387.45
trans-4-methyl-2-heptene	3.44	2.29	1.27	0.88	5.14	112.22	0.060230	387.15
2,4-dimethyl-2-hexene	3.31	2.63	1.26	0.69	3.94	112.22	0.337380	387.75
2-methyl-2-heptene	3.40	2.42	1.11	0.59	5.14	112.22	0.325760	395.75
4-methyl-3,5-octadiene	3.64	2.17	1.14	0.69	6.13	124.23	0.387970	422.65
5-methyl-3-heptene	3.47	2.22	1.34	0.50	5.14	112.22	0.034980	385.15
2,7-nonadiene	3.64	2.07	1.21	0.69	8.00	124.23	0.032850	424.65
2,6-dimethyl-2,5-heptadiene	3.39	2.80	1.05	0.53	4.84	124.23	0.570300	423.65
2-methyl-1-octene	3.91	2.77	1.46	0.85	6.13	126.24	0.463150	417.95
3,7-dimethyl-1-octene	4.29	3.54	1.84	0.97	5.76	140.27	0.417780	427.15
4-decene	4.56	2.76	1.60	0.91	9.00	140.27	0.001120	443.75
1-butene	1.52	0.70	0.29	0.00	3.00	56.11	0.399050	266.85
trans-5-decene	4.56	2.76	1.62	0.88	9.00	140.27	0.011240	443.75
cis-1,3-pentadiene	1.65	0.76	0.33	0.14	4.00	68.12	0.383060	317.22
trans-1,3,5-hexatriene	1.82	0.86	0.38	0.16	5.00	80.13	0.0	351.65
cis-2-hexene	2.53	1.36	0.69	0.33	5.00	84.16	0.583630	342.00
2-methyl-2-pentene	2.40	1.69	0.61	0.41	3.20	84.16	0.336030	340.50
cis-2-heptene	3.03	1.71	0.96	0.49	6.00	98.19	0.590510	372.15
cis-2-methyl-3-hexene	2.94	2.08	0.76	0.41	4.17	98.19	0.616080	359.15
3-methyl-1-hexene	2.93	2.00	1.15	0.57	4.17	98.19	0.452840	357.05
3,4-dimethyl-1-pentene	2.81	2.34	1.42	0.27	3.06	98.19	0.449450	353.95
2,3-dimethyl-2-pentene	2.81	2.13	1.38	0.35	3.06	98.19	0.051280	370.65
1,4-pentadiene	1.63	0.81	0.33	0.12	4.00	68.12	0.252930	299.10
1,2-hexadiene	2.26	1.20	0.58	0.25	5.00	82.15	0.493800	349.15
cis-3-heptene	3.06	1.67	0.86	0.45	6.00	98.19	0.625370	368.95
trans-2-methyl-3-hexene	2.94	2.08	0.76	0.41	4.17	98.19	0.038780	359.05
1,4-heptadiene	2.67	1.40	0.71	0.36	6.00	96.17	0.408330	366.15
6-methyl-2-heptene	3.38	2.54	1.15	0.60	5.14	112.22	0.056860	390.15
4-methyl-3-heptene	3.46	2.21	1.20	0.84	5.14	112.22	0.343690	394.15
3,5-dimethyl-2,4-heptadiene	3.50	2.38	1.49	0.57	4.84	124.23	0.747150	403.15
2,6-dimethyl-1,5-heptadiene	3.40	2.78	1.12	0.61	4.84	124.23	0.649970	416.15
2,6-dimethyl-2-octene	4.30	3.37	1.91	0.86	5.76	140.27	0.319550	413.42

[a] Valance connectivity indices: $^1\chi^v$,$^2\chi^v$,$^3\chi^v$,$^4\chi^v$; second-order kappa index: $^2\kappa$. Reported data obtained from Reid et al. (1977); DIPPR (1996); and Properties of Organic Compounds, CRC Press, Inc. (1996).

Table 3. Molecular Descriptors and Experimental and Predicted Boiling Points for alkynes[a] Using Back-Propagation and Fuzzy ARTMAP Algorithms

	inputs parameters							BPs, K				
									estimated			
name	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	$^2\kappa$	MW	dipole moment (Debye)	reported[b]	back-propagation[c]		Fuzzy ARTMAP[d]	
acetylene	0.33	0.00	0.00	0.00	0.00	26	0.0	189.15	tr	169.81	tr	189.14
propyne	0.79	0.29	0.00	0.00	2.00	40	0.7750	250.00	tr	231.40	te	249.00
1-butyne	1.35	0.56	0.20	0.00	3.00	54	0.8353	281.20	tr	286.54	tr	281.21
2-butyne	1.25	0.50	0.25	0.00	3.00	54	0.0010	300.20	s	281.30	tr	300.21
1,3-pentadiyne	1.29	0.52	0.20	0.07	4.00	64	1.3206	328.15	tr	323.57	te	327.03
1-pentyne	1.85	0.95	0.39	0.14	4.00	68	0.9199	313.30	tr	315.43	tr	313.29
2-pentyne	1.81	0.78	0.35	0.18	4.00	68	0.1035	329.25	s	314.11	tr	327.03
3-methyl-1-butyne	1.73	1.32	0.33	0.00	2.25	68	1.5836	299.45	tr	308.98	tr	299.10
1-hexyne	2.35	1.31	0.67	0.28	5.00	82	0.9246	344.45	tr	337.87	tr	343.16
2-hexyne	2.31	1.18	0.55	0.25	5.00	82	0.1129	357.65	s	345.29	tr	356.55
3-hexyne	2.37	1.06	0.48	0.25	5.00	82	0.1725	354.15	tr	351.14	te	351.05
3-methyl-1-pentyne	2.27	1.48	0.90	0.12	3.20	82	0.9549	330.85	tr	334.54	tr	330.84
4-methyl-1-pentyne	2.21	1.80	0.53	0.24	3.20	82	0.9176	334.35	s	335.35	te	334.33
3,3-dimethyl-1-butyne	2.04	2.39	0.43	0.00	1.63	82	0.9726	310.85	tr	309.21	tr	309.24
4-methyl-2-pentyne	2.19	1.55	0.43	0.29	3.20	82	0.1203	346.25	tr	336.93	te	343.|16
3-heptyne	2.87	1.46	0.68	0.34	6.00	96	0.1814	380.35	tr	384.75	tr	377.73
2-heptyne	2.81	1.53	0.83	0.39	6.00	96	0.1164	385.15	tr	375.65	te	382.27
3-methyl-1-hexyne	2.77	1.86	1.01	0.52	4.17	96	0.9597	358.15	tr	361.90	tr	356.55
5-methyl-2-hexyne	2.67	2.02	0.69	0.31	4.17	96	0.1205	375.65	v	368.12	te	374.65
5-methyl-1-hexyne	2.71	2.14	0.87	0.37	4.17	96	0.9246	365.15	tr	360.91	tr	365.00
3,4-dimethyl-1-pentyne	2.64	2.20	1.26	0.19	3.06	96	0.9530	353.15	tr	354.08	tr	351.05
3-ethyl-1-pentyne	2.81	1.68	1.22	0.52	4.17	96	0.9622	369.32	tr	362.63	tr	368.15
1-heptyne	2.85	1.66	0.92	0.48	6.00	96	0.9210	372.85	tr	365.55	tr	370.64
3,3-dimethyl-1-pentyne	2.60	2.38	1.27	0.10	2.34	96	0.9778	343.15	tr	350.00	tr	343.16
4,4-dimethyl-1-pentyne	2.50	2.94	0.63	0.31	2.34	96	0.9186	349.25	s	362.60	tr	349.15
4-methyl-1-hexyne	2.74	1.92	1.16	0.35	4.17	96	0.9221	364.15	tr	358.53	tr	360.74
2-methyl-3-hexyne	2.75	1.83	0.57	0.31	4.17	96	0.1413	368.35	tr	376.04	tr	368.15
4,4-dimethyl-2-pentyne	2.50	2.63	0.50	0.38	2.34	96	0.1255	356.15	tr	352.61	te	354.75
1-octyne	3.35	2.01	1.17	0.65	7.00	110	0.9247	399.45	tr	392.58	tr	398.15
2-octyne	3.31	1.88	1.08	0.59	7.00	110	0.1185	410.75	tr	401.42	te	410.20
3-octyne	3.37	1.81	0.96	0.48	7.00	110	0.1772	406.25	s	406.61	tr	404.56
3-methyl-3-ethyl-1-pentyne	3.16	2.41	1.91	0.45	3.11	110	0.9827	374.65	tr	379.74	tr	374.65
4-octyne	3.37	1.85	0.88	0.43	7.00	110	0.1894	404.75	tr	410.19	tr	404.56
1-nonyne	3.85	2.37	1.42	0.83	8.00	124	0.9211	423.95	tr	420.93	tr	420.65
2-nonyne	3.81	2.24	1.33	0.77	8.00	124	0.1125	435.05	tr	426.83	tr	435.05
4-nonyne	3.87	2.21	1.16	0.57	8.00	124	0.1853	427.15	s	429.07	tr	426.95
3-nonyne	3.87	2.16	1.21	0.68	8.00	124	0.1831	418.35	tr	428.31	tr	417.53
1-decyne	4.35	2.72	1.67	1.01	9.00	138	0.9246	447.15	tr	442.96	tr	447.16
2-decyne	4.31	2.59	1.58	0.94	9.00	138	0.1149	457.75	tr	443.85	tr	457.75
3-decyne	4.37	2.52	1.46	0.85	9.00	138	0.1770	450.15	s	445.67	tr	447.16
4-decyne	4.37	2.56	1.41	0.77	9.00	138	0.1948	448.05	v	444.83	tr	447.16
5-decyne	4.37	2.56	1.44	0.71	9.00	138	0.1802	450.15	tr	444.26	tr	447.16
8-methyl-4-nonyne	4.23	3.04	1.35	0.66	7.11	138	0.1009	377.75	tr	379.12	tr	377.73

[a] tr = training set, te = test set, v = validation set; valance connectivity indices: $^1\chi^v$, $^2\chi^v$, $^3\chi^v$, $^4\chi^v$; second-order kappa index: $^2\kappa$. [b] Reid et al. (1977); DIPPR (1996); and Properties of Organic Compounds, CRC Press, Inc. (1996). [c] Boiling points estimated from back-propagation composite model 7−9−1. [d] Boiling points estimated from composite Fuzzy ARTMAP model.

propagation network models. Separate training, testing, and validation data sets were generated for each of the boiling point models. For the alkane model, the training, testing, and validation data sets numbered 92, 28, and 20 compounds, respectively. For the alkene model, the training set contained 97 compounds with 26 and 21 compounds for testing and validation, respectively. The alkyne data set, which contained only 43 compounds, was deemed too small for developing an independent QSPR. However, a model based on the composite data set of 327 compounds was developed with 228 compounds in the training set and 67 and 32 compounds in the testing and validation sets, respectively. The training set consisted of 97 alkanes, 101 alkenes, and 30 alkynes, the test set included 28 alkanes, 30 alkenes, and 9 alkynes, and the validation set consisted of 15 alkanes, 13 alkenes, and 4 alkynes. Training, testing, and validation sets were selected randomly after normalizing the data from 0 to 1 using the NeuralSim software.[38]

Model building proceeded by using a back-propagation neural network with an initial architecture of one input layer, with seven inputs, one hidden layer, and one output layer, with one output. Subsequently, using the optimization capability of NeuralSim, a cascade method of network construction, together with an adaptive gradient learning rule were used to build an initial range of possible architectures. The performance of neural network models generated based on a number of possible architectures was evaluated by testing and validating the models. Of the top performing models, the one with the least number of hidden units was selected for final optimization using the NeuralWorks Professional II/PLUS system.[39] To prevent over fitting, the final model was built using a train/test method. The

Table 4. Neural Network/QSPR Performance for Boiling Point Prediction of Aliphatic Hydrocarbons (327 Compounds) Using Back-Propagation and Fuzzy ARTMAP Networks

data set	no. of records	contribution to total absolute av error		standard dev		av absolute error		max. absolute error	
		K	percent, %	K	percent, %	K	percent, %	K	percent, %
Back-Propagation Model: 7-4-1 with 140 Alkanes									
all data	140			1.53	0.35	1.54	0.37	10.71	2.51
training	92			1.44	0.33	1.47	0.56	8.17	1.86
testing	28			1.96	0.46	1.65	0.40	10.71	2.51
validation	20			1.24	0.28	1.73	0.42	5.01	1.17
Fuzzy ARTMAP Model: with 140 Alkanes									
all data	140			1.13	0.27	1.30	0.31	3.36	0.88
training	112			1.10	0.27	1.39	0.31	3.36	0.88
testing	28			1.26	0.27	0.81	0.19	3.3	0.77
Back-Propagation Model: 7-10-1 with 144 Alkene Hydrocarbons									
all data	144			3.76	1.14	4.42	1.25	19.57	7.11
training	97			3.00	8.84	3.34	0.93	16.33	4.37
testing	26			4.17	1.50	6.79	1.96	19.57	7.11
validation	21			4.18	1.21	6.45	1.83	18.18	4.39
Fuzzy ARTMAP Model: with 144 Alkene Hydrocarbons									
all data	144			0.92	0.25	0.95	0.25	2.73	0.91
training	116			0.91	0.25	0.99	0.26	2.73	0.91
testing	26			0.94	0.24	0.73	0.19	2.71	0.71
Back-Propagation Composite Model: 7-9-1 with 327 Aliphatic Hydrocarbons									
all data	327	total = 4.85 K	total = 1.37%	4.45	1.46	4.85	1.37	20.92	9.75
alkanes	140	1.22	0.32	2.46	0.76	2.86	0.75	15.5	5.18
alkenes	144	2.72	0.78	5.06	1.59	6.16	1.77	20.92	7.05
alkynes	43	0.89	0.27	4.76	1.89	6.85	2.04	18.99	9.74
training	228	total = 4.51 K	total = 1.28%	4.20	1.41	4.51	1.28	20.92	9.74
alkanes	97	1.10	0.30	2.05	0.72	2.60	0.71	9.45	5.18
alkenes	101	2.61	0.75	5.06	1.61	5.85	1.68	20.92	7.05
alkynes	30	0.80	0.23	3.69	1.65	6.08	1.79	18.33	9.74
testing	67	total = 6.09 K	total = 1.74%	5.33	1.74	6.09	1.74	19.99	7.44
alkanes	28	1.43	0.36	2.76	0.77	3.41	0.87	10.35	3.25
alkenes	30	3.45	0.99	5.40	1.67	7.72	2.21	19.99	6.95
alkynes	9	1.21	0.38	7.14	2.60	9.00	2.88	18.89	7.44
validation	32	total = 4.68 K	total = 1.25%	3.68	1.00	4.68	1.25	15.54	3.78
alkanes	15	1.67	0.41	3.68	0.90	3.56	0.88	15.54	3.78
alkenes	13	2.03	0.59	3.16	0.94	5.00	1.45	10.27	3.03
alkynes	4	0.98	0.25	3.24	0.97	7.82	2.02	12.36	3.45
Fuzzy ARTMAP Model Composite Model: with 327 Aliphatic Hydrocarbons									
all data	327	total = 1.35 K	total = 0.35%	1.15	0.27	1.35	0.35	3.45	0.98
alkanes	140	0.54	0.15	1.06	0.29	1.25	0.34	3.42	0.98
alkenes	144	0.66	0.16	1.10	0.26	1.51	0.37	3.45	0.83
alkynes	43	0.16	0.04	1.16	0.31	1.20	0.32	3.41	0.94
training	228	total = 1.44 K	total = 0.35%	1.09	0.27	1.44	0.35	3.45	0.92
alkanes	112	0.66	0.17	1.03	0.27	1.34	0.35	3.42	0.98
alkenes	114	0.78	0.19	1.11	0.26	1.56	0.37	3.45	0.83
alkynes	34	.20	0.05	1.19	0.31	1.33	0.35	3.41	0.94
testing	67	total = 1.15 K	total = 0.84%	1.08	0.31	1.15	0.84	3.42	0.98
alkanes	28	0.39	0.13	1.10	0.35	0.94	0.30	2.99	0.98
alkenes	30	0.56	0.16	1.02	0.29	1.24	0.36	2.98	0.83
alkynes	9	0.08	0.02	0.8	0.24	0.56	0.17	2.2	0.67

hyperbolic tangent transfer functions were chosen to correlate weighted inputs and outputs of the hidden layer. To improve the separation among data points, the data sets were rescaled to fall between −1 and +1. The extended-delta-bar-delta rule was used in building the final model with a momentum rate set to 0.4. The above approach was also conducted, for comparison, using back-propagation and cascade correlation codes developed in our laboratories.

To improve results obtained with back-propagation, a neural system based on Fuzzy ARTMAP neural network was implemented to provide a prediction of the boiling point as output, instead of simply classifying data.[25,26] In developing the Fuzzy ARTMAP models, boiling points and molecular descriptors were divided into two data sets: training and testing. In this case the test set served to validate the resulting model since training is considered acceptable when the number of classes generated is sufficient to explain the data within experimental error. For the alkane model, these two sets contained 112 and 28 compounds, respectively. In the alkene model a number of 116 compounds were selected for training and 26 for testing. Finally 228 compounds were selected for training and 67 for testing in the composite model. The corresponding compounds for each set are respectively reported in Tables 5, 7, and 3. Molecular descriptors and their corresponding boiling points were preprocessed (normalization and complement coding) prior to its processing by modules ART_a and ART_b (as input and output vectors) for training. The training process evolved according to the set fuzzy rules of classification of input and output patterns in each ART module until stability of classes was reached. Once the network was trained, the ART_b module, i.e., the module that classified boiling points, was

Table 5. Experimental and Predicted Boiling Points for Alkanes Using Back-Propagation and Fuzzy ARTMAP Algorithms

	BPs,[a] K								
		estimated: back-propagation				estimated: Fuzzy ARTMAP			
name	reported points[b]		alkanes (7−4−1)		composite (7−9−1)		alkanes		composite
methane	111.66	tr	111.70	tr	113.95	tr	111.66	tr	111.66
ethane	182.52	te	182.40	tr	191.98	te	182.54	te	182.50
propane	231.08	tr	231.10	tr	236.05	tr	231.10	tr	231.10
butane	272.65	tr	274.00	te	281.52	tr	272.64	tr	272.66
2-methylpropane	261.42	v	260.10	tr	269.29	tr	261.43	tr	261.41
pentane	309.24	tr	308.00	tr	304.86	tr	309.26	tr	309.24
2-methylbutane	301.00	tr	302.00	tr	299.83	tr	301.00	tr	300.21
2,2-dimethylpropane	282.68	tr	283.90	tr	280.16	tr	282.68	tr	281.21
hexane	341.89	tr	341.00	tr	343.86	tr	341.87	tr	339.59
3-methylpentane	336.43	tr	336.00	v	341.20	tr	336.44	tr	334.33
2-methylpentane	333.42	v	331.90	tr	332.31	tr	331.14	tr	330.84
2,2-dimethylbutane	322.89	te	324.30	te	316.88	te	322.88	tr	321.39
2,3-dimethylbutane	331.14	tr	329.70	tr	330.80	tr	331.14	tr	330.84
heptane	371.58	tr	371.70	tr	378.62	tr	371.58	te	370.64
2-methylhexane	363.20	tr	362.50	tr	361.60	tr	362.95	tr	360.74
3-methylhexane	365.00	v	364.40	te	362.49	tr	365.00	tr	365.00
3-ethylpentane	366.63	tr	365.70	tr	369.07	tr	365.00	te	365.00
2,2-dimethylpentane	352.34	te	351.20	v	347.64	te	352.35	tr	351.05
2,3-dimethylpentane	362.93	te	361.00	tr	360.72	te	362.95	tr	360.74
3,3-dimethylpentane	359.21	tr	359.90	te	359.86	tr	359.22	te	356.55
2,4-dimethylpentane	353.65	tr	353.70	tr	351.42	tr	352.35	tr	351.05
2,2,3-trimethylbutane	354.03	tr	352.40	tr	351.26	tr	352.35	tr	351.05
octane	398.82	tr	401.10	tr	395.72	tr	398.82	tr	398.15
2-methylheptane	390.80	te	390.50	tr	385.42	te	390.80	tr	390.16
3-methylheptane	392.08	tr	393.10	te	389.60	tr	390.80	tr	390.16
4-methylheptane	390.85	tr	390.00	tr	383.84	tr	390.80	te	390.16
3-ethylhexane	391.68	tr	392.30	tr	384.65	tr	390.80	te	390.16
2,2-dimethylhexane	379.99	te	378.50	tr	384.32	te	379.63	te	377.73
2,3-dimethylhexane	388.76	te	387.30	tr	388.05	te	386.61	te	385.94
2,4-dimethylhexane	382.58	tr	384.50	v	379.46	tr	379.63	tr	382.27
2,5-dimethylhexane	382.25	tr	382.80	te	383.37	tr	379.63	tr	382.27
3,4-dimethylhexane	390.88	te	390.70	tr	385.95	te	390.80	tr	390.16
2,2,3-trimethylpentane	382.99	tr	383.10	tr	380.27	tr	379.63	tr	382.27
2,2,4-trimethylpentane	372.39	tr	372.10	tr	374.28	tr	372.38	tr	370.64
2,3,3-trimethylpentane	387.91	te	387.00	tr	384.41	te	386.61	tr	385.94
2,3,4-trimethylpentane	386.62	tr	385.10	te	384.01	tr	386.61	te	385.94
3,3-dimethylhexane	385.12	tr	385.20	tr	382.79	tr	385.13	tr	382.27
2-methyl-3-ethylpentane	388.80	tr	388.80	tr	388.75	tr	386.61	te	385.94
3-methyl-3-ethylpentane	391.41	v	392.40	v	390.98	tr	390.80	te	390.16
2,2,3,3-tetramethylbutane	379.62	tr	377.70	tr	378.56	tr	379.63	te	377.73
nonane	423.95	tr	423.30	tr	417.88	tr	423.96	tr	420.65
2-methyloctane	416.41	tr	416.50	tr	414.90	tr	414.66	te	413.76
3-methyloctane	417.55	te	418.80	te	421.08	te	414.66	te	417.53
4-methyloctane	415.57	tr	416.20	tr	417.12	tr	414.66	tr	413.76
3-ethylheptane	416.15	tr	418.00	tr	413.78	tr	414.66	tr	413.76
4-ethylheptane	414.35	tr	416.20	tr	412.77	tr	411.05	te	413.76
2,2-dimethylheptane	405.84	te	403.70	tr	396.91	te	404.56	tr	404.56
3,3-dimethylheptane	410.45	tr	410.10	tr	410.77	tr	408.36	te	410.20
4,4-dimethylheptane	408.35	tr	407.50	te	402.76	tr	408.36	tr	408.36
2,3-dimethylheptane	413.65	tr	413.30	tr	412.63	tr	411.05	tr	410.20
2,4-dimethylheptane	406.05	te	407.60	te	401.67	te	404.56	tr	408.36
3,4-dimethylheptane	413.75	te	414.00	tr	417.44	te	411.05	tr	413.76
2-methyl-4-ethylhexane	406.95	tr	410.70	tr	411.04	tr	404.56	tr	404.56
2-methyl-3-ethylhexane	411.15	tr	412.00	tr	413.89	tr	411.05	te	410.20
3-methyl-3-ethylhexane	413.12	tr	414.60	te	419.05	tr	411.05	tr	410.20
3-methyl-4-ethylhexane	413.15	te	415.30	tr	416.87	te	411.05	te	410.20
2,3,4-trimethylhexane	412.25	tr	412.90	tr	416.41	tr	411.05	te	410.20
2,3,5-trimethylhexane	404.55	tr	406.40	v	404.91	tr	404.56	tr	404.56
2,2,3-trimethylhexane	406.75	tr	406.70	tr	407.85	tr	404.56	te	404.56
2,2,4-trimethylhexane	399.69	te	401.10	te	401.70	te	398.82	tr	398.15
2,2,5-trimethylhexane	397.15	te	398.30	tr	401.67	te	395.33	te	394.14
2,3,3-trimethylhexane	410.85	tr	410.00	te	411.61	tr	408.36	te	410.20
3,3,4-trimethylhexane	413.65	tr	414.80	v	417.16	tr	411.05	tr	410.20
3,3-diethylpentane	419.45	te	420.80	tr	420.19	te	419.46	tr	417.53
2,3,3,4-tetramethylpentane	414.65	te	413.50	tr	412.49	te	414.66	tr	413.76
2,2,3,3-tetramethylpentane	413.35	te	411.70	te	408.07	te	411.05	te	410.20
2,2,4,4-tetramethylpentane	395.35	tr	391.20	tr	395.26	tr	395.33	te	394.14
2,2,3,4-tetramethylpentane	406.15	tr	407.00	v	407.00	tr	404.56	tr	404.56
2,2-dimethyl-3-ethylpentane	406.95	tr	410.40	tr	407.53	tr	404.56	te	404.56
2,3-dimethyl-3-ethylpentane	417.85	te	417.50	te	420.83	te	414.66	tr	417.53
2,4-dimethyl-3-ethylpentane	409.85	tr	411.10	tr	411.20	tr	408.36	te	408.36
decane	447.25	tr	440.60	tr	442.58	tr	447.25	te	447.16

Table 5 (Continued)

name	BPs,[a] K: reported points[b]	estimated: back-propagation alkanes (7−4−1)		estimated: back-propagation composite (7−9−1)		estimated: Fuzzy ARTMAP alkanes		estimated: Fuzzy ARTMAP composite	
2-methylnonane	440.15	tr	438.60	tr	436.89	tr	438.86	tr	438.85
3-methylnonane	440.95	v	440.50	tr	439.07	tr	438.86	te	438.85
4-methylnonane	438.85	v	437.90	te	437.25	tr	438.86	tr	438.85
5-methylnonane	438.30	tr	439.00	tr	437.22	tr	435.03	tr	435.05
3-ethyloctane	439.65	tr	439.50	tr	439.33	tr	438.86	tr	438.85
4-ethyloctane	436.85	tr	438.50	te	437.85	tr	435.03	tr	435.05
4-isopropylheptane	432.05	tr	432.60	tr	433.44	tr	431.14	tr	431.14
4-propylheptane	435.15	tr	436.50	tr	431.64	tr	430.64	tr	430.65
2,2-dimethyloctane	428.15	v	426.50	tr	424.41	tr	426.95	tr	426.95
2,3-dimethyloctane	437.45	v	435.40	v	435.45	tr	435.03	tr	435.05
2,4-dimethyloctane	429.15	tr	430.10	tr	430.68	tr	426.95	te	426.95
2,5-dimethyloctane	431.65	tr	431.80	tr	432.13	tr	431.14	tr	431.14
2,6-dimethyloctane	433.55	v	433.80	te	433.53	tr	431.14	te	431.14
2,7-dimethyloctane	433.05	tr	431.10	tr	429.96	tr	431.14	tr	431.14
3,3-dimethyloctane	434.35	tr	432.10	tr	430.80	tr	431.14	tr	431.14
3,4-dimethyloctane	436.58	tr	436.20	v	438.42	tr	435.03	tr	435.05
3,5-dimethyloctane	432.55	v	432.90	tr	435.31	tr	431.14	te	431.14
3,6-dimethyloctane	433.95	tr	436.80	tr	437.51	tr	431.14	tr	431.14
4,4-dimethyloctane	430.65	tr	428.80	te	428.28	tr	430.64	te	430.65
4,5-dimethyloctane	435.25	te	434.10	tr	434.50	te	435.03	te	435.05
2-methyl-3-ethylheptane	436.15	tr	433.70	tr	436.64	tr	435.03	tr	435.05
3-methyl-4-ethylheptane	438.15	v	435.30	te	438.60	tr	435.03	tr	435.05
2-methyl-4-ethylheptane	431.15	tr	431.20	tr	432.99	tr	431.14	te	431.14
2-methyl-5-ethylheptane	432.85	tr	433.90	tr	435.33	tr	431.14	te	431.14
3-methyl-5-ethylheptane	431.35	te	435.60	v	438.17	te	431.14	te	431.14
4-methyl-3-ethylheptane	438.15	tr	434.70	tr	437.23	tr	435.03	te	435.05
3-methyl-3-ethylheptane	436.95	tr	435.50	tr	436.99	tr	435.03	tr	435.05
4-methyl-4-ethylheptane	433.95	tr	433.20	tr	434.45	tr	431.14	tr	431.14
2,3,4-trimethylheptane	434.15	tr	432.90	te	433.61	tr	431.14	te	431.14
2,4,5-trimethylheptane	429.65	v	430.70	tr	431.13	tr	426.95	te	426.95
2,3,6-trimethylheptane	429.15	v	430.50	tr	428.52	tr	426.95	tr	426.95
2,2,3-trimethylheptane	430.95	tr	429.20	tr	427.59	tr	430.64	tr	430.65
2,2,4-trimethylheptane	421.45	tr	421.90	tr	420.98	tr	419.46	tr	420.65
2,2,5-trimethylheptane	423.95	te	425.50	te	424.32	te	423.96	te	420.65
2,2,6-trimethylheptane	421.35	tr	421.20	tr	421.66	tr	419.46	tr	420.65
2,3,3-trimethylheptane	433.35	te	431.80	te	431.06	te	431.14	te	431.14
3,3,4-trimethylheptane	435.05	tr	434.40	tr	434.71	tr	435.03	tr	435.05
3,3,5-trimethylheptane	428.85	v	430.20	tr	430.68	tr	426.95	te	426.95
3,4,4-trimethylheptane	434.25	te	433.80	v	433.39	te	431.14	te	431.14
2,3,5-trimethylheptane	433.85	tr	432.50	tr	432.30	tr	431.14	te	431.14
2,4,6-trimethylheptane	420.75	tr	423.70	tr	422.54	tr	419.46	tr	420.65
2,5,5-trimethylheptane	425.95	tr	428.20	v	427.13	tr	423.96	tr	424.15
3,4,5-trimethylheptane	435.65	tr	436.50	tr	438.38	tr	435.03	te	435.05
2,4,4-trimethylheptane	424.15	te	424.60	tr	422.41	te	423.96	tr	424.15
3,3-diethylhexane	439.45	tr	438.20	tr	437.21	tr	438.86	tr	438.85
2-methyl-3-isopropylhexane	438.15	tr	430.00	tr	430.33	tr	435.03	te	435.05
2,2-dimethyl-3-ethylhexane	429.25	tr	429.10	tr	427.85	tr	426.95	te	426.95
3,3-dimethyl-4-ethylhexane	426.05	te	436.80	te	436.41	te	423.96	tr	424.15
2,3-dimethyl-3-ethylhexane	439.15	te	435.80	tr	435.91	te	438.86	te	438.85
2,3-dimethyl-4-ethylhexane	435.15	tr	434.70	tr	436.32	tr	435.03	te	435.05
2,4-dimethyl-4ethylhexane	434.25	tr	431.60	te	431.73	tr	431.14	te	431.14
3,4-dimethyl-4-ethylhexane	435.35	tr	440.00	tr	440.56	tr	435.03	te	435.05
2,2-dimethyl-4-ethylhexane	420.15	tr	425.30	tr	424.98	tr	419.46	tr	417.53
2,2,3,3-tetramethylhexane	433.45	tr	431.40	v	430.05	tr	431.14	te	431.14
2,2,4,4-tetramethylhexane	426.95	v	421.90	tr	425.65	tr	426.95	tr	426.95
2,2,5,5-tetramethylhexane	410.55	v	413.60	v	426.09	tr	408.36	te	410.20
2,3,4,5-tetramethylhexane	429.35	tr	433.10	tr	431.44	tr	426.95	tr	426.95
2,2,4,5-tetramethylhexane	421.05	tr	422.20	te	424.73	tr	419.46	te	420.65
2,3,3,4-tetramethylhexane	437.75	tr	437.50	tr	435.59	tr	435.03	te	435.05
2,2,3,4-tetramethylhexane	430.15	v	432.00	tr	433.04	tr	426.95	te	426.95
2,2,3,5-tetramethylhexane	421.56	tr	424.70	v	425.67	tr	419.46	te	420.65
3,3,4,4-tetramethylhexane	443.15	v	439.70	te	434.38	tr	442.62	tr	442.14
2,3,4,4-tetramethylhexane	434.75	v	435.40	tr	434.39	tr	434.77	tr	434.74
2,4-dimethyl-3-isopropylpentane	430.25	tr	430.10	tr	428.38	tr	426.95	te	426.95
2-methyl-3,3-diethylpentane	445.15	tr	442.20	tr	440.62	tr	442.62	tr	442.14
2,3,4-trimethyl-3-ethylpentane	442.62	tr	439.80	tr	438.25	tr	442.62	tr	442.14
2,2,3-trimethyl-3-ethylpentane	442.65	v	439.00	te	435.92	tr	442.62	te	442.14
2,2,4-trimethyl-3-ethylpentane	428.45	tr	431.50	tr	428.97	tr	426.95	tr	426.95

[a] tr = training set, te = test set, v = validation set. [b] Reid et al. (1977); DIPPR (1996); and Properties of Organic Compounds, CRC Press, Inc. (1996).

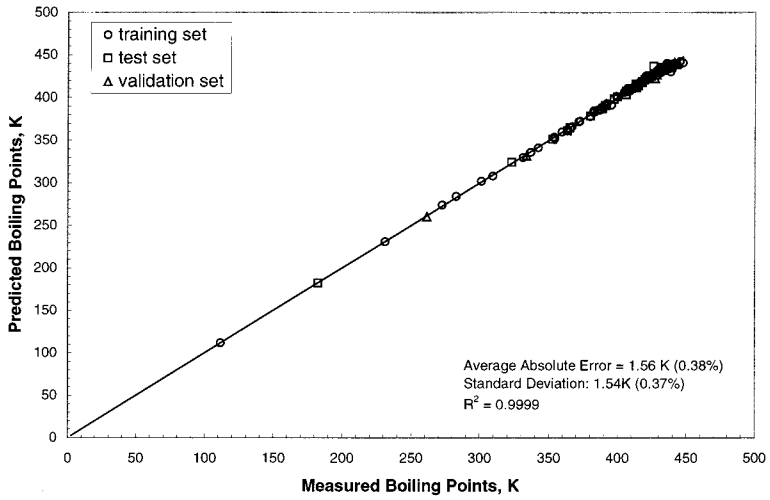

Figure 2. Measured and predicted boiling points of alkanes (7-4-1 back-propagation neural network model)

Figure 2. Measured and predicted boiling points of alkanes (7−4−1 back-propagation neural network model).

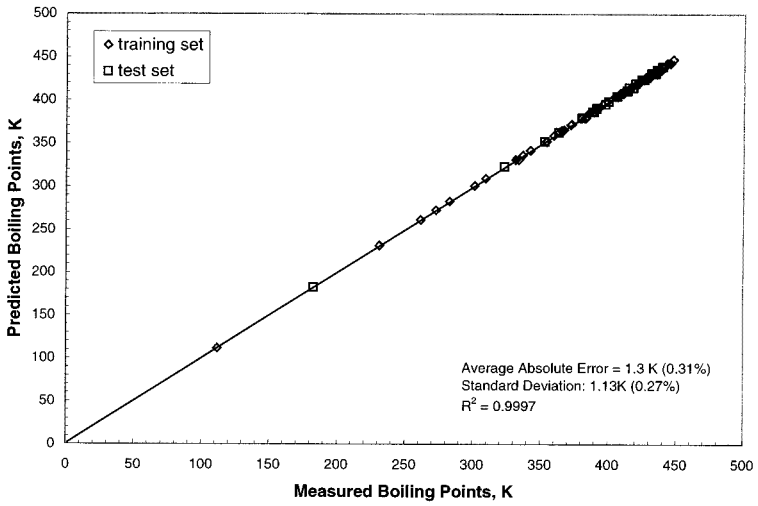

Figure 3. Measured and predicted boiling points of alkanes (alkane Fuzzy ARTMAP neural network model).

disconnected, and predictions were generated for each input vector of descriptors via the associative memory established between ART_a and ART_b during training.

III. RESULTS AND DISCUSSION

Overall performance of the back-propagation and Fuzzy ARTMAP neural network/QSPR models is summarized in Table 4. Boiling points of alkanes predicted using back-propagation and Fuzzy ARTMAP neural networks are given in Table 5. Graphical representations of the performance of the alkane neural network models are shown in Figures 2 and 3. The optimal back-propagation QSPR model for alkanes, based on a 7−4−1 architecture, predicted boiling points within an accuracy of 99.6% ± 0.4% (based on test set) with a maximum absolute error of 2.51% (10.71 K). The average absolute error when the entire data set is considered was about 0.37% (1.54 K). It is worth noting that the average percent errors in the training, test, and validation sets were about the same. The Fuzzy ARTMAP network predicted boiling points of alkanes within a slightly higher accuracy of 99.7% ± 0.3% for the entire data set, with a maximum absolute error of 0.88% (3.36 K). It is emphasized that experimental error associated with experimental boiling point temperatures is typically about 1%.[36] This is well within the range of the accuracy of the alkane model predictions. Performances of the alkane models were compared to the composite neural network/QSPR models of Ivanciuc[8] and Gakh et al.[19] as shown in Table 6. The Ivanciuc[8] DEG and RDS descriptor-based models predicted boiling points, for a 25 alkanes test set, with average absolute errors of 0.74% (3.0 K) and 0.42% (1.71 K), respectively.[7] The composite model of Gakh et al.[19] predicted six different physicochemical properties using an architecture consisting of seven inputs, eight hidden units, and six outputs, arranged in three layers. For the same set of 25 alkanes, the Gakh et al.[19] model predicts normal boiling points within an average absolute error of 1.19%.

The present back-propagation (7−4−1) and Fuzzy ARTMAP alkane models predicted boiling points with average absolute errors of 0.40% (1.65 K) and 0.30% (1.3 K), respectively. The back-propagation and Fuzzy ARTMAP alkane models were comparable to or better than the RDS MolNet model of Ivanciuc.[8] When predictions of the present pack-propagation and Fuzzy ARTMAP alkane models were compared to the DEG-MolNet, RDS-MolNet, and Gakh et al.[19] models (Table 6), for the same 125 compounds present

Table 6. Comparison of Neural Network/QSPR Models Predictions for Boiling Points of Alkanes[a]

		estimated BPs, K				
				Ivanciuc (1998)		
name	exptl BPs, K	alkane model: back propagation (7−4−11)	alkane model: Fuzzy ARTMAP	DEG	RDS	Gakh et al. (1994)
3-methylpentane	336.43	335.96	336.44	331.24	329.10	341.20
2-methylpentane	333.42	331.93	331.14	328.18	327.42	337.82
2,2-dimethylbutane	322.89	324.29	322.88	316.23	317.61	328.15
2,3-dimethylbutane	331.14	329.70	331.14	328.08	327.58	335.40
heptane	371.58	371.68	371.58	373.52	370.56	368.98
2-methylhexane	363.20	362.53	362.95	364.53	363.71	363.80
3-methylhexane	365.00	364.39	365.00	365.32	365.02	366.70
3-ethylpentane	366.63	365.73	365.00	366.83	366.62	367.68
2,2-dimethylpentane	352.34	351.23	352.35	352.01	355.76	355.01
2,3-dimethylpentane	362.93	361.03	362.95	364.33	365.33	364.62
3,3-dimethylpentane	359.21	359.89	359.22	358.35	360.23	362.36
2,4-dimethylpentane	353.65	353.74	352.35	351.84	354.86	357.75
2,2,3-trimethylbutane	354.03	352.37	352.35	354.38	356.45	358.08
octane	398.82	401.11	398.82	404.74	399.09	393.21
2-methylheptane	390.80	390.45	390.80	396.08	391.52	387.14
3-methylheptane	392.08	393.06	390.80	395.98	392.86	390.38
4-methylheptane	390.85	390.05	390.80	394.14	391.90	389.93
3-ethylhexane	391.68	392.30	390.80	394.97	392.90	390.63
2,2-dimethylhexane	379.99	378.53	379.63	383.76	382.82	379.39
2,3-dimethylhexane	388.76	387.27	386.61	392.66	390.79	388.21
2,4-dimethylhexane	382.58	384.54	379.63	383.11	382.49	385.41
2,5-dimethylhexane	382.25	382.83	379.63	386.28	383.52	383.10
3,4-dimethylhexane	390.88	390.71	390.80	393.44	392.66	391.56
2,2,3-trimethylpentane	382.99	383.09	379.63	384.65	384.35	382.33
2,2,4-trimethylpentane	372.39	372.09	372.38	372.97	376.25	375.32
2,3,3-trimethylpentane	387.91	387.02	386.61	390.16	389.39	386.84
2,3,4-trimethylpentane	386.62	385.08	386.61	389.27	389.74	386.62
3,3-dimethylhexane	385.12	385.20	385.13	387.35	387.14	386.37
2-methyl-3-ethylpentane	388.80	388.82	386.61	391.76	390.73	388.48
3-methyl-3-ethylpentane	391.41	392.38	390.80	393.57	393.29	391.42
2,2,3,3-tetramethylbutane	379.62	377.70	379.63	379.80	377.20	391.56
2-methyloctane	416.41	416.46	414.66	422.99	415.61	412.53
3-methyloctane	417.55	418.82	414.66	422.56	417.54	414.54
4-methyloctane	415.57	416.17	414.66	420.09	414.94	414.30
3-ethylheptane	416.15	418.01	414.66	420.82	416.97	413.12
4-ethylheptane	414.35	416.20	411.05	418.88	414.63	412.65
2,2-dimethylheptane	405.84	403.74	404.56	411.17	406.95	401.93
3,3-dimethylheptane	410.45	410.14	408.36	413.82	410.91	409.66
4,4-dimethylheptane	408.35	407.55	408.36	410.95	408.20	409.02
2,3-dimethylheptane	413.65	413.28	411.05	418.22	413.93	404.90
2,4-dimethylheptane	406.05	407.56	404.56	407.02	403.77	408.13
3,4-dimethylheptane	413.75	414.04	411.05	415.91	413.22	414.08
2-methyl-4-ethylhexane	406.95	410.69	404.56	408.30	404.92	409.10
2-methyl-3-ethylhexane	411.15	411.99	411.05	414.75	410.52	410.83
3-methyl-3-ethylhexane	413.12	414.56	411.05	416.21	413.81	414.41
3-methyl-4-ethylhexane	413.15	415.29	411.05	415.51	412.70	413.68
2,3,4-trimethylhexane	412.25	412.86	411.05	413.57	411.92	413.29
2,3,5-trimethylhexane	404.55	406.39	404.56	405.53	402.83	407.59
2,2,3-trimethylhexane	406.75	406.74	404.56	408.75	405.15	405.26
2,2,4-trimethylhexane	399.69	401.08	398.82	403.64	402.72	404.12
2,2,5-trimethylhexane	397.15	398.27	395.33	400.37	397.16	399.12
2,3,3-trimethylhexane	410.85	409.99	408.36	413.15	409.79	409.68
3,3,4-trimethylhexane	413.65	414.83	411.05	424.09	421.76	417.27
3,3-diethylpentane	419.45	420.82	419.46	424.38	421.77	417.06
2,3,3,4-tetramethylpentane	414.65	413.50	414.66	415.83	413.05	410.95
2,2,3,3-tetramethylpentane	413.35	411.71	411.05	415.69	410.90	412.39
2,2,4,4-tetramethylpentane	395.35	391.20	395.33	399.01	401.54	395.68
2,2,3,4-tetramethylpentane	406.15	407.01	404.56	406.16	405.11	405.06
2,2-dimethyl-3-ethylpentane	406.95	410.37	404.56	409.04	405.38	405.78
2,3-dimethyl-3-ethylpentane	417.85	417.54	414.66	421.73	418.89	413.88
2,4-dimethyl-3-ethylpentane	409.85	411.14	408.36	412.77	410.15	410.04
3-ethyloctane	439.65	439.47	438.86	443.54	442.17	439.02
4-ethyloctane	436.85	438.47	435.03	440.85	437.67	436.64
4-isopropylheptane	432.05	432.57	431.14	435.75	431.23	433.12
4-propylheptane	435.15	436.49	430.64	435.41	431.46	432.54
2,2-dimethyloctane	428.15	426.51	426.95	432.01	430.69	428.79
2,3-dimethyloctane	437.45	435.38	435.03	440.66	439.73	436.63
2,4-dimethyloctane	429.15	430.06	426.95	428.49	426.06	434.11
2,5-dimethyloctane	431.65	431.78	431.14	434.21	431.27	435.30
2,6-dimethyloctane	433.55	433.77	431.14	436.46	435.00	435.22
2,7-dimethyloctane	433.05	431.06	431.14	437.44	434.36	434.60

Table 6 (Continued)

name	exptl BPs, K	estimated BPs, K: alkane model: back propagation (7−4−11)	alkane model: Fuzzy ARTMAP	Ivanciuc (1998) DEG	Ivanciuc (1998) RDS	Gakh et al. (1994)
3,3-dimethyloctane	434.35	432.08	431.14	437.12	437.94	434.03
3,4-dimethyloctane	436.58	436.21	435.03	437.69	437.36	438.65
3,5-dimethyloctane	432.55	432.86	431.14	432.76	432.08	437.35
3,6-dimethyloctane	433.95	436.79	431.14	434.75	434.27	437.48
4,4-dimethyloctane	430.65	428.85	430.64	432.12	430.17	433.67
4,5-dimethyloctane	435.25	434.05	435.03	436.59	435.11	438.74
2-methyl-3-ethylheptane	436.15	433.71	435.03	439.46	437.21	433.22
3-methyl-4-ethylheptane	438.15	435.28	435.03	441.43	439.01	436.47
2-methyl-4-ethylheptane	429.35	431.17	431.14	431.61	428.44	431.15
2-methyl-5-ethylheptane	432.85	433.89	431.14	436.55	434.97	432.75
3-methyl-5-ethylheptane	431.35	435.56	431.14	430.47	429.79	434.58
4-methyl-3-ethylheptane	438.15	434.73	435.03	441.02	439.82	436.66
3-methyl-3-ethylheptane	436.95	435.48	436.95	439.93	440.37	437.43
4-methyl-4-ethylheptane	433.95	433.17	435.03	436.59	434.87	436.81
2,3,4-trimethylheptane	434.15	432.88	431.14	433.37	433.68	435.69
2,4,5-trimethylheptane	429.65	430.68	426.95	429.05	429.21	434.71
2,3,6-trimethylheptane	429.15	430.50	426.95	430.79	430.32	430.78
2,2,3-trimethylheptane	430.95	429.16	430.644	432.33	432.19	427.79
2,2,4-trimethylheptane	421.45	421.89	419.46	420.10	420.45	424.70
2,2,5-trimethylheptane	423.95	425.53	423.96	425.85	425.59	425.84
2,2,6-trimethylheptane	421.35	421.20	419.46	424.46	424.90	421.27
2,3,3-trimethylheptane	433.35	431.84	431.14	433.89	433.70	432.20
3,3,4-trimethylheptane	435.05	434.44	435.03	435.29	435.23	434.98
3,3,5-trimethylheptane	428.85	430.22	426.95	427.49	428.77	432.83
3,4,4-trimethylheptane	434.25	433.82	431.14	434.21	433.48	436.33
2,3,5-trimethylheptane	433.85	432.46	431.14	436.58	436.71	436.27
2,4,6-trimethylheptane	420.75	423.67	419.46	417.38	417.16	427.58
2,5,5-trimethylheptane	425.95	428.22	423.96	426.50	426.69	430.01
3,4,5-trimethylheptane	435.65	436.52	435.03	434.43	435.30	440.24
2,4,4-trimethylheptane	424.15	424.61	423.96	422.43	422.52	429.32
3,3-diethylhexane	439.45	438.16	438.86	445.53	443.16	439.55
2-methyl-3-isopropylhexane	438.15	429.98	435.03	450.20	446.33	436.40
2,2-dimethyl-3-ethylhexane	429.25	429.05	426.95	432.28	428.12	428.84
3,3-dimethyl-4-ethylhexane	426.05	436.77	423.96	428.48	426.14	435.81
2,3-dimethyl-3-ethylhexane	439.15	435.84	438.86	441.59	439.33	437.76
2,3-dimethyl-4-ethylhexane	435.15	434.67	435.03	436.43	434.91	437.07
2,4-dimethyl-4-ethylhexane	434.25	431.60	431.14	437.40	437.43	436.75
3,4-dimethyl-4-ethylhexane	435.35	440.04	435.03	432.65	431.39	439.33
2,2-dimethyl-4-ethylhexane	420.15	425.26	419.46	417.69	416.22	427.30
2,2,3,3-tetramethylhexane	433.45	431.38	431.14	433.56	430.39	429.67
2,2,4,4-tetramethylhexane	426.95	421.94	426.95	430.26	432.48	428.20
2,2,5,5-tetramethylhexane	410.55	413.58	408.36	409.62	409.56	415.95
2,3,4,5-tetramethylhexane	429.35	433.12	426.95	424.85	425.35	435.28
2,2,4,5-tetramethylhexane	421.05	422.19	419.46	421.05	421.42	433.08
2,3,3,4-tetramethylhexane	437.75	437.46	435.03	436.50	435.52	437.84
2,2,3,4-tetramethylhexane	430.15	432.01	426.95	429.47	429.35	432.77
2,2,3,5-tetramethylhexane	421.56	424.69	419.46	418.59	417.97	425.43
3,3,4,4-tetramethylhexane	443.15	439.70	442.62	445.88	443.25	439.48
2,3,4,4-tetramethylhexane	434.75	435.41	434.77	433.14	434.61	433.62
2,4-dimethylisoisopropylpentane	430.25	430.14	426.95	431.87	429.10	432.19
2-methyl-3,3-diethylpentane	445.15	442.17	442.62	451.73	448.10	440.56
2,3,4-trimethyl-3-ethylpentane	442.62	439.78	442.62	446.24	443.61	435.62
2,2,3-trimethyl-3-ethylpentane	442.65	438.97	442.62	447.93	443.51	434.99
2,2,4-trimethyl-3-ethylpentane	428.45	431.51	426.95	429.16	425.93	428.99

[a] Comparison is based on 125 alkanes common to the compared models.

in both studies, the average absolute errors were found to be 0.38% (1.59 K), 0.31% (1.29 K), 0.65% (2.66 K), 0.43% (1.75 K), and 11.03% (2.63 K), respectively. The above comparison suggests that independent QSPRs for boiling points are clearly more accurate than the composite QSPR for multiple properties.

Boiling points of alkenes predicted using back-propagation and Fuzzy ARTMAP neural networks are given in Table 7. Graphical representation of the performance of the alkene models is shown in Figures 4 and 5. The back-propagation QSPR model for boiling point of alkenes, with a 7−10−1 architecture, was able to distinguish between cis and trans geometric isomers suggesting that the selected set of molecular descriptors was adequate for alkenes. When all the data were considered, boiling points were predicted with an average absolute error of less than 1.25% (4.4 K) and a maximum absolute error of 7.11% (19.57 K). We note that the standard deviation for the entire data set is 1.13% (3.73 K). A slightly higher average absolute error of 1.83% (6.45 K) and a maximum absolute error of 4.39% (18.18 K) were obtained for the validation data set. The standard deviation for the combined test and validation sets is 1.21% (4.18 K).

Table 7. Experimental and Predicted Boiling Points for Alkenes Using Back-Propagation and Fuzzy ARTMAP Algorithms

	BPs,[a] K								
		estimated: back-propagation				estimated: Fuzzy ARTMAP			
name	reported[b]	alkenes (7−10−1)		composite (7−9−1)		alkenes[c]		composite	
ethylene	169.40	tr	172.60	tr	179.51	tr	169.40	tr	169.39
propene	225.75	tr	230.22	tr	238.71	tr	225.75	tr	225.77
1,3-butadiene	268.70	tr	269.64	te	287.39	tr	266.25	te	266.26
cis-2-butene	276.90	tr	280.30	v	282.99	tr	276.91	tr	276.88
trans-2-butene	274.00	tr	271.58	tr	284.99	tr	274.00	tr	272.66
2-methyl-1-propene	266.85	tr	265.64	tr	271.91	tr	266.25	tr	266.26
trans-1,3-pentadiene	315.18	tr	314.41	te	320.69	tr	315.18	tr	313.29
2,3-pentadiene	321.40	tr	317.05	v	319.20	tr	321.39	tr	321.39
3-methyl-1,2-butadiene	314.00	tr	308.82	tr	307.09	tr	311.72	tr	313.29
2-methyl-1,3-butadiene	307.20	tr	312.82	te	310.33	tr	307.19	tr	304.29
1-pentene	303.12	tr	305.73	tr	312.34	tr	303.12	te	300.21
trans-2-pentene	309.50	tr	306.76	tr	307.80	tr	309.50	te	309.24
3-methyl-1-butene	293.30	tr	305.90	v	299.11	tr	293.31	tr	293.29
2-methyl-1-butene	304.30	tr	309.93	tr	298.27	tr	303.12	te	304.29
cis-1,3,5-hexatriene	351.15	tr	356.79	tr	368.82	tr	349.14	te	351.05
1,3-hexadiene	346.15	tr	348.34	tr	350.70	tr	344.49	tr	343.16
trans,*trans*-2,4-hexadiene	353.15	tr	354.41	te	359.92	tr	353.15	tr	351.05
1,4-hexadiene	266.25	tr	345.24	v	348.42	tr	336.64	tr	338.14
2,3-dimethyl-1,3-butadiene	341.93	tr	340.86	tr	349.54	tr	339.61	tr	339.59
trans-2-hexene	314.00	tr	339.17	tr	343.14	tr	339.61	tr	339.59
cis-3-hexene	339.60	tr	343.34	te	344.37	tr	339.61	tr	339.59
1-hexene	336.64	tr	332.61	tr	337.09	tr	336.64	te	334.33
trans-4-methyl-2-pentene	331.76	tr	330.77	tr	333.97	tr	331.75	tr	330.84
4-methyl-1-pentene	327.02	tr	327.97	te	333.57	tr	327.02	tr	327.03
trans-3-methyl-2-pentene	343.59	tr	341.40	tr	339.57	tr	343.59	tr	339.59
cis-3-methyl-2-pentene	340.88	tr	341.50	tr	339.65	tr	339.61	te	339.59
2-methyl-1-pentene	335.25	tr	334.03	tr	336.96	tr	335.24	tr	334.33
2,3-dimethyl-1-butene	328.75	tr	331.17	tr	316.56	tr	327.02	tr	327.03
3,3-dimethyl-1-butene	314.45	tr	317.84	te	326.62	tr	311.72	tr	313.29
1,3,5-heptatriene	396.80	tr	383.88	tr	390.43	tr	394.94	te	394.14
1,5-heptadiene	367.15	tr	369.92	te	377.26	tr	365.56	te	365.00
1,6-heptadiene	362.15	tr	364.25	tr	373.78	tr	360.76	te	360.74
2,4-heptadiene	381.15	tr	381.31	tr	386.09	tr	380.05	tr	377.73
2-methyl-2,4-hexadiene	384.65	tr	372.90	tr	373.25	tr	383.65	tr	384.50
2,4-dimethyl-1,3-pentadiene	366.35	tr	365.84	tr	371.09	tr	365.56	tr	368.15
trans-3-heptene	368.82	tr	369.74	te	376.63	tr	368.54	tr	368.15
1-heptene	366.75	tr	360.67	tr	366.71	tr	365.56	tr	365.00
trans-2-heptene	371.15	tr	362.95	tr	375.67	tr	368.54	tr	370.64
2-methyl-2-hexene	368.55	tr	365.35	tr	368.97	tr	368.54	tr	368.15
4,4-dimethyl-1-pentene	345.65	tr	345.49	tr	366.29	tr	344.49	te	343.16
5-methyl-1-hexene	358.45	tr	357.98	te	362.10	tr	355.82	te	356.55
cis-5-methyl-2-hexene	362.65	tr	361.81	tr	365.78	tr	360.76	tr	360.74
trans-5-methyl-2-hexene	361.25	tr	362.05	tr	365.76	tr	360.76	tr	360.74
3-ethyl-1-pentene	357.25	tr	362.74	v	367.33	tr	355.82	tr	356.55
cis-3-methyl-3-hexene	368.55	tr	369.13	tr	368.71	tr	368.54	te	368.15
3-ethyl-2-pentene	369.15	tr	367.34	tr	370.61	tr	368.54	tr	368.15
4-methyl-1-hexene	359.85	tr	359.67	tr	361.90	tr	359.06	tr	356.55
trans-4-methyl-2-hexene	360.75	tr	365.34	tr	368.00	tr	360.76	tr	360.74
trans-3-methyl-3-hexene	366.65	tr	369.23	tr	368.83	tr	365.56	te	365.00
2-methyl-1-hexene	365.15	tr	363.40	tr	365.52	tr	365.16	tr	365.00
2,4-dimethyl-1-pentene	354.75	tr	357.97	tr	359.92	tr	353.15	tr	354.75
2,3,3-trimethyl-1-butene	351.05	tr	346.79	tr	351.15	tr	349.14	tr	351.05
3,3-dimethyl-1-pentene	350.65	tr	355.94	te	365.89	tr	349.14	te	349.15
cis-4,4-dimethyl-2-pentene	353.55	tr	353.96	tr	345.16	tr	353.15	te	351.05
2,4-dimethyl-2-pentene	356.55	tr	357.70	tr	361.74	tr	355.82	tr	356.55
2,3-dimethyl-1-pentene	357.45	tr	360.27	v	359.07	tr	355.82	tr	356.55
cis-3,4-dimethyl-2-pentene	362.45	tr	363.09	tr	360.55	tr	360.76	tr	360.74
trans-3,4-dimethyl-2-pentene	365.65	tr	362.13	tr	359.75	tr	365.56	te	365.00
2,4,6-octatriene	420.65	tr	413.29	tr	415.41	tr	417.94	tr	420.65
5-methyl-1,3,6-heptatriene	390.15	tr	398.08	tr	399.55	tr	387.44	te	390.16
2,6-octadiene	397.65	tr	395.49	te	406.09	tr	394.94	tr	397.63
2,5-dimethyl-2,4-hexadiene	407.65	tr	400.18	te	408.74	tr	407.66	tr	404.56
3-methyl-1,5-heptadiene	373.15	tr	389.48	tr	391.02	tr	372.14	tr	370.64
cis-2-octene	398.75	tr	394.36	tr	397.12	tr	398.16	te	398.15
trans-2-octene	398.15	tr	388.88	v	402.81	tr	398.16	tr	398.15
1-octene	394.35	tr	391.30	tr	395.77	tr	391.63	te	394.14
6-methyl-1-heptene	386.35	tr	386.40	tr	382.76	tr	385.96	tr	385.94
6-methyl-3-heptene	388.15	tr	390.58	tr	396.44	tr	387.44	tr	385.94
6,6-dimethyl-2-hexene	380.05	tr	382.05	tr	375.51	tr	380.05	tr	377.73
2,3-dimethyl-2-hexene	394.95	tr	393.58	te	392.81	tr	394.94	te	394.14
2-methyl-1-heptene	392.37	tr	388.65	tr	385.69	tr	391.63	te	390.16
cis-3-methyl-2-heptene	385.15	tr	389.64	tr	387.02	tr	383.65	te	382.27
3-methyl-3-heptene	394.15	tr	392.11	te	388.89	tr	391.63	tr	394.14
4-methyl-1-heptene	385.95	tr	389.06	tr	385.30	tr	385.96	tr	385.94

Table 7 (Continued)

	BPs,[a] K								
		estimated: back-propagation				estimated: Fuzzy ARTMAP			
		alkenes		composite		alkenes[c]		composite	
name	reported[b]	(7−10−1)		(7−9−1)					
5-methyl-1-heptene	386.45	tr	384.64	v	383.06	tr	385.96	te	385.94
5-methyl-2-heptene	391.15	tr	389.51	tr	393.68	tr	391.15	tr	390.16
2,3-dimethyl-1-hexene	383.65	tr	388.29	tr	387.41	tr	383.65	te	382.27
2,4-dimethyl-1-hexene	384.35	tr	389.19	te	386.17	tr	383.65	tr	382.27
1,8-nonadiene	415.65	tr	416.31	tr	424.66	tr	414.15	te	413.76
7-methyl-2,4-octadiene	422.15	tr	416.15	tr	409.02	tr	422.14	tr	420.65
2,4-dimethyl-2,4-heptadiene	411.15	tr	408.87	tr	415.72	tr	411.15	te	410.20
2,6-dimethyl-1,3-heptadiene	414.15	tr	414.05	te	411.59	tr	414.15	tr	413.76
2,6-dimethyl-2,4-heptadiene	413.15	tr	416.79	v	410.81	tr	411.15	te	410.20
3-ethyl-2-methyl-1,5-hexadiene	418.15	tr	412.18	tr	415.91	tr	417.94	tr	417.53
1-nonene	420.05	tr	419.41	tr	424.83	tr	417.94	te	417.53
trans-3-nonene	420.65	tr	417.35	te	408.38	tr	417.94	tr	420.65
trans-4-nonene	416.15	tr	419.61	tr	417.68	tr	414.15	te	413.76
7-methyl-3-octene	415.15	tr	415.24	tr	420.84	tr	414.15	te	413.76
2-merthyl-4-octene	412.15	tr	414.01	te	421.50	tr	411.15	tr	410.20
3-methyl-2-octene	418.15	tr	412.35	tr	420.28	tr	417.94	te	417.53
2,3-dimethyl-2-heptene	418.35	tr	414.92	tr	409.30	tr	417.94	tr	417.53
1,3-decadiene	442.15	tr	442.18	tr	445.77	tr	441.54	tr	442.14
1-decene	443.65	tr	441.15	tr	444.11	tr	441.54	tr	442.14
cis-5-decene	444.15	tr	447.87	tr	444.75	tr	441.54	te	442.14
2-methyl-1-nonene	441.55	tr	434.36	tr	438.69	tr	441.54	te	438.85
4-proply-3-heptene	433.65	tr	433.04	tr	445.01	tr	433.65	tr	431.14
2-methyl-3-nonene	434.15	tr	434.99	te	414.15	tr	433.65	te	431.14
propadiene	238.70	te	221.72	te	228.56	te	238.69	tr	238.71
cis-2-pentene	310.09	te	310.86	tr	313.50	te	309.50	tr	309.24
2-methyl-2-butene	311.72	te	305.25	tr	295.78	te	311.72	te	309.24
cis,*trans*-2,4-hexadiene	356.65	te	350.48	tr	352.38	te	355.82	tr	356.65
1,5-hexadiene	332.60	te	344.56	te	351.17	te	331.75	tr	330.84
trans-3-hexene	340.30	te	345.48	tr	339.22	te	339.61	tr	339.59
3-methyl-1-pentene	327.31	te	334.79	v	334.98	te	327.02	tr	327.03
2,3-dimethyl-2-butene	346.45	te	326.88	tr	339.27	te	344.49	tr	343.16
1,2-butadiene	284.00	te	274.16	tr	284.62	te	284.00	tr	281.21
2,5-heptadiene	381.15	te	375.05	v	381.58	te	380.05	te	377.73
2-ethyl-1-pentene	368.15	te	365.40	te	366.94	te	365.56	tr	368.15
1,2-pentadiene	318.00	te	312.64	tr	318.75	te	317.22	tr	317.20
trans-4,4-dimethyl-2-pentene	349.85	te	346.03	tr	360.86	te	349.14	te	349.15
2,5-dimethyl-1,3,5-hexatriene	419.15	te	410.76	te	414.88	te	417.94	tr	417.53
1,7-octadiene	388.65	te	386.87	tr	406.57	te	387.44	te	385.94
2,5-dimethyl-1,5-hexadiene	387.45	te	391.43	tr	401.28	te	387.44	tr	385.94
trans-4-methyl-2-heptene	387.15	te	395.14	tr	391.47	te	385.96	tr	385.94
2,4-dimethyl-2-hexene	387.75	te	392.42	tr	386.87	te	387.44	tr	385.94
2-methyl-2-heptene	395.75	te	390.21	tr	388.78	te	394.94	te	394.14
4-methyl-3,5-octadiene	422.65	te	417.11	te	424.78	te	422.14	te	420.65
5-methyl-3-heptene	385.15	te	392.44	tr	377.28	te	383.65	tr	382.27
2,7-nonadiene	424.65	te	416.23	tr	422.67	te	422.14	tr	424.15
2,6-dimethyl-2,5-heptadiene	423.65	te	414.34	tr	413.77	te	422.14	tr	420.65
2-methyl-1-octene	417.95	te	413.04	tr	416.71	te	417.94	te	417.53
3,7-dimethyl-1-octene	427.15	te	430.49	tr	421.47	te	427.14	tr	426.95
4-decene	443.75	te	440.63	tr	444.85	te	441.54	tr	442.14
1-butene	266.85	v	276.95	tr	285.67	tr	266.25	tr	266.26
trans-5-decene	443.75	v	441.42	te	439.41	tr	441.54	tr	442.14
cis-1,3-pentadiene	317.22	v	314.16	tr	320.58	tr	317.22	tr	317.20
trans-1,3,5-hexatriene	351.65	v	356.89	te	367.43	tr	349.14	tr	351.05
cis-2-hexene	342.00	v	338.15	tr	341.12	tr	339.61	tr	339.59
2-methyl-2-pentene	340.50	v	333.93	tr	339.86	tr	339.61	tr	339.59
cis-2-heptene	372.15	v	365.39	v	369.88	tr	372.14	tr	370.64
cis-2-methyl-3-hexene	359.15	v	367.91	tr	370.70	tr	359.06	tr	356.55
3-methyl-1-hexene	357.05	v	362.77	te	364.50	tr	355.82	tr	356.55
3,4-dimethyl-1-pentene	353.95	v	356.94	te	359.08	tr	353.15	tr	351.05
2,3-dimethyl-2-pentene	370.65	v	362.51	tr	372.69	tr	368.54	tr	370.64
1,4-pentadiene	299.10	v	311.58	tr	320.03	tr	299.11	tr	299.10
1,2-hexadiene	349.15	v	340.85	tr	345.04	tr	349.14	tr	349.15
cis-3-heptene	368.95	v	372.59	te	375.31	tr	368.54	tr	368.15
trans-2-methyl-3-hexene	359.05	v	369.63	tr	349.85	tr	359.06	tr	356.55
1,4-heptadiene	366.15	v	375.58	tr	381.74	tr	365.56	tr	365.00
6-methyl-2-heptene	390.15	v	386.78	tr	393.40	tr	387.44	tr	390.16
4-methyl-3-heptene	394.15	v	394.85	tr	390.31	tr	391.63	tr	394.14
3,5-dimethyl-2,4-heptadiene	403.15	v	403.11	tr	414.27	tr	403.16	tr	403.14
2,6-dimethyl-1,5-heptadiene	416.15	v	410.85	tr	412.18	tr	414.15	tr	413.76
2,6-dimethyl-2-octene	413.42	v	431.60	v	421.68	tr	411.15	tr	410.20

[a] tr = training set, te = test set, v = validation set. [b] Reid et al. (1977); DIPPR (1996); and Properties of Organic Compounds, CRC Press, Inc. (1996). [c] 119 ARTA classes and 49 ARTB classes.

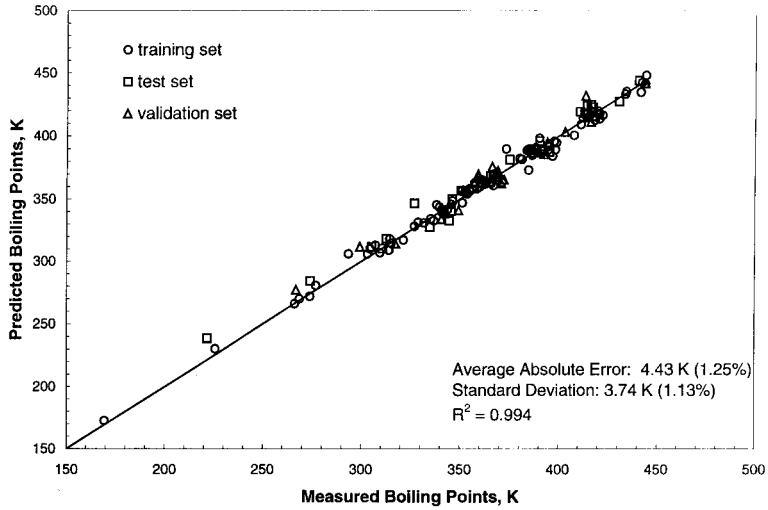

Figure 4. Measured and predicted boiling points of alkenes (7−10−1 back-propagation neural network model).

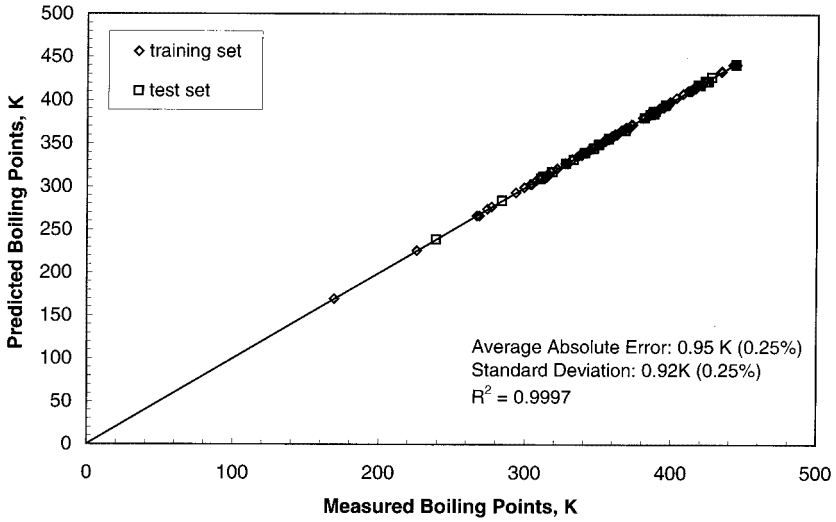

Figure 5. Measured and predicted boiling points of alkenes (alkene Fuzzy ARTMAP neural network model).

We note that boiling points predictions were significantly improved using the modified Fuzzy ARTMAP network. Average absolute errors based on the test and entire data sets for the Fuzzy ARTMAP alkene model are 0.19% (0.73 K) and 0.25% (0.95 K), respectively. It should be noted that the Fuzzy ARTMAP network categorizes classes of compounds (including isomers) with boiling points within the experimental error of 1 K for similar values of the molecular descriptors.

Results from the back-propagation alkene model were compared with the study of Zhang et al.[20] who developed a neural network/QSPR boiling point model, based on 85 alkenes, with a reported average absolute error of 2.3% (2 K) and a maximum absolute error of 10% (5.7 K). Clearly, the present back-propagation alkene model, with an average absolute error (for the overall data set) of 1.25% (4.42 K) and the ability to distinguish quantitatively among diastereomers, is acceptable given the errors (up to 1.0% or higher) associated with experimental boiling point measurements.[36]

An independent back-propagation/QSPR model was also attempted for alkynes. However, since a small set of only 43 compounds was available, the average absolute error of 1.55% (4.95 K) is higher than for the alkanes and alkene models, although it is lower than the error of 3.35% (8.23 K) for standard group contribution methods.[40] The alkynes were included in the composite back-propagation and Fuzzy ARTMAP/QSPR models developed using the complete aliphatic hydrocarbon data set (Tables 3, 5, and 7). Predicted boiling points for alkynes from the back-propagation and Fuzzy ARTMAP composite models are depicted in Figures 6 and 7, and Table 3, with a summary error analysis provided in Table 4.

The optimum back-propagation architecture (7−9−1) for the composite model had an average absolute boiling point error of 1.37% (4.85 K) for the entire data set, with a maximum error of 9.75% (20.92 K) and standard deviation of 1.46% (4.45 K). The average absolute error for the training, testing, and validation sets were 1.28% (4.51 K), 1.74% (6.1 K), and 1.25% (4.68 K), respectively; the corresponding standard deviations were 1.41% (4.2 K), 1.74% (5.33 K), and 1.0% (3.68 K), respectively. It is noted that the average and maximum percent absolute errors associated with the 30 alkynes in the training set were 1.79% (6.08 K) and 9.74% (18.33 K), respectively, with a standard deviation of 1.65% (3.69 K). Although errors from alkynes in the training set were relatively higher, the alkynes represented only 13% of the training set. Indeed, the alkynes

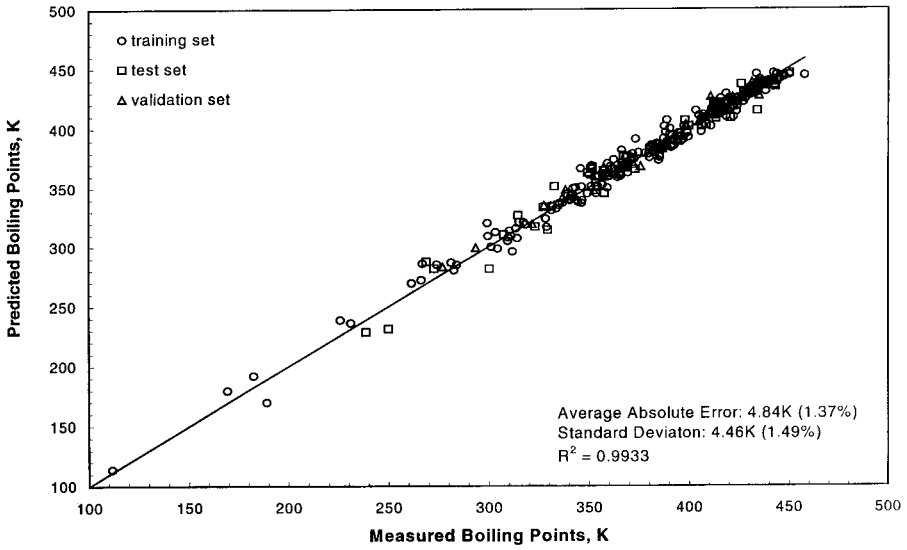

Figure 6. Measured and predicted boiling points of aliphatic hydrocarbons (7−9−1 back-propagation neural network composite model).

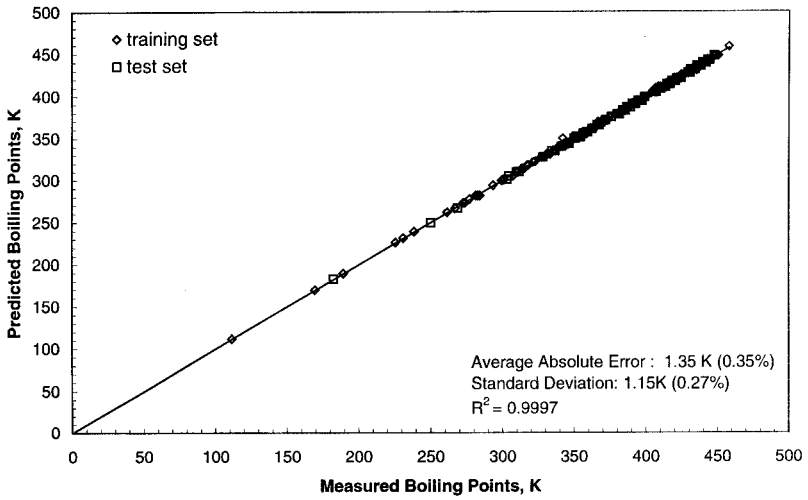

Figure 7. Measured and predicted boiling points of aliphatic hydrocarbons (composite Fuzzy ARTMAP neural network model).

contributed about 17.7% of the total average absolute error in the training (about 0.8 K). The alkenes represented 44.5% of the data in the training set and contributed about 57.9% of the average absolute error in the training. Alkanes contributed about 24.4% to the absolute average error in the composite model training set.

Overall, the back-propagation composite model did not perform as well as the individual alkane and alkene models. The error in boiling point prediction of alkanes and alkenes was higher with the composite model than with the individual alkane or alkene models. For example, the composite model resulted in an average absolute error of 1.77% (6.16 K) for alkenes compared to the error of 1.25% (4.44 K) from the individual model. The poorer performance of the composite model could be due to the relatively smaller training set for the alkynes and possibly the failure to adequately represent the complexity of the triple bond with the present set of descriptors. Nonetheless, the average absolute error for the alkanes was 0.75% (2.86 K), which is within the acceptable accuracy relative to the typical uncertainty (absolute error of 1.0%) in experimentally reported boiling points.

Improvements in predicting boiling points of the composite set of aliphatic hydrocarbons were made using the Fuzzy ARTMAP network (Tables 3, 5, and 7). As shown in Table 4, the average absolute error, standard deviation, and maximum errors from the composite model based on the entire data set reduce to 0.35% (1.35 K), 0.27% (1.15 K), and 0.98% (3.45 K), respectively. The average absolute error for the training and test sets were 0.35% (1.44 K), and 0.84% (1.15 K), respectively. Clearly, the accuracy of the Fuzzy ARTMAP composite model developed in this study is exceptional with an error lower than the 1% error associated with experimental boiling point temperatures.[36] It is worth noting that for the composite data set (alkanes + alkenes + alkynes) the entire set of molecular descriptors influences the classification. For the homogeneous alkanes the resulting classification appears to be dominated by the number of carbon atoms. For example, two compounds with the same molecular weight, e.g., 2-methyl-1-pentane and 2−3-dimethyl-butane, that have similar boiling points (within the specified tolerance) were placed in the same class. The same classification pattern not only was revealed for alkenes but also was influenced by the position of the double bonds. In this latter case, the resulting classification was weighted by nearly 80% of the classification weight attributable to the number of carbon atoms and about 20% to the position of double bonds.

Table 8. Comparison of Neural Network/QSPR Models Predictions for Boiling Points of Aliphatic Hydrocarbons (Alkanes and Alkenes)[a]

		estimated BPs, K			
		composite models			
aliphatic compound	exptl BPs, K	back propagation (7−9−1)	Fuzzy ARTMAP	Hall and Story (1966)	Egolf et al. (1994)
propane	231.08	236.05	231.10	258.10	242.11
butane	272.65	281.52	272.66	280.70	271.94
2-methylpropane	261.42	269.29	261.43	274.60	270.68
pentane	309.24	304.86	309.24	308.20	301.85
2-methylbutane	301.00	299.83	301.21	302.60	300.44
2,2-dimethylpropane	282.68	280.16	281.21	286.80	297.92
hexane	341.89	343.86	339.59	336.20	331.42
3-methylpentane	336.43	341.20	334.33	336.40	330.07
2-methylpentane	333.42	332.31	330.84	331.50	329.93
2,2-dimethylbutane	322.89	316.88	321.39	317.40	327.42
2,3-dimethylbutane	331.14	330.80	330.84	327.40	328.30
heptane	371.58	378.62	370.64	365.10	359.95
2-methylhexane	363.20	361.60	360.74	360.60	358.76
3-methylhexane	365.00	362.49	365.00	360.60	359.04
2,2,3-trimethybutane	354.03	351.26	351.05	348.10	354.60
octane	398.82	395.72	398.82	393.60	388.09
2,3-dimethylhexane	388.76	388.05	385.94	387.60	385.67
2,2,3-trimethylpentane	382.99	380.27	382.27	383.80	383.23
2,2,4-trimethylpentane	372.39	374.28	370.64	380.70	382.94
2,3,3-trimethylpentane	387.91	384.41	385.94	380.70	383.29
2-methyl-3-ethylpentane	388.80	388.75	385.94	387.60	386.19
nonane	423.95	417.88	420.65	421.20	415.19
2,2,3,3-tetramethylpentane	413.35	408.07	411.20	409.50	408.34
decane	447.25	442.58	447.16	447.70	441.12
propene	225.75	238.71	225.77	252.80	257.27
1,3-butadiene	268.70	287.39	266.26	275.20	296.44
cis-2-butene	276.90	282.99	276.88	273.80	278.13
1-butene	266.85	285.67	266.26	274.90	283.11
trans-2-butene	274.00	284.99	272.66	273.80	278.01
2-methyl-1-propene	266.25	271.91	266.26	279.30	281.56
1-pentene	303.12	312.34	303.21	300.60	311.75
3-methyl-1-butene	293.30	299.11	293.29	291.90	308.83
2-methyl-2-butene	311.72	295.78	311.72	304.20	303.76
2-methyl-1-butene	304.30	298.27	304.29	306.10	308.42
1,5-hexadiene	332.60	351.17	330.84	329.80	348.36
2,3-dimethyl-1,3-butadiene	341.93	349.54	339.59	339.70	340.93
cis-2-hexene	342.00	341.12	339.59	337.90	333.97
trans-2-hexene	341.00	343.14	339.59	337.90	333.48
1-hexene	336.64	337.09	334.33	331.70	340.64
2-methyl-2-pentene	340.50	339.86	339.59	336.90	331.22
4-methyl-1-pentene	327.02	333.57	327.03	322.60	336.72
3-methyl-1-pentene	327.31	334.98	327.03	322.20	337.12
2-methyl-1-pentene	335.25	336.96	334.33	333.80	336.78
2,3-dimethyl-1-butene	328.75	316.56	327.03	329.00	333.70
2,3-dimethyl-2-butene	346.45	339.27	343.16	344.80	328.75
3,3-dimethyl-1-butene	314.45	326.62	313.29	309.90	333.86
1-heptene	366.75	366.71	365.00	360.50	368.54
1-octene	394.35	395.77	394.14	389.40	394.87
1-decene	443.65	444.11	442.14	444.50	445.50
av absolute error		5.63	1.16	4.97	7.99
av percent error		1.84	0.33	1.69	2.61
max. absolute error		18.82	3.30	27.05	31.52
max. percent error		7.05	0.95	11.98	13.96

[a] Comparison is based on 49 compounds common to the three compared models.

Boiling point predictions from the Fuzzy ARTMAP and back-propagation composite models were also compared with the models of Hall and Story[22] and Egolf et al.,[21] for 49 hydrocarbons (alkanes and alkenes) common to these studies, revealing average absolute errors of 0.24% (0.85 K), 1.8% (5.6 K), 1.7% (4.9 K), and 2.6% (7.9 K), respectively, and maximum absolute errors of 0.91% (3.4 K), 7.1% (18.8 K), 12% (27.1 K), and 14% (31.5 K), respectively (Table 8 and Figure 8). Although the Hall and Story[22] model had a slightly lower average absolute boiling point error than the present back-propagation composite model, it did not differentiate between cis and trans isomers of alkenes (i.e., equal boiling points were predicted for such isomers). In contrast, the present back-propagation models and the model of Egolf et al.[21] were able to distinguish between geometric isomers.

IV. CONCLUSIONS

Back-propagation and Fuzzy ARTMAP QSAR/QSPR models for estimating the boiling points of aliphatic hydrocarbons (alkanes, alkenes, and alkynes) were studied using seven molecular descriptors. The molecular descriptors

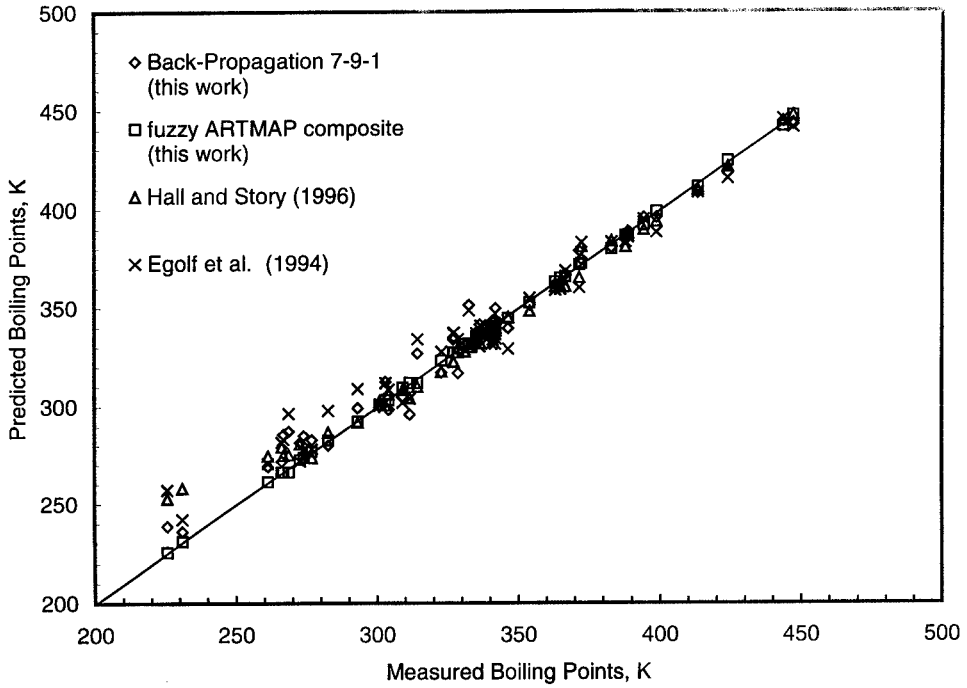

Figure 8. Comparison of estimated boiling points for selected alkanes and alkenes.

included the first, second, third, and fourth order valance molecular connectivity indices, a second-order kappa shape index, dipole moment, and molecular weight. The addition of a dipole moment as an input descriptor enabled the back-propagation network to distinguish between cis and trans isomers; however, it had little effect with the Fuzzy ARTMAP models. The Fuzzy ARTMAP and 7−4−1 back-propagation alkane models predicted boiling points with an average absolute error of 0.31% (1.30 K), and 0.4% (1.54 K), respectively. The absolute error for boiling point predictions based on the overall set of alkenes using a Fuzzy ARTMAP and a 7−10−1 back-propagation network architecture was 0.25% (0.95 K) and 1.3% (6 K). A composite back-propagation model for the three aliphatic types using a 7−9−1 architecture had an average absolute error of 0.75% (2.86 K), 1.77% (6.16 K), and 2.04% (6.85 K) for alkanes, alkenes, and alkynes, respectively, and an overall absolute error of 1.37% (4.85 K) for the entire data set. A substantial improvement in the accuracy of boiling point predictions was obtained with the Fuzzy ARTMAP composite model, with overall absolute error and maximum error of 0.35% (1.35 K) and 0.98% (3.45 K), respectively, although differentiation among isomers was inconsistent.

The present study with boiling points of aliphatic hydrocarbons demonstrated that Fuzzy ARTMAP neural network-based models lead to QSPRs of high accuracy. Recognition of geometric isomers using back-propagation was possible with the addition of dipole moment as a molecular descriptor. For both types of neural network-based QSPRs we utilized a modest set of descriptors which are simple and readily calculable. Current work is underway to expand and test our neural network/QSPR approach for the prediction of multiple physicochemical properties and with an expanded set of chemical descriptors, if necessary, to obtain higher resolution of chemical classification and greater accuracy.

REFERENCES AND NOTES

(1) Sabljić, A.; Horvatić, D. Graph III: A Computer Program from Calculation Molecular Connectivity Indices on Microcomputers. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 837−843.
(2) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(3) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
(4) Weiner, H. Prediction of Isomeric Difference in Paraffin Properties. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(5) Kier, L. B. A Shape Index from Molecular Graphs. *Quant. Struct-Act. Relat.* **1985**, *4*, 109−116.
(6) Kier, L. B.; Hall, L. H. An Electrotopological State Index for Atoms in Molecules. *Pharm. Res.* **1990**, *7*, 801−807.
(7) Medir, M.; Giralt, F. Correlation of Activity Coefficients of Hydrocarbons in Water at Infinity Dilution with Molecular Parameters. *AIChE J.* **1982**, *28*, 341−343.
(8) Ivanciuc, O. Artificial Neural Networks Applications. Part 9. MolNet Prediction of Alkane Boiling Points. *Rev. Roum. Chim.* **1998**, *43*, 885−894.
(9) Arupjyoti, S.; Iragavarapu, S. New Electrotopological Descriptors for Prediction of Boiling Points of Alkanes and Aliphatic Alcohols Through Artificial Neural Networks and Multiple Linear Regression Analysis. *Computers Chem.* **1998** 22, No. 6, 515-522.
(10) Ivanciuc, O.; Balaban, A. Design of Topological Indices. Part 6. A New Topological Parameter for the Steric Effect of Alkyl Substituents. *Croat. Chem. Acta* **1996**, *69*, 75−83.
(11) Basak, S.; Gute, B.; Grunwald, G. A Comparative Study of Topological and Geometric Parameters in Estimating Normal Boiling Points and Octanol/Water Partition Coefficient. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1054−1060.
(12) Dowdy, D.; McKone, J.; Hsieh, D. Prediction of Chemical Biotransfer of Organic Chemicals from Cattle Diet into Beef and Milk using the Molecular Connectivity Index. *Environ. Sci. Technol.* **1996**, *30*, 984−989.
(13) Katritzky, A.; Mu, L. A QSPR study of the solubility of gases and vapors in water. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1162−1168.
(14) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis;* John Wiley & Sons Inc.: New York, 1985.
(15) Pogliani, P. Molecular Modeling by Linear Combinations of Connectivity Indexes. *J. Phy. Chem.* **1995**, *99*, 925−937.
(16) Pogliani, L. A Strategy for Molecular Modeling of a Physicochemical Properties using a Linear Combination of Connectivity Indexes. *Croat. Chem. Acta* **1996**, *69*, 95−109.
(17) Balaban, A.; Basak, C.; Colburn, T.; Grunwald, G. Correlation between Structure and Normal Boiling Points of Haloalkanes C1−C4 using Neural Networks. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1118−1121.

(18) Estrada, E.; Ramirez, A. Edge Adjacency Relationships and Molecular Topographic Descriptors. Definitions and QSAR Applications. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 837−843.
(19) Gakh, A.; Gakh, E.; Sumpter, B.; Noid, D. Neural Network-Graph Theory Approach to the Prediction of the Physical Properties of Organic Compounds. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 832−839.
(20) Zhang, R., Kiu, S.; Liu, M.; Hu, Z. Neural Networks-Molecular Descriptors Approach to the Prediction of Properties of Alkenes. *Computers Chem.* **1997**, *21*, No. 5, 335−342.
(21) Egolf, L.; Wessel, M.; Jurs, P. Prediction of Boiling Points and Critical Temperatures of Industrially Important Organic Compounds from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 947−956.
(22) Hall, L. H.; Story, T. Boiling Points and Critical Temperature of a Heterogeneous Data Set: QSAR with Atom Type Electrotopological State Indices using Artificial Neural Networks. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1004−1014.
(23) Carpenter, G.; Grossberg, S. Fuzzy ARTMAP: A Synthesis of Neural Networks and Fuzzy Logic for Supervised Categorization and Non-stationary Prediction. In *Fuzzy Sets, Neural Networks, and Soft Computing*; Yager and Zadeh, Eds.; Van Nostrand Reinhold: 1994.
(24) Carpenter, G. A.; Grossberg, S.; Marcuzon, N.; Reynolds, J. H.; Rosen, D. B. Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analogue Multidimensional Maps. *IEEE Trans. Neural Networks* **1992**, *3*(5), 698−713.
(25) Ferre-Gine, J.; Rallo, R.; Arenas, A.; Giralt, F. Extraction of structures embedded in the velocity field of a turbulent wake, in Solving Engineering Problems with Neural Networks, Proceedings of the International Conference on Engineering Applications of Neural Networks (EANN'96); Bulsari, A. B., Kallio, S., Tsaptsinos, D., Turku, Eds.; 1996; Vol. 1, pp 17−20.
(26) Giralt, F.; Arenas, A.; Ferre-Giné, J.; Rallo, R. The Simulation and Interpretation of Turbulence with a Cognitive Neural System. *Phys. Fluids* (submitted 1999).
(27) Ivanciuc, O. MolNet Neural Network Application in Structure−Property Studies. *The 23rd Chemistry Conference*; Romania, October 8−10, 1997.
(28) Diudea, M. V.; Ivanciuc, O. *Molecular Topology*; Comprex: Cluj, Romania, 1995.
(29) Ivanciuc, O.; Balaban, A. T. Graph Theory in Chemistry. In *Encyclopedia of Computational Chemistry*; Schleyer, P. V. R., Ed.; Wiley: 1998.
(30) Ivanciuc, O. *Rev. Roum. Chim.* **1989**, *34*, 1361−1368.
(31) Ivanciuc, O.; Balaban, T. S.; Balaban, A. T. *J Math Chem.* **1993**, *12*, 309−318.
(32) Stanton, D. T.; Jurs, P. C. Development and Use of Charged Partial Surface Area Structural Descriptors for Quantitative Structural-Property Relationship Studies. *Anal. Chem.* **1990**, *62*, 2323.
(33) Randić, M. On Molecular Identification Numbers. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164.
(34) *POC* (*Properties of Organic Compounds*) − Personal Edition, Version 5.1; CRC Press: Boca Raton, FL, 1996.
(35) Reid, R. C.; Prausnitz, J. M.; Sherwood, T. K. *The Properties of Gases and Liquids*, 3rd ed.; McGraw-Hill: New York, 1977.
(36) Design Institute for Physical Property Data (DIPPR) − Physical and thermodynamic property database from PRO/II Version 4.02; SIMSCI Simulation Sciences Inc.: 1996.
(37) Molecular Modeling Pro.− Revision 3.14; ChemSM Inc.: 1998.
(38) NeuralSim Software; Aspen Technology, Inc.: 1999.
(39) NeuralWorks Professional II/PLUS Neural Networks Software Version 5.30: Aspen Technology, Inc.: 1997.
(40) Meissner, H., P. Critical Constants from Parachor and Molar Refraction. *Chem Eng. Prog.* **1949**, *45*, 149−53.

CI000442U