

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/231390335>

Prediction of the θ (UCST) of Polymer Solutions: A Quantitative Structure–Property Relationship Study

ARTICLE *in* INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH · SEPTEMBER 2009

Impact Factor: 2.59 · DOI: 10.1021/ie9000426

CITATIONS

17

READS

11

2 AUTHORS:



Farhad Gharagheizi

Texas Tech University

167 PUBLICATIONS 2,907 CITATIONS

SEE PROFILE



Mehdi Sattari

University of KwaZulu-Natal

40 PUBLICATIONS 324 CITATIONS

SEE PROFILE

CORRELATIONS

Prediction of the θ (UCST) of Polymer Solutions: A Quantitative Structure–Property Relationship Study

Farhad Gharagheizi*[†] and Mehdi Sattari[‡]

Department of Chemical Engineering, Faculty of Engineering, University of Tehran,
P.O. Box 11365-4563, Tehran, Iran, and Division of Polymer Science and Technology, Research Institute of
Petroleum Industry (RIPI), P.O. Box 14665-1998, Tehran, Iran

One of the industrially important thermodynamic properties of polymer solutions is the upper critical solution temperature at the limit of infinite chain length of polymers, which is used to realize the usage limits of polymer solutions; this property is denoted as the θ (UCST). In this study, the quantitative structure–property relationship technique (QSPR) was used to correlate the θ (UCST) of polymer solutions. Based on molecular descriptors calculated from the chemical structures of the polymer and the solvent, a nine-parameter multilinear equation was obtained. This correlation can predict the θ (UCST) of 107 polymer solutions with a mean relative error of 6.12% and squared correlation coefficient of 0.912.

Introduction

Knowledge of the phase behavior of polymer solutions is a necessity for industrial and laboratory processes, and its understanding allows optimization of the design and processing conditions. Partially miscible polymer solutions usually have two boundaries: an upper critical solution temperature (UCST) and a lower critical solution temperature (LCST). These two characteristics are dependent on the molar mass of polymer and also the pressure.

The θ temperature is defined as a critical miscibility temperature at the limit of infinite molar mass. A solution may exhibit two critical miscibility temperatures; therefore, a polymer solution may display two θ temperatures.

In the case of exothermic solutions, an increasing in temperature leads to less-positive second virial coefficients and the θ temperature corresponds to a LCST in the limit of infinite molar mass (θ (LCST)). The θ temperature corresponds to a UCST in the case of endothermic solutions (θ (UCST)).¹

θ (UCST) is the upper critical solution temperature at infinite molar mass; as a result, it is not affected by the molar mass of polymers. Also, it can serve as a lower temperature limit for polymer processing in the form of solution. Solvent systems that exhibit this UCST behavior have been suggested for applications where partial miscibility above and below UCST offers advantages.^{2,3}

The ability to predict the physicochemical properties of materials from their molecular structure has been one of the wishes of scientists and engineers for a long time. One of the methods applied for this purpose is the quantitative structure–property relationship (QSPR). QSPR is defined as a mathematical model that predicts the physical, mechanical, or chemical properties of materials from their chemical structures. The main goal of QSPR studies is to find a relationship between the

structure of a compound expressed in terms of numeric characteristics associated with its chemical structure (called molecular descriptors) and the properties of interest.

Once a correlation between structure and a desired property has been found, any number of compounds, including those not yet synthesized, or not yet prepared, can be readily screened using a computer to select structures with the desired properties. Thus, the QSPR approach conserves resources and accelerates the development process of new molecules to use for any purpose.

There are more than 3000 molecular descriptors in the literature that can be used to represent the molecular structure, and the usual problem in this context is to select those which are the most representative for the property under consideration.

In comparison with numerous methods presented to estimate θ (LCST) of polymer solutions,^{4–15} there is no computational method for estimation of the properties. Perhaps, this problem is related to the rarity of the experimental data for θ (UCST) in the literature, in comparison with θ (UCST).

In the present study, the genetic algorithm-based multivariate linear regression (GA-MLR) technique is applied to select the most statistically effective molecular descriptors on θ (UCST) from a pool of 3328 molecular descriptors. The result is a multilinear equation that can be used to predict the θ (UCST) of polymer solutions.

Materials and Methods

Dataset Preparation. The source of experimental data in this study was the compilation provided by Wohlfarth.¹⁶ Consequently, a set of 107 experimental θ (UCST) data was found in this reference and used in this study. These experimental data have been collected from other original literature.^{17–51} This dataset includes 23 polymers and 92 solvents, which is presented in Table 1.

Each polymer solution is a combination of a polymer and a solvent. Therefore, the phase behavior of each binary system is the result of interactions between both polymer and solvent. To consider these interactions, all molecular descriptors of both parts were examined to develop a molecular-based model.

* To whom correspondence should be addressed. Fax: +98 21 66957784. E-mail addresses: fghara@ut.ac.ir, fghara@gmail.com.

[†] Department of Chemical Engineering, Faculty of Engineering, University of Tehran.

[‡] Division of Polymer Science and Technology, Research Institute of Petroleum Industry (RIPI).

Table 1. Predicted θ (UCST) for 107 Polymer–Solvent Combinations

No.	polymer	solvent	status	UCST [K]		Molecular Descriptors										ref
				exp.	pred.	X4A _p	IVDE _p	SPAM _p	FDI _p	K _{Sp}	BE _{Lm} 7 _s	Mor16u _s	Mor1Im _s	H-052 _s		
1	cellulose diacetate	2-butanone	training	310	285.08	0.119	1.58	0.39	0.927	0.412	0	0.029	-0.12	5	40	
2	2-propanone	2-propanone	test	280	278.17	0.119	1.58	0.39	0.927	0.412	0	0.013	-0.089	6	27	
3	cellulose triacrylate	N,N-dimethylformamide	training	413	401.36	0.135	1.318	0.372	0.873	0.383	0	-0.049	0.034	0	18	
4	cellulose triacetate	benzyl alcohol	training	341	367.94	0.119	1.581	0.414	0.929	0.481	0.311	0.084	0.329	0	38	
5	2-propanone	2-propanone	test	300	278.93	0.119	1.581	0.414	0.929	0.481	0	0.013	-0.089	6	27	
6	hydroxypropylcellulose	water	training	343.85	382.27	0.131	1.557	0.443	0.945	0.731	0	0.128	0	0	28	
7	methylcellulose	water	training	332.63	349.89	0.139	1.585	0.415	0.938	0.533	0	0.128	0	0	30	
8	poly(1-butene) (isotactic)	anisole	test	362.3	396.86	0	1	0.449	0.941	0.909	0.238	0.041	0.201	0	19	
9	poly(butyl methacrylate)	2-propanol	training	297.3	338.58	0.175	1.522	0.487	0.975	0.818	0	0.157	0.036	0	42	
10	poly(4-chlorostyrene)	benzene	training	274	255.72	0.145	1.436	0.525	1	0.821	0.188	0.413	0.248	0	23	
11		ethylbenzene	test	258.45	270.61	0.145	1.436	0.525	1	0.821	0.311	0.396	0.338	0	29	
12		ethyl chloroacetate	training	271.35	280.44	0.145	1.436	0.525	1	0.821	0	0.036	-0.127	0	29	
13		isopropylbenzene	training	332.15	288.76	0.145	1.436	0.525	1	0.821	0.464	0.353	0.28	0	29	
14		isopropyl chloroacetate	training	264.95	279.71	0.145	1.436	0.525	1	0.821	0	0.124	-0.232	0	29	
15		methyl chloroacetate	test	337.75	295.45	0.145	1.436	0.525	1	0.821	0	-0.159	-0.18	0	29	
16		tetrachloroethene	training	317.55	285.09	0.145	1.436	0.525	1	0.821	0	0.161	-0.557	0	29	
17		tetrachloromethane	training	323.85	317.11	0.145	1.436	0.525	1	0.821	0	-0.013	-0.912	0	29	
18	poly(ethyl acrylate)	1-butanol	training	318.05	318.2	0.232	1.449	0.487	0.966	0.73	0	0.171	0.109	0	25	
19		ethanol	test	310.55	326.62	0.232	1.449	0.487	0.966	0.73	0	0.082	0.061	0	25	
20		methanol	training	293.65	340.31	0.232	1.449	0.487	0.966	0.73	0	-0.067	0.087	0	25	
21		1-propanol	training	312.65	326.14	0.232	1.449	0.487	0.966	0.73	0	0.093	0.079	0	25	
22	polyethylene (linear)	anisole	training	426.65	414.49	0	0	0.442	0.894	1	0.238	0.041	0.201	0	24	
23		benzyl phenyl ether	training	462.65	481.98	0	0	0.442	0.894	1	0.981	0.332	0.349	0	24	
24		biphenyl	test	400.65	445.02	0	0	0.442	0.894	1	0.802	0.395	0.529	0	24	
25		1-decanol	training	426.45	431.84	0	0	0.442	0.894	1	0.935	0.706	0.706	0	24	
26		diphenyl ether	training	437.05	436.03	0	0	0.442	0.894	1	0.918	0.627	0.676	0	24	
27		diphenylmethane	training	415.35	423.67	0	0	0.442	0.894	1	0.918	0.729	0.812	0	24	
28		1-dodecanol	training	410.15	431.24	0	0	0.442	0.894	1	1.116	0.88	0.979	0	44	
29		4-nonylphenol	training	435.55	437.53	0	0	0.442	0.894	1	1.082	0.663	1.03	0	24	
30		1-octanol	training	453.25	426.5	0	0	0.442	0.894	1	0.728	0.555	0.428	0	24	
31		4-octylphenol	training	447.65	450.85	0	0	0.442	0.894	1	1.082	0.562	0.896	0	24	
32		pentyl acetate	test	434	404.78	0	0	0.442	0.894	1	0.55	0.377	-0.019	3	33	
33	poly(p-hexylstyrene)	4-tert-pentylphenol	training	463.35	459.34	0	0	0.442	0.894	1	0.936	0.569	0.294	0	24	
34	poly(2-hydroxyethyl methacrylate)	2-butanone	training	302.6	268.22	0.147	1.149	0.464	0.999	0.908	0	0.029	-0.12	5	43	
35		1-propanol	test	326.5	344.98	0.172	1.53	0.468	0.949	0.695	0	0.093	0.079	0	26	
36	polyisobutylene	diphenyl ether	training	306	301.11	0	0.811	0.406	0.934	0.465	0.918	0.627	0.676	0	17	
37		ethylbenzene	training	249	262.3	0	0.811	0.406	0.934	0.465	0.311	0.396	0.338	0	17	
38		ethyl heptanoate	test	306	314.21	0	0.811	0.406	0.934	0.465	0.854	0.394	0.069	2	17	
39		ethyl hexanoate	training	330	311.17	0	0.811	0.406	0.934	0.465	0.854	0.394	0.069	2	17	
40	polyisobutylene	toluene	training	260	262.53	0	0.811	0.406	0.934	0.465	0.232	0.254	0.298	0	17	
41	poly(dl-lactide)	dibutyl phthalate	training	358.8	389.83	0.131	1.522	0.422	0.953	0.506	1.141	0.909	0.213	0	51	
42		dipentyl phthalate	test	415.8	366.36	0.131	1.522	0.422	0.953	0.506	1.193	0.862	0.785	0	51	
43	poly(methyl methacrylate)	acetonitrile	training	314	272.29	0.236	1.379	0.441	0.944	0.459	0	-0.082	0.002	3	37	
44		1-chlorobutane	training	320	301.44	0.236	1.379	0.441	0.944	0.459	0	0.075	-0.045	0	37	
45		2,2-dimethyl-3-pentanone	test	308.15	316.35	0.236	1.379	0.441	0.944	0.459	0.362	0.071	0.116	2	20	
46		2,4-dimethyl-3-pentanone	training	319.15	341.65	0.236	1.379	0.441	0.944	0.459	0.441	-0.017	-0.043	2	20	
47		2-ethylbutanal	training	295.15	327.55	0.236	1.379	0.441	0.944	0.459	0.332	0.149	-0.006	1	20	
48		3-heptanone	training	307.7	316.78	0.236	1.379	0.441	0.944	0.459	0.582	0.162	0.052	4	42	
49		4-heptanone	training	309	323.12	0.236	1.379	0.441	0.944	0.459	0.622	0.161	0.037	4	37	
50		2-octanone	test	325.15	296.02	0.236	1.379	0.441	0.944	0.459	0.669	0.154	0.154	5	20	
51		3-octanone	training	346.85	319.42	0.236	1.379	0.441	0.944	0.459	0.724	0.251	0.121	4	47	
52	poly(methyl methacrylate) (isotactic)	acetonitrile	training	301	273.56	0.236	1.379	0.441	0.944	0.459	0	-0.081	0.003	3	37	
53		1-chlorobutane	training	309	302.84	0.236	1.379	0.441	0.944	0.459	0	0.075	-0.045	0	37	
54		4-heptanone	test	319	321.2	0.236	1.379	0.441	0.944	0.459	0.622	0.161	0.037	4	37	
55		3-hexanone	training	279	263.64	0.236	1.379	0.441	0.944	0.459	0.126	0.113	0.012	4	37	
56	poly(4-methyl-1-pentene) (isotactic)	biphenyl	training	467.8	463.28	0.289	1.459	0.431	0.947	0.745	0.802	0.395	0.529	0	32	
57		diphenyl ether	training	483.2	445.31	0.289	1.459	0.431	0.947	0.745	0.918	0.627	0.676	0	32	
58		diphenylmethane	test	449.8	442.33	0.289	1.459	0.431	0.947	0.745	0.918	0.728	0.811	0	32	

Table 1. Continued

No.	polymer	solvent	status	UCST [K]		Molecular Descriptors										ref
				exp.	pred.	X4A _p	IVDE _p	SPAM _p	FDI _p	K _{sp}	BELm7 _s	Mor16u _s	Mor1Im _s	H-052 _s		
59	poly(α-methylstyrene)	methylcyclohexane	training	357	342.92	0.145	1.436	0.47	0.99	0.744	0.568	0.406	0.089	0	48	
60	poly(2-methyl-5-vinylpyridine)	butyl acetate	training	294.95	318.8	0.157	0.986	0.463	0.977	0.791	0.403	0.185	-0.09	3	22	
61		ethyl butyrate	training	323.15	314.01	0.157	0.986	0.463	0.977	0.791	0.349	0.28	-0.05	2	22	
62		ethyl propionate	training	298.55	284.92	0.157	0.986	0.463	0.977	0.791	0	0.069	-0.025	2	22	
63		3-methylbutyl acetate	test	322.15	320.43	0.157	0.986	0.463	0.977	0.791	0.41	0.158	-0.095	3	22	
64		2-methylpropyl acetate	training	326.15	311.19	0.157	0.986	0.463	0.977	0.791	0.364	0.251	-0.179	3	22	
65		pentyl acetate	training	321.35	318.42	0.157	0.986	0.463	0.977	0.791	0.55	0.377	-0.019	3	22	
66	poly(2-methyl-5-vinylpyridine)	propionitrile	training	269.55	284.91	0.145	1.436	0.47	0.99	0.744	0	0.005	-0.013	2	22	
67		propyl acetate	training	292.45	264.43	0.145	1.436	0.47	0.99	0.744	0	0.131	-0.101	3	22	
68		propyl propionate	test	331.15	317.79	0.145	1.436	0.47	0.99	0.744	0.352	0.138	-0.029	2	22	
69		tetrahydronaphthalene	training	322.65	314.7	0.145	1.436	0.47	0.99	0.744	0.561	0.645	0.326	0	22	
70	polypropylene (isotactic)	benzyl phenyl ether	training	455	443.64	0	0.918	0.449	0.922	0.81	0.981	0.332	0.349	0	24	
71		benzyl propionate	training	430.7	415.98	0	0.918	0.449	0.922	0.81	0.794	0.167	0.166	2	24	
72		biphenyl	training	398.3	412.03	0	0.918	0.449	0.922	0.81	0.802	0.359	0.529	0	24	
73		4-tert-butylphenol	test	439.2	442.57	0	0.918	0.449	0.922	0.81	0.914	0.328	0.232	0	24	
74		dibenzyl ether	training	456.4	438.86	0	0.918	0.449	0.922	0.81	1.082	0.568	0.351	0	24	
75		diphenyl ether	training	416	400.2	0	0.918	0.449	0.922	0.81	0.918	0.627	0.676	0	24	
76		4-tert-pentylphenol	training	414	426.57	0	0.918	0.449	0.922	0.81	0.936	0.57	0.295	0	24	
77	polystyrene	1-chloro- <i>n</i> -decane	training	279.7	312.74	0.16	1.061	0.471	0.992	0.688	0.91	0.573	0.452	0	46	
78		1-chloro- <i>n</i> -dodecane	training	331.7	308.57	0.16	1.061	0.471	0.992	0.688	1.094	0.733	0.701	0	46	
79		1-chloro- <i>n</i> -undecane	training	305.9	313.85	0.16	1.061	0.471	0.992	0.688	1.006	0.63	0.566	0	46	
80		1-nitropropane	training	272	247.95	0.16	1.061	0.471	0.992	0.688	0	0.045	0.179	0	46	
81		2,5-dimethyl furan	test	145	162.62	0.16	1.061	0.471	0.992	0.688	0	0.185	0.298	6	46	
82		2-methylpropyl acetate	training	227	261.81	0.16	1.061	0.471	0.992	0.688	0.364	0.251	-0.179	3	34	
83		2-propyl acetate	training	250	237.54	0.16	1.061	0.471	0.992	0.688	0	-0.017	-0.258	3	35	
84		3-methylbutyl acetate	training	224	270.54	0.16	1.061	0.471	0.992	0.688	0.41	0.157	-0.095	3	34	
85		cyclodecane	training	289	286.19	0.16	1.061	0.471	0.992	0.688	0.758	0.762	0.293	0	41	
86		cycloheptane	training	290	294.94	0.16	1.061	0.471	0.992	0.688	0.615	0.438	0.229	0	41	
87		cyclohexane	training	306.51	282.77	0.16	1.061	0.471	0.992	0.688	0.484	0.495	0.017	0	50	
88		cyclooctane	training	286	308.65	0.16	1.061	0.471	0.992	0.688	0.583	0.182	0.266	0	41	
89		cyclopentane	training	293	273.96	0.16	1.061	0.471	0.992	0.688	0	-0.142	-0.101	0	31	
90		diethyl malonate	training	309	308.65	0.16	1.061	0.471	0.992	0.688	0.501	0.131	-0.433	2	35	
91		dl-menthol	training	388.1	347.69	0.16	1.061	0.471	0.992	0.688	1.029	0.307	0.264	0	46	
92		dl-terpineol	training	351.6	362	0.16	1.061	0.471	0.992	0.688	0.97	0.141	0.193	0	46	
93		dodecauterocyclohexane	test	312.5	284.55	0.16	1.061	0.471	0.992	0.688	0.484	0.495	0.017	0	36	
94		dodecyl acetate	training	285.2	284.23	0.16	1.061	0.471	0.992	0.688	1.305	0.943	0.724	3	49	
95		ethyl acetate	training	229	218.93	0.16	1.061	0.471	0.992	0.688	0	0.138	-0.087	3	34	
96		ethyl chloroacetate	training	255	265.3	0.16	1.061	0.471	0.992	0.688	0	0.036	-0.127	0	46	
97		ethylcyclohexane	training	342.95	324.47	0.16	1.061	0.471	0.992	0.688	0.666	0.211	0.031	0	21	
98		hexyl- <i>n</i> -xylene	training	285.5	306.22	0.16	1.061	0.471	0.992	0.688	1.102	0.675	0.971	0	46	
99		isoamyl acetate	training	224	270.46	0.16	1.061	0.471	0.992	0.688	0.41	0.158	-0.095	3	46	
100		isobutyl acetate	training	227	261.81	0.16	1.061	0.471	0.992	0.688	0.364	0.251	-0.179	3	46	
101		isopropyl acetate	training	246	237.6	0.16	1.061	0.471	0.992	0.688	0	-0.01	-0.265	3	46	
102		methylcyclopentane	training	348	308.59	0.16	1.061	0.471	0.992	0.688	0.353	-0.11	-0.036	0	45	
103		<i>n</i> -butyl formate	test	264	256.54	0.16	1.061	0.471	0.992	0.688	0	0.072	-0.007	0	46	
104		<i>n</i> -propyl acetate	training	193	222.51	0.16	1.061	0.471	0.992	0.688	0	0.131	-0.101	3	46	
105		o-dichlorobenzene	training	220	211.97	0.16	1.061	0.471	0.992	0.688	0	0.15	0.804	0	46	
106		propyl acetate	training	193	222.53	0.16	1.061	0.471	0.992	0.688	0	0.13	-0.1	3	34	
107		trans-decahydronaphthalene	training	293	266.9	0.16	1.061	0.471	0.992	0.688	0.758	0.949	0.332	0	39	

Table 2. Molecular Descriptors of eq and Their Physical Meanings

ID	molecular descriptor	type	definition
1	X4A _p	connectivity indices	average connectivity index chi-4
2	IVDE _p	information indices	mean information content on vertex degree equality
3	SPAM _p	geometrical descriptors	average span R
4	FDI _p	geometrical descriptors	fouling degree index
5	Ks _p	WHIM descriptors	K global shape index/weighted by atomic electrotopological states
6	BELm _{6s}	Burden eigenvalues	lowest eigenvalue n.7 of Burden matrix/weighted by atomic masses
7	Mor16u _s	3D-MORSE descriptors	3D-MORSE signal 16/unweighted
8	Mor11m _s	3D-MORSE descriptors	3D-MORSE signal 11/weighted by atomic masses
9	H-052 _s	atomic-centered fragments	H attached to C0(sp3) with 1X attached to next C

Table 3. Correlation Matrix for Nine Molecular Descriptors

	X4A _p	IVDE _p	SPAM _p	FDI _p	Ks _p	BELm _{6s}	Mor16u _s	Mor11m _s	H-051 _s
X4A _p	1								
IVDE _p	0.530954	1							
SPAM _p	0.0696902	0.0344682	1						
FDI _p	0.304256	0.289409	0.510212	1					
Ks _p	0.231933	0.397736	0.180189	0.0130519	1				
BELm _{6s}	0.150714	0.219079	0.0636215	0.119922	0.0535737	1			
Mor16u _s	0.113193	0.155216	0.0195326	0.0469949	0.0758474	0.676258	1		
Mor11m _s	0.115135	0.19526	0.0800664	0.17373	0.0679118	0.485606	0.535392	1	
H-051 _s	0.0626947	0.0385447	0.0422783	0.0058291	0.132254	0.0688008	0.105682	0.149223	1

In this work, the hydrogen-saturated repeating unit of polymers (instead of polymers structure) was used to determine the molecular descriptors of polymers. After providing the dataset, all molecules (polymers and solvents) were drawn into Hyperchem software⁵² and optimized using the MM+ molecular mechanics force field and, thereafter, molecular descriptors were calculated using Dragon software.⁵³ Dragon can calculate 1664 molecular descriptors for every chemical structure. Molecular descriptors of ~234 000 pure compounds have been calculated by Dragon are freely accessible from Milano Chemometrics and QSAR research group Web site (http://michem.disat.unimib.it/mol_db). For more information about the types of these molecular descriptors, please refer to Dragon software user's guide.⁵³

In this part, a total of 3328 molecular descriptors for each polymer solution were computed (1664 molecular descriptors for polymer and 1664 molecular descriptors for solvent). Note that some complementary works should be performed over the pool of descriptors obtained from Dragon software. After calculation of molecular descriptors, the pool of molecular descriptors was reduced by removing descriptors with an essentially constant value for all the structures and those could not be calculated for every structure in the dataset. The pair correlation cutoff selection then was performed. In this method, for each pair of variables with a squared correlation coefficient (R^2) of >0.8 (the cutoff value), one of two descriptors is arbitrary eliminated. After these operations, our pool of descriptors decreased from 3328 descriptors to 710 (227 for polymer and 483 solvent); hence, we used a pool of descriptors that contained 710 molecular descriptors for subset variable selection.

GA-MLR Calculations. Generally, in QSPR studies, after calculation of molecular descriptors, the problem is to find a linear equation that can predict the desired property with the least number of variables, as well as the highest accuracy. In other words, the problem is to find a subset of variables (most statistically effective molecular descriptors on θ (UCST)) from all available variables (all molecular descriptors) to predict θ (UCST) with minimum error, in comparison with the available data.

A generally accepted method for this problem is genetic algorithm-based multivariate linear regression (GA-MLR). In this method, a genetic algorithm is used to select the best subset

variables, with respect to an objective function, which was presented by Leardi et al. for the first time.⁵⁴

In this study, the GA-MLR technique presented by Leardi et al., with a couple of RQK objective functions presented by Todeschini et al.,^{55,56} were used for subset variable selection. This methodology has been extensively presented in our previous works, and the results are satisfactory.^{14,57–73}

Before performing the GA-MLR technique, the dataset must be divided into two new collections. First one is allocated for training and second one for testing. Using the training set, the best model is found and then the predictive power of the obtained model is checked by the test set as external dataset. In this work, 80% of the dataset was selected randomly for the training set and remainder was used for the test set (from 107 polymer solutions, 86 polymer solutions in the training set and 21 polymer solutions in the test set).

The inputs of our program were the pool of molecular descriptors, the θ (UCST) of polymer solutions, and the number of molecular descriptors that we wanted to enter into our final model.

To obtain the best multivariate linear equation, all molecular descriptors must be introduced to the program and the minimum number of possible variables must be tested at the starting point. Therefore, the program was started with one variable. We then obtained the best multivariate linear model. In the next steps, we increased the number of desired variables to two, three, four, and so on, and all calculations were repeated for them.

When we saw that increasing in the number of variables had no considerable effect on the accuracy of the best-obtained model, the calculations was stopped, because the best multivariate linear model already had been obtained.

Results and Discussion

Using the presented procedure, the best multivariate linear equation was obtained. We increased the number of descriptors to develop the best model. This work was continued until the change in the R^2 values of the consequent models was <0.01; hence, the best obtained multivariate linear model in this study has 9 parameters. Note that the difference between the best 8-parameter multivariate linear model and the best 9-parameter multivariate linear model was 0.0101; on the other hand, the difference between the best 9-parameter multivariate linear

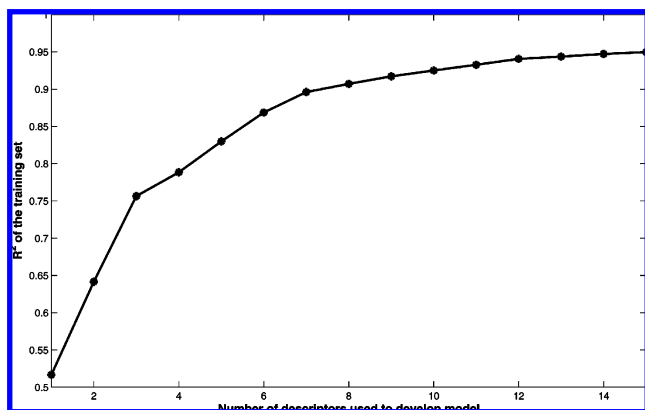


Figure 1. Squared correlation coefficient (R^2) for the best multivariate linear model with 1–15 parameters.

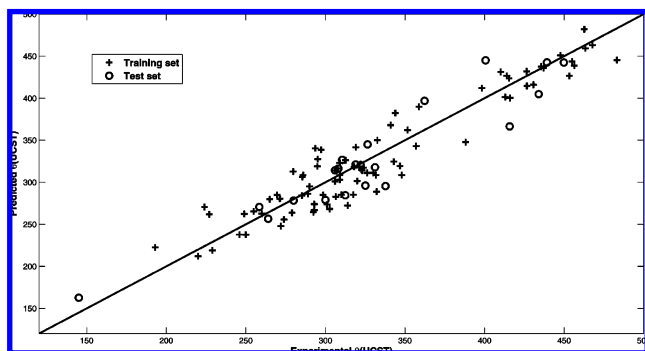


Figure 2. Comparison between the $\theta(\text{UCST})$ obtained by eq and experimental data.

model and the best 10-parameter multivariate linear model was 0.0079. The R^2 value of the best multivariate linear models with 1–15 parameters is shown in Figure 1.

The best obtained equation is

$$\begin{aligned} \theta(\text{UCST}) [\text{K}] = & 1525.4392(\pm 82.6631) + \\ & 178.1497(\pm 47.5914)X4A_p + 73.1814(\pm 10.5220)IVDE_p - \\ & 801.022(\pm 189.2475)SPAM_p - 1225.8276(\pm 130.9757)FDI_p + \\ & 324.0948(\pm 29.4254)KS_p + 121.3236(\pm 10.6136)BELm6_s - \\ & 70.0030(\pm 16.7255)Mor16u_s - 45.78383(\pm 11.58467)Mor11m_s - \\ & 12.0077(\pm 1.87561)H-051_s \quad (1) \end{aligned}$$

$$n_{\text{training}} = 86; n_{\text{test}} = 21;$$

$$R_{\text{training}}^2 = 0.9172; R_{\text{test}}^2 = 0.8824;$$

$$Q_{\text{LOO}}^2 = 0.8955; Q_{\text{BOOT}}^2 = 0.8866; Q_{\text{EXT}}^2 = 0.8816;$$

$$s_{\text{training}} = 21.34; s_{\text{training}} = 25.42;$$

$$a = -0.071; F = 93.59$$

RQK function parameters:

$$(\Delta K = 0.006; \Delta Q = 0.002; R^P = 0.004; R^N = 0.000)$$

The predicted values of $\theta(\text{UCST})$ using eq, in comparison with the experimental data, are presented in Figure 2 and are listed in Table 1. Also, the values of the selected molecular descriptors and the status of all of the polymer solutions (training set or test set) are presented in Table 1.

The molecular descriptors and their physical meanings are presented in Table 2. The subscripts “p” and “s” indicate that

the molecular descriptor is related to the polymer and the solvent, respectively.

Table 3 presents the correlation matrix. The nine selected descriptors clearly are not highly correlated.

The numbers of polymer solutions in the training set and in the test set are given as n_{training} and n_{test} , respectively. To check the validity of the model more, bootstrap techniques, y-scrambling, and external validation techniques were used.⁵⁶ The bootstrapping was repeated 5000 times. Also, y-scrambling was repeated 300 times. As can be seen, the difference between each pair of Q_{LOO}^2 , Q_{BOOT}^2 , Q_{EXT}^2 , and R^2 show that the obtained model is a good one and has good predictive power.⁵⁶ Also, the intercept value of the y-scrambling technique has a low value ($a = -0.071$), which reveals the validity of the model. The y-scrambling, bootstrapping, and external validation techniques have been presented extensively by Todeschini et al.⁵⁶

All the validation techniques show that the obtained model is a valid model and can be used to predict the $\theta(\text{UCST})$ of polymer solutions.

Conclusion

In this study, a molecular-based model was presented to predict the $\theta(\text{UCST})$ of polymer solutions. The parameters of the model can be easily calculated only from the chemical structure of the polymer and the solvent (the monomers of polymers are used for polymers).

Also, the validity and predictive power of the model was checked by several techniques. The relative error and squared correlation coefficient (R^2) of the obtained model over 107 polymer solutions are 6.12% and 0.912, respectively. Therefore, the obtained model has predictive power and can be used to predict the $\theta(\text{UCST})$ of polymer solutions.

Literature Cited

- (1) Elias, H. G. Theta Solvents. In *Polymer Handbook*; Brandrup, J., Immergut, H., Grulke, E. A., Eds.; Wiley: New York, 1999; pp VII-291–VII-326.
- (2) Kavanagh, C. A.; Rochev, Y. A.; Gallanger, W. M.; Dawson, K. A.; Keenan, A. K. Local Drug Delivery In Restenosis Injury: Thermoresponsive Co-Polymers As Potential Drug Delivery Systems. *Pharmacol. Ther.* **2004**, *102*, 1.
- (3) Kopecek, J. Smart and Genetically Engineered Biomaterials and Drug Delivery Systems. *Eur. Pharm. Sci.* **2003**, *20*, 1–16.
- (4) Chang, B. H.; Bae, C. Y. Liquid–Liquid Equilibria of Binary Polymer Solutions with Specific Interactions. *Polymer* **1998**, *39*, 6449.
- (5) Pappa, G. D.; Voutsas, E. C.; Tassios, D. P. Correlation and Prediction of the Polymer Molecular Weight and the Pressure Effect. *Ind. Eng. Chem. Res.* **2001**, *40*, 4654.
- (6) Bogdanic, G.; Vidal, J. A Segmental Interaction Model For Liquid–Liquid Equilibrium Calculations for Polymer Solutions. *Fluid Phase Equilib.* **2000**, *173*, 241.
- (7) Imre, A. R.; Bae, Y. C.; Chang, B. H.; Kraska, T. Semiempirical Method for the Prediction of the Theta (Lower Critical Solution Temperature) in Polymer Solutions. *Ind. Eng. Chem. Res.* **2004**, *43*, 237.
- (8) Vetere, A. Empirical Method To Correlate and To Predict the Vapor–Liquid Equilibrium and Liquid–Liquid Equilibrium of Binary Amorphous Polymer Solutions. *Ind. Eng. Chem. Res.* **1998**, *37*, 2864.
- (9) Vetere, A. An Empirical Method to Predict the Liquid–Liquid Equilibria of Binary Polymer Systems. *Ind. Eng. Chem. Res.* **1998**, *37*, 4463.
- (10) Wang, F.; Saeki, S.; Yamaguchi, T. Absolute prediction of upper and lower critical solution temperatures in polymer/solvent systems based on corresponding state theory. *Polymer* **1999**, *40*, 2779.
- (11) Liu, H.; Zhong, C. Modeling of the Θ (Lower Critical Solution Temperature) in Polymer Solutions Using Molecular Connectivity Indices. *Eur. Polym. J.* **2005**, *41*, 139.
- (12) Melagraki, G.; Afantitis, A.; Sarimves, H.; Koutenis, P. A.; Markopoulos, J.; Iglessi-Markopoulou, O. A Novel QSPR Model For

Predicting Θ (Lower Critical Solution Temperature) in Polymer Solutions Using Molecular Descriptors. *J. Mol. Model.* **2007**, *13*, 55.

(13) Xu, J.; Chen, B.; Liang, H. Accurate Prediction of (Lower Critical Solution Temperature) in Polymer Solutions with Topological Indices. *J. Mol. Graph. Model.* **2007**, *26*, 352.

(14) Gharagheizi, F. A New Neural Network Quantitative Structure-Property Relationship for Prediction of θ (Lower Critical Solution Temperature) of Polymer Solutions. e-Polym. 2007; Article No. 114.

(15) Xu, J.; Chen, B.; Liang, H. Accurate Prediction of (Lower Critical Solution Temperature) in Polymer Solutions Based on 3D Descriptors and Artificial Neural Networks. *Macromol. Theor. Simul.* **2008**, *17*, 109.

(16) Wohlfarth, C. *Handbook of Liquid-liquid Equilibrium Data of Polymer Solutions*; CRC Press/Taylor-Francis Group; Boca Raton, FL, 2008.

(17) Fox, T. G.; Flory, P. J. Intrinsic Viscosity-Temperature Relationships for Polyisobutylene in Various Solvents. *J. Am. Chem. Soc.* **1951**, *73*, 1909.

(18) Mandelkern, L.; Flory, P. J. Molecular Dimensions of Cellulose Triesters. *J. Am. Chem. Soc.* **1952**, *74*, 2517.

(19) Krigbaum, W. R.; Kurz, J. E.; Smith, P. Conformation of Polymer Molecules. IV. Poly-(1-Butene). *J. Phys. Chem.* **1961**, *65*, 1984.

(20) Fox, T. G. Properties of Dilute Polymer Solutions III: Intrinsic Viscosity/Temperature Relationships for Conventional Polymethyl Methacrylate. *Polymer* **1962**, *3*, 111.

(21) Debye, P.; Woermann, D.; Chu, B. Critical Opalescence of Polystyrene in Ethylcyclohexane. *J. Polym. Sci., Part A* **1963**, *1*, 255.

(22) Gechele, G. B.; Crescentini, L. Phase Separation, Viscosity, and Thermodynamic Parameters for Poly-2-Methyl-5-Vinylpyridine Diluent System. *J. Polym. Sci. A* **1965**, *3*, 3599.

(23) Kubo, K.; Ogino, K. Solution Properties of Poly(P-Chlorostyrene). *Sci. Pap. Coll. Art. Sci.; Univ. Tokyo* **1966**, *16*, 193.

(24) Nakajima, A.; Fujiwara, H.; Hamada, F. Phase Relationships And Thermodynamic Interactions In Linear Polyethylene-Diluent Systems. *J. Polym. Sci. A-2* **1966**, *4*, 507.

(25) Llopis, J.; Albert, A.; Usobinaga, P. Studies on Poly(ethyl acrylate) in Θ Solvents. *Eur. Polym. J.* **1967**, *3*, 259.

(26) Dusek, K. Solubility of Poly(2-Hydroxyethyl Methacrylate) in Some Aliphatic Alcohols. *Collect. Czech. Chem. Commun.* **1969**, *34*, 3309.

(27) Cowie, J. M. G.; Maconnachie, A.; Ranson, R. J. Phase Equilibria of Cellulose Acetate-Acetone Solutions. The Effect of Degree of Substitution and Molecular Weight on Upper and Lower Critical Solution Temperature. *Macromolecules* **1971**, *4*, 57.

(28) Kagemoto, A.; Baba, Y. Phase Diagrams of Polymer Solutions (in Jpn.). *Kobunshi Kagaku* **1971**, *28*, 784.

(29) Izumi, Y.; Miyake, Y. Study of Linear Poly(P-Chlorostyrene)-Diluent Systems. I. Solubilities, Phase Relationships, and Thermodynamic Interactions. *Polym. J.* **1971**, *3*, 647.

(30) Kagemoto, A.; Baba, Y.; Fujishiro, R. Phase Equilibrium of the Cellulose Derivatives Determined by Differential Thermal Analysis. I. Methylcellulose/Water. *Makromol. Chem.* **1972**, *154*, 105.

(31) Saeki, S.; Kuwahara, N.; Konno, S.; Kaneko, M. Upper and Lower Critical Solution Temperatures in Polystyrene Solutions II. *Macromolecules* **1973**, *6*, 589.

(32) Tani, S.; Hamada, F.; Nakajima, A. Unperturbed Chain Dimensions of Poly(4-Methyl-Pentene-1) in Theta Solvents. *Polym. J.* **1973**, *5*, 86.

(33) Kuwahara, N.; Saeki, S.; Chiba, T.; Kaneko, M. Upper and Lower Critical Solution Temperatures in Polyethylene Solutions. *Polymer* **1974**, *15*, 777.

(34) Saeki, S.; Konno, S.; Kuwahara, N.; Nakata, M.; Kaneko, M. Upper and Lower Critical Solution Temperatures in Polystyrene Solutions III. Temperature Dependence of the X_1 Parameter. *Macromolecules* **1974**, *7*, 521.

(35) Konno, S.; Saeki, S.; Kuwahara, N.; Nakata, M.; Kaneko, M. Upper and Lower Critical Solution Temperatures in Polystyrene Solutions. IV. Role of Configurational Heat Capacity. *Macromolecules* **1975**, *8*, 799.

(36) Strazielle, C.; Benoit, H. Some Thermodynamic Properties of Polymer-Solvent Systems. Comparison Between Deuterated and Undeuterated Systems. *Macromolecules* **1975**, *8*, 203.

(37) Cowie, J. M. G.; McEwen, I. J. Influence of Microstructure on the Upper and Lower Critical Solution Temperatures of Poly(Methyl Methacrylate) Solutions. *J. Chem. Soc. Faraday Trans. I* **1976**, *72*, 526.

(38) Panina, N. I.; Lozacheva, V. P.; Aver'yanova, V. M. Macromolecular Parameters of Cellulose Acetates Near Theta-Temperature (in Russ.). *Vysokomol. Soedin. Ser. B* **1977**, *19*, 786.

(39) Wolf, B. A.; Jend, R. Über Die Möglichkeiten Zur Bestimmung Von Mischungsenthalpien Und -Volumina Aus Der Molekulargewichtabhängigkeit Der Kritischen Entmischungstemperaturen Und -Drücke Am

Beispiel Des Systems Trans- Decahydronaphthalin/Polystyrol. *Makromol. Chem.* **1977**, *178*, 1811.

(40) Suzuki, H.; Muraoka, Y.; Saitoh, M.; Kamide, K. Upper and Lower Critical Solution Temperatures in 2-Butanone Solutions of Cellulose Diacetate. *Br. Polym. J.* **1982**, *14*, 23.

(41) Cowie, J. M. G.; McEwen, I. J. A Comparison of the Phase Behaviour of Polystyrene in Cycloalkanes and *N*-Alkanes. *Br. Polym. J.* **1986**, *18*, 387.

(42) Herold, F. K.; Wolf, B. A. Poly(*N*-alkylmethacrylate)s: Characterization, Good and Poor Solvents, Densities and Intrinsic Viscosities. *Mater. Chem. Phys.* **1986**, *14*, 311.

(43) Magarik, S. Y.; Filippov, A. P.; D'yakonova, N. V. Temperature Dependence of the Intrinsic Viscosity of Polystyrene and Poly(Alkylstyrene) Solutions (in Russ.). *Vysokomol. Soedin. Ser. A* **1987**, *29*, 698.

(44) Chiu, G.; Mandelkern, L. Effect of Molecular Weight on the Phase Diagram of Linear Polyethylene in 1-Dodecanol. *Macromolecules* **1990**, *23*, 5356.

(45) Imre, A.; Van Hook, W. A. Demixing in Polystyrene/Methylcyclohexane Solutions. *J. Polym. Sci. B. Polym. Phys.* **1996**, *34*, 751.

(46) Imre, A.; Van Hook, W. A. Liquid-Liquid Demixing From Solutions of Polystyrene. 1. A Review. 2. Improved Correlation with Solvent Properties. *J. Phys. Chem. Ref. Data* **1996**, *25*, 637.

(47) Xia, K. Q.; An, X. Q.; Shen, W. G. Measured Coexistence Curves of Phase Separated Polymer Solutions. *J. Chem. Phys.* **1996**, *105*, 6018.

(48) Pruessner, M. D.; Retzer, M. E.; Greer, S. C. Phase Separation Curves of Poly(A-methylstyrene) in Methylcyclohexane. *J. Chem. Eng. Data* **1999**, *44*, 1419.

(49) Imre, A.; van Hook, W. A. End Group Effects on Liquid-Liquid Demixing of Polystyrene/Oligomethylene Solutions. Polystyrene/Dodecyl Acetate Solubility. *Macromolecules* **2000**, *33*, 5308.

(50) Siporska, A.; Szydłowski, J.; Rebelo, L. P. N. Solvent H/D Isotope Effects on Miscibility and Theta-Temperature in the Polystyrene-Cyclohexane System. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2996.

(51) Lee, J. S.; Lee, H. K.; Kim, S. C. Thermodynamic Parameters of Poly(lactic acid) Solutions in Dialkyl Phthalate. *Polymer* **2004**, *45*, 4491.

(52) HyperChem Release 7.5 for Windows, Molecular Modeling System, Hypercube, Inc., 2002.

(53) Talete srl, Dragon for windows (Software for Molecular Descriptor Calculations). Version 5.4, 2006; <http://www.talete.mi.it/>.

(54) Leardi, R.; Boggia, R.; Terrile, M. Genetic Algorithm as a Strategy for Feature-Selection. *J. Chemometr.* **1992**, *6*, 281.

(55) Todeschini, R.; Consonni, V. In *Handbook Of Molecular Descriptors*; Manhold, R., Kubinyi, H., Temmerman, H., Series Eds.; Wiley-VCH: Weinheim, Germany, 2000.

(56) Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. Detecting "Bad" Regression Models: Multicriteria Fitness Functions in Regression Analysis. *Anal. Chim. Acta* **2004**, *515*, 199.

(57) Gharagheizi, F. QSPR Analysis For Intrinsic Viscosity of Polymer Solutions by Means of GA-MLR and RBFNN. *Comput. Mater. Sci.* **2007**, *40*, 159.

(58) Gharagheizi, F. Quantitative Structure-Property Relationship for Prediction of the Lower Flammability Limit of Pure Compounds. *Energy Fuels* **2009**, *22*, 3037.

(59) Gharagheizi, F.; Mehrpooya, M. Prediction of Standard Chemical Exergy by a Three Descriptors QSPR Model. *Energy Convers. Manage.* **2007**, *48*, 2453.

(60) Gharagheizi, F.; Alamdari, R. F. A Molecular-Based Model for Prediction of Solubility of C₆₀ Fullerene in Various Solvents. *Fullerenes, Nanotubes, Carbon Nanostruct.* **2008**, *16*, 40.

(61) Gharagheizi, F. QSPR Studied For Solubility Parameter By Means of Genetic Algorithm-Based Multivariate Linear Regression and Generalized Regression Neural Network. *QSAR Comb. Sci.* **2008**, *27*, 165.

(62) Gharagheizi, F. A Simple Equation For Prediction of Net Heat of Combustion of Pure Chemicals. *Chemometr. Intell. Lab. Syst.* **2008**, *91*, 177.

(63) Gharagheizi, F. A New Molecular-Based Model for Prediction of Enthalpy of Sublimation of Pure Components. *Thermochim. Acta* **2008**, *469*, 8.

(64) Gharagheizi, F.; Alamdari, R. F. Prediction of Flash Point Temperature of Pure Components Using a Quantitative Structure-Property Relationship Model. *QSAR Comb. Sci.* **2008**, *27*, 679.

(65) Gharagheizi, F.; Fazeli, A. Prediction of the Watson Characterization Factor of Hydrocarbon Compounds from Molecular Properties. *QSAR Comb. Sci.* **2008**, *27*, 758.

(66) Sattari, M.; Gharagheizi, F. Prediction of Molecular Diffusivity of Pure Components Into Air: A QSPR Approach. *Chemosphere* **2008**, *72*, 1298.

(67) Vatani, A.; Mehrpooya, M.; Gharagheizi, F. Prediction of Standard Enthalpy of Formation by a QSPR Model. *Int. J. Mol. Sci.* **2007**, *8*, 407.

(68) Gharagheizi, F.; Sattari, M. Estimation of Molecular Diffusivity of Pure Chemicals in Water: A Quantitative Structure–Property Relationship Study. *SAR QSAR Environ.* **2009**, *20*, 267.

(69) Gharagheizi, F.; Mehrpooya, M. Prediction of Some Important Physical Properties of Sulfur Compounds Using QSPR Models. *Mol. Divers.* **2008**, *12*, 143.

(70) Gharagheizi, F.; Tirandazi, B.; Barzin, R. Estimation of Aniline Point Temperature of Pure Hydrocarbons: A Quantitative Structure–Property Relationship Approach. *Ind. Eng. Chem. Res.* **2009**, *48*, 1678.

(71) Gharagheizi, F. A QSPR Model for Estimation of Lower Flammability Limit Temperature of Pure Compounds Based on Molecular Structure. *J. Hazard. Mater.* **2009**, *169*, 217.

(72) Gharagheizi, F. Prediction of Upper Flammability Limit Percent of Pure Compounds from Their Molecular Structures. *J. Hazard. Mater.* **2009**, *167*, 507.

(73) Gharagheizi, F. Prediction of Standard Enthalpy of Formation of Pure Compounds Using Molecular Structure. *Aust. J. Chem.* **2009**, *62*, 374.

Received for review January 10, 2009
Revised manuscript received June 23, 2009
Accepted August 20, 2009

IE9000426