Colloid Polym Sci (2009) 287:811–818
DOI 10.1007/s00396-009-2035-y

ORIGINAL CONTRIBUTION

Artificial neural network prediction of glass transition temperature of polymers

Wanqiang Liu · Chenzhong Cao

Received: 8 January 2009 / Revised: 2 April 2009 / Accepted: 6 April 2009 / Published online: 5 May 2009

Abstract In this article, the molecular average polarizability α, the energy of the highest occupied molecular orbital E_{HOMO}, the total thermal energy $E_{thermal}$, and the total entropy S were used to correlate with glass transition temperature T_g for 113 polymers. The quantum chemical descriptors obtained directly from polymer monomers can represent the essential factors that are governing the nature of glass transition in polymers. Stepwise multiple linear regression (MLR) analysis and back-propagation artificial neural network (ANN) were used to generate the model. The final optimum neural network with 4–[4–2]$_2$–1 structure produced a training set root mean square error (RMSE) of 11 K (R=0.973) and a prediction set RMSE of 17 K (R=0.955). The results show that the ANN model obtained in this paper is accurate in the prediction of T_g values for polymers.

Keywords Artificial neural network · Density function theory · Glass transition temperature · QSPR

Introduction

The most familiar and important property of polymeric and composite material is the glass transition temperature, T_g. T_g determines temperature windows for processing and utilizing these material and is a prerequisite for the prediction and understanding of the mechanical and other properties, such as heat capacity, coefficient of thermal expansion, and viscosity [1]. T_g is generally determined from two main types of experiments: thermodynamic versus dynamic measurements. The former include differential scanning calorimetry (DSC), dilatometry, lateral force microscopy, and ellipsometry. Dynamic measurements, on the other hand, generally measure the viscosity or relaxation time as a function of temperature or frequency using techniques, such as dielectric spectroscopy, neutron scattering, solvation dynamics, and nuclear magnetic resonance [2]. T_g is difficult to determine experimentally and still is an unresolved problem [1–3].

Numerous researchers have attempted to predict T_gs for polymers on the basis of quantitative structure–property relationships (QSPRs). According to the view of Katritzky et al., there are two kinds of approaches, the empirical and the theoretical [4]. Empirical methods correlate the target property with other physical or chemical properties of the polymers, for example, group additive properties (GAP) [5]. The GAP method is purely empirical approach and limited to systems composed only of functional groups that have been previously investigated. Furthermore, it is only approximate, since this approach fails to account for the presence of neighboring groups or conformational influences.

The most widely referenced model of the theoretical estimations has been produced by Bicerano [6]. Bicerano produced a regression model (R=0.9749, s=24.65 K) to relate T_g with the solubility parameter and the weighted sum of 13 structural parameters for the data set of 320 polymers. But he has not used external data set compounds to validate this model.

Katritzky et al. [7] introduced a mode with R^2 of 0.928 for 22 medium molecular weight polymers using four parameters. After that work, Katritzky et al. [4] applied the CODESSA method to predict T_gs for 88 linear homopol-

W. Liu (✉) · C. Cao
School of Chemistry and Chemical Engineering,
Hunan University of Science and Technology,
Xiangtan, Hunan 411201, China
e-mail: wanqiangliu2008@sina.com.cn

ymers using five parameters and generated a QSPR model with a standard error of 32.9 K for T_gs. Cao and Lin [8] tested the same set of 88 polymers using five descriptors in an attempt to derive a more physically meaningful QSPR with a coefficient of determination of R^2=0.9056 and a standard error of 20.86 K. Mattioni and Jurs [9] developed a ten-descriptor model and an 11-descriptor model, which were used to predict T_g values for two diverse sets of polymers, respectively. The test sets *RMSEs* of the two models are above 21 K. Chen et al. [10] introduced a comprehensive neural network model with 28 descriptors. The network trained with the 65 polymers was tested with six polymers and had a training RMSEs of 17 K (R^2=0.95) and prediction average error of 17 K (R^2=0.85). The model is accurate, but there are too many descriptors included. Generally, the number of descriptors n and the number of samples m should satisfy following criterion: $n \geq 5m$.

The quantum chemical descriptors encode information about the electronic structure of a molecule and thus implicitly account for the cooperative effects between functional groups, charge redistribution, and possible hydrogen bonding in the polymer. Furthermore, the quantum chemical descriptors have a clearly physical meaning. The goal of this article is to produce a robust QSPR model that could predict T_g values for 113 polymers by four quantum chemical descriptors, which are calculated from the monomers of polymers.

Materials and methods

Data set

The experimental T_g values for 113 polyacrylates and polystyrenes are listed in Table 1 [6, 11]. The entire set contains a wide range of T_g values (198 K~389 K) and is characterized by a high degree of structural variety. The functional groups presented in the side chains include halides, acetates, ethers, hydrocarbon chains, aromatic, nonaromatic rings, and so on. The experimental T_g value in Table 1 are divided into a training set and a prediction set. The training set includes 58 polyacrylates; while the prediction set includes 23 polyacrylates and 32 polystyrenes.

Quantum chemical descriptors

It is impossible directly to calculate descriptors for the entire molecule because all polymers have wide molecular weight distributions and possess high molecular weights. All the properties depend on the chemical structure of the polymer molecule, and all this structure is conditioned by the monomer structure. The properties of polymers are correlated with their monomer structure [9, 13–16]. To calculate the descriptors, the polymers were represented by their corresponding monomers. For example, the structure used to calculate descriptors for poly(benzyl acrylate) is the benzyl acrylate molecule. The calculated model is shown in Fig. 1.

Density functional theory (DFT) has become increasingly popular in the calculation of quantum chemical descriptors used for the QSPR studies [12–16]. There are two important reasons for this phenomenon. One is its lower computational cost, formally scaling as N^3 (with Coulomb fitting), where N is the number of basis functions. The other is the fact that DFT includes the effects of electron correlation at some reasonable level. The combination of lower computational cost with reasonable accuracy compared to other approaches has led to the successful application of the DFT method to the prediction of a broad range of properties of molecules in the ground state. Thus, we chose DFT method to calculate the quantum chemical descriptors with the Gaussian 03 [17] program, at the B3LYP level of theory with 6-31G(d,p) basis set.

All of the geometries of the monomers were fully optimized without applying symmetry or structural constraints. This was accomplished by using the default Gaussian convergence criteria. All of the optimized structures were characterized as true local energy minima on the potential energy surfaces, without imaginary frequencies. Vibrational frequencies and thermodynamic properties of the monomers were calculated applying the ideal gas, rigid rotor, and harmonic oscillator approximations [18–20].

Thermodynamic parameters, such as the total energy E_T, the total thermal energy $E_{thermal}$, the heat capacity at constant volume C_v, and the total entropy S can express the size of the molecular. The energy of the lowest unoccupied molecular orbital (E_{LUMO}) and the energy of the highest occupied molecular orbital (E_{HOMO}) are very popular quantum chemical descriptors. The molecular average polarizability α and the molecular dipole moment μ are used widely in the QSPR studies. The most positive net atomic charge on hydrogen atoms in a molecule (q^+) and the net charge of the most negative atom (q^-) were also calculated. In addition, the molecular weight of a monomer (M_{MON}) should also be considered [21, 22]. Therefore, we calculated the 11 quantum chemical descriptors, E_T, $E_{thermal}$, C_v, S, α, μ, E_{LUMO}, E_{HOMO}, q^+, q^-, and M_{MON}, to correlate with T_gs of polymers.

The parameters $E_{thermal}$ ($=E^t+E^r+E^v+E^e$), C_v ($=C_v^t+C_v^r+C_v^v+C_v^e$) and S ($=S^t+S^r+S^v+S^e$) are the total internal thermal energy, the total heat capacity, and the total entropy, respectively [20], and all are the contributions from molecular translation, rotational motion, vibrational motion, electronic motion, at the condition T=298.150 K and P= 1.00000 atm. A larger of the descriptors S or $E_{thermal}$ stands for a larger atomic numbers in the calculated model. The molecular average polarizability α is defined as:

$$\alpha = \left(\alpha_{xx} + \alpha_{yy} + \alpha_{zz}\right)/3 \quad (1)$$

Table 1 Quantum chemical descriptors and T_g values for 113 polymers [a]

No. of polymers		α/au	E_{HOMO}/Hartree	$E_{thermal}$/kcal mol^{-1}	S/Cal mol^{-1} K^{-1}	T_g/K (exp)	T_g/K (calc)
The training set							
1	Poly(benzyl acrylate)	105.042	−0.24907	117.824	108.925	279	268
2	Poly(4-biphenyl acrylate)	167.259	−0.22024	152.449	123.241	383	374
3	Poly(butyl acrylate)	82.019	−0.26938	120.240	102.732	219	239
4	Poly(*sec*-butyl acrylate)	81.396	−0.26855	119.907	100.848	250	241
5	Poly(2-tertbutylphenyl acrylate)	138.779	−0.23613	172.916	122.584	345	335
6	Poly(4-tertbutylphenyl acrylate)	143.255	−0.23026	172.831	124.891	344	343
7	Poly(2-chlorophenyl acrylate)	105.860	−0.24497	93.469	105.414	326	322
8	Poly(4-chlorophenyl acrylate)	108.600	−0.24021	93.441	105.773	331	337
9	Poly(4-cyanobenyl acrylate)	123.453	−0.26401	118.033	117.383	317	335
10	Poly(2-cyanoisobutyl acrylate)	92.828	−0.28625	120.003	111.422	324	320
11	Poly(2-cyanoethyl acrylate)	71.311	−0.28828	83.085	98.117	277	293
12	Poly(2-cyanohexyl acrylate)	115.466	−0.28603	157.911	127.462	358	353
13	Poly(4-cyanophenyl acrylate)	115.141	−0.25700	98.956	108.112	363	349
14	Poly(2-cyanoisopropyl acrylate)	81.903	−0.28686	100.872	100.909	339	336
15	Poly(1,3-dimethylbutyl acrylate)	103.263	−0.26772	157.118	114.092	258	240
16	Poly(dodecyl acrylate)	172.621	−0.26895	270.224	161.452	270	276
17	Poly(2-ethoxyl-carbonyl-phenyl acrylate)	134.102	−0.25498	147.054	129.216	303	286
18	Poly(3-ethoxyl-carbonyl-phenyl acrylate)	138.312	−0.24790	147.136	129.732	297	306
19	Poly(3-ethoxypropyl acrylate)	97.854	−0.25572	142.485	117.604	218	213
20	Poly(ethyl acrylate)	59.445	−0.27021	82.710	87.519	249	250
21	Poly(fluromethyl acrylate)	48.156	−0.28569	59.951	86.983	288	276
22	Poly(1H,1H-heptafluorobutyl acrylate)	80.571	−0.28980	88.568	132.397	243	240
23	Poly(2,2,3,3,5,5,5-heptafluoro-4-oxapentyl acrylate)	84.415	−0.28998	91.937	137.245	218	236
24	Poly(heptafluoro-2-propyl acrylate)	69.691	−0.30094	69.417	120.562	283	278
25	Poly(2-heptyl acrylate)	115.371	−0.26786	176.112	123.610	235	237
26	Poly(hexadecyl acrylate)	216.907	−0.26987	345.101	191.160	308	311
27	Poly(hexyl acrylate)	104.625	−0.26913	157.731	117.443	216	236
28	Poly(isobutyl acrylate)	81.468	−0.27020	119.910	101.768	249	242
29	Poly(6-cyano-4-thiahexyl acrylate)	124.673	−0.23985	141.211	133.654	215	221
30	Poly(3-methoxybutyl acrylate)	96.879	−0.25474	142.256	116.236	217	213
31	Poly(3-methoxycarbonylphenyl acrylate)	126.873	−0.24862	128.504	122.656	311	308
32	Poly(4-methoxycarbonylphenyl acrylate)	129.709	−0.24862	128.520	122.928	340	322
33	Poly(4-methoxyphenyll acrylate)	114.657	−0.21329	120.781	111.930	324	312
34	Poly(3-methoxypropyl acrylate)	86.310	−0.25777	123.826	109.713	198	218
35	Poly(2-methylbutyl acrylate)	92.579	−0.26986	138.728	109.260	241	239
36	Poly(3-methylbutyl acrylate)	92.725	−0.26921	138.716	108.970	228	238
37	Poly(2-methylpentyl acrylate)	103.950	−0.26970	157.420	117.561	235	235
38	Poly(2-naphthyl acrylate)	143.909	−0.21552	129.696	111.443	358	375
39	Poly(1H,1H-nonafluoro-4-oxahexyl acrylate)	95.506	−0.29174	101.507	151.564	224	228
40	Poly(1H,1H-nonafluoropentyl acrylate)	91.339	−0.29042	98.068	144.930	236	231
41	Poly(1H,1H,5H-octafluoropentyl acrylate)	91.589	−0.29030	103.037	140.941	238	235
42	Poly(1H,1H-pentafluropropyl acrylate)	70.629	−0.29156	78.397	110.190	247	272
43	Poly(3-pentyl acrylate)	91.837	−0.26821	138.740	108.111	267	236
44	Polyphenylethyl acrylate)	116.393	−0.24572	136.641	114.711	270	287
45	Poly(propyl acrylate)	70.708	−0.26980	101.479	95.313	236	244
46	Poly(tetradecyl acrylate)	195.337	−0.26893	307.721	176.002	297	293
47	Poly(4,4,5,5-tetrafluoro-3-oxapentyl acrylate)	85.830	−0.28477	106.279	127.895	251	237
48	Poly(3-thiabutyl acrylate)	88.221	−0.22558	103.134	107.470	213	209
49	Poly(5-thiahexyl acrylate)	110.871	−0.21928	140.699	122.195	203	208
50	Poly(3-thiapentyl acrylate)	100.519	−0.22342	122.006	115.396	202	209

Table 1 (continued)

No. of polymers		α/au	E_{HOMO}/Hartree	$E_{thermal}$/kcal mol^{-1}	S/Cal mol^{-1} K^{-1}	T_g/K (exp)	T_g/K (calc)
51	Poly(*m*-totyl acrylate)	110.144	−0.23498	117.089	108.110	298	311
52	Poly(*o*-totyl acrylate)	109.332	−0.23431	117.153	106.245	325	318
53	Poly(2,2,2trifluoroethyl acrylate)	70.325	−0.28188	88.342	108.485	263	250
54	Poly(5,5,5-trifluoro-3-oxapentyl acrylate)	86.228	−0.28028	110.515	123.554	235	233
55	Poly(1H,1H-tridecafluoro-4-oxaoctyl acrylate)	119.584	−0.29223	120.364	184.283	205	212
56	Poly(1H,1H-undecafluorohexyl acrylate)	104.518	−0.29211	107.429	159.717	234	224
57	Poly(5-cyano-3-thiapentyl acrylate)	111.987	−0.23890	122.367	125.773	223	217
58	Poly(4-butoxycarbonylphenyl acrylate)	164.925	−0.24715	184.658	146.279	286	289
	The prediction set						
59	Poly(2,4-dichlorophenyl acrylate)	119.488	−0.24737	88.144	112.731	333	352
60	Poly(4-cyanobutyl acrylate)	93.987	−0.27900	120.706	113.004	233	263
61	Poly(5-cyano-3-oxapentyl acrylate)	98.335	−0.28148	124.151	121.154	250	253
62	Poly(3-dimethylaminophenyl acrylate)	128.426	−0.19992	147.182	117.786	320	341
63	Poly(4-ethoxyl-carbonyl-phenyl acrylate)	141.467	−0.24785	147.156	129.816	310	321
64	Poly(2-ethoxyethyl acrylate)	86.517	−0.25952	123.705	110.322	223	218
65	Poly(2-ethylbutyl acrylate)	102.837	−0.26968	157.649	115.518	223	239
66	Poly(5,5,6,6,7,7,7-heptafluoro-3-oxaheptyl acrylate)	107.060	−0.27957	129.652	153.698	228	212
67	Poly(heptyl acrylate)	115.931	−0.26904	176.478	125.055	213	237
68	Poly(1H,1H,3H-hexaflurobutyl acrylate)	79.567	−0.28963	93.358	124.149	251	251
69	Poly(isopropyl acrylate)	70.355	−0.26905	101.103	93.350	270	246
70	Poly(2-methoxycarbonylphenyl acrylate)	123.122	−0.25377	128.438	121.514	319	292
71	Poly(2-methoxyethyl acrylate)	74.981	−0.26178	105.049	102.458	223	225
72	Poly(methyl acrylate)	47.822	−0.27302	64.106	79.862	283	252
73	Poly(nonyl acrylate)	138.586	−0.26899	213.972	139.628	215	248
74	Poly(*n*-pentyl acrylate)	93.319	−0.26917	138.977	110.297	216	236
75	Poly(phenyl acrylate)	95.377	−0.23990	98.731	98.585	330	281
76	Poly(7,7,8,8-tetrafluoro-3,6-dioxaoctyl acrylate)	112.556	−0.27684	147.337	150.245	233	211
77	Poly(4-thiahexyl acrylate)	111.787	−0.21962	140.789	122.912	197	209
78	Poly(4-thiapentyl acrylate)	99.486	−0.22164	121.910	115.321	208	207
79	Poly(*p*-totyl acrylate)	111.916	−0.22982	117.097	109.081	316	318
80	Poly(8-cyano-7-thiaoctyl acrylate)	146.820	−0.24237	178.685	146.668	214	228
81	Poly(2-cyanoheptyl acrylate)	126.692	−0.28594	176.635	134.705	389	368
82	Poly(4-benzoylstyrene)	167.495	−0.23162	149.183	116.701	371	375
83	Poly(2-butoxycarbonylstyrene)	146.767	−0.22952	173.984	127.503	339	343
84	Poly(4-butoxycarbonylstyrene)	152.810	−0.23153	173.941	129.775	349	349
85	Poly(5-tert-butyl-2-methystyrene)	139.467	−0.21569	180.527	114.543	360	364
86	Poly(4-sec-butylstyrene)	131.441	−0.21512	162.550	111.917	359	361
87	Poly(4-butyrylstyrene)	136.877	−0.23317	151.187	116.568	347	361
88	Poly(4-diethylcarbamoylstyrene)	154.879	−0.22362	181.707	125.817	375	362
89	Poly(4-ethoxycarbonylstyrene)	129.386	−0.23184	136.461	113.511	367	360
90	Poly(2-ethoxymethylstyrene)	117.898	−0.22439	147.678	107.352	347	338
91	Poly(4-ethoxymethylstyrene)	114.655	−0.20129	128.734	103.134	359	353
92	Poly(4-hexanoylstyrene)	160.337	−0.23303	188.674	131.435	339	351
93	Poly(2-hexyloxycarbonylstyrene)	162.703	−0.23653	211.184	142.475	318	288
94	Poly(4-hexyloxycarbonylstyrene)	175.436	−0.23145	210.769	139.390	339	344
95	Poly[4-(2-hydroxybutoxymethyl)styrene]	149.245	−0.21723	188.484	133.091	319	321
96	Poly(4-[(1-hydroxyimino)-2phenethyl]styrene)	186.354	−0.21733	178.752	133.980	384	371
97	Poly(4-isobutoxycarbonylstyrene)	149.367	−0.23506	173.457	128.704	363	343
98	Poly(2-isopentyloxycarbonylstyrene)	150.620	−0.23669	192.289	131.436	341	320
99	Poly(2-isopentyloxymethylstyrene)	152.544	−0.22367	203.669	128.782	351	339
100	Poly(4-isopropoxycarbonylstyrene)	140.645	−0.23152	154.847	119.668	368	360

Table 1 (continued)

No. of polymers		α/au	E_{HOMO}/Hartree	$E_{thermal}$/kcal mol^{-1}	S/Cal mol^{-1} K^{-1}	T_g/K (exp)	T_g/K (calc)
101	Poly(2-isopropoxymethylstyrene)	128.632	−0.22340	166.050	113.885	361	344
102	Poly(4-methoxymethylstyrene)	112.976	−0.21589	128.728	105.414	350	335
103	Poly(4-octanoylstyrene)	181.743	−0.23296	226.163	146.375	323	320
104	Poly(4-phenylacetylstyrene)	176.319	−0.23442	167.585	127.103	351	372
105	Poly(4-propoxycarbonylstyrene)	137.523	−0.23482	155.003	123.233	365	340
106	Poly(4-propoxystyrene)	126.401	−0.20094	147.459	110.933	343	358
107	Poly(4-*p*-toluoylostyrene)	182.921	−0.22986	167.558	129.788	372	372
108	Poly(4-valerylstyrene)	148.448	−0.23311	169.937	124.098	343	357
109	Poly(4-*p*-anisoylstyrene)	188.513	−0.22706	171.276	129.483	376	373
110	Poly[4-(1-hydroxy-1-methylhexyl)styrene]	169.517	−0.21363	221.931	137.887	364	343
111	Poly(4-methoxy-2-methylstyrene)	113.771	−0.20135	128.720	102.278	358	352
112	Poly(2-methoxystyrene)	99.001	−0.20710	110.176	94.187	348	314
113	Poly(4-phenoxystyrene)	155.237	−0.20724	144.733	114.455	373	375

[a] The atomic unit of polarizability: 1au=1.648777×10^{-41} C^2 m^2 J^{-1}; the unit of energy: 1 Hartree=4.3597482×10^{-18} J; 1 Cal=4.184 J.

where α_{xx}, α_{yy}, and α_{zz} reflect electronic perturbation in the *x*-, *y*-, and *z*-coordinates, respectively.

Variable selection

Stepwise multiple linear regression has proved to be an extremely useful computational technique in seeking an optimum linear combination of variables from the subsets of the *N* variables. The technique only adds one parameter to a model at a time and always in the order from most significant to least significant. Some important statistical parameters were used to valuate the variables. The *t* test measures the statistical significance of the regression coefficients. The larger *t* test absolute values indicate the relatively more significant regression coefficients. The variance inflation factors (VIF) computed as VIF=$(1-R^2)^{-1}$ (where R^2 is the coefficient of determination) can be used to identify whether excessively high multicollinearity coefficients exist among the descriptors; VIF<10 indicates tolerable colinearity among the descriptors, i.e., multicollinearity coefficients for descriptors do not exceed 0.90. In general, a descriptor is a significant descriptor and can be acceptable when the *Sig*. value of the descriptor is less than or equal to 0.05 (default level of significance). The detailed theory of multiple linear regression can be found in the mathematical textbook [23].

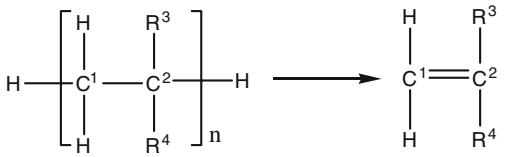

Fig. 1 The calculated models of polymers

Artificial neural network

Artificial neural network (ANN) has some remarkable properties such as self-learning and adaptation, a resistance to noise, a high degree of fault tolerance, and is suitable for nonlinear problems with complex factors. It is powerful for exploiting information from a vast amount of experimental data through learning, and especially useful for quantitative prediction [21, 22, 24–27].

Figure 2 shows the architecture of back-propagation neural network (BPNN) with the three-layer, i.e., input, output, and hidden layers [21, 22]. Hidden layers can contain one or several layers for its practical application. Each layer has different numbers of neurons (or nodes). Each neuron (or node) in the network is influenced by those

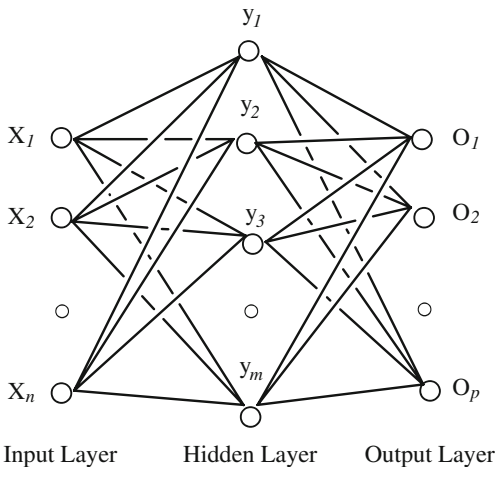

Fig. 2 Artificial neural network model (*n*, *m*, and *p* are the number of input nodes, hidden nodes, and output nodes, respectively) [21, 22]

neurons to which it is connected. For nodes in the input layer, the output (X_O) is equal to its input (X_i).

$$X_{O_i}^L = X_i^L \qquad (i = 1, 2, \ldots, n; L = 1, 2, \ldots, s) \tag{2}$$

where i is ith input layer node, n is the number of input nodes, s is the number of trained samples. For the Lth trained sample, the net input of the jth hidden layer node y_j^L is expressed as:

$$y_j^L = \sum_i^n W_{ij}^L X_{Oi}^L \qquad (j = 1, 2, \ldots, m;\ L = 1, 2, \ldots, s) \tag{3}$$

Where m is the number of hidden nodes, W_{ij}^L is a weight from unit i to unit j in Lth trained samples. The relationship between the input y_k^L and output Y_{Ok}^L of a neural element in hidden layer can be described as follows:

$$Y_{Ok}^L = f\left(Y_k^L\right) \qquad (k = 1, 2, \ldots, m; L = 1, 2, , \ldots, s) \tag{4}$$

where f(x) is basic sigmoid function, which possesses the distinctive properties of nonlinearity, continuity, and differentiability in (−∞, +∞). The function is expressed as:

$$f(x) = \frac{1}{1 + e^{-x/Q}} \tag{5}$$

where Q is the sigmoid parameter ($0.9 \le Q \le 1.0$). Similarly, the output of the last layer node O_1^L can be obtained with following equation.

$$O_l^L = f\left(\sum_j^m W_{jl}^L y_{Oj}^L\right) \qquad (l = 1, 2, \ldots, p; L = 1, 2, \ldots, s) \tag{6}$$

The learning procedure is based on a gradient search, with a least sum squared optimality criterion of errors (E) between the target output (predicted) values (O_1^L) for the Lth trained sample and the actual output (measured) values (Y^L):

$$E = \frac{1}{2}\sum_{L=1}^{s}\left(O_l^L - Y^L\right)^2 \quad (L = 1, 2, \ldots, s) \tag{7}$$

The learning algorithm for the weights between the lth output node and kth hidden node at the tth learning step is

$$W_{kl}^L(t+1) = W_{kl}^L(t) + \eta\delta_l^L Y_{Ok}^L + \alpha_0 W_{kl}^L(t) - \alpha_0 W_{kl}^L(t-1)$$
$$(k = 1, 2, \ldots, m; l = 1, 2, \ldots, p; L = 1, 2, \ldots, s) \tag{8}$$

where η is the learning rate ($0 < \eta < 1$), t is the number of learning steps, and α_0 is the momentum parameter ($0 < \alpha_0 < 1$). The error δ_j^L is defined as:

$$\delta_l^L = \left(Y_l^L - O_l^L\right) f'(x)\left(O_l^L\right) \qquad (l = 1, 2, \ldots, p) \tag{9}$$

Where $f'(x)$ is the derivative with respect to x of the sigmoid function. Similarly, the learning algorithm between the ith input node and jth hidden node is

$$W_{ij}^L(t+1) = W_{ij}^L(t) + \eta\delta_j^L X_{Oi}^L + \alpha_0 W_{ij}^L(t) - \alpha_0 W_{ij}^L(t-1)$$
$$(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m; L = 1, 2, \ldots, s) \tag{10}$$

where δ_j^L is

$$\delta_j^L = \left(\sum_{\alpha=1}^{p} \delta_\alpha^L W_{\alpha j}^L\right) f'(x)\left(y_{ji}^L\right) \qquad (j = 1, 2, \ldots, m) \tag{11}$$

In this work, the output parameter is the glass transition temperature, and the number of output layer nodes P is equal to 1. Appropriate values of these parameters aid network learning. The training of ANNs by back-propagation involves three stages [28]: (1) the feed forward of the input training pattern, (2) the calculation and back-propagation of the associated error, and (3) the adjustment of the weights. When the network error E is less than the permission error E_0 ($0.001 \le E_0 \le 0.00001$), or some limit is reached in the number of training iterations, the training process is over. Then, the trained neural network can be used to predict for the test set.

The ANN architecture is described with the code [22]: N_{in}–[N_{h1}–N_{h2}]$_e$–N_{out}, where N_{in} and N_{out} are the element numbers of input and output nodes, respectively; N_{h1} and N_{h2} are numbers of nodes in the first hidden and the second hidden, respectively; e is the number of hidden layers.

The sum of root mean square errors (E_{sum}) of the training set and the prediction set is used to valuate the accuracy of an ANN model. E_{sum} can be expressed as

$$E_{\text{sum}} = \text{RMSE}_\text{T} + \text{RMSE}_\text{P} \tag{12}$$

where RMSE_T and RMSE_P are root mean square errors (RMSEs) of the training set and the prediction set, respectively. The definition of RMSE can be found in [24, 25].The smaller the E_{sum} is, the higher is the predictive quality.

Results and discussion

By carrying out the correlation between the 11 descriptors and T_gs of 58 samples in the training set (see Table 1) with stepwise multiple linear regression (MLR) analysis [23], the best subset of descriptors is obtained. The descriptors include the molecular average polarizability α, the energy of the highest occupied molecular orbital E_{HOMO}, the total thermal energy E_{thermal}, and the total entropy S, which are listed in Table 1. The characteristics of four descriptors in MLR model are shown in Table 2.

The four descriptors are then fed to ANN as input parameters. Some good ANN models are obtained by adjusting various parameters by trial and error and listed in Table 3. Table 3 shows the model of No. 5 has the lowest E_{sum} value (E_{sum}=28 K). Thus, the architecture of the final optimum neural network is 4–[4–2]$_2$–1, with the permission error being 0.00001, the maximum number of epochs being 5,000, the momentum being 0.6, and the sigmoid parameter being 0.9. The results from the optimum ANN are listed in

Table 2 The characteristics of descriptors α, E_{HOMO}, $E_{thermal}$, and S in MLR model

Descriptors	Coefficients	SE	Sig. test	t Test	VIF
Constant	−87.274	59.321	0.147	−1.471	–
α	3.643	0.310	0.000	11.751	8.559
E_{HOMO}	−1615.569	238.356	0.000	−6.778	2.661
$E_{thermal}$	−0.988	0.139	0.000	−7.115	4.225
S	−2.689	0.292	0.000	−9.216	3.402

Table 1 and depicted in Fig. 3, which indicate that the predicted T_g values are close to the experimental ones. Only two polymers, poly(phenyl acrylate) and poly(2-methoxystyrene), present absolute errors in prediction greater than 30 K (49 K and 34 K, respectively). Root mean square errors (RMSEs) are 11 K (R=0.973) for the training set and 17 K (R=0.955) for the prediction set. In comparison with previous models [4, 6–10], the present ANN model shows better statistical quality.

Our results indicate that the ANN model developed from the training set of polyacrylates could make prediction for polyacrylates and polystyrenes. By convention, the structures of the prediction set cover the range of the structures of the training set. However, in this paper, the unusual principle for separation into training set and test sets are carried out, and the results demonstrate the ANN model could be extrapolated.

The four descriptors are significant descriptors from *Sig*.-test (see Table 2). Furthermore, all the VIF values in this paper are less than 10, which show the descriptors are acceptable without "mixing" or contamination from other descriptors.

According to the t test (in Table 2), the most significant descriptor appearing in MLR model is the molecular average polarizability α. Generally, the major factors affecting the T_gs of polymers are intermolecular forces and the chain stiffness (or mobility). Since the description α can reflect the polarity of a molecular, the parameter T_g increases with increasing α.

Table 3 Root mean square errors (RMSEs) of training and prediction using different ANN architectures

No.	Architectures	$RMSE_T$ (K)	$RMSE_P$ (K)	E_{sum} (K)
1	4–[8–4]$_2$–1	10	21	31
2	4–[6–4]$_2$–1	11	20	31
3	4–[4]$_2$–1	11	18	29
4	4–[4–3]$_2$–1	11	19	30
5	4–[4–2]$_2$–1	11	17	28
6	4–[3]$_2$–1	12	18	30
7	4–[3–2]$_2$–1	12	18	30

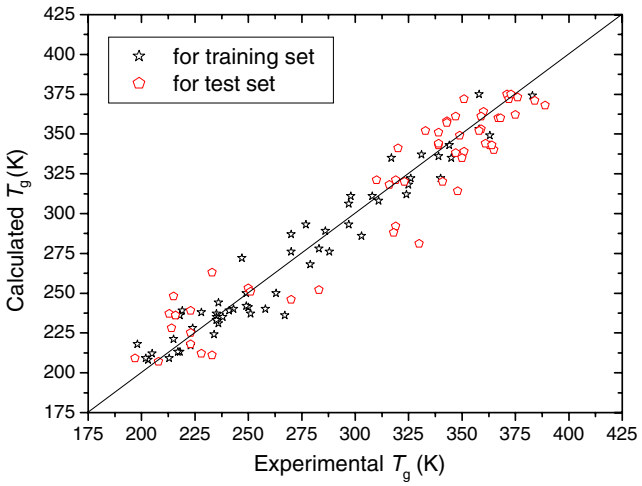

Fig. 3 Plot of experimental T_g(K) versus calculated T_g(K)

The second significant descriptor included into the model is the total entropy S. A larger S stands for an increase spatial conformations for the calculation model and may lead to a larger free volume, which results in a smaller T_g value. Therefore, T_g decreases with increasing S.

The next significant descriptor is the total thermal energy $E_{thermal}$. The negative term associated with the descriptor in MLR model suggests that an increase in the number of atoms in molecular branching will lead to a corresponding increase in free volume. The descriptor $E_{thermal}$ is negatively correlated with T_g, too.

The last one is the energy of the highest occupied molecular orbital E_{HOMO}, which carries a negative weight in the MLR model. The HOMO and the LUMO play a major role in governing many chemical reactions and determining electronic band gaps in solids [29]. The energy of the HOMO is directly related to the ionization potential and characterizes the susceptibility of the molecule toward attack by electrophiles. It has been shown that the small difference between HOMO and LUMO energies usually means that the molecular is easily polarized, which suggests that E_{HOMO} relates to the molecular polarity [30, 31]. Thus, the descriptor E_{HOMO} is correlated with T_g.

Despite many different factors affecting the T_g values of polymers, intermolecular forces and the chain stiffness (or mobility) are the main ones. The molecular average polarizability α and the energy of the highest occupied molecular orbital E_{HOMO} can describe molecular polarity and intermolecular forces; while the total thermal energy $E_{thermal}$, and the total entropy S can reflect molecular free volume and the chain stiffness (or mobility). Therefore, the four descriptors can represent the essential factors governing the nature of glass transition in polymers.

Conclusions

A QSPR model was developed to predict the T_g values for polymers. Stepwise MLR analysis and back-propagation ANN were used to generate the model after descriptors generation. Four quantum chemical descriptors, α, E_{HOMO}, $E_{thermal}$, and S, obtained directly from the monomer structures by density function theory (DFT) calculation were selected to produce the mode. Simulated with the final optimum ANN model, the results show that the predicted T_g values are in good agreement with the experimental ones, with the root mean square errors (RMSEs) being 11 K (R=0.973) for the training set and 17 K (R=0.955) for the prediction set. The results encourage the further application of the ANN model to other classes of polymer.

Acknowledgments This study is financially supported by the Natural Science Foundation of China (nos. 20772028, 20772027).

References

1. Bicerano J (2003) Encyclopedia of polymer science and technology. Wiley, New York
2. Tracht U, Wilhelm M, Heuer A, Feng H, Schmidt-Rohr K, Spiess HW (1998) Phys Rev Lett 81:2727
3. Anderson PW (1995) Science 267:1615
4. Katritzky AR, Sild S, Lobanov V, Karlson M (1998) J Chem Inf Comput Sci 38:300
5. van Krevelen DW (1976) Properties of polymers, their estimation and correlation with chemical Structure, 2nd edn. Elsevier, Amsterdam
6. Bicerano J (1996) Prediction of polymers properties, 2nd edn. Marcel Dekker, New York
7. Katrizky AR, Rachwal P, Law KW, Karelson M, Lobanov VS (1996) J Chem Inf Comput Sci 36:879
8. Cao CZ, Lin YB (2003) J Chem Inf Comput Sci 43:643
9. Mattioni BE, Jurs PC (2002) J Chem Inf Comput Sci 42:232
10. Chen X, Sztandera L, Cartwright HM (2008) Int J Intell Syst 23:22
11. Brandrup J, Immergut EH, Grulke EA (1999) Polymer Handbook, 4th edn. Wiley, New York
12. Parr RG, Yang W (1989) Density-Functional Theory of Atoms and Molecules. Oxford University Press, Oxford
13. Yu XL, Xie ZM, Yi B, Wang XY, Liu F (2007) Euro Polym J 43:818
14. Yu XL, Yi B, Yu WH, Wang XY (2008) Chem Pap 62:623
15. Yu XL, Yi B, Xie ZM, Wang XY, Liu F (2007) Chemometr Intell Lab Syst 87:247
16. Yu XL, Yi B, Wang XY (2007) J Comput Chem 28:2336
17. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, MonTgomery JA, Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (2003) Gaussian 03, Revision B.05. Gaussian Inc, Pittsburgh PA
18. Davidson M (1962) Statistical mechanics. McGraw-Hill, New York
19. McQuarrie A (1973) Statistical Thermodynamics. Harper and Row, New York, 1973
20. Ochterski JW (2000) Thermochemistry in Gaussian. Gaussian Inc, Pittsburgh PA
21. Zhang LX, Zhao DL, Huang YX (2002) Chin J Polym Sci 20:25
22. Zhang Z, Friedrich K (2003) Compos Sci Technol 63:2029
23. Montgomery DC, Peck EA, Vining GG (2001) Introduction to linear regression analysis, 3rd edn. Wiley, New York
24. Yu XL, Yi B, Liu F, Wang XY (2008) React Funct Polym68:1557
25. Yu XL, Yi B, Wang XY (2008) Euro Polym J 44:3997
26. Liu WQ, Yi PG, Tang ZL (2006) QSAR Comb Sci 25:936
27. Zhang LX, Li J, Jiang ZT, Xia A (2003) Polymer 44:1751
28. Fausett LV (1994) Fundamentals of neural networks, 1st edn. Prentice Hall, Englewood Cliffs, New Jersey
29. Zhou Z, Parr RG (1990) J Am Chem Soc 112:5720
30. Katritzky AR, Sild S, Karelson M (1998) J Chem Inf Comput Sci 38:1171
31. Hamerton I, Howlin BJ, Larwood V (1995) J Mol Graphics 13:14