J Mater Sci (2009) 44:3156–3164
DOI 10.1007/s10853-009-3420-0

# Artificial neural network prediction of glass transition temperature of fluorine-containing polybenzoxazoles

**Liwei Ning**

Received: 19 December 2008 / Accepted: 17 March 2009 / Published online: 2 April 2009


**Abstract** Fluorine-containing polymers belong to high-performance polymers with unique chemical and physical properties that are not observed with other organic polymers. In this article, three structural parameters were used to correlate with glass transition temperature $T_g$ values for 52 fluorine-containing polybenzoxazoles. The descriptors obtained directly from the structures of repeating units can reflect the chain stiffness (or mobility). Back propagation artificial neural network (ANN) and multiple linear regression (MLR) analysis were used in the study. The final optimum neural network with [3-1-1] structure produced a training set root mean square (*rms*) error of 2.35 K ($R = 0.980$) and a test set *rms* error of 2.30 K ($R = 0.978$). The statistical results indicate that the ANN model given here has better predictive capability than other existing models.

## Introduction

The glass transition temperature ($T_g$) of amorphous polymer is the most important property, which affects many other polymer properties, such as heat capacity, coefficient of thermal expansion, and viscosity. $T_g$ is difficult to determine experimentally because the transition takes place over a comparatively wide temperature range and it is dependent on conditions such as the method of measurement, duration of the experiment, and pressure. For this reason, the discrepancies in reported $T_g$ values in the literature can be large [1, 2].

L. Ning (✉)
College of Mechanical Engineering, Hunan Institute of Engineering, Xiangtan, Hunan 411101, China
e-mail: ningliwei2002@sina.com.cn

Numerous researchers have attempted to predict $T_g$ values for polymers on the basis of quantitative structure–property relationship (QSPR) models. Many group contribution and empirical/semiempirical methods have been used to build models for prediction of $T_g$ values [3]. The group additive properties method is purely empirical approach and limited to systems composed only of functional groups that have been previously investigated.

The most widely referenced model of the theoretical estimations has been produced by Bicerano [4]. Bicerano produced a regression model ($R = 0.9749$, $s = 24.65$ K) to relate $T_g$ values with the solubility parameter and the weighed 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. [5] introduced a mode with $R^2$ of 0.928 for 22 medium molecular weight polymers using four parameters. After that work, Katritzky et al. [2] applied the CODESSA method to predict $T_g$ values for 88 linear homopolymers using five parameters and generated a QSPR model with a standard error of 32.9 K for $T_g$ values. Yu et al. [6] developed a linear model with only two descriptors. The model was tested to be aureate, with correlation coefficient of 0.953 for the training set and 0.952 for the test set. Mattioni and Jurs [7] developed a 10-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 *rms* errors of the two models were above 21 K. Chen et al. [8] introduced a comprehensive neural network model with 28 descriptors. The network trained with the 65 polymers was tested with 6 polymers and had a training *rms* error of 17 K ($R^2 = 0.95$) and prediction average error of 17 K ($R^2 = 0.85$). The model were accurate, but there were two many descriptors included.

Fluorine-containing polymers have been synthesized and developed as high-performance polymers with unique

chemical and physical properties that are not observed with other organic polymers [9]. For example, they possess improved thermal stability, high chemical stability, low adhesion, low refractive index, and so on. An additional benefit of the presence of the C–F bond is the decreased dielectric constant, which fluorinated materials exhibit [10]. The goal of this article is to produce a robust QSPR model that could predict $T_g$ values for 52 fluorine-containing polybenzoxazoles.

## Materials and methods

Table 1 shows 52 fluorine-containing polybenzoxazoles and their respective experimental $T_g$ values, which are taken from literatures [11]. It is impossible to calculate descriptors directly for an entire molecule because all the polymers have wide distribution of molecular weights and possess high molecular weights. Since the polymer chain is long enough, the terminal groups hold only a very small proportion in a polymer and its effect on the parameters $T_g$ is ignored. Thus, the repeating unit can be used as representative of the polymeric material [2, 6–8]. This method is adopted in this article. The calculated models of fluorine-containing polybenzoxazoles are shown in Fig. 1.

Generally, the stiffness (or mobility) of the main chain has effects on $T_g$ values. The flexibility of group $R_1$ can be expressed with the descriptor $n_A$, which is equal to the number of atoms in the main chain of groups $R_1$. For example, the descriptor $n_A$ for the compound of No. 25 is 14, since there are two O atoms and 12 C atoms in the main chain of $R_1$. The correlation coefficient between the $T_g$ values of fluorine-containing polybenzoxazoles and the descriptor $n_A$ is −0.696, which shows that the descriptor $n_A$ is an important variable to fit the $T_g$ values (see Fig. 2).

Table 1 shows that many polymer compounds have $–CF_2CF_2O–$ groups. The descriptor $m$ is used to express the total number of the $–CF_2CF_2O–$ groups in the repeating unit. Here taking the compound of No. 35 as the calculation example. There are two $–CF_2CF_2O–$ groups in $R_1$ and one $–CF_2CF_2O–$ group in $R_2$, respectively, thus its $m = 3$. From Fig. 3, it is shown that there is good correlation between the $T_g$ values and the descriptor $m$ ($R = -0.893$).

In addition, the descriptor $n_{CF_2}$ is defined. $n_{CF_2}$ is equal to the number of $–CF_2–$ groups in $R_2$. But the entire $–CF_2–$ in $–CF_2CF_2O–$ group is not accounted repeatedly, because these groups have been taken into account in the descriptor $m$. For example, the descriptor $n_{CF_2}$ for the compound of No. 2 is 2, there are four $–CF_2–$groups in $R_2$, but the two $–CF_2–$groups in $–CF_2CF_2O–$ is not accounted.

Figure 4a shows the correlation between the $T_g$ values and the descriptor $n_{CF_2}$ ($R = 0.280$). There is a turning point at the near of $n_{CF_2} = 3$. Thus, $n_{CF_2}$ is substituted with the descriptor $N_{CF_2}$, which is used to express the rigidness of main chain and defined with following equation:

$$N_{CF_2} = \mathrm{ABS}(n_{CF_2} - 3) \tag{1}$$

Figure 4b shows the correlation coefficient $R$ between $T_g$ values and $N_{CF_2}$ is 0.463, which is slightly higher than the former ($R = 0.280$).

In this work, a three-layered back propagation artificial neural network (ANN) [12–14] was employed to fit and predict $T_g$ values for the samples. Back propagation ANN is based on searching an error surface using gradient descent for the point with minimum error. The transfer function is the basic sigmoid which possesses the distinctive properties of nonlinearity, continuity, and differentiability in $(-\infty, +\infty)$. The sigmoid function is expressed as:

$$f(x) = \frac{1}{1 + e^{-x/Q}} \tag{2}$$

where $Q$ is the sigmoid parameter ($0.9 \le Q \le 1.0$). The error function can be expressed by:

$$E = \frac{1}{2}\sum_{k=1}^{N}(y_k - \bar{y}_k)^2 \tag{3}$$

where $N$ is the number of the samples $(x_k, y_k)$ ($k = 1, 2, 3, \ldots, N$), $\bar{y}_k$ is the output of the neuron. 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. The three descriptors stated above were fed to ANN, as input vectors. To avoid a local minimum, a simulated annealing technique was used. The single output neuron represents the output value to be compared to the experimental value. The number of neurons in the hidden layer was optimized by trial and validation until no improvement was seen for that model. The sum of root mean square (*rms*) errors of the training set and the test set was used to valuate the accuracy of a model.

## Results and discussion

The experimental data of $T_g$ in Table 1 are randomly divided into a training set and a test set. The former includes 35 polymers, the latter includes 17 polymers. The three descriptors ($n_A$, $m$, and $N_{CF_2}$) are fed to ANN as input vectors. The optimal condition of the neural network is obtained by adjusting various parameters by trial and error. The architecture of the final optimum neural network is [3-1-1], with the permission error being 0.0001, the maximum number of epochs being 5000, the momentum being 0.6, and the sigmoid parameter being 0.9. The results from ANN method are listed in Table 1, which indicate that the predicted $T_g$ values are close to the experimental ones. The root mean square (*rms*) errors are 2.35 K ($R = 0.980$) for

**Table 1** Experiment and calculated $T_g$ values for 52 fluorine-containing polybenzoxazoles

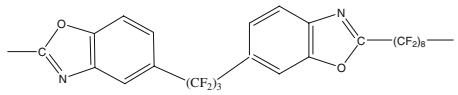
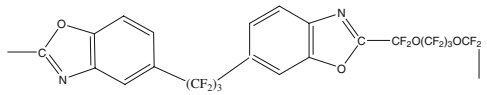
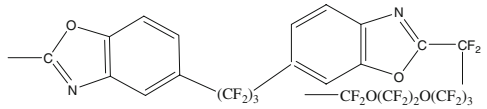
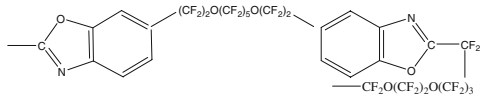
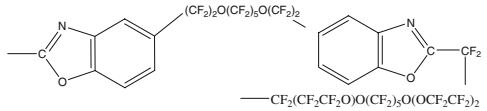
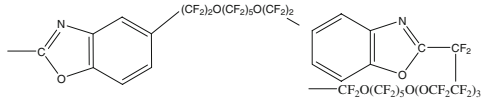
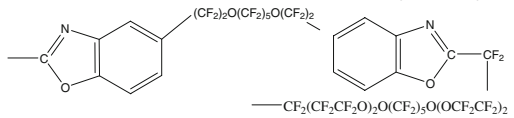
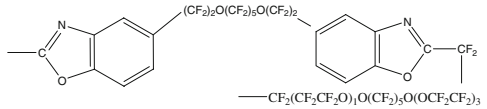
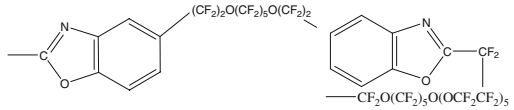
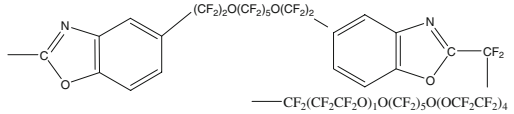
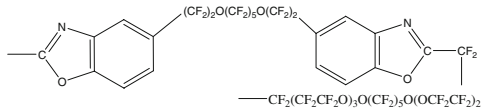
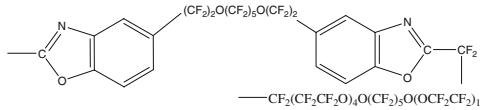
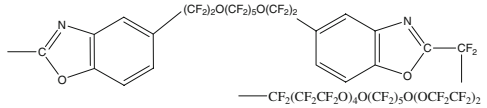
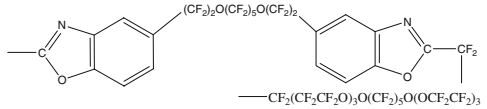
| No. | Polymer | $T_g$ K(exp) | $T_g$ K(calc) | $\Delta T_g$ K[a] |
|---|---|---|---|---|
| *Training set* | | | | |
| 1 |  | 335 | 330 | 5 |
| 2 |  | 318 | 314 | 4 |
| 3 |  | 300 | 303 | −3 |
| 4 |  | 263 | 259 | 4 |
| 5 |  | 236 | 236 | 0 |
| 6 |  | 236 | 236 | 0 |
| 7 |  | 232 | 228 | 4 |
| 8 |  | 232 | 228 | 4 |
| 9 |  | 228 | 224 | 4 |
| 10 |  | 228 | 224 | 4 |
| 11 |  | 228 | 224 | 4 |
| 12 |  | 228 | 224 | 4 |
| 13 |  | 221 | 221 | 0 |
| 14 |  | 221 | 221 | 0 |

**Table 1** continued

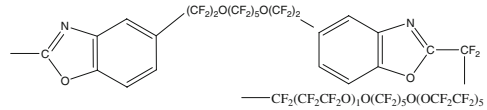
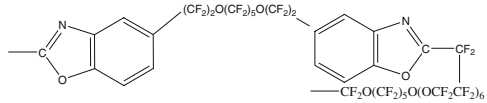
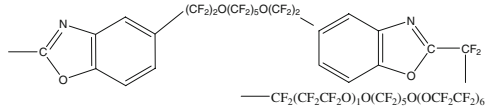
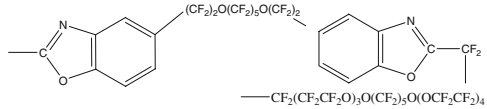
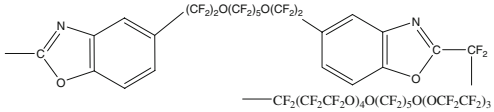
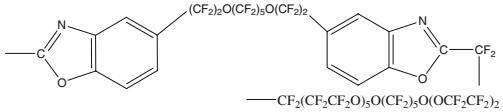
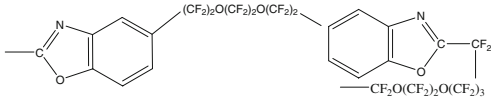
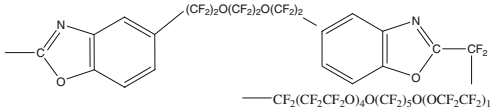
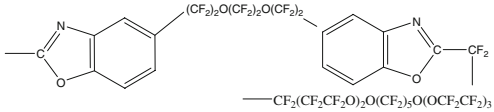
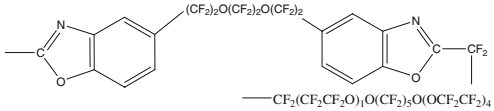
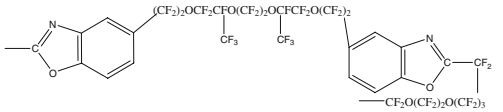
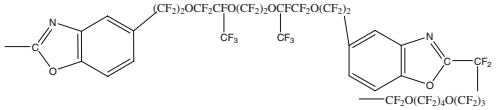
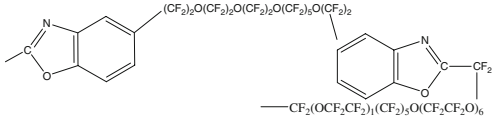
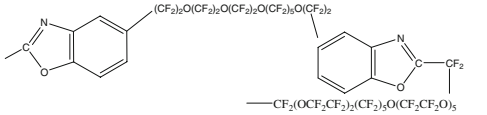
| No. | Polymer | $T_g$ K(exp) | $T_g$ K(calc) | $\Delta T_g$ K[a] |
|---|---|---|---|---|
| 15 |  | 221 | 221 | 0 |
| 16 |  | 221 | 221 | 0 |
| 17 |  | 218 | 221 | −3 |
| 18 |  | 218 | 221 | −3 |
| 19 |  | 218 | 221 | −3 |
| 20 |  | 218 | 221 | −3 |
| 21 |  | 268 | 267 | 1 |
| 22 |  | 233 | 234 | −1 |
| 23 |  | 233 | 234 | −1 |
| 24 |  | 233 | 234 | −1 |
| 25 |  | 248 | 241 | 7 |
| 26 |  | 249 | 240 | 9 |
| 27 |  | 215 | 222 | −7 |
| 28 |  | 215 | 222 | −7 |

**Table 1** continued

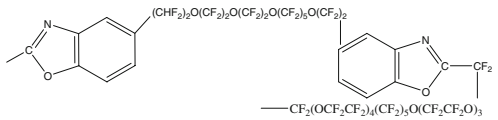
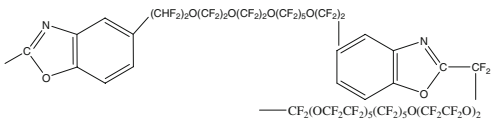
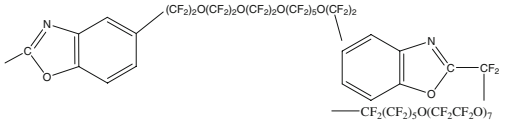
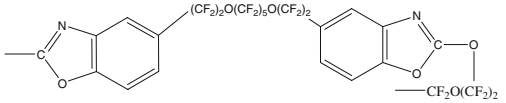
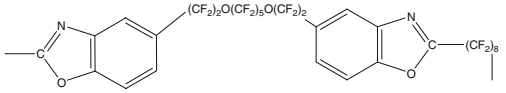
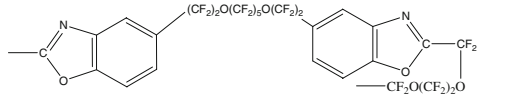
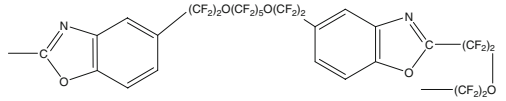
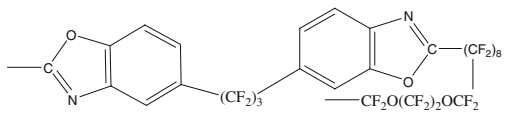
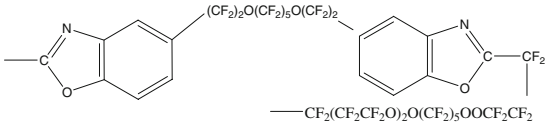
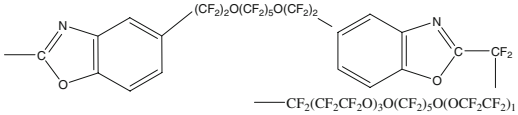
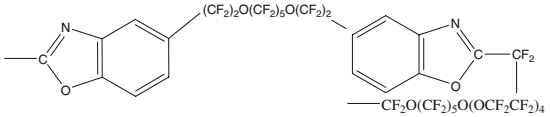
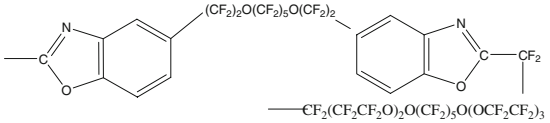
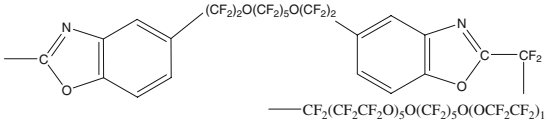
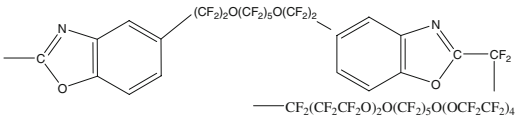
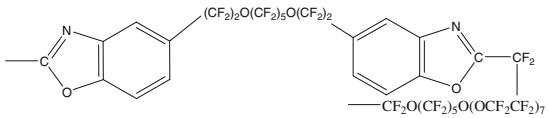
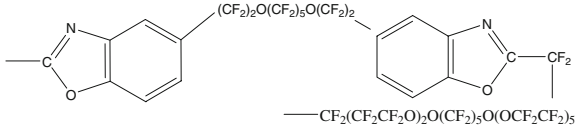
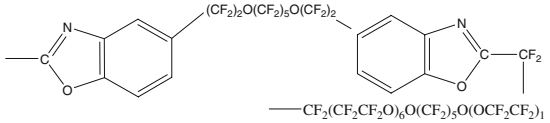
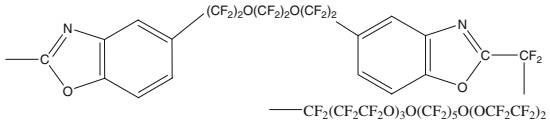
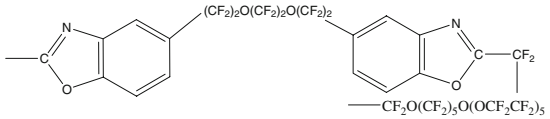
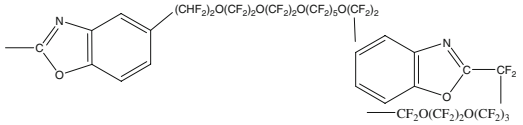
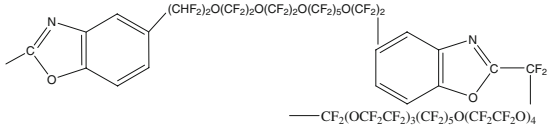
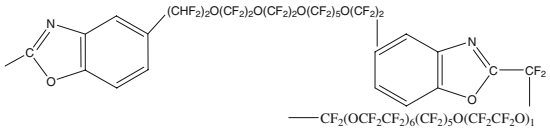
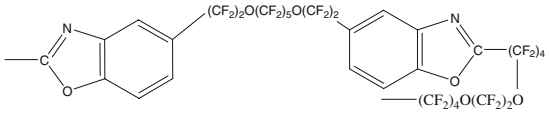
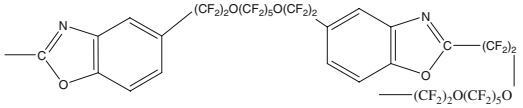
| No. | Polymer | $T_g$ K(exp) | $T_g$ K(calc) | $\Delta T_g$ K[a] |
|---|---|---|---|---|
| 29 |  | 215 | 222 | –7 |
| 30 |  | 215 | 222 | –7 |
| 31 |  | 215 | 222 | –7 |
| 32 |  | 276 | 258 | 18 |
| 33 |  | 253 | 259 | –6 |
| 34 |  | 253 | 258 | –5 |
| 35 |  | 244 | 259 | –15 |
| *Test set* | | | | |
| 36 |  | 323 | 324 | –1 |
| 37 |  | 236 | 236 | 0 |
| 38 |  | 232 | 228 | 4 |
| 39 |  | 232 | 228 | 4 |
| 40 |  | 228 | 224 | 4 |
| 41 |  | 221 | 221 | 0 |

**Table 1** continued

| No. | Polymer | $T_g$ K(exp) | $T_g$ K(calc) | $\Delta T_g$ K[a] |
|---|---|---|---|---|
| 42 |  | 221 | 221 | 0 |
| 43 |  | 218 | 221 | −3 |
| 44 |  | 218 | 221 | −3 |
| 45 |  | 218 | 221 | −3 |
| 46 |  | 233 | 234 | −1 |
| 47 |  | 233 | 234 | −1 |
| 48 |  | 244 | 245 | −1 |
| 49 |  | 215 | 222 | −7 |
| 50 |  | 215 | 222 | −7 |
| 51 |  | 267 | 259 | 8 |
| 52 |  | 241 | 258 | −17 |

[a] $\Delta T_g$ K = $T_g$ K(exp) − $T_g$ K(calc)

the training set and 2.30 K ($R = 0.978$) for the test set. The statistical results are more accurate than that from other existing models [2, 4–8]. Weights matrixes of neuron links in the networks are shown in Table 2.

The three descriptors are also used to develop a multiple linear regression (MLR) model from the training set with MLR analysis [15]. Statistical parameters corresponding to the MLR model of $T_g$ are the following:

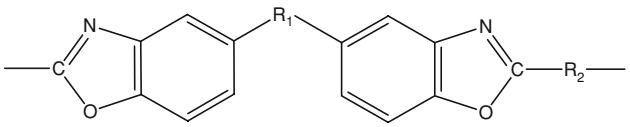



**Fig. 1** The calculated models of fluorine-containing polybenzoxazoles

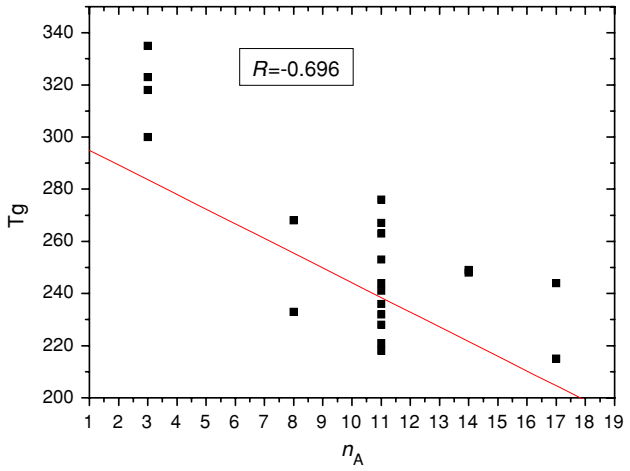



**Fig. 2** Plot of experimental $T_g$ (K) versus the parameter $n_A$

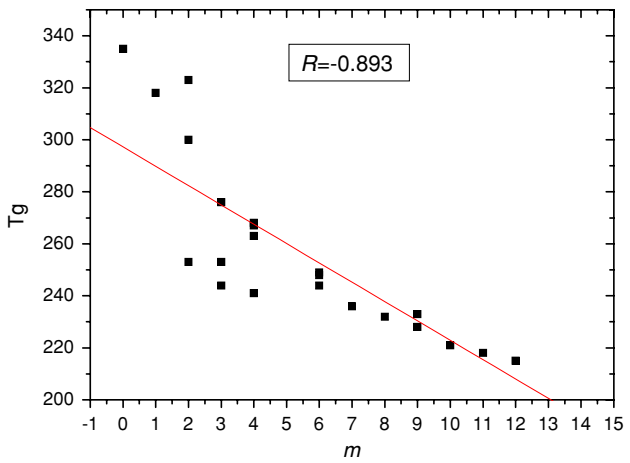



**Fig. 3** Plot of experimental $T_g$ (K) versus the parameter $m$

$$T_g = 308.567 - 3.106n_A - 4.917m + 4.110N_{CF_2}$$
$$R = 0.939,\ R^2 = 0.883,\ s = 10.519,\ F = 77.624,\ n = 35 \quad (4)$$

where $R$ is correlation coefficient, $s$ is the standard error, $F$ is the Fisher ratio, and $n$ is the number of samples in the training set. The MLR model is used to predict the $T_g$ values for test set in Table 1. The correlation coefficient between the experimental $T_g$ values and predicted values is 0.954. The results from ANN model are better than that of the MLR model, which can also be seen from the plots of experimental $T_g$ (K) versus calculated $T_g$ (K) in Fig. 5a: ANN model; b: MLR model. This indicates that the correlation between $T_g$ values and structural parameters outlined above is nonlinear rather than linear.

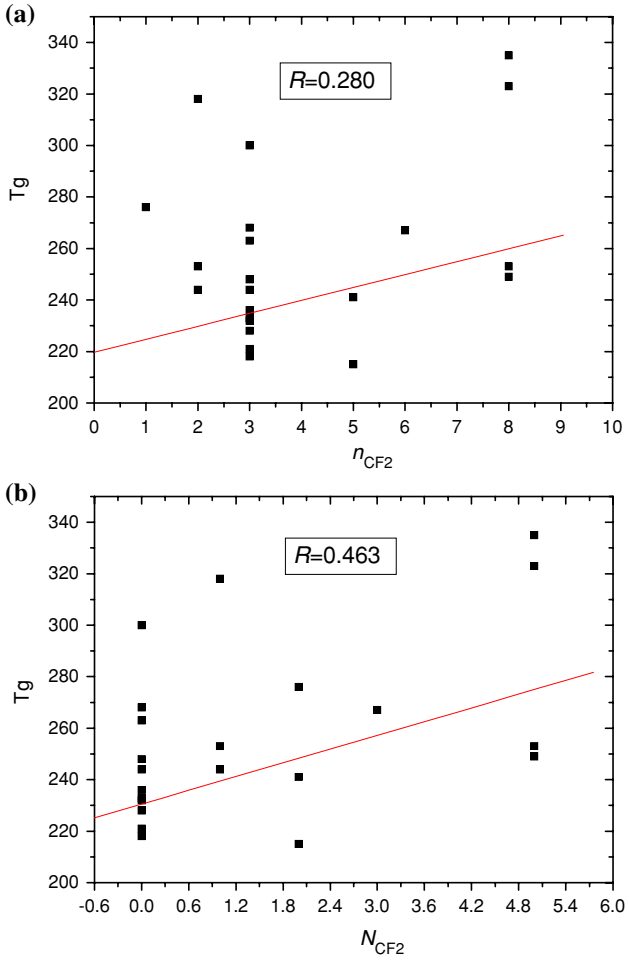



**Fig. 4** Plot of experimental $T_g$ (K) versus parameters, **a** $n_{CF_2}$; **b** $N_{CF_2}$

**Table 2** Weights matrixes of neuron links in the networks

| The first hidden layer | | | The output layer |
|---|---|---|---|
| −2.7747 | 0.8210 | −6.1752 | 11.7154 |
| 0.9375 | −0.2193 | 5.6301 | −3.3960 |
| 0.3651 | 1.9026 | −1.0099 | −5.1125 |

The characteristics of descriptors used in the MLR model are shown in Table 3. The three descriptors all are significant descriptors from *Sig*-test (see Table 3). According to the statistical theory, the variance inflation factors (VIF) values can be used to identify whether excessively high multicollinearity coefficients exist among the descriptors; $VIF < 10$, indicating tolerable collinearity among the descriptors, i.e., multicollinearity coefficients for descriptors do not exceed 0.90. All the VIF values in this article are less than 3 (see Table 3), which show the descriptors are

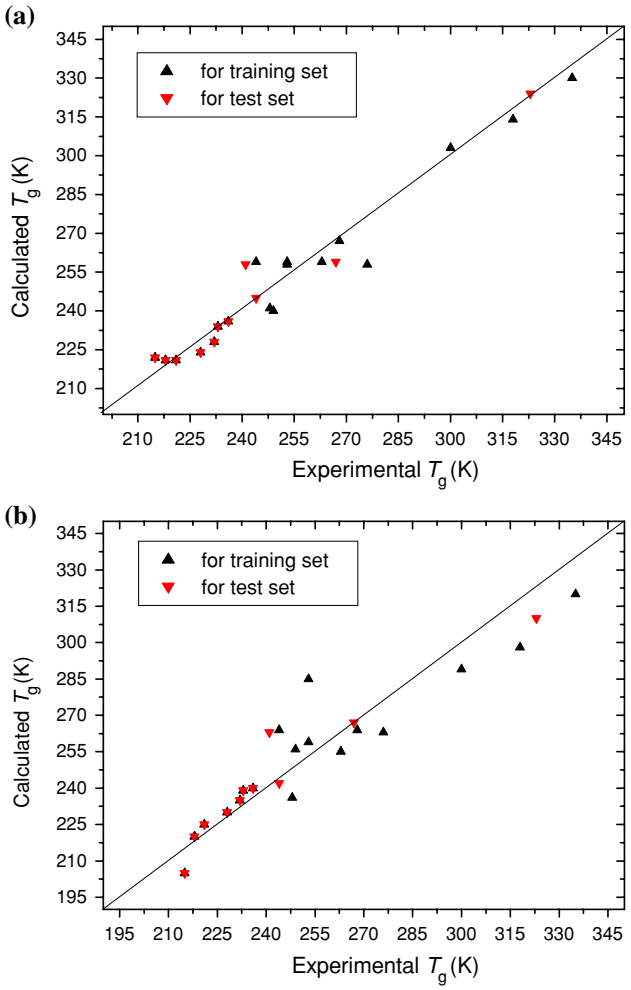



**Fig. 5** Plots of experimental $T_g$ (K) versus calculated $T_g$ (K), **a** ANN model; **b** MLR model

**Table 3** The characteristics of descriptors used in the MLR model

| Descriptor | Unstandardized coefficients | Std error | *t*-test | *Sig*-test | VIF |
|---|---|---|---|---|---|
| Constant | 308.567 | 5.918 | 52.141 | 0.000 | / |
| $n_A$ | −3.106 | .787 | −3.948 | 0.000 | 2.383 |
| $m$ | −4.917 | .826 | −5.951 | 0.000 | 2.601 |
| $N_{CF_2}$ | 4.110 | 1.479 | 2.779 | 0.009 | 1.508 |

acceptable without "mixing" or contamination from other descriptors.

According to the *t*-test (in Table 3), the most significant descriptor appearing in MLR model is the descriptor $m$, which is equal to the total number of the $–CF_2CF_2O–$ groups in the repeating unit. It has been shown that –O– groups can result in flexibility of the chain [16]. A larger descriptor $m$ expresses the more –O– groups. Thus, the $T_g$ value decreases with the increasing descriptor $m$.

The second significant descriptor included into the model is the descriptor $n_A$. As the descriptor $n_A$ increases, the benzoxazoles density decreases, and the chain stiffness also decreases. Therefore, the $T_g$ value decreases with the increasing $n_A$, too.

The last one is the descriptor $N_{CF_2}$, which carries a positive weight in the MLR model. When there is $n_{CF_2} \leq 3$, the $T_g$ value decreases with the increasing descriptor $n_{CF_2}$. It is easy to understand because the long flexible chain such as $–CF_2–$ can decrease $T_g$ values; When there is $n_{CF_2} \geq 3$, the longer flexible polymer chains may appear fold. This phenomenon not only decreases the effective length of flexible chain, but also increases the hindrance to the rotation around the main chain bonds. The $T_g$ value then increases with the increasing descriptor $n_{CF_2}$. This may be the reason why there is a turning point at the near of $n_{CF_2} = 3$ in Fig. 4a. Therefore, the descriptor $N_{CF_2}$, defined as the absolute value of $(n_{CF_2} - 3)$, can express the rigidness of the polymer chain, and the larger $N_{CF_2}$ expresses the higher $T_g$ value.

## Conclusion

An ANN model was developed to predict the $T_g$ values for fluorine-containing polybenzoxazoles. Back propagation ANN and MLR analysis were used to generate the model and analyze these descriptors. The results from ANN model are better than that of the MLR model, which indicates that the correlation between $T_g$ and structural parameters is nonlinear rather than linear. The three structural parameters were obtained easily from the structures of repeating units. The final optimum ANN model is tested to be most accurate, with the root mean square (*rms*) errors 2.35 K ($R = 0.980$) for the training set and 2.30 K ($R = 0.978$) for the test set.

## References

1. Bicerano J (2003) Encyclopedia of polymer science and technology. Wiley, New York
2. Katritzky AR, Sild S, Lobanov V, Karlson M (1998) J Chem Inf Comput Sci 38:300
3. Van Krevelen DW (1990) Properties of polymers. Elsevier Science, New York
4. Bicerano J (1996) Prediction of polymers properties, 2nd edn. Marcel Dekker, New York
5. Katrizky AR, Rachwal P, Law KW, Karelson M, Lobanov VS (1996) J Chem Inf Comput Sci 36:879
6. Yu XL, Yi B, Wang XY, Xie ZM (2007) Chem Phys 332:115
7. Mattioni BE, Jurs PC (2002) J Chem Inf Comput Sci 42:232
8. Chen X, Sztandera L, Cartwright HM (2008) Int J Intell Syst 23:22
9. Hougham G, Cassidy PE, Johns K, Davidson T (1999) Fluoropolymers I. Kluwer Academic/Plenum Publishers, New York
10. Lee JK, Kim JH, Kim YJ (2003) Bull Korean Chem Soc 24:1029

11. Приалко ВП (1995) Handbook of polymer physical chemistry, vol 2. China Petrochemical Press, Beijing (trans: Yan J, Zhang Y)
12. Yu XL, Yi B, Liu F, Wang XY (2008) React Funct Polym 68:1557
13. Yu XL, Yi B, Wang XY (2008) Eur Polym J 44:3997
14. Yu XL, Yi B, Wang XY (2008) J Theor Comput Chem 7:953
15. Jurs PC (1996) Computer software applications in chemistry, 2nd edn. Wiley, New York
16. Yu XL, Wang XY, Li XB, Gao JW, Wang HL (2006) Macromol Theory Simul 15:94