RSC Advances

RSC Advances

www.rsc.org/advances

ROYAL SOCIETY
OF CHEMISTRY

www.rsc.org/advances

Conductive polymer composite/CNT
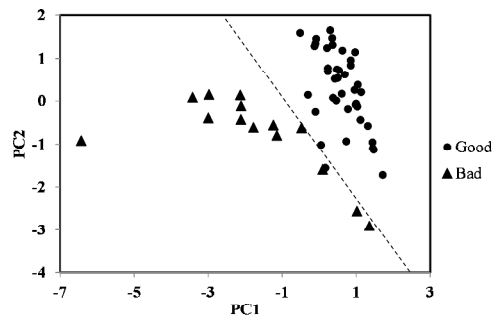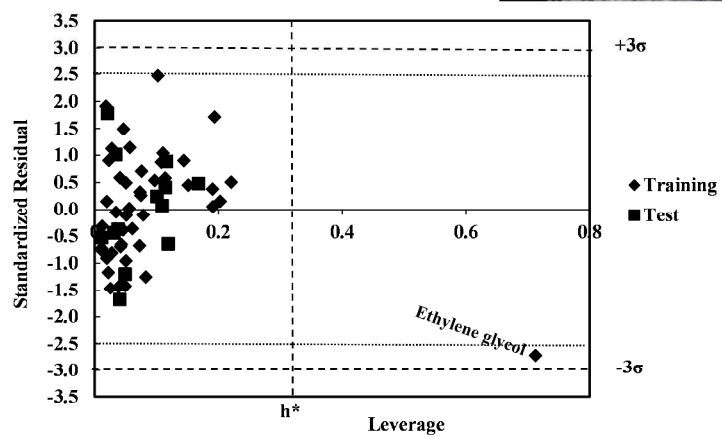
Linear solvent structure-polymer solubility and solvation energy relationships to study conductive polymer/carbon nanotube composite solutions

Saeed Yousefinejad [a, b*], **Fatemeh Honarasa** [c], **Hanieh Montaseri** [d]

[a] Department of Chemistry, Shiraz University, Shiraz, Iran
[b] Department of Chemistry, Farhangian University, Tehran, Iran
[c] Department of Chemistry, Shiraz Branch, Islamic Azad University, Shiraz, Iran
[d] Department of Chemistry, Yasuj University, Yasuj, Iran

ABSTRACT

The solvation and solvent selectivity of polymer composites in different solvents is an important subject in colloid and polymer chemistry. Two multiparameter linear models based on theoretical and empirical parameters were constructed / validated, for 59 and 54 solvents respectively, to predict relative energy difference (RED) of solvents and a conductive polymer composite containing carbon nanotube. In addition to excellent external prediction ability, models 1 (QSPR) and 2 (LSER) covered 87% and 93% of cross-validated variance respectively. Different statistical methods were applied to test and validate models. From the description view, it was showed by model 1 that the compactness of solvent structure, mass and polar interactions are important in the resistance of polymer and its RED in the desired solvent. In addition to the Hildebrand solubility parameter, acidity of solvent and hydrogen bonding interactions has direct relationship with RED. Both models confirmed the moderate and complex effect of polar interactions in the solvation of desired polymer composite in different solvents.

Keywords: Linear solvation energy relationship, Polymer composite/carbon nanotube, QSPR, Solvent empirical scale, Structural parameters.

---

* Corresponding author, Tel.: +98 917 704 2635

E-mail address: yousefinejad.s@gmail.com, yousefinejad@cfu.ac.ir (S. Yousefinejad)

1. Introduction

Electrically conductive polymer composites (CPCs) are a kind of smart materials which have come into the focus of research because of their highly applicable characteristics. Smart materials are solid-state transducers that have different sensing and actuating properties like electroactivity, piezoelectric electrostrictivity, pyroelectricity, magnetostrictivity and piezoresistivity.[1] The most promising and viable strategy for the successful development of electrically conductive polymer composites (CPCs) is the blending of a customary polymer with electrically conductive fillers.

The exceptional intrinsic properties of carbon nanotubes (CNTs), such as thermal and electrical conductivity or mechanical properties, has made them one of the most promising nanomaterials used to modify the properties of polymer and conductive polymer composites [2]. Moreover, CNTs have high length/diameter ratio (aspect ratio) of about 100-1000 which leads to moderately low percolation threshold in composite materials in comparison with carbon fibers or carbon black [3]. Conductive polymer/nanofiller composites have been widely investigated in industry and academia because of their outstanding multifunctional properties compared to conventional conductive polymer composites (CPCs). Polymer/carbon nanofiller composites can be used for the electrostatic dissipation (ESD) [4], electromagnetic interference- shielding (EMI-shielding) [1,5,6], electrostatic painting [7], and mechanical reinforcement [8,9]. Some of the important aspects related to CPCs are their relative resistance change and solvent selectivity as a key property of these sensory polymers [10].

Few reports have been published on the establishment of correlation between relative resistance change of CPCs and solvents' solubility parameters. Chen et al. reported that waterborne polyurethane composites filled with carbon black show a maximum relative resistance change that correlates with the polar component of Hansen solubility (δ_P) [11]. Fan et al. found the same result for thermoplastic polyurethane multifilament covered with carbon nanotube networks [12]. The study on the role of difference of the solubility parameters between polymer matrix and tested solvents ($\Delta\delta$) in polystyrene/carbon black composites were also done by Li et al [13]. Flory-Huggins interaction parameter, χ_{12}, was also utilized to correlate relative resistance change of CPCs/CNTs against some solvent vapors [14,15], however this approach is almost not sufficiently accurate [16]. Villmow et al. investigated the relative resistance change of a

CPC with 1.5 wt.% CNT immersed in different organic solvents and calculated the Hansen solubility parameters of the CPC [16]. They also described the selectivity of the CPC by using two parameters, namely the distance in Hansen space R_a and the solvents' molar volume V_{mol}.

In the current study, we focused on CPCs based on polycarbonate (PC) filled with 1.5 wt.% carbon nanotubes [16] in different solvents in order to scrutinize the details of the interactions between solvent and CPCs and the role of solvent characteristics in the dispersibility of CPCs and their solvent selectivities. We used quantitative structure property relationship (QSPR) and linear solvation energy relationship (LSER), one of the well-known approaches of QSPR, to predict the relative energy difference (RED) of CPC/CNT and different solvents. In the proposed approach for the first time, some solvent empirical parameters were utilized to clear solvent-polymer interaction in addition to the prediction purpose. Here, we tried to emphasize the role of solvent in solvent-sensory polymers using a predictive and descriptive method.

2. Materials and Methods

2.1. Data set preparation and processing

In this work, two kinds of models were used for the examination of the dispersibility of PC/CNT composite in different solvents. In the first model, the theoretical and structural descriptors of solvents were used to construct a QSPR model to predict the RED numbers of PC/CNT composite in 59 solvents adapted from literature [16]. RED number of 0 is found for no energy difference. RED numbers less than 1.0 indicate high affinity; RED equal to or close to 1.0 is a boundary condition; and progressively higher RED numbers indicate lower affinities [17]. The name of solvents and their RED values are presented in Table 1. In the first step, the chemical structures of these 59 solvents were drawn in Hyperchem software (Version 7, Hypercube Inc., http://www.hyper.com, USA) and optimized using semi-empirical method of AM1. Then, molecular descriptors were calculated for desired solvents from the optimized chemical structures by Dragon software (Milano Chemometrics and QSAR research group; http://michem.disat.unimib.it/chm/). Furthermore, it was crucially important to do several complementary works. In other words, numerous molecular descriptors were lessened by removing descriptors that could not be calculated for every structure in the dataset and those descriptors with an essentially constant or near constant value for all structures. Additionally, to

decrease the redundancy in the descriptors data matrix, the correlations among X-variables and with the RED vector were examined, and among the detected collinear variables (i.e., $R^2 > 0.95$) one with the highest correlation with RED was retained and the others were removed from the data matrix. After these steps 834 descriptors were remained for each solvent.

In the second model based on empirical solvent scales, 127 scales of three principal classes (Equilibrium-kinetic, Spectroscopic and Multiparameters of other measurements) [18] were used for LSER modeling the RED of polymer-solvents. Among 59 solvents used in the previous model, the experimental parameters of 54 solvents were available and were used in the model construction and validation. Here again, the correlation between solvent scales were checked to delete the redundant parameters.

2.2. Data Modeling

The main goal of QSPR is to establish a significant relation based on few structure-based molecular descriptors which can accurately predict an experimental property/activity (here RED of polymer/solvent samples). As it is feasible to produce a great number of molecular descriptors for each solvent in the dataset, a thorny problem is the selection of the set of molecular descriptors in order to build an accurate relationship. In the presented work, multiple linear regression analysis (MLR) with stepwise selection of variables (using SPSS software, SPSS Inc., version 15.0) was applied to relate the solvent structural parameters with the solubility of PC/CNT composite.

In order to better performance, all selected-independent and dependent variables, theoretical descriptors/empirical solvent scales and RED vector were auto scaled. For model evaluation, the data of both models (QSPR and LSER) were divided to two parts; training set for model building and test set for checking the model's predictability. Cross validation and y-scrambling were also done to test the stability and significance of models.

All calculations were run on a laptop computer under the Windows XP operating system. MATLAB (version 7, Math work, Inc., http://mathworks.com, USA) was used for the MLR analysis and other statistical calculations.

3. Results and Discussion

3.1. Model construction based on structural parameters of solvents (Model 1)

The study was conducted based on 834 solvent structural parameters for 59 solvents with known RED number for PC/CNT composite [16]. Hence, a data matrix with the size of 59×834 was achieved. Randomly, 13 solvents out of 59 (about 20%) were chosen as test samples, while the remaining 46 solvents (about 80%) were utilized as the training set. It is noteworthy to mention that the RED numbers of solvents in training set covered the RED number of the whole data set. The range of 0.18 to 3.45 of RED numbers was selected for the training set and the values of test set were in the range of 0.26 to 1.90. It is clear that the concept of RED in PC/CNT as a composite material is defined as similar as this concept in regular polymers.

Seven different models were created step by step through the stepwise regression run but some of them may be over fitted. Cross-validation method was applied for each model so that the most convenient correlation equation was selected. The plot of variation of the calibration-squared correlation coefficient (R^2_{cal}) and cross validation (Q^2), are illustrated in Fig. 1a. As can be seen, the model performance refined up to four variables by introducing each new variable and after that no drastic change was observed. The resultant four-parametric equation is:

$$RED = 0.895(\pm0.024) - 0.283(\pm0.025)FDI + 0.236(\pm0.30)GATS1e - 0.148(\pm.033)BEHe8 + 0.106(\pm0.034)RDF010m \qquad (1)$$

$N = 59$, $N_{train} = 46$, $N_{test} = 13$, $R^2_{train} = 0.91$, $Q^2_{LOO} = 0.87$, $F = 102.58$, $F_{crit.} = 2.60$

In this equation, FDI is related to folding degree index, GATS1e is Geary autocorrelation coefficient lag1 which is weighted by atomic Sanderson electronegativities, BEHe8 shows the highest eigenvalue (number 8) of Burden matrix weighted by atomic Sanderson electronegativities and RDF010m is a radial distribution function descriptor weighted by atomic Sanderson electronegativities and atomic masses respectively (Table 2). With the relative amount of contribution of each of these parameters, it is possible to estimate relevant information about solvent-PC/CNT composite interaction in the solution phase and will be discussed more in the next parts of manuscript.

Moreover, N is the number of solvent, R^2_{train} is the squared correlation coefficient of calibration (training set) and Q^2_{LOO} is the squared correlation coefficients for leave-one-out

cross-validation and their values show the goodness and stability of the proposed model. F is the Fischer F-statistic and F_{crit} is the F-critical value. The high value of calculated F (in comparison with F_{crit}) verified the statistical significance of the resultant model. Also, the values in parentheses are the coefficients' standard deviation and their low values indicate the significance of the selected parameters. On the other hand t-test showed the significance of the model with p-values almost equal to zero (Table S1, Supporting information). It is worth mentioning that this equation explained more than 91% of the variances in the PC/CNT composite solubility data with an excellent statistical quality.

After auto scaling of the selected parameters, the obtained standardized MLR-coefficients were used to calculate values of RED number for training and test set which the predicted REDs are shown in Table 1. To check the stability of the proposed model, leave-one-out and leave-two out cross validations were also done and the squared correlation coefficients for cross-validation (Q^2_{LOO} and Q^2_{L2O}) were obtained equal to 0.87 and 0.86 respectively. These parameters were indicators of model stability and prediction ability.

Furthermore, it has been proposed that the true predictive power of a QSPR model should be evaluate by an external test set of compounds that were not used in the model construction [19]. Hence, the squared correlation coefficient (R^2_{test}) and root mean square error of the test set (RMSEP) were calculated for further checking. It is worth mentioning that the R^2_{test} and RMSEP were obtained equal to 0.91 and 0.31 respectively, which are another evidence of the excellent predictability of the obtained QSPR model. For better visualization, the predicted values were plotted versus the experimental values of the model in Fig. 1b. This figure displayed an outstanding agreement between experimental and predicted values of RED based on proposed model.

Model validation is by far the most crucial step of QSPR and a great number of procedures have been generated for the determination of the quality of QSPR model [20]. As it was explained, cross-validation and external validation procedures were established in this work for this goal. As a general rule of thumb, some statistical parameters including cross-validation square correlation coefficient (Q^2_{cv}), prediction residual sum of squares and root mean square error in cross-validation (RMSEcv) are the most powerful parameters to check the cross-validation predictability [21]. In comparison with R^2_{cal}, which can be soared by adding more

6

parameters, Q^2_{cv} declined in the present of an over parameterized model [22]. Therefore, Q^2_{cv} value is far more meaningful for measuring the average predictive power of a model and also it can be considered to select the optimum number of parameters (four-parametric linear model in this study).

The permutation test named y-randomization or y-scrambling (randomization of response, i.e., RED in this study) is another procedure for checking the significance of the Q^2_{cv} value [23]. In the current study, RED in the training set was scrambled 50 times and as the statistic parameter of the permutation test, a maximum of the cross- validated squared correlation coefficient of the scrambled data (Q^2_{MS}) was calculated, which was 0.32 and was so far from the Q^2_{cv} of the original model. This parameter could confirm that the model 1 was not obtained by chance.

3.2. Model construction based on empirical solvent scale

The data of 54 solvents out of 59 solvents for solvent empirical parameters were available in literature and the data of 5 solvents (isophorone, mesityl oxide, diethylene glycol monobuthyl ether, diethylene glycol monoethyl ether acetate and ethylene glycol monoethyl ether acetate) were not found. Therefore, this model was constructed based on the solvent empirical parameters of 54 solvents and after deleting the collinear variables, the data matrix with the size of 54×121 was obtained.

42 solvents were selected randomly for training set, whereas 12 solvents were dedicated to test set. As mentioned earlier, the RED number of training set covered the RED number of the whole data set.

Again, stepwise MLR was employed for selecting the most relevant subset of scales among the solvent scales as independent variables. Similar to which was noted for QSPR model, cross validation were used to select final model without the risk of over fitting. According to Fig. 2a the performance of model was not improved very much after including 4 parameters. So, a four-parametric equation was selected:

$$\text{RED} = -0.360(\pm 0.182) + 0.062(\pm 0.007)\text{Ap} - 1.697(\pm 0.199)\text{SPP}^N + 0.031(\pm 0.006)\delta +$$
$$0.751(\pm 0.239)X_e^R \qquad\qquad\qquad\qquad (2)$$

$N = 54$, $N_{train} = 42$ $N_{test} = 12$, $R^2_{train} = 0.95$, $Q^2_{LOO} = 0.93$, $F = 180.744$, $F_{crit.} = 2.6261$

Because of using the solvent empirical parameters as independent variables of the model, this model could be considered as a linear solvation energy relationship (LSER). In the LSER approach, the effects of solvent–solute interactions on physicochemical properties and reactivity parameters are studied [24]. So from this aspect, this equation is similar to the LSERs which has been proposed by Kamlet and Taft [25,26].

This model confirmed an acceptable relationship between the selected variables and the RED number of the solvents which explained %95 of data variance. It is noteworthy to mention that the model is statistically significant based on the F-statistics and t-test (See Table S2, Supporting information). Besides, the model represented a cross-validation statistics similar to the calibration set. The high value of correlation coefficient of leave-one-out and leave-two-out cross validation for the LSER model ($Q^2_{LOO} = Q^2_{L2O} = 0.93$) and acceptable value of cross validation root mean square error (RMSEcv = 0.27) proved the predictive power and the stability of the model based on interstice error in experimental RED data (Table 1).

The root mean square error and correlation coefficient of the test set (RMSEtest and R^2_{test}) were 0.15 and 0.98 respectively which confirmed the high external predictive ability of the model. The y-randomization test also showed this model was not a chancy model ($Q^2_{MP} = 0.29$ compared to Q^2_{cv} of original model = 0.93). Fig. 2b displays the predicted values versus the experimental values of the model which illustrates an outstanding agreement between experimental and predicted values of RED.

3.3. Applicability Domain

To show the scope and limitation of a QSAR model, the applicability domain (AD) is widely utilized. The concept for the applicability domain of a model is closely pertained to the term model validation. AD is defined as the substantiation that a model possesses a satisfactory range of accuracy within the intended application of the model [27].

The leverage of a compound produces a check for multivariate normality of observation [28]. In other words, it provides a measure of the distance of the compound from the centroid of the model space and in the model building; the chemical compounds close to the centroid are less influential and beneficial. The so-called influence matrix or hat matrix (H) is given by:

8

$$\mathbf{H} = \mathbf{X} (\mathbf{X}^{\mathrm{T}} \mathbf{X})^{-1} \mathbf{X}^{\mathrm{T}} \tag{3}$$

where X is the descriptor matrix which is derived from the training set, $\mathbf{X}^{\mathrm{T}}$ is the transpose of $\mathbf{X}$ and $(\mathbf{X}^{\mathrm{T}} \mathbf{X})^{-1}$ is the inverse of matrix $(\mathbf{X}^{\mathrm{T}} \mathbf{X})$. The leverage or hat values (h_i) of the ith compound in the descriptor space are the diagonal elements of $\mathbf{H}$, which can be calculated by:

$$h_i = \boldsymbol{x}_i (X^{\mathrm{T}}X)^{-1} \boldsymbol{x}_i^{\mathrm{T}} \tag{4}$$

where $\boldsymbol{x}_i$ is the descriptor row vector of the interest compound. h^* or warning leverage is defined as $h^* = 3m/n$, where n is the number of training compounds and m is the number of model parameters plus one [28] (here 4+1 in both models 1 and 2). A compound that possesses a high leverage (bigger than a warning value) in the training set enjoys a remarkable influence on the regression line and can force the fitted line to be close to the observed of that compound. The leverage values for all polymer/solvent samples in training and test sets were calculated for both models.

In addition to the high leverage values, it should be noted that some compound may also be outside the AD due to their large standardized residuals [28]. As a result, Williams plots of the models can be utilized to identify the compounds outside the applicability domain [24] which are shown in Fig. 3a and 3b for QSPR and LSER models respectively. In this famous plot, both leverage and standardized residuals are taken into consideration to visualize the applicability domain. It is crystal clear that all solvents are in the applicability domain except water in model 2 and Ethylene glycol in model 1 and 2 and most of solvents have a standardized residual in the acceptable range of $\pm 3\sigma$.

3.4. Interpretation of models

Simplicity of the use and interpretability are considered as outstanding merits of MLR. In addition, in modeling of the target property (RED number), the magnitudes of coefficients show the relative significance of the descriptors and their signs indicate the positive or negative contribution of the molecular descriptors to the RED. It is notably to mention that the molecular descriptors should be mathematically independent of or orthogonal to each other and collinear descriptors can cause coefficients larger than expected or produce the wrong signs [20].

As it is shown in Table 3, which indicates the correlation matrix of variables of both models, the correlation coefficients amongst parameters are insignificant and negligible. The correlation matrix shows the inter-correlations between parameters of each model (QSPR and LSER). To check the multicollinearity of the parameters in each models, the variance inflation factor (VIF) was calculated for each variable [29] and is shown in the last column of Table 3.

If R^2_j is the multiple correlation coefficient of one variable's effect regressed on the remaining molecular variables, the VIF of each theoretical descriptors in model 1 and each solvent empirical parameter in model 2 can be calculated according to the following equation:

$$VIF = \frac{1}{1-R^2_j} \tag{5}$$

It has been suggested that 5.0 is the cut-off value for VIF and if VIF would be larger than the cut-off value, the main information of descriptors can be concealed by correlation of descriptors [29]. According to Table 3, the VIF values of four parameters of two models were less than 5.0. So, the correlation matrix and VIF of parameters show that the statistical significance of the proposed models is acceptable and the sign and attribute of the coefficients could be used to describe the models.

3.4.1. Description of model 1

Standardized regression coefficient can be applied for the assessment of the relative importance of the variables included in an MLR model. The order of significance of the variables of Eq. (1) is FDI > GATS1e > BEHe8 > RDF010m which is shown in Fig. S1-a (supporting information). As it can be seen, the most significant parameter is FDI, while RDF010m is the less important variable of the model in the dispersibility of PC/CNT composite. FDI is the "folding degree index" which is more applicable for peptides and proteins [30], but is used for other molecules as a structural descriptor for other small molecules [31]. A molecule with more folded structure (e.g. lower structural volume) has higher FDI value. The negative sign of the MLR coefficient related to FDI in the proposed model shows that solvents with more unfolded and higher structural volume (lower FID) resulted in higher RED for PC/CNT. It is worth mentioning that better solvents for PC/CNT have lower RED. This finding was in accordance with literature [16].

Two next descriptors in the model GATS1e and BEHe8 "Geary autocorrelation coefficient lag1" and "Highest eigenvalue number 8 of Burden matrix" respectively and both of them are weighted with atomic Sanderson electronegativity [31]. Therefore, it can be concluded that the electronegativity of solvents and consequently the polar interactions have a significant role in solvation or dispersibility of PC/CNT. On the other hand, the positive sign of GATS1e and negative sign of BEHe8, imply that polar interactions may have an optimum value (not high, not low) to decrease the RED.

The last parameter of the model was RDF10m which is a Radial Distribution Function/ weighted by atomic masses with a positive coefficient [31]. It shows that the solvents with lower atomic mass possess lower RED and could be better solvents for PC/CNT.

3.4.2. Description of model 2

According to the standardized regression coefficient of model 2 and as it is illustrated in Fig. S1-b (supporting information), the order of importance of the variables of Eq. (2) is $Ap > SPP^N > \delta^H > X_e^R$.

Ap is the "acidity parameter" calculated from the data for the Gibbs solvation energy for the alkali metal cations and halide ions [32]. According to the thermodynamic studies of the properties of electrolyte solutions, it is well-known that acidity and basicity are important in determining solution properties. The positive sign of Ap in Eq. (2) shows the direct relationship of acidity parameter of solvents and RED related to PC/CNT. So the good solvents, with lower RED, have lower acidity parameter.

SPP^N is a solvent empirical parameters calculated from the UV-Vis spectra of 2-(dimethyl-amino)-7- nitrofluorene and its homomorph 2-fluoro-7-nitrofluorene [33]. SPP^N is a solvent dipolarity-polarizability scale that combines the medium dipolarity and polarizability into a single parameter. This scale is potentially useful for assessing the polarity of solution medium and so it can be assumed as an indicator of polar interactions. The negative sign of SPP^N in the proposed model shows that increasing the solvent polarity and polar interaction cause decreasing the RED of solvents and increasing solvation of PC/CNT.

The 3rd parameter in the proposed model was δ which is the squared root of cohesive energy density and is known as total solubility or Hildebrand solubility parameter [17,34]. According to the positive contribution of this parameter in the model (Eq. (2)), decreasing the squared root of

11

cohesive energy leads to decreasing the RED of a solvent. It should be emphasized that the values of δ (used for model construction) were adapted from the work of Katritzky et. al. [18], in which the δ of some solvents have been measured experimentally and some others have been calculated computationally (based on experimental results). The well-known routine way to calculate the RED values of solvent/polymer systems is using Hansen solubility parameters. The basic equation to define Hansen parameters is that the total cohesion energy, E, has three parts[17]:

$$E = E_D + E_P + E_H \tag{6}$$

Square of the total (or Hildebrand) solubility could be obtained by dividing this by the molar volume gives the parameter as the sum of the squares of the Hansen D, P, and H components (Dispersive, Polar and Hydrogen bonding):

$$E/V = E_D/V + E_P/V + E_H/V \tag{7}$$

$$\delta^2 = \delta^2_D + \delta^2_P + \delta^2_H \tag{8}$$

The interpretation of similarity between two materials (indices 1 and 2) is possible by calculating their solubility parameters "distance" R_a based on their respective partial solubility parameters [17]:

$$R_a^2 = 4(\delta_{D1} - \delta_{D2})^2 + (\delta_{P1} - \delta_{P2})^2 + (\delta_{H1} - \delta_{H2})^2 \tag{9}$$

A convenient single parameter to characterize solvent quality in this model is the relative energy difference, RED number:

$$RED = R_a/R_o \tag{10}$$

Which R_o is the interaction radius and R_a is the solubility parameter distance.

In this work, we showed that some other parameters could be important in the prediction and description of RED values in addition to the dependency of RED of solvent/PC/CNT to cohesive energy components.

The last parameter of the proposed LSER model is X_e^R with positive effect on RED. This parameter is defined as selectivity parameter: reflects a composite of solvent's dipolarity-

dipolarizability, hydrogen bond acidity and hydrogen bond basicity [35]. According to the discussion about SPP^N, it could be concluded that dipolarity-dipolarizability has a moderate effect on RED but hydrogen bonding interactions shows positive effects on RED. The results about the moderate effect of polar interaction, was in agreement with model 1. In conclusion, better solvents for solvation of PC/CNT might have lower hydrogen bonding interactions.

3.5. Comparison of the QSPR and LSER models

All statistical data of QSPR and LSER models are gathered in Table 4. The non-significant values of Q^2_{cv} of permutation test (Q^2_{MP}) confirm that the models are reasonable and they have not obtained by chance. More importantly, although both proposed models resulted in very good statistics, LSER provides better statistical data for both calibration and prediction sets. Hence, experimental solvent scale supplies remarkable ability to relate the dispersibility of PC/CNT composite and solvent type. In the LSER approach, the effects of solvent-solute interactions on physiochemical properties and reactivity parameters could be studied. However, a crucially important constraint of solvent scale is that they are experimental parameters and it is difficult to gather these data for new solvents. As a matter of fact, the application of the model is in need of previous experiments to measure the solvent scale. For instance, in our study we could not find solvent scales of five solvents. Another great advantage of solvent empirical parameters, in comparison with solvent theoretical descriptors, is their lower initial population which decreases the risk of obtaining chancy model in the case of limited number of compounds available for modeling [36–38]. While, the descriptors of QSPR model are calculated theoretically with the contribution of the available software and this model can be applied for the prediction of the dispersibility in new solvents, even for virtual solvents that have not been synthesized or checked experimentally yet.

3.6. Exploratory Data Analysis

Because of the success of proposed models in prediction the RED of polymer/solvents, principal component analysis (PCA) was utilized on the parameters of each model (#1 and #2) for an exploratory data analysis of relative resistance change (R_{rel}) of PC/CNT in different discussed solvents. R_{rel} could be calculated using the time dependent resistance, R (t), of the

samples over immersion time in solvent and initial resistance, R_i, according to the following equation:

$$R_{rel}\% = \frac{R(t) - R_i}{R_i} \times 100 \qquad (11)$$

Information about the sensory composite's response to the solvent contact (Good: $R_{rel} > 0$ or Bad: $R_{rel} \sim 0$) were adapted from the work of Villmow et al [16] and is represented in Table S3 (supporting information).

2D-plot of PC1 vs. PC2 of the parameters of model 1 based on the theoretical descriptors and model 2 constructed from empirical scales are shown in Fig. 4a and 4b respectively. As it is clear from Fig. 4, both set of used parameters shows acceptable discrimination between good and bad solvents and can be an indicator of solvent selectivity.

3.7. Test on another kind of polymer

To show the generality of the proposed model, the utilized parameters in model 1 and model 2 was used to check the behavior of another kind of polymer. The original RED data used for construction of models in the above was belonging to a polymer composite based on PC Lexan 141R (SABIC Innovative Plastics) [16]. We used anther data set based on R polycarbonate polymer [17]. The RED of this polymer in the same solvents noted for PC Lexan 141R were used as the dependent variable for model construction base on the parameters in two proposed model (Eqs. 1 and 2). The RED values of this polymer in the discussed 59 solvents is represented in Table S4 (Supporting information). Here again, similar to which was done for original data, RED data of R polycarbonate polymer was divided to two parts i.e. training and test sets. It was observed both series of parameters used in model 1 and model 2 shows good ability in the prediction of RED of the new kind of polymer. This ability was checked using different statistical parameters related to training, cross validation and test set which is presented in Table 5. The agreement between predicted RED and their experimental values are also shown graphically in Fig. 5. The MLR coefficient of the models for this R polycarbonate polymer using the parameters of model 1 and model 2 and Williams plots of these two new models are represented in supporting information (Tables S5, Table S6 and Fig. S2). The results show the ability of proposed models to predict the behavior of other kind of similar polymers.

14

4. Conclusion

New approaches were proposed based on QSPR and also a kind of LSER to predict RED of solvents-PC/CNT and the validity, stability and prediction ability of models were verified using different statistical methods. It was shown that the conductive polymer composites/ carbon nanotube has better selectivity (lower RED) in solvents with more folded molecular structure (lower structural volume) and lower structural mass. According to the both positive and negative effects of structural electronegativity on RED, it was concluded that polar interactions have moderate effect on the solvation of the desired PC/CNT.

Some other aspects of solvent-PC/CNT interactions were also emerged using the 4 parametric LSER model. The suggested model showed that in addition to the total (or Hildebrand) solubility parameter, acidity parameter of solvent and hydrogen bonding interactions have direct relationship with RED. But again, polarity-polarizability interactions show two-side (reverse and direct) effect on the RED and selectivity. So the role of this kind of interactions in our polymer-solvent system is more complex.

The combination of using theoretical descriptors and solvent empirical parameters resulted in successful experience in the investigation of solvation of PC/CNT in different organic solvents. Extension of this study on conductive polymer composites with different amount of CNT is in progress of our future researches.

References

1. H. M. Kim, K. Kim, C. Y. Lee, J. Joo, S. J. Cho, H. S. Yoon, D. A. Pejaković, J. W. Yoo, and A. J. Epstein, *Appl. Phys. Lett.*, 2004, **84**, 589.

2. P. Pötschke, A. R. Bhattacharyya, and A. Janke, *Carbon N. Y.*, 2004, **42**, 965–969.

3. B. Krause, T. Villmow, R. Boldt, M. Mende, G. Petzold, and P. Pötschke, *Compos. Sci. Technol.*, 2011, **71**, 1145–1153.

4. F. H. Ferguson DW, Bryant EWS, in *SPE/ANTEC '98 Proceedings*, 1998, pp. 1219–1222.

5. C.-C. M. Ma, Y.-L. Huang, H.-C. Kuan, and Y.-S. Chiu, *J. Polym. Sci. Part B Polym. Phys.*, 2005, **43**, 345–358.

6.   Y. Yang, M. C. Gupta, K. L. Dudley, and R. W. Lawrence, *J. Nanosci. Nanotechnol.*, 2005, **5**, 927–931.

7.   O. Breuer and U. Sundararaj, *Polym. Compos.*, 2004, **25**, 630–645.

8.   K. Ryan, M. Cadek, V. Nicolosi, D. Blond, M. Ruether, G. Armsrong, H. Swan, A. Fonseca, J. Nagy, and W. Maser, *Compos. Sci. Technol.*, 2007, **67**, 1640–1649.

9.   B. K. Satapathy, R. Weidisch, P. Pötschke, and A. Janke, *Macromol. Rapid Commun.*, 2005, **26**, 1246–1252.

10.   T. Villmow, S. Pegel, P. Pötschke, and G. Heinrich, *Polymer (Guildf).*, 2011, **52**, 2276–2285.

11.   S. G. Chen, J. W. Hu, M. Q. Zhang, M. W. Li, and M. Z. Rong, *Carbon N. Y.*, 2004, **42**, 645–651.

12.   Q. Fan, Z. Qin, T. Villmow, J. Pionteck, P. Pötschke, Y. Wu, B. Voit, and M. Zhu, *Sensors Actuators B Chem.*, 2011, **156**, 63–70.

13.   J. R. Li, J. R. Xu, M. Q. Zhang, and M. Z. Rong, *Carbon N. Y.*, 2003, **41**, 2353–2360.

14.   J. F. Feller, J. Lu, K. Zhang, B. Kumar, M. Castro, N. Gatt, and H. J. Choi, *J. Mater. Chem.*, 2011, **21**, 4142.

15.   J. Lu, J. F. Feller, B. Kumar, M. Castro, Y. S. Kim, Y. T. Park, and J. C. Grunlan, *Sensors Actuators B Chem.*, 2011, **155**, 28–36.

16.   T. Villmow, A. John, P. Pötschke, and G. Heinrich, *Polymer (Guildf).*, 2012, **53**, 2908–2918.

17.   C. Hansen, *Hansen solubility parameters: a user's handbook*, CRC Press LLC, Boca Raton, 2000.

18.   A. R. Katritzky, D. C. Fara, M. Kuanar, E. Hur, and M. Karelson, *J. Phys. Chem. A*, 2005, **109**, 10323–41.

19.   I. V. Tetko, D. J. Livingstone, and A. I. Luik, *J. Chem. Inf. Model.*, 1995, **35**, 826–833.

20.   L. Eriksson, J. Jaworska, A. P. Worth, M. T. D. Cronin, R. M. McDowell, and P. Gramatica, *Environ. Health Perspect.*, 2003, **111**, 1361–1375.

21.   V. Consonni, D. Ballabio, and R. Todeschini, *J. Chemom.*, 2010, **24**, 194–201.

22.   D. M. Hawkins, S. C. Basak, and D. Mills, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 579–86.

23.   C. Rücker, G. Rücker, and M. Meringer, *J. Chem. Inf. Model.*, 2007, **47**, 2345–57.

24.   R. Todeschini, V. Consonni, and P. Gramatica, in *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, ed. S. D. Tauler, R., Walczak, B., & Brown, Elsevier B.V., Amsterdam, 2009, pp. 140–141.

25.   M. J. Kamlet and R. W. Taft, *J. Chem. Soc. Perkin Trans. 2*, 1979, 337.

26.   R. W. Taft, J.-L. M. Abboud, M. J. Kamlet, and M. H. Abraham, *J. Solution Chem.*, 1985, **14**, 153–186.

27.   P. Gemperline, Ed., *Practical Guide to Chemometrics*, Taylor & Francis Group, Boca Raton, 2nd edn., 2006.

28.   T. I. Netzeva, A. P. Worth, T. Aldenberg, R. Benigni, M. D. . Cronin, P. Gramatica, J. S. Jaworska, S. Kahn, G. Klopman, C. A, G. Myatt, N. Nikolova-jeliazkova, G. Y. Patlewicz, and R. Perkins, *ATLA*, 2005, **2**, 1–19.

29.   T. a. Craney and J. G. Surles, *Qual. Eng.*, 2002, **14**, 391–403.

30.   E. Estrada, *Comput. Biol. Chem.*, 2003, **27**, 305–313.

31.   R. Todeschini and V. Consonni, *Molecular Descriptors for Chemoinformatics*, WILEY-VCH, Weinheim, Second., 2009.

32.   W. R. Fawcett, *J. Phys. Chem.*, 1993, **97**, 9540–9546.

33.   J. Catalán, V. López, P. Pérez, R. Martin-Villamil, and J.-G. Rodríguez, *Liebigs Ann.*, 1995, **1995**, 241–252.

34.   C. Reichardt and T. Welton, *Solvents and Solvent Effects in Organic Chemistry*, WILEY-VCH, Marburg, 3rd edn., 2011.

35.   S. C. Rutan, P. W. Carr, W. J. Cheong, J. H. Park, and L. R. Snyder, *J. Chromatogr. A*, 1989, **463**, 21–37.

36.   S. Yousefinejad and B. Hemmateenejad, *Colloids Surfaces A Physicochem. Eng. Asp.*, 2014, **441**, 766–775.

37.   S. Yousefinejad, F. Honarasa, F. Abbasitabar, and Z. Arianezhad, *J. Solution Chem.*, 2013, **42**, 1620–1632.

38.   S. Yousefinejad, F. Honarasa, and N. Saeed, *J. Sep. Sci.*, 2015, DOI: 10.1002/jssc.201401427

Figures' captions

Fig. 1. The plot of model performance vs. number of variables included in model 1 (a) and plot of predicted RED number using the model 1 versus experimental values

Fig. 2. The plot of model performance vs. number of variables included in model 2 (a) and plot of predicted RED number using the model 2 versus experimental values (b)

Fig. 3. Williams plot of the entire set of solvents in model 1 (a) and model 2 (b). Cut off values of leverage (h^*) and standardized residual (±3 times the standards deviation) are depicted by vertical and horizontal dashed lines, respectively. Ethylene glycol is out of the applicable domain in model 1; Ethylene glycol and water are out of the applicable domain in model 2.

Fig. 4. Discrimination of "good" and "bad" solvents for PC/CNT using principal component analysis on parameters entered in model 1 (a) and model 2 (b).

Fig 5. Plot of predicted RED number using the parameters in model 1(a) and in model 2 (b) versus experimental values
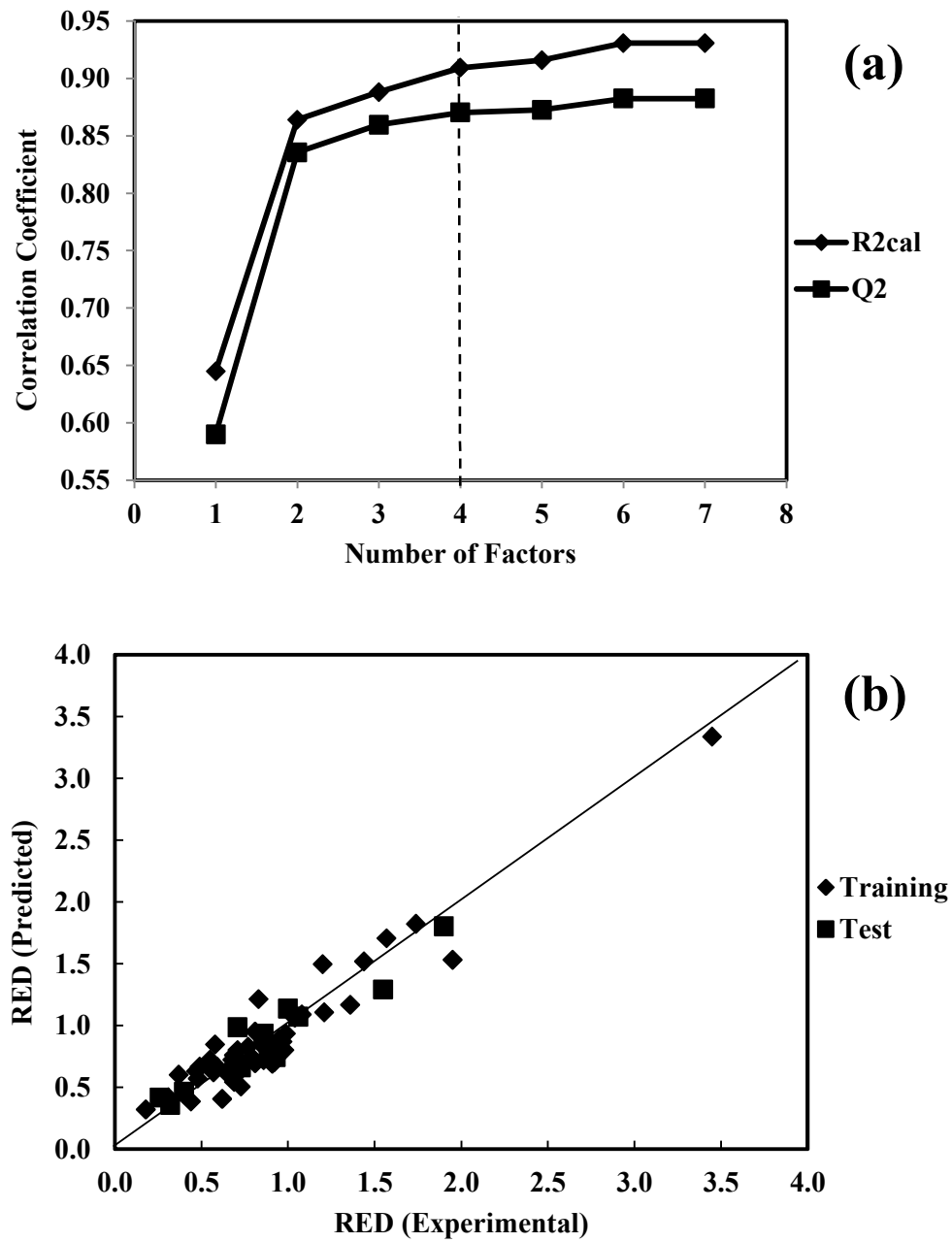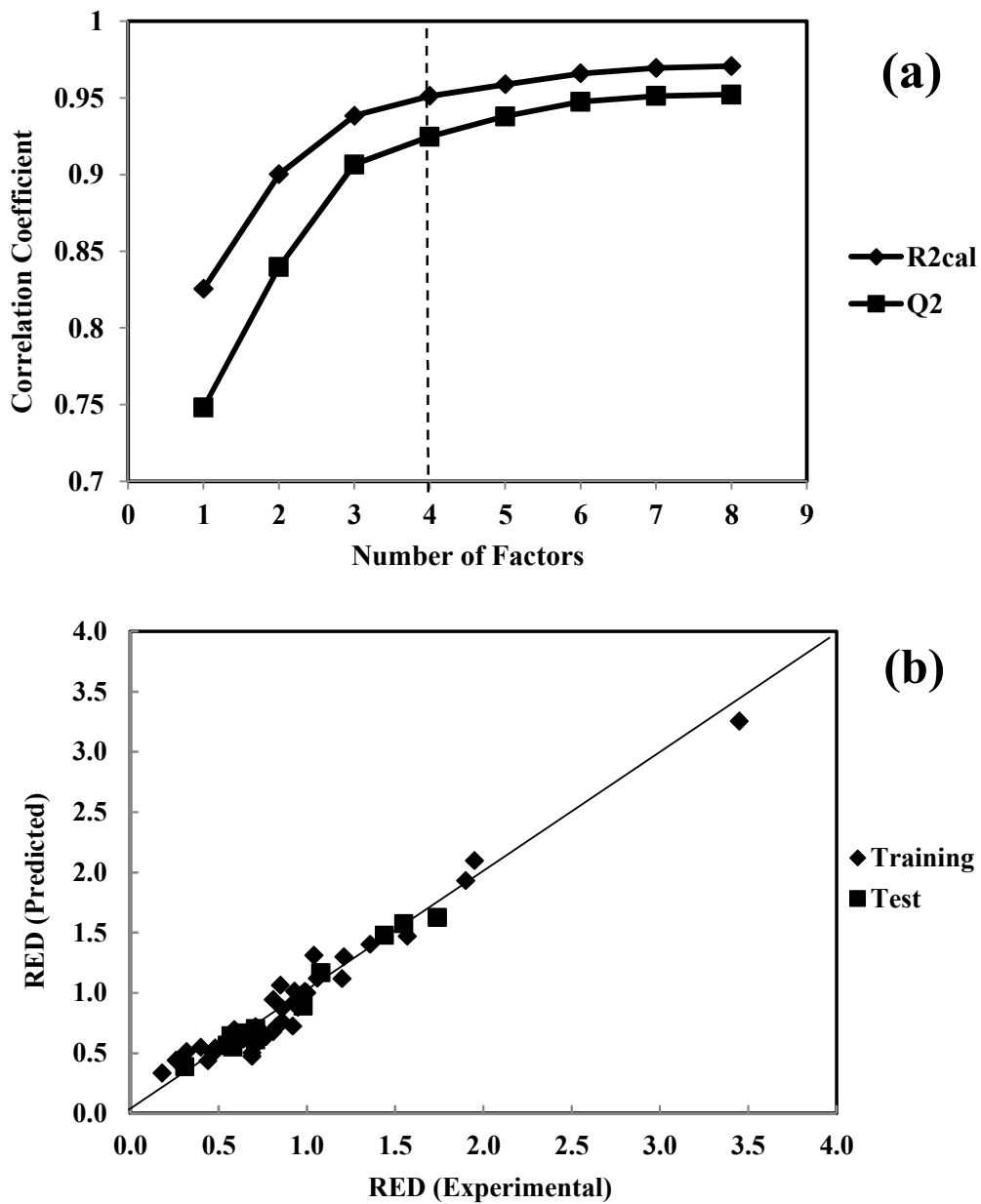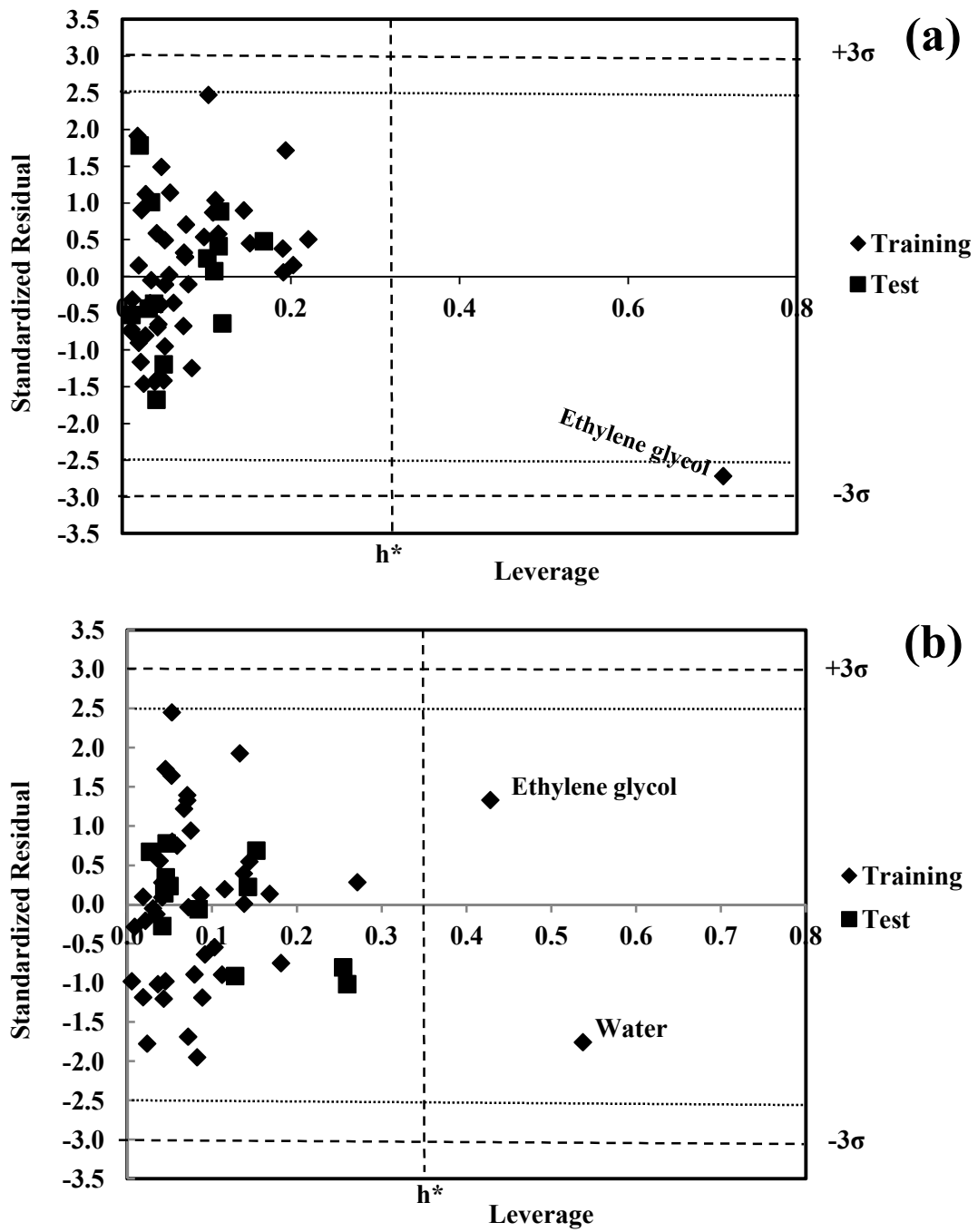
Fig. 1.

Fig. 2.

Fig. 3.
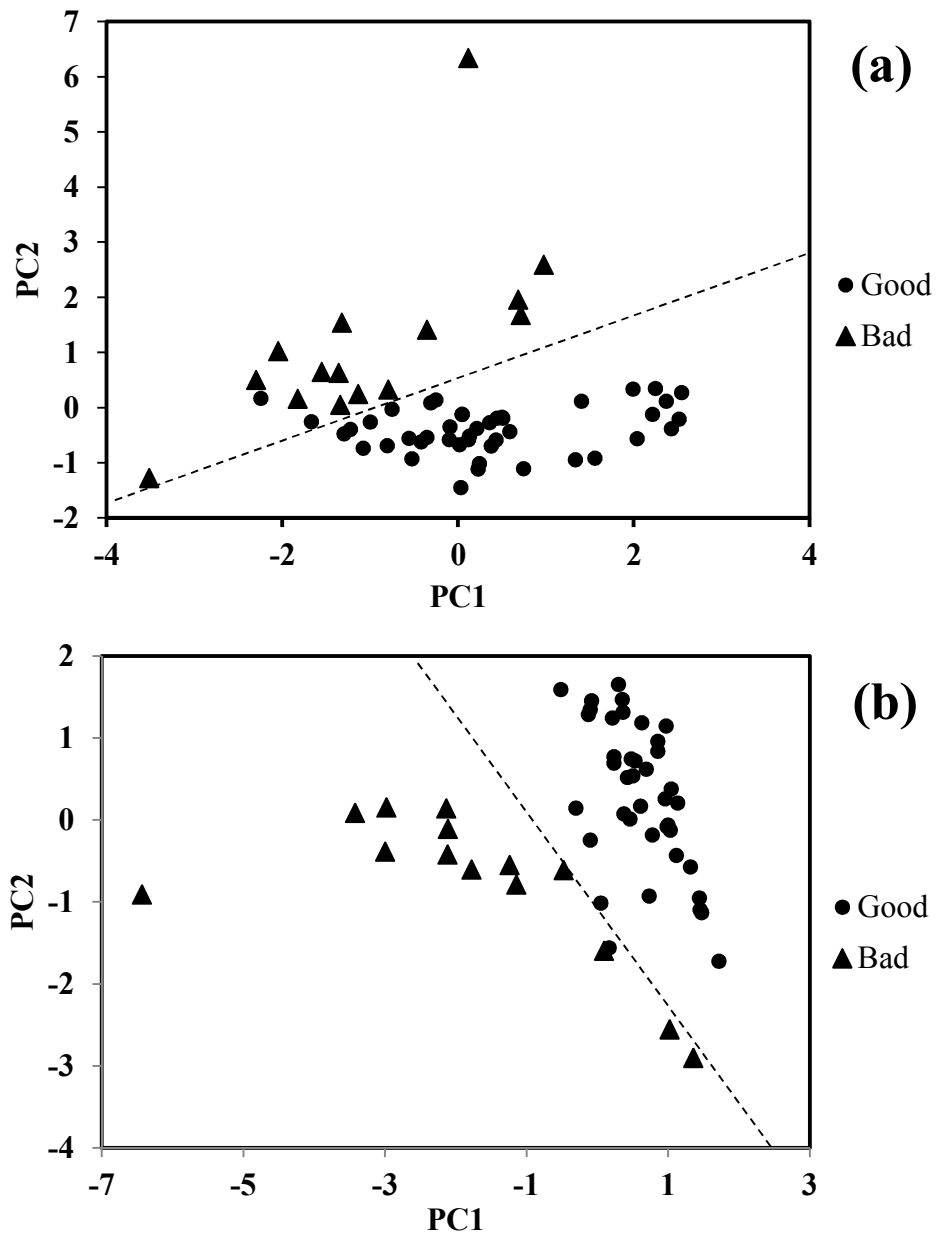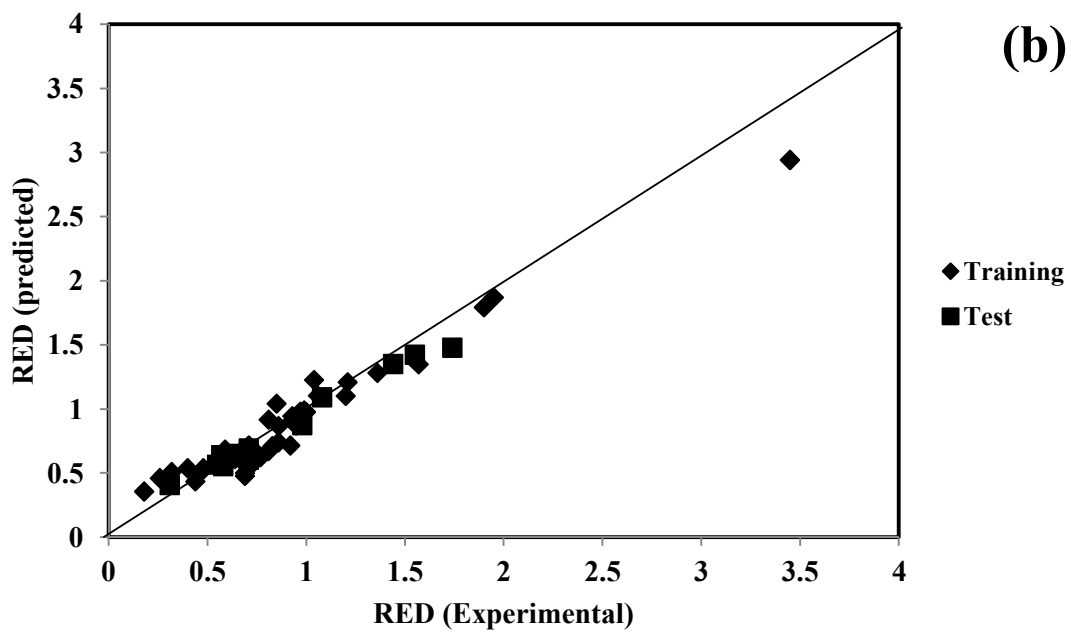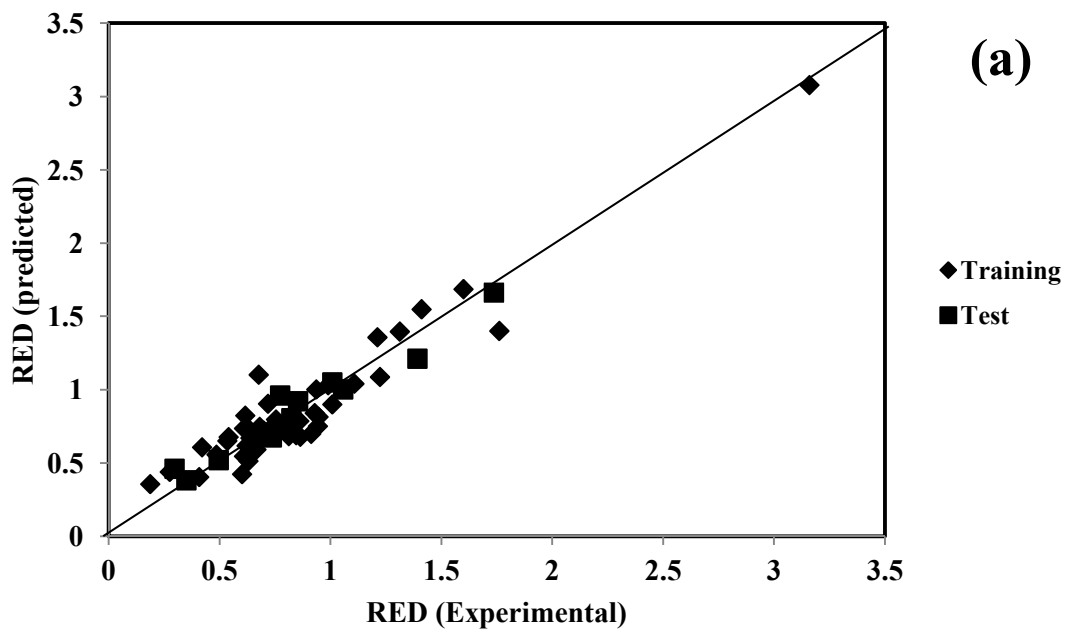

(a)

(b)

Fig. 4.

Fig. 5.

(a)

(b)

24

Table 1 Experimental and predicted RED number of PC/CNT composite in various solvents

No.	Solvent	RED (Exp)	RED (Pred1)[a]	RED (Pred2)[b]	No.	Solvent	RED (Exp)	RED (Pred1)[a]	RED (Pred2)[b]
1	Acetophenon	0.18	0.32	0.33	**31**	n-Butyl acetate	0.81	0.69	0.68
2[c]	Cyclohexanone	0.26	0.42	0.44	**32**	Morpholin	0.81	0.95	0.95
3[d]	n-Metyl-2-pyrrolidone	0.31	0.42	0.39	**33**	Aniline	0.83	1.21	0.72
4[c]	Ethylene dichloride	0.32	0.36	0.51	**34**	Furan	0.85	0.84	1.06
5	Isophorone	0.37	0.60	-	**35[c]**	Nitromethane	0.86	0.93	0.76
6[c]	2-Nitropropane	0.40	0.46	0.55	**36**	Toluene	0.86	0.72	0.87
7	o-Dichlorobenzene	0.44	0.38	0.44	**37[c]**	Diethylene glycol monomethyl ether	0.89	0.83	-
8	Methyl ethyl ketone	0.47	0.63	0.50	**38**	Ethylene glycol monoethyl ether acetate	0.91	0.69	-
9	Methylene dichloride	0.48	0.57	0.54	**39**	Isoamyl acetate	0.92	0.70	0.72
10	Mesityl oxide	0.49	0.67	-	**40**	1,4-dioxane	0.93	0.81	0.92
11[d]	Diethyl ketone	0.55	0.72	0.57	**41[c]**	m-Cresol	0.93	0.75	1.01
12[d]	Butyronitrile	0.57	0.62	0.64	**42**	Ethyl benzene	0.95	0.89	0.88
13[d]	Acetone	0.58	0.84	0.55	**43**	Mesitylene	0.97	0.87	0.98
14	Nitroethane	0.59	0.67	0.69	**44[d]**	Benzene	0.98	0.80	0.89
15	Chlorobenzene	0.62	0.41	0.61	**45**	Diethyl ether	0.99	0.93	1.01
16[c]	Anisole	0.64	0.56	0.62	**46[c]**	Dipropyl amine	1.00	1.14	1.00
17[d]	Tetrahydrofuran	0.64	0.62	0.67	**47**	Cyclohexanol	1.04	1.06	1.31
18	Methyl acetate	0.68	0.72	0.69	**48[c]**	Cyclohexane	1.06	1.07	1.12
19	Propylene carbonate	0.69	0.54	0.63	**49[d]**	Ethylene glycol monomethyl ether	1.08	1.09	1.17
20	Acetic anhydride	0.69	0.58	0.58	**50**	n-Hexane	1.20	1.50	1.12
21	Methyl butyl ketone	0.69	0.67	0.47	**51**	1-Butanol	1.21	1.11	1.30
22	Methyl isobutyl ketone	0.69	0.76	0.50	**52**	Diethylene glycol	1.36	1.17	1.40
23	Trichloroethylene	0.70	0.70	0.59	**53[d]**	Ethanol	1.44	1.52	1.48
24[d]	Dimethyl formamide	0.71	0.73	0.61	**54[c,d]**	Propylene glycol	1.55	1.29	1.57
25[d]	Chloroform	0.71	0.80	0.70	**55**	Ethanolamine	1.57	1.70	1.47
26[c]	Diethyl carbonate	0.71	0.98	0.72	**56[d]**	Methanol	1.74	1.82	1.63
27	Dimethyl sulfoxide	0.73	0.50	0.60	**57[c]**	Formamide	1.90	1.80	1.93
28[c]	Ethyl acetate	0.73	0.66	0.70	**58**	Ethylene glycol	1.95	1.53	2.10
29	Di-(2-methoxyethyl) ether	0.77	0.83	0.64	**59**	Water	3.45	3.34	3.26
30	Diethylene glycol monobutyl ether	0.79	0.74	-					

[a] The predicted values using model 1, Eq. (1)

[b] The predicted values using model 2, Eq. (2)

[c] Compounds in the test set using model 1

[d] Compounds in the test set using model 2

Table 2 The brief definitions of variables of the proposed models.

Solvent scale	Definition	Property of scale
FDI	Folding degree index	Molecular Size and Volume
GATS1e	Geary autocorrelation coefficient lag1/ weighted by atomic Sanderson electronegativities	Topology and electronegativity/polarity
BEHe8	Highest eigenvalue n. 8 of Burden matrix / weighted by atomic Sanderson electronegativities	Topology and electronegativity/polarity
RDF010m	Radial Distribution Function-1.0/ weighted by atomic masses	Topology and structural mass
Ap	Acidity parameter calculated from the data for the Gibbs solvation energy for the alkali metal cations and halide ions.	Acidity
SPP^N	Calculated from the UV-Vis spectra of 2-(dimethyl-amino)-7- nitrofluorene and its homomorph 2-fluoro-7-nitrofluorene	Dipolarity and dipolarizability
δ^H	Square root of cohesive energy density	Total Cohesive energy density
X_e^R	Selectivity parameter: reflects a composite of solvent dipolarity-dipolarizability , hydrogen bond acidity and hydrogen bond basicity	dipolarity-dipolarizability and hydrogen bonding

Table 3 Correlation coefficient between parameters of QSPR and LSER models and their VIF values

	FDI	GATS1e	BEHe8	RDF010m	Ap	SPP^N	δ^H	X_e^R	VIF [a]
FDI	1.000								1.202
GATS1e	0.094	1.000							1.358
BEHe8	0.065	0.000	1.000						1.830
RDF010m	0.002	0.141	0.348	1.000					1.995
Ap					1.000				3.654
SPP^N					0.006	1.000			1.245
δ^H					0.492	0.137	1.000		2.773
X_e^R					0.386	0.001	0.197	1.000	2.000

[a] Variation inflation factor

27

Table 4 Various statistics parameter of the developed model

	nv^a	N_{train}^b	N_{test}^c	$R^2_{train}^d$	$RMSE_{train}^e$	$R^2_{LOO\,cv}^f$	$R^2_{L2O\,cv}^f$	$RMSEcv^g$	$R^2_{test}^h$	$RMSEtest^i$	$Q^2_{MP}^j$
Model 1 (QSPR)	4	46	13	0.91	0.30	0.87	0.86	0.36	0.91	0.31	0.32
Model 2 (LSER)	4	42	12	0.95	0.22	0.93	0.93	0.27	0.98	0.15	0.29

[a] Number of descriptors applied for the model development

[b] Number of molecules in training set

[c] Number of molecules in test set

[d] Training correlation coefficient

[e] Training root mean square error

[f] Leave-one-out and leave-two-out cross-validation correlation coefficient

[g] Leave-one-out cross-validation root-mean-square errors

[h] Correlation coefficient of the test set

[i] Test root-mean-square errors

[j] Maximum cross-validation correlation coefficient for Y-randomization test

Table 5 Various statistics parameter of the developed model for RED of R polycarbonate polymer to show the applicability of proposed model for other kind of polymers.

	nv^a	N_{train}^b	N_{test}^c	$R^2_{train}^d$	$RMSE_{train}^e$	$R^2_{LOO\,cv}^f$	$R^2_{L2O\,cv}^f$	$RMSEcv^g$	$R^2_{test}^h$	$RMSEtest^i$	$Q^2_{MP}^j$
Model 1 (QSPR)	4	46	13	0.91	0.30	0.88	0.87	0.35	0.94	0.25	0.35
Model 2 (LSER)	4	42	12	0.94	0.25	0.90	0.91	0.32	0.97	0.18	0.33