

*Full Paper*

## **High Accurate Prediction of Carbonate Selectivity of PVC-Plasticized Membranes Sensors by Genetic Algorithm-Support Vector Machine**

**Eslam Pourbasheer**

*Department of Chemistry, Faculty of Science, University of Mohaghegh Ardabili, P.O. Box 179, Ardabil, Iran*

\*Corresponding Author,

E-Mail: [e.pourbasheer@uma.ac.ir](mailto:e.pourbasheer@uma.ac.ir)

*Received: 20 December 2022 / Received in revised form: 7 February 2023 /*

*Accepted: 20 February 2023 / Published online: 28 February 2023*

---

**Abstract-** The quantitative structure-property relationship (QSPR) method has been used for the prediction of carbonate potentiometric selectivity of plasticized polymeric membrane sensors. The variable selection tools of genetic algorithm (GA) combined with the multiple linear regressions (MLR) as linear and support vector machine (SVM) as nonlinear regression methods have been used. The K-means clustering method has been used for dividing the data set into the training set and test set. The validation of the models was done by the internal cross-validation and external test set. The results showed that the GA-SVM was a very accurate method in predicting of carbonate potentiometric selectivity with high correlation coefficients of 0.983 and 0.965 for the training and test sets. The results of this study and the interpretation of entered descriptors in the model can help to design new selective ligands.

**Keywords-** Ion-selective electrode; Carbonate sensing; QSPR; Genetic algorithm; Support vector machine

---

### **1. INTRODUCTION**

One of the popular analytical devices for measuring ion concentrations in aqueous solutions is a potentiometric sensor. It measures the potential difference between two electrodes under the conditions of no current flow. In terms of instrumentation, it is simple, portable, and not very expensive, yet quite accurate [1]. The ion-selective sensors, with the base of ionophores

are widely employed for the quantitative determination of ions in different samples [2]. The search for finding of novel ionophores, which are sensitive for different ions, is time-consuming and tedious process. For assay the sensitivity and selectivity of the novel membranes sensors, it need to chemical synthesis of candidate substances, their purification and characterization, sensor membrane preparation, and potentiometric measurements [1].

Overcoming these limitations, techniques based on quantitative structure-property relationships (QSPR) have emerged as a suitable and useful alternative tool in various disciplines [3,4]. Previous studies have shown that QSPR can be applied to modeling the sensing properties of ionophore-based potentiometric sensors [5]. QSPR models based on different approaches essentially seek to correlate and establish a logical relationship between the studied composite structures and their related properties. To begin the QSPR method, it is necessary to calculate the theoretical parameters called descriptors, which are described by the algebraic value of the structure or shapes of each of the selected molecules[6-8]. Among the large number of descriptors calculated for each molecule, only a few descriptors play an important role in molecular properties. Therefore, employing a technique to select the respective variables is one of the essential steps in QSPR method. Over the last few decades, the genetic algorithm (GA) feature selection approach as an effective variable selection method wildly used in the development of QSAR/QSPR models[3,9-13]. The genetic algorithms feature selection method is based on Darwin's evolutionary hypothesis and includes some critical genetic-based functions such as mutation, cross-over, and so forth [14,15].

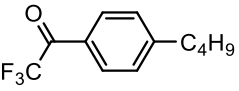
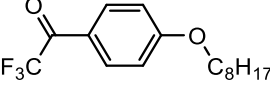
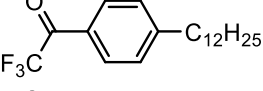
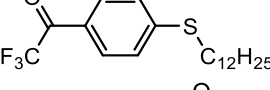
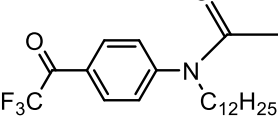
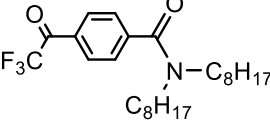
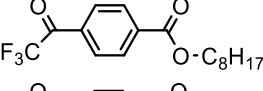
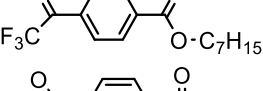
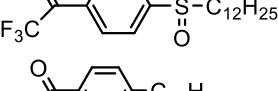
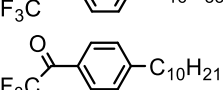
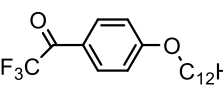
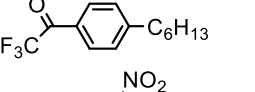
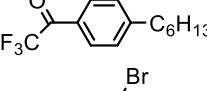
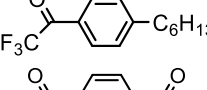
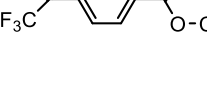

The multiple linear regressions (MLR) are one of the best linear modeling methods, because of its computational efficiency, simplicity and interpretability. Also, the support vector machine (SVM) is one of the most commonly used machine learning algorithms for non-linear data modeling. In this work, the GA-MLR and GA-SVM as linear and nonlinear methods, were used to construct the QSPR models and prediction of carbonate potentiometric selectivity of plasticized polymeric membrane sensors and finally, the results of two models were compared.

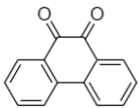
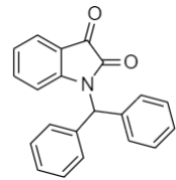
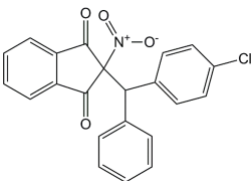
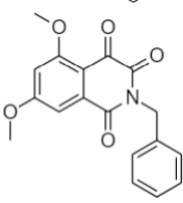
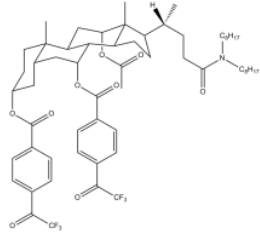
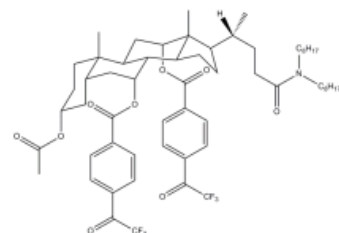
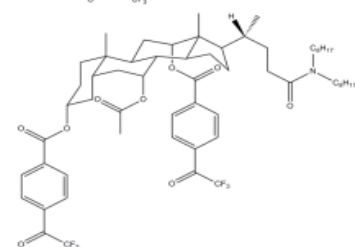
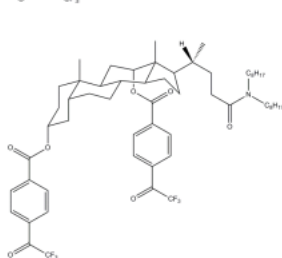
## 2. EXPERIMENTAL SECTION

### 2.1. Data Set

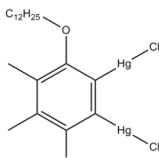
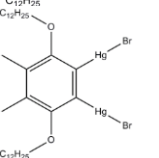
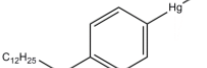
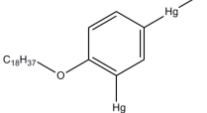
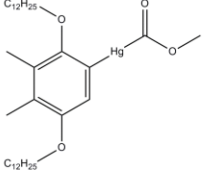
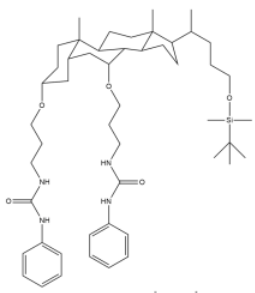
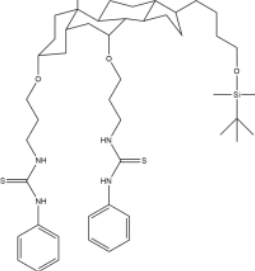
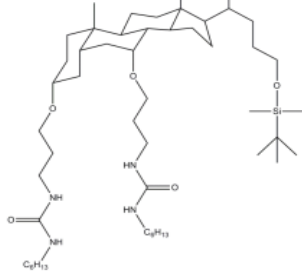
The data set containing of the 40 carbonate ionophores which was compiled from the literature[16-23]. In this study, the selectivity values were the  $\log K^{\text{sel}}(\text{HCO}_3^-/\text{Cl}^-)$ . In this work, from the different data on the literature, only the data that were obtained in the pH range of 7.0–8.6 was considered, to ensure that the  $\text{HCO}_3^-$  is a dominant ionic form in the solutions. The chemical structures of ionophores and the respective selectivity's are listed in Table 1.

**Table 1.** Chemical structures of ionophores, and respective experimental and predicted selectivity

No.	Structure of ionophore	Exp. logK(HCO <sub>3</sub> <sup>-</sup> /Cl <sup>-</sup> )	Pred. GA-MLR	Pred. GA-SVM
M1		-2	-2.2	-2.5
M2		-1.8	-2.8	-2.3
M3		-3.2	-2.7	-3.2
M4		-3.8	-3.2	-3.3
M5		-4	-2.8	-3.4
M6		-5	-5.0	-4.6
M7a		-4	-4.5	-3.7
M8		-5	-4.5	-3.6
M9a		-5	-3.8	-4.6
M10		-2.1	-3.1	-3.1
M11		-3.7	-2.6	-3.4
M12		-3	-3.2	-3.0
M13		-2.8	-2.3	-2.8
M14		-3	-2.0	-2.9
M15		-3.2	-2.4	-3.2
M16		-3.2	-4.4	-3.4
		-3.6	-	-

No.	Structure of ionophore	Exp. $\log K(\text{HCO}_3^-/\text{Cl}^-)$	Pred. GA-MLR	Pred. GA-SVM
M17b				
M18		-0.7	-0.9	-0.7
M19		-1	-1.9	-1.0
M20a		-2.5	-3.4	-2.5
M21		-5.2	-4.3	-5.2
M22		-4.7	-4.0	-4.7
M23a		-5.7	-4.1	-4.9
M24		-3.1	-3.8	-3.7
		-5.8	-4.7	-5.8

No.	Structure of ionophore	Exp. $\log K(\text{HCO}_3^-/\text{Cl}^-)$	Pred. GA-MLR	Pred. GA-SVM
M25				
M26		-4.2	-5.4	-4.2
M27		-2.6	-3.7	-2.6
M28		1.7	1.3	1.7
M29		1.6	0.9	1.6
M30		5.5	6.3	5.5
M31a		4.9	5.9	5.2
M32		6.2	6.5	6.0
		4.8	4.4	4.8

No.	Structure of ionophore	Exp. $\log K(\text{HCO}_3^-/\text{Cl}^-)$	Pred. GA-MLR	Pred. GA-SVM
M33				
M34a		2.8	1.1	1.6
M35		1.4	1.6	1.4
M36		4.5	3.2	4.5
M37a		0.2	0.7	1.5
M38		2.9	2.6	2.9
M39a		1.1	2.4	2.0
M40		1.3	2.1	1.3

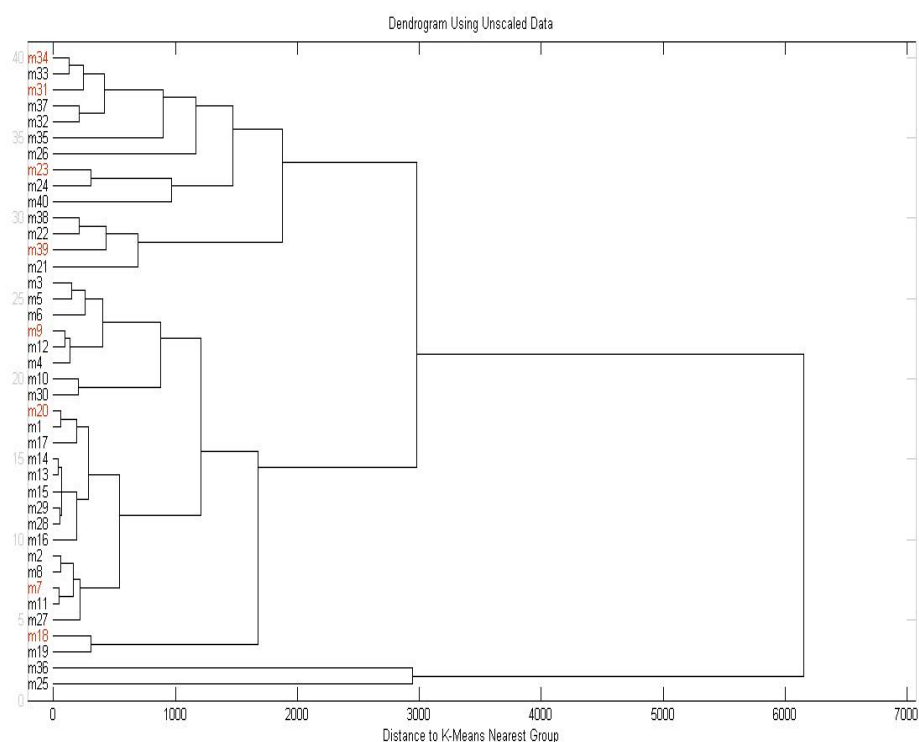
<sup>a</sup> test set<sup>b</sup> Outlier

## 2.2. Descriptors calculation

The 2D chemical structures of the 40 studied ligands were first drawn in Hyperchem 7.5 software. The pre-optimization and also, the final optimization have been done using the molecular mechanics force field (MM+) and semi-empirical methods (AM1), respectively [24]. Optimizing the molecular structures was carried out until the root mean square gradient amounted to 0.01 kcal/mol. Next, the molecular descriptors were obtained using DRAGON v5.5, which uses minimum-energy molecular geometries. Thus for each molecule in the data set, 3224 descriptors were calculated [25,26]. A set of descriptors with constant values and also with almost constant values were removed from the calculated descriptors. After they had been analyzed for constant or near constant variables, among the descriptors with correlation coefficients more than 0.9, only the one with the highest correlation with the selectivity remains for development of the QSPR models. Finally, 302 molecular descriptors remained.

## 2.3. Dividing data set

In order to avoiding any information lost during creation of models and fitting step in the QSPR study, dividing of the data set into the training and test sets is one of the most important steps. In this work, the hierarchical clustering method was used to dividing of the data set. The hierarchical clustering is a statistical method for finding relatively homogeneous clusters of cases based on measured characteristics [27].



**Figure 1.** A dendrogram results of the hierarchical clustering for the training and test sets

The hierarchical clustering can be illustrated by a dendrogram that each step of clustering process is showed by a linkage. Figure 1 shows dendrogram of used data set. The selection of the compounds to dividing of data set to the training and test sets was done randomly from each cluster with close investigation of the property of each selection and also, by the following manner. a) the range of the property values ( $\log K^{\text{sel}}(\text{HCO}_3^-/\text{Cl}^-)$ ) of the training and the test sets should be covered from the lowest to the highest; b) each selected data point for the test set, should show the high distance linkage of dendrogram from the previously chosen one. The entire data set consisted of 40 ionophores divided into the training (32 ionophores) and test set (8 ionophores) considering the ratio of 80% and 20% of the data set, respectively.

## 2.4. Genetic algorithms

Using a genetic algorithm (GA), the most relevant descriptors were selected with respect to the objective function [28,29]. In this project, a genetic algorithm method was written in Matlab 6.5 program [30] and used as a selection tool. The details of genetic algorithms can be found in our previous works [31,32]. Here, the cross-validation correlation coefficient of leave-one-out ( $Q^2_{\text{LOO}}$  derived using MLR), was the fitness function for the genetic algorithm [33].

## 3. RESULTS AND DISCUSSION

### 3.1. Variable Selection

Suitable descriptors were chosen by the genetic algorithm. Only three descriptors (Mor06m, BEHp1, B07[O-O]) were selected based on the genetic algorithm. Using these descriptors, multiple linear regression analysis was applied to the training data and evaluated the results by the test data. It is necessary to construct a correlation matrix involving the correlation coefficients between selected descriptors to ensure the independent behavior of the implemented descriptors. The low correlation coefficients between each pair of the descriptors confirm that these variables behave independently in the models. Also, the multi-collinearity of descriptors were determined using the variation of inflation factors (VIF) [34]. In this case, models constructed with numerical values of VIF between 1 and 5 are considered acceptable and predictive. There is no inter-correlation if it equals 1. If the VIF value exceeds 10.0, there is an unacceptable and unstable model. The correlation coefficients and VIF values of selected descriptors based genetic algorithm are shown in Table 2. Here, the highest numerical correlation coefficient between each pair of descriptors is 0.30. Also, as shown in Table 2, the selected descriptors have VIF values below 2 which confirms the adequate predictiveness of the suggested models based on these descriptors.



**Table 2.** The correlation coefficient of selected descriptors and corresponding VIF values based on GA-MLR

	Mor06m	BEHp1	B07[O-O]	VIF
Mor06m	1	0	0	1.17
BEHp1	-0.30	1	0	1.09
B07[O-O]	-0.24	-0.06	1	1.14

### 3.2. Linear model construction

Based on the genetic algorithm-multiple linear regression analysis, a predictive QSPR model was developed with three descriptors as the followings:

$$\log K(\text{HCO}_3^-/\text{Cl}^-) = -43.402 - 0.275 (\text{Mor06m}) + 10.638(\text{BEHp1}) - 2.087(\text{B07[O-O]}) \quad (1)$$

$$N_{\text{train}}=32, R^2_{\text{train}}=0.916, R^2_{\text{test}}=0.903, R^2_{\text{adj}}=0.907, F_{\text{train}}=102.046, F_{\text{test}}=12.306, Q^2_{\text{Loo}}=0.896, Q^2_{\text{LGO}}=0.821$$

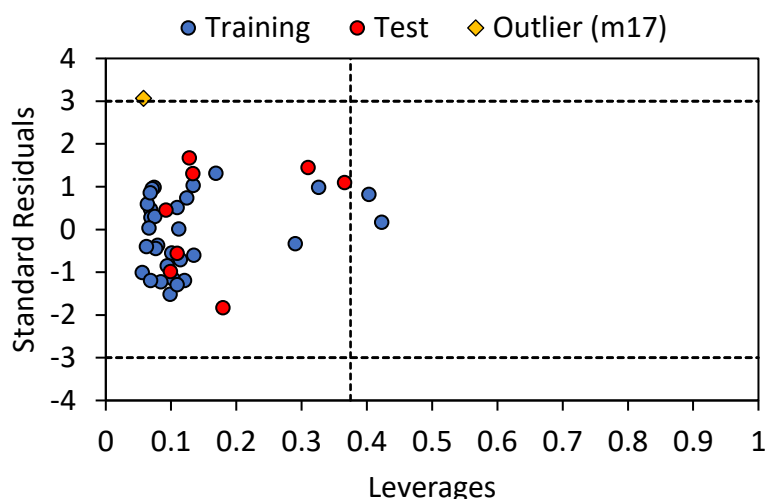
N represents the number of molecules in the training set, and  $Q^2_{\text{Loo}}$  and  $Q^2_{\text{LGO}}$  are cross-validation coefficients for leaving one out and leaving a group out (usually, 20% are excluded), respectively. Based on the value for  $Q^2_{\text{Loo}}$  (0.896), the built model demonstrates remarkable reliability. A squared correlation coefficient, an adjusted correlation coefficient, and a Fisher F statistic are known as  $R^2_{\text{adj}}$ ,  $R^2$  and F.

### 3.3. Outlier detection

To visualize the applicability domain and assess any possible outliers within the data set, the William plot was used. Figure 2, shows the Williams plot. The warning leverage ( $h^*$ ) is defined by the following formula:

$$h^* = 3p/n \quad (2)$$

Here, the number of calibration compounds is n, and the number of model variables plus one is p. The leverage (h) greater than the warning leverage ( $h^*$ ) indicates that the compound is highly influential. Furthermore, standardized residuals of 3 are commonly used as a cut-off value for accepting predictions, since they cover about 99% of all normally distributed data points. Based on the Williams plot (Figure 2), two compounds (m30 and m32) have leverage (h) greater than the warning  $h^*$  value of 0.375. As a result, they should be considered structural outliers. However, these compounds have the low standardized residual value and can be retained in the model. Also, compound m17, showed the large standardized residual (more than 3), and so, only compound m17, was detected as outlier compound. In the next step, this compound should be removed from the data set, and the new model should be generated.



**Figure 2.** The Williams plot of GA-MLR model for the training and test sets

The new GA-MLR model after removing of outlier compound (m17) was developed and the linear equation was as follow:

$$\log K(\text{HCO}_3^-/\text{Cl}^-) = -42.570 - 0.285 (\text{Mor06m}) + 10.473(\text{BEHp1}) - 2.275(\text{B07}[\text{O-O}]) \quad (3)$$

$$N_{\text{train}}=31, R^2_{\text{train}}=0.944, R^2_{\text{test}}=0.907, R^2_{\text{adj}}=0.937, F_{\text{train}}=150.444, F_{\text{test}}=12.914, Q^2_{\text{LOO}}=0.926, Q^2_{\text{LGO}}=0.888$$

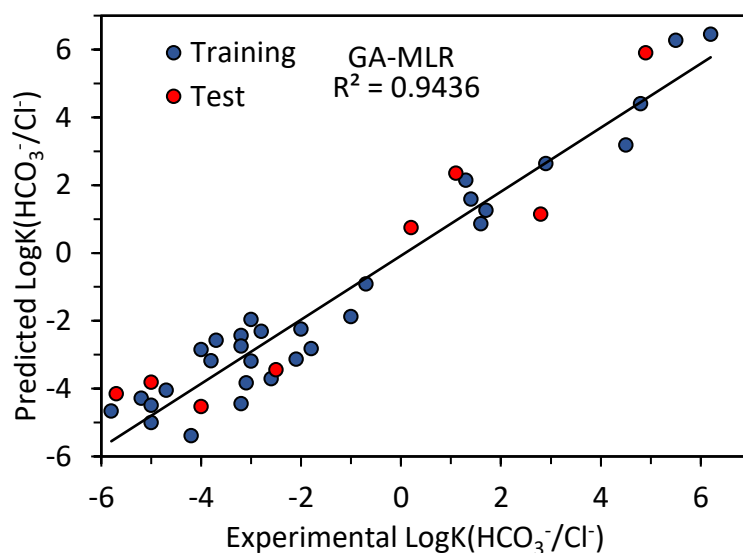
Compared to calculated  $R^2$  values for both sets, the built model after removing of the outlier compound could improve the results. Table 3 shows the statistical parameters of the GA-MLR model. Lower root means square error values ( $\text{RMSE}_{\text{train}} = 0.798$  and  $\text{RMSE}_{\text{test}} = 1.153$ ) and higher  $R^2$  and  $F$  values demonstrate the model's predictive capability. Table 1 shows the predicted carbonate potentiometric selectivity of plasticized polymeric membrane sensors for whole molecules based on GA-MLR model. The predicted values of  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  for the compounds in the training and test sets using equation 3 were plotted against the experimental values in Figure 3.

**Table 3.** Statistical results of different QSPR models

	Method	Training/Calibration			Test/Validation		
		$R^2$	RMSE	F	$R^2$	RMSE	F
This work	GA-MLR	0.944	0.798	150.444	0.907	1.153	12.914
This work	GA-SVM	0.986	0.398	617.692	0.962	0.780	26.707
Reference [16]	PLS	0.93	0.83	-	0.82	1.40	-

In order to evaluate the robustness of the constructed model, a Y-randomization test was conducted. Using this approach,  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  values are shuffled and then a new model is

developed based on the randomized data. It is imperative that the new built models have lower  $R^2$  and  $Q^2_{\text{LOO}}$  values to validate the main derived model's efficiency. This means that the goodness of the built model cannot be attributed to chance, as shown in Table 4 by the values of  $R^2$  and  $Q^2_{\text{LOO}}$  (the values are less than 0.2).



**Figure 3.** The plot of predicted vs. experimental selectivity by GA-MLR

**Table 4.** The  $Q^2_{\text{LOO}}$  and  $R^2$  training values after several Y-randomization tests

No	$Q^2$	$R^2$
1	0.009	0.135
2	0.032	0.174
3	0.008	0.119
4	0.181	0.032
5	0.017	0.150
6	0.108	0.041
7	0.008	0.114
8	0.008	0.082
9	0.009	0.092
10	0.154	0.035

### 3.4. Genetic algorithm-support vector machine

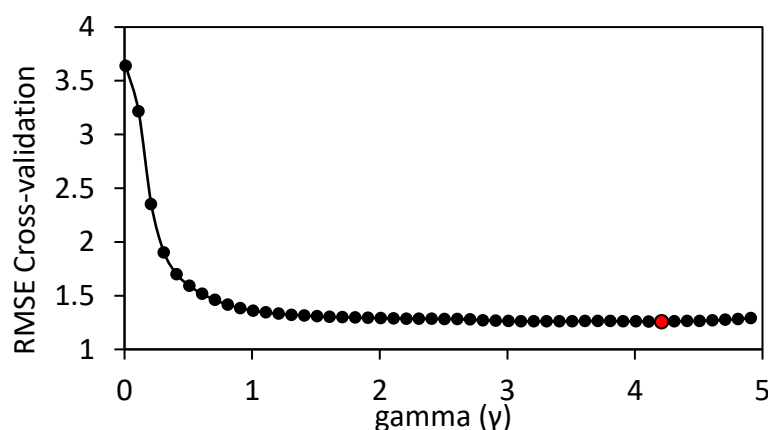
In addition to the GA-MLR model developed as a linear model, the SVM method was used to develop the nonlinear model based on the same selected descriptors and then compared its performance to the GA-MLR method. The details of support vector machine can be found in

our previous works [35-37]. In SVM regression, different factors are taken into account, including capacity parameter, kernel function type,  $\varepsilon$ -insensitive loss function and also its corresponding parameters [38].

First of all, the kernel function type should be decided, which determines the sample distribution in the mapping space. The radial basis function (RBF) is commonly used in many studies because of its good general performance and few parameters to be adjusted. The RBF is given below:

$$\exp(-\gamma * |u-v|^2) \quad (4)$$

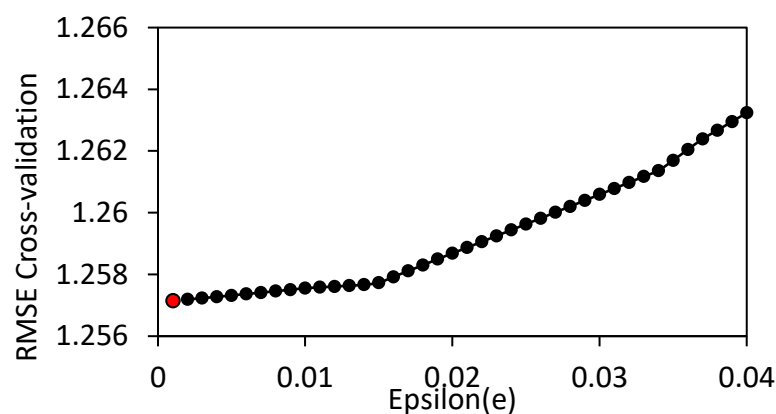
According to this formula,  $u$  and  $v$  are independent variables, and  $\gamma$  is a kernel parameter.  $\gamma$  regulates the RBF function and is directly responsible for the SVM performance and training time. To optimize the  $\gamma$  parameter, the leave one out cross-validation was used on the original training set to perform a grid search. It was checked from 0.01 to 5 with incremental steps of 0.01. RMSEs of cross-validation were also determined. A plot of gamma ( $\gamma$ ) parameter values against cross-validation RMSE is presented in Figure 4 which shows that 4.21 is the optimal value for gamma ( $\gamma$ ) parameter.



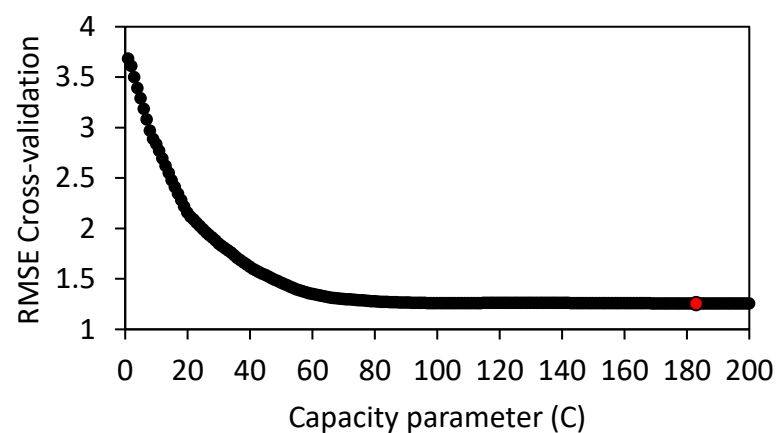
**Figure 4.** The gamma( $\gamma$ ) vs. RMSE for the training set

As a result of  $\varepsilon$ -insensitive parameter, the entire training set cannot meet boundary conditions, hence sparsity is allowed in the dual formulation's solution. The optimal value for this parameter depends on the type of noise present in the data. For the different values of  $\varepsilon$ , the RMSE of cross-validation varied from 0.001 to 0.04 in increments of 0.001. Figure 5 shows the values of  $\varepsilon$ -insensitive as a function of the obtained RMSE of cross-validation, and the optimal value for this parameter is 0.001.

SVM modelling concludes with parameter  $C$ , which controls the trade-off between maximizing margins and minimizing training errors. An optimal value for parameter  $C$  was found by incrementally increasing it by 1 from 1 to 200, as shown in Figure 6. Based on the results of the analysis in Figure 6, it is evident that 183 is the optimal capacity parameter.

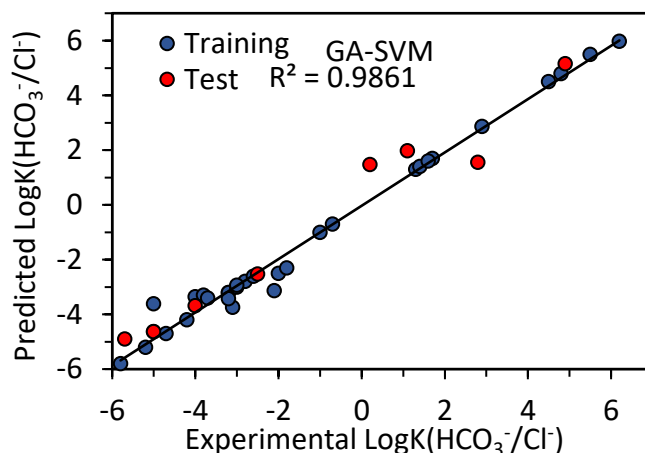


**Figure 5.** The epsilon ( $\varepsilon$ ) vs. RMSE for the training set



**Figure 6.** The capacity parameter(C) vs. RMSE for the training set

Table 1 and Figure 7 present the results for predicting the  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  value using the GA-SVM. In the study above,  $C = 183$ ,  $\varepsilon = 0.0001$ ,  $\gamma = 4.21$  were found to be the optimal value for developing an SVM model. Statistical results of the optimal model for training ( $R^2 = 0.986$ ,  $F = 617.692$ ,  $\text{RMSE} = 0.398$ ) and test ( $R^2 = 0.962$ ,  $F = 26.707$ ,  $\text{RMSE} = 0.780$ ) sets indicate the appropriate predictive ability of built model. In comparison with GA-MLR, both of the training set and test set compounds performed better on prediction (Table 3). The GA-SVM shows superiority over GA-MLR because of lower RMSE, and higher  $F$  and  $R^2$ . Also the results are compared with the previous reported work [16]. In the reported work, Vladimirova et al. predicted the sensitivity of the same data set by partial least squares (PLS) method and substructural molecular fragments were used as molecular descriptors. The comparison of our work and their results are shown in Table 3. As can be seen, in all the cases, the GA-MLR and GA-SVM were superior than the previously reported PLS method and GA-SVM provided high accurate prediction model for carbonate potentiometric selectivity.



**Figure 7.** The plot of predicted vs. experimental selectivity by GA-SVM

### 3.5. Interpreting of molecular descriptors

By interpreting the descriptors contained in the model, it is possible to gain some insights into factors which are related to the potentiometric selectivity of the studied carbonate ionophores. The first selected descriptor is Mor06m which implies 3D-Morse descriptor weighted by atomic masses. Based on the distance distribution in the geometrical display of the molecules, the 3D-MORSE descriptors contribute to the formation of the radial distribution function code and are processed in light of the sum of the atomic weights when using divergent angular scattering [39]. According to equation 3, the Mor06m descriptor has a negative sign, indicating that the  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  value is inversely related to this descriptor.

The second descriptor is BEHp1 (highest eigenvalue n. 1 of Burden matrix / weighted by atomic polarizabilities) which is belong to the Burden eigenvalue descriptors. The atomic polarizabilities play main role in this descriptor. This descriptor has a positive sign, in equation 3, which indicates that the  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  value is directly related to this descriptor.

The third descriptor is B07[O-O]. This descriptor belongs to the 2D binary fingerprint descriptors. 2D binary fingerprint descriptors have the value one (1) if there is a relevant bond at a topological distance, and have the value zero (0) if there is not. This descriptor represents, the presence or absence of O-O at topological distance 07. According to equation 3, the B07[O-O] descriptor has a negative sign indicating that the  $\log K(\text{HCO}_3^-/\text{Cl}^-)$  value is inversely related to this descriptor.

## 4. CONCLUSION

In this study, an accurate QSPR model derived from GA-SVM analysis for predicting the carbonate potentiometric selectivity of plasticized polymeric membrane sensors. A simple model with three descriptors was obtained. The results obtained show that the GA-SVM model was able to establish a satisfactory relationship between the molecular descriptors and the

potentiometric selectivity of different carbonate ionophores. Good results with high statistical quality and low prediction errors were obtained. In comparison; the GA-SVM method predicted both of the training and test set compounds more accurately than the GA-MLR method and also the previously reported PLS method. The QSPR model developed in this study can provide a useful tool to predict the potentiometric selectivity of new carbonate ionophores.

### Declarations of interest

The authors declare no conflict of interest in this reported work.

### REFERENCES

- [1] N. Vladimirova, E. Puchkova, D. Dar'in, A. Turanov, V. Babain, and D. Kirsanov, *Membranes* 12 (2022) 953.
- [2] U. Schaller, E. Bakker, U.E. Spichiger, and E. Pretsch, *Anal. Chem.* 66 (1994) 391.
- [3] E. Pourbasheer, R. Aalizadeh, J.S. Ardabili, and M.R. Ganjali, *J. Mol. Liq.* 204 (2015) 162.
- [4] S. Riahi, E. Pourbasheer, M.R. Ganjali, P. Norouzi, and A.Z. Moghaddam, *J. Chin. Chem. Soc.* 55 (2008) 1086.
- [5] E. Martynko, V. Solov'ev, A. Varnek, A. Legin, and D. Kirsanov, *Electroanalysis* 32 (2020) 792.
- [6] E. Pourbasheer, S. Ahmadpour, R. Zare-Dorabei, and M. Nekoei, *Arab. J. Chem.* 10 (2017) 33.
- [7] E. Pourbasheer, and R. Aalizadeh, *Sar Qsar Environ. Res.* 27 (2016) 385.
- [8] E. Pourbasheer, S.S. Tabar, V.H. Masand, R. Aalizadeh, and M.R. Ganjali, *Sar Qsar Environ. Res.* 26 (2015) 461.
- [9] H. Rafiei, M. Khanzadeh, S. Mozaffari, M.H. Bostanifar, Z.M. Avval, R. Aalizadeh, and E. Pourbasheer, *Excli J.* 15 (2016) 38.
- [10] A. Beheshti, E. Pourbasheer, M. Nekoei, and S. Vahdani, *J. Saudi Chem. Soc.* 20 (2016) 282.
- [11] E. Pourbasheer, S. Vahdani, R. Aalizadeh, A. Banaei, and M.R. Ganjali, *J. Chem. Sci.* 127 (2015) 1243.
- [12] E. Pourbasheer, A. Banaei, R. Aalizadeh, M.R. Ganjali, P. Norouzi, J. Shadmanesh, and C. Methenitis, *J. Ind. Eng. Chem.* 21 (2015) 1058.
- [13] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, and P. Norouzi, *Fullerenes, Nanotubes, Carbon Nanostruct.* 23 (2015) 290.
- [14] H. Khajehsharifi, and E. Pourbasheer, *J. Chin. Chem. Soc.* 55 (2008) 163.
- [15] S. Riahi, M.R. Ganjali, P. Norouzi, and F. Jafari, *Sens. Actuators B* 132 (2008) 13.
- [16] N. Vladimirova, V. Polukeev, J. Ashina, V. Babain, A. Legin, and D. Kirsanov, *Chemosensors* 10 (2022) 43.
- [17] J.H. Shim, I.S. Jeong, M.H. Lee, H.P. Hong, J.H. On, K.S. Kim, H.S. Kim, B.H. Kim, G.S. Cha, and H. Nam, *Talanta* 63 (2004) 61.

- [18] S. Makarychev-Mikhailov, O. Goryacheva, J. Mortensen, A. Legin, S. Levitchev, and Y. Vlasov, *Electroanalysis* 15 (2003) 1291.
- [19] M. Maj-Żurawska, T. Sokalski, J. Ostaszewska, D. Paradowski, J. Mieczkowski, Z. Czarnocki, A. Lewenstam, and A. Hulanicki, *Talanta* 44 (1997) 1641.
- [20] Y.K. Hong, W.J. Yoon, H.J. Oh, Y.M. Jun, H.-J. Pyun, G.S. Cha, and H. Nam, *Electroanalysis* 9 (1997) 865.
- [21] T. Sokalski, D. Paradowski, J. Ostaszewska, M. Maj-Żurawska, J. Mieczkowski, A. Lewenstam, and A. Hulanicki, *Analyst* 121 (1996) 133.
- [22] M. Rothmaier, U. Schaller, W.E. Morf, and E. Pretsch, *Anal. Chim. Acta* 327 (1996) 17.
- [23] C. Behringer, B. Lehmann, J.-P. Haug, K. Seiler, W.E. Morf, K. Hartman, and W. Simon, *Anal. Chim. Acta* 233 (1990) 41.
- [24] HyperChem, Molecular modeling system, 7.03rd edn. Hypercube, Gainesville (2002).
- [25] R. Todeschini, and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000.
- [26] R. Todeschini, V. Consonni, A. Mauri, and M. Pavan, *DRAGON software for the calculation of molecular descriptors*, 5.3<sup>rd</sup> edn. Talete SRL, Milan (2005).
- [27] F. Zhou, F.D.I. Torre, and J.K. Hodgins, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 582.
- [28] C.L. Waller, and M.P. Bradley, *J. Chem. Inf. Comput. Sci.* 39 (1999) 345.
- [29] J. Aires-de-Sousa, M.C. Hemmer, and J. Gasteiger, *Anal. Chem.* 74 (2002) 80.
- [30] T. MathWorks, *Genetic algorithm and direct search toolbox user's guide*, The Mathworks Inc. USA (2005).
- [31] Z.M. Avval, E. Pourbasheer, M.R. Ganjali, and P. Norouzi, *J. Serb. Chem. Soc.* 80 (2015) 187.
- [32] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, P. Norouzi, and J. Shadmanesh, *J. Saudi Chem.Soc.* 18 (2014) 681.
- [33] R. Leardi, R. Boggia, and M. Terrile, *J. chemom.* 6 (1992) 267.
- [34] V. Agrawal, and P. Khadikar, *Bioorg. Med. Chem.* 9 (2001) 3035.
- [35] Z. Rostami, and E. Pourbasheer, *Eurasian Chem. Commun.* 1 (2019) 79.
- [36] E. Pourbasheer, R. Aalizadeh, and M.R. Ganjali, *Arab. J. Chem.* 12 (2019) 2141.
- [37] E. Pourbasheer, R. Aalizadeh, M.R. Ganjali, and P. Norouzi, *Struct. Chem.* 25 (2014) 355.
- [38] V. Vapnik, and V. Vapnik, *Statistical learning theory*, Wiley New York (1998).
- [39] S. Gosav, M. Praisler, and D. Dorohoi, *J. Mol. Struct.* 834 (2007) 188.