Indian Journal of Chemical Technology
Vol. 27, January, 2020 pp. 43-50

Support vector regression: A novel soft computing technique for predicting the removal of cadmium from wastewater

Nusrat Parveen[1], Sadaf Zaidi*[, 2] & Mohammad Danish[3]

Department of Chemical Engineering, Aligarh Muslim University, India

E-mail: sadaf63in@yahoo.com

Received 25 August 2017; accepted 13 Novermber 2018

The presence of toxic heavy metals in the wastewater coming from industries is of great concern across the world. In the present work, a novel soft computing technique support vector regression (SVR)technique has been used to predict the removal of cadmium ions from wastewater with agricultural waste 'rice polish' as a low-cost adsorbent, with contact time, initial adsorbate concentration, *p*H of the medium, and temperature as the independent parameters. The developed SVR-based model has been compared with the widely used multiple regression (MR) model based on the statistical parameters such as coefficient of determination (R^2), average relative error (AARE) etc. The prediction performance of SVR-based model has been found to be more accurate and generalized in comparison to MR model with low AARE values of 0.67% and high R^2 values of 0.9997 while MR model gives an AARE value of 29.27% and 0.2161 as coefficient of determination (R^2). Furthermore, it has also been observed that the SVR model effectively predicts the behavior of the complex interaction process of cadmium ions removal from waste water under various experimental conditions.

Keywords: Heavy metals, Low cost adsorbent, Support vector regression (SVR), Coefficient of determination (R^2), Average relative error (AARE)

Cadmium being the toxic heavy metal is of great concern across the world as its long-term exposer in the environment can cause kidney damage, high blood pressure, renal disorder, bone degradation, anemia and destruction of red blood cells. The major source of cadmium discharge in the environment include mining industries, metal plating, pigments, phosphate fertilizer, cadmium nickel batteries, stabilizers and alloys[1,2]. Various conventional technologies have been used in the past for the removal of heavy metals like membrane process, filtration, ion exchange, precipitation, solvent extraction etc[3-5]. Adsorption process has offered an ideal alternative to these technologies, being economical and efficient treatment method for wastewater treatment[6,7]. However, its performance is mostly affected by the type of adsorbent used. A lot of research is going on natural materials to be used as adsorbent such as saw dust, rice husk, fly ash, tea waste etc., since these are environmental friendly, low cost, biodegradable and having high metal recovery[8,9].

There are many factors that influence the heavy metal removal in an adsorption process such as initial concentration of adsorbate, operating temperature, pressure, initial *p*H, contact time, and other coexisting substances and adsorbent structure[10]. Several techniques like multiple regression (MR), artificial neural network (ANN), genetic algorithm (GA) etc. are being used to predict the heavy metal removal efficiency[11-13]. Recently, support vector machine (SVM), an artificial intelligence (AI) technique have been developed as a support vector classification (SVC) and support vector regression (SVR). The support vector machines (SVMs)based on the superior structural risk minimization (SRM) principle offer a lot of advantages over the traditional techniques ANN and MR such as unique, global and optimal solution[14,15]. The most commonly used multiple regression (MR) technique suffers from multiple local solutions, over-fit to data and often leads to poor generalizability. This means good performance for training dataset and poor performance for unseen test dataset. Thus, in the present work, the issue has been address to explore the use of SVR to predict and analyze the adsorption capacity of Cd (II) ions. The applicability of SVR-based models in the field of chemical engineering has been well demonstrated[16-21].

In the present study, SVR-based model has been developed to predict the adsorption capacity of Cd(II). The experimental data from literature was first preprocessed. Using this data, the objectives were twin-fold: to develop a unified MR model from the statistically sound experimental data followed by the formulation of a unified SVR-based model from the same data. Furthermore, the performance of SVR model has also been comprehensively studied for predicting the behavior of the complex interaction process of cadmium ions removal from waste water under various experimental conditions.

Experimental Section

Abridged theory of SVR-based modeling

A detailed description of the theory of SVM can be referred to in several excellent works[22-26]. In a typical regression problem, a training dataset is given as:

$$\{(x_1, y_1) \dots \dots (x_i, y_i)\} \subset X \times \Re \quad \dots (1)$$

where, X denotes the space of the input patterns.

The objective of a ε-SVR model is to fit a regression function, such that it accurately predicts the outputs y_i corresponding to a new set of input samples, x_i. The SVR algorithm attempts to place a tube around the data as shown in Fig. 1.

In SVR, the objective is to minimize empirical error as well as the model should be as flat as possible (means a small value for weight vector w). The regression function in the feature space is approximated as:

$$f(x) = (w \bullet \phi(x) + b) \quad \dots (2)$$

where, the function of x termed feature and $w \bullet \phi(x)$ is the dot product in the high dimensional feature space, F. Here, SVR first maps the input data into this feature space using a nonlinear mapping, and subsequently regression is carried out linearly.

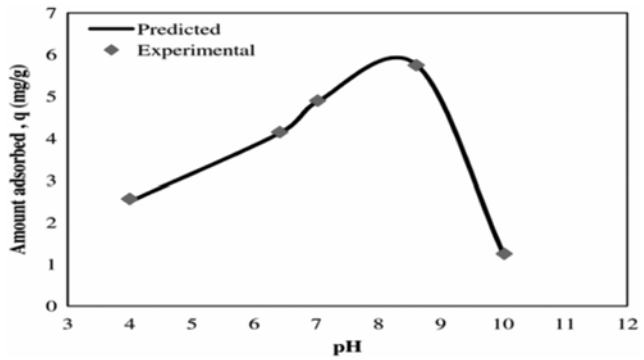

Fig. 1 — The illustration of SVM for regression problem using ε-insensitive loss function

The coefficients are estimated by minimizing the regularized risk function given as:

$$R = C\frac{1}{N}\sum_{i=1}^{N} L_\epsilon\left(y_i, f(x_i)\right) + \frac{1}{2}||w^2|| \quad \dots (3)$$

where, the first part of this equation is the empirical risk or error and is measured by a loss function called ε-insensitive loss function proposed by Vapnik[27] as:

$$L_\epsilon\big(y, f(x)\big) = \begin{cases} 0 \; if |y - f(x) \le \varepsilon \\ |y - f(x)| - \varepsilon otherwise \end{cases} \quad \dots (4)$$

The second part of Eq. (3) is used as a measure of function flatness (or less complexity of function fitness). C is a regularized constant determining the trade-off between the training error and model flatness. After introducing slack variables, the SVR formulation can be expressed mathematically in the form of a convex optimization problem as:

$$\text{Minimize } \frac{1}{2}\left|\left|w^2\right|\right| + C\sum_{i=1}^{N}\xi_i + \xi_i^* \quad \dots (5)$$

$$\text{Subject to } (y_i - w \bullet \emptyset(x) - b) \le \varepsilon + \xi_i$$

$$(w \bullet \emptyset(x) + b - y_i) \le \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \ge 0$$

The above-mentioned convex optimization problem (Eq. (5)) can be solved by transforming it into its dual form by introducing Lagrange multipliers and exploiting optimality constraints. The final decision function takes the following form:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N_{sv}}(\alpha_i - \alpha_i^*)\left(\emptyset(x_i) \bullet \emptyset\left(x_j\right)\right) + b \dots (6)$$

where, α and α^* are the introduced Lagrange multipliers. Only the non-zero coefficients, and the corresponding input vectors, $\mathbf{x}_i$, are called support vectors (SVs).

The problem of contradiction between high dimensions and computational complexity can be overcome by defining appropriate kernel functions in place of the dot product of the input vectors in high-dimensional feature space[28] as:

$$K\left(x_i, x_j\right) = \left(\emptyset(x_i) \bullet \emptyset\left(x_j\right)\right) \quad \dots (7)$$

Any function that satisfies Mercer's condition (kernel function must be symmetric, and it must be positive semi-definite) can be used as the kernel function[29]. The most commonly used kernel function

is the Gaussian radial basis function (RBF) defined below:

$$K(x_i, x_j) = \left(\emptyset(x_i) \bullet \emptyset(x_j)\right) = exp\left(-\frac{1}{2\sigma^2}||x_i - x_j||^2\right) exp(-\gamma||x_i - x_j||^2) \quad \ldots (8)$$

Thus, when using a kernel function all the necessary computations related to can be performed implicitly in the input space rather than in the high dimensional feature space. The basic SVR decision function modeling the data takes the following form:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{N_{sv}} (\alpha_i - \alpha_i^*) \left(K(x_i, x_j)\right) + b \quad \ldots (9)$$

The bias parameter, b, can be determined by applying Karush–Kuhn–Tucker (KKT) conditions which state that at the optimal solution the product between dual variables and constraints has to vanish.

SVR modeling procedure

The following steps have been carried out for the development of SVR-based model[30, 31]:

i. First, the whole dataset is organized into dependent and independent variables. Then it is normalized between 0 to 1 using the formula as given below:

$$X_{normalize} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

ii. After that using simple random sampling (SRS) technique, the whole dataset is then divided into training dataset (80% dataset) and test dataset (20% dataset) respectively.

iii. Among various kernels such as linear, polynomial, sigmoid, radial basis function (RBF) etc., the appropriate kernel function is chosen.

iv. 10-fold cross validation technique is usedto optimize the model parameters: C, ε and the kernel parameter (Υ for RBF kernel).

v. Using the optimum values of model parameters, training and testing of SVR simulation is done.

vi. The developed SVR model performance is then evaluated based on the statistical parameters like coefficient of determination (R^2), average absolute relative error (AARE), standard deviation (SD), root mean square error (RMSE) etc.

Evaluation parameters

The statistical analysis of the developed models prediction is based on the following performance criteria[18, 19]:

1. The average absolute relative error (AARE) of the developed models on test dataset should be minimum:

$$AARE = \frac{1}{N}\left[\sum_{i=1}^{N}\left|\frac{(y_{pred,i} - y_{exp,i})}{y_{exp,i}}\right|\right]$$

2. The cross-correlation co-efficient (R) of the models between input and output (target) should be around unity:

$$R = \frac{\sum_{i=1}^{N}(y_{exp,i} - \overline{y_{exp}})(y_{pred,i} - \overline{y_{pred}})}{\sqrt{\sum_{i=1}^{N}(y_{exp,i} - \overline{y_{exp}})^2 \sum_{i=1}^{N}(y_{pred,i} - \overline{y_{pred}})^2}}$$

3. The root mean square (RMSE) used for assessing the model should be minimum:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{pred,i} - y_{exp,i})^2}{N}}$$

4. The standard deviation (SD) on test data (unknown sample) should be minimum:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}\left((y_{pred,i} - y_{exp,i}) \times 100\right)^2}{N-1}}$$

5. The mean relative error (MRE) of the models should be as minimum as possible:

$$MRE = \frac{1}{N}\sum_{i=1}^{N}\left|\left[\frac{y_{exp,i} - y_{pred,i}}{y_{pred,i}}\right]\right|$$

6. Leave-one-out cross validationon the training dataset (Q_{LOO}^2)was used to evaluate the internal predictive capability of the models:

$$Q_{LOO}^2 = 1 - \frac{\sum_{i=1}^{n_{training}}(y_{exp,i} - y_{pred,i})^2}{\sum_{i=1}^{n_{training}}(y_{exp,i} - y_{exp,mean})^2}$$

7. Leave-one-out cross validation on the test dataset (Q_{ext}^2)was used to evaluate the external predictive capability of the models:

$$Q_{ext}^2 = 1 - \frac{\sum_{i=1}^{n_{test}}(y_{exp,i} - y_{pred,i})^2}{\sum_{i=1}^{n_{test}}(y_{exp,i} - y_{exp,mean})^2}$$

Results and Discussion

In this study, two models SVR and MR model have been developed using the data available from the published literature[32] to predict the removal of cadmium ions from wastewater with agricultural waste 'rice polish' as a low cost adsorbent, based on the contact time (min), initial adsorbate concentration (mg/L), *p*H of the medium, and temperature (°C). LIBSVM package[33] on MATLAB platform has been used to develop the SVR-based models and a comparison of the developed SVR-based model has

been made with the widely used classical technique, multiple regression.

Multiple regression (MR)-based model

The whole dataset of 95 samples is being divided into training and test dataset using the simple random sampling (SRS) technique, as 80% (76 sample) and 20% (19 sample), respectively. Thereafter, training dataset is used to develop a MR model for the predicting the adsorption capacity of Cd(II), q that depends on temperature (x_1), pH (x_2), initial adsorbate concentration (x_3), and the contact time (x_4).

The following MR model equation has been obtained:

$$q = 0.0155 \times x_1^{-0.1567} \times x_2^{1.1201} \times x_3^{0.6988} \times x_4^{0.1379} \quad \ldots (10)$$

Prediction capability of the MR model has been compared with the experimental values as shown in the Fig. 2. This shows that the most of the predicted data points for both training as well as test dataset are away from the ideal fit line. Furthermore, the performance of MR model has also been evaluated based on statistical parameters given in Table 1.

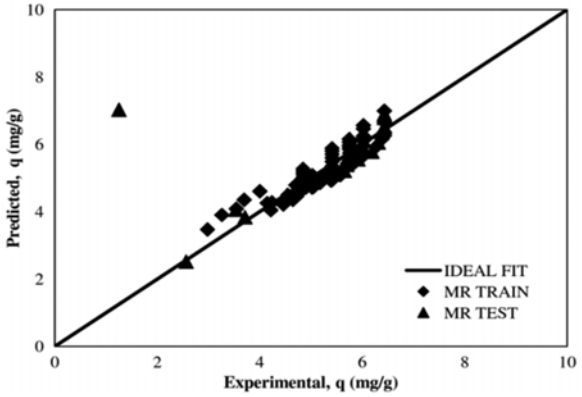

Fig. 2 — Performance of MR model using training data and test data for prediction of cadmium removal

Table 1 — Model evaluation parameters for MR-based model using training and test data set

MR Model evaluation parameter	Train data	Test data
AARE (%)	5.3013	29.2753
R^2	0.8258	0.2161
RMSE	0.0622	0.1973
SD	0.7377	1.4512
MRE	0.0520	0.0929
Q^2LOO (Train data),Q^2ext(Test data)	0.8213	0.0753

Average absolute relative error (AARE) values for the training data set is only 5.3% whereas a high AARE value of 29.27% for the test dataset is obtained. Similarly, other statistical parameters have also much greater discrepancy between training and the test dataset. Since, MR and other classical regression techniques are based on empirical risk minimization (ERM) principle which minimizes only empirical error or training error and do not consider the capacity of the machine. This result in overtraining i.e. high accuracy for the training dataset and low for the test dataset. It can be deduced from this table that the performance of MR model is poor, particularly for the test dataset (unknown sample).

Analysis of the SVR-based model

In the present study, the RBF kernel has been used because of its good generalization performance and only few parameters are required to be adjusted[18]. The optimal values of the SVR model parameters (C, ε and γ) has been obtained using grid search methodology with 10-fold cross validation with a wide range of these parameters: C [2^5, 2^{15}], γ [2^{-15}, 2^2] and ε [2^{-15}, 2^4]. Table 2 gives the obtained optimal values of the SVR model parameters.

After optimization of the SVR parameters, the training and test course have been plotted in Fig. 3 and 4 respectively. It is observed that the predicted values of the adsorption capacity, q for Cd(II) are very close to the experimental values of the

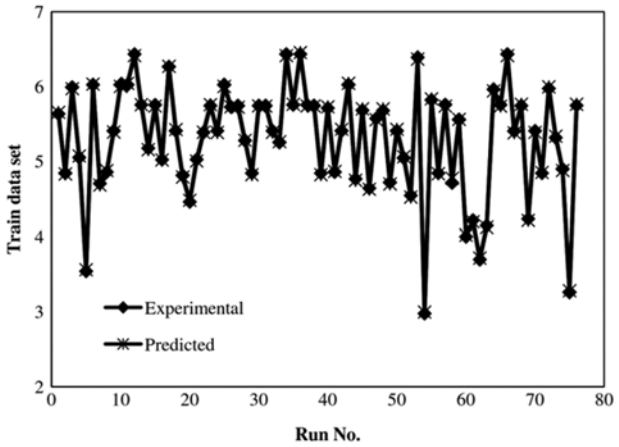

Fig. 3 — Training course curve for predicting the adsorption capacity in cadmium removal

Table 2 — Optimal parameters of SVR-based model for the adsorption capacity of Cd(II)

Model	C	$\gamma = 1/2\sigma^2$	ε	Kernel type	Type of Loss function	Number of support vectors	Number of training points
Adsorption capacity (q)	32768	0.5	0.028	RBF	ε - insensitive	42	76

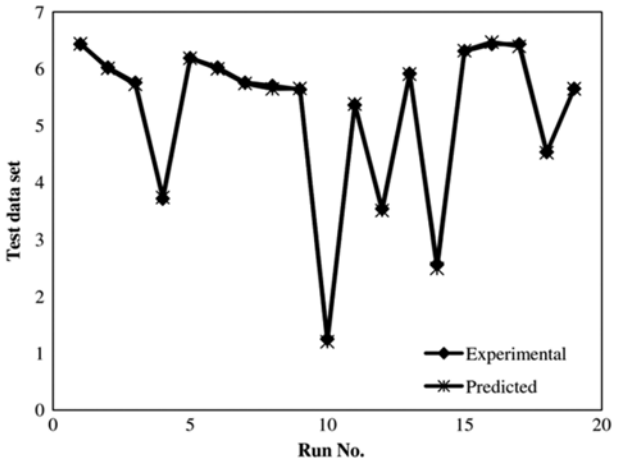

Fig. 4 — Test course curve for predicting the adsorption capacity in cadmium removal

Table 3 — Model evaluation parameters for SVR-based model using the training and test data

SVR Model evaluation parameter	Train data	Test data
AARE (%)	0.3366	0.6784
R^2	0.9993	0.9997
RMSE	0.0038	0.0116
SD	0.7353	1.4634
MRE	0.0034	0.0070
Q^2LOO (Train data),Q^2ext(Test data)	0.9993	0.9995

adsorption capacity, q both for training and test dataset. Table 3 gives the model evaluation parameters of the SVR-based model for both the training dataset as well as the test dataset and the obtained values of statistical parameters shows a close proximity between the training and the test dataset. It is found in Fig. 5 that both the predicted values for the training data as well as the test data lie close to the ideal fit line.Thus, it can be said that SVR-based mode gives the excellent prediction performance.

Comparison of SVR-based model with the MR model

The prediction performance of the developed SVR-based model has been compared with MR model using the test dataset. Table 4 shows that average absolute relative error values (AARE) for the SVR-based model is only 0.68% while for MR model a very high value of AARE of 29.28% is obtained. Similarly, other statistical parameters values for the case of SVR model are much improved and acceptable. Figure 6 has been plotted between experimental values of adsorption capacity of Cd(II) versus predicted values of adsorption capacity of Cd(II) using the test dataset. Fig. 6 clearly shows that the predicted values lies on the ideal fit line whereas most the data points in case of MR model lies away

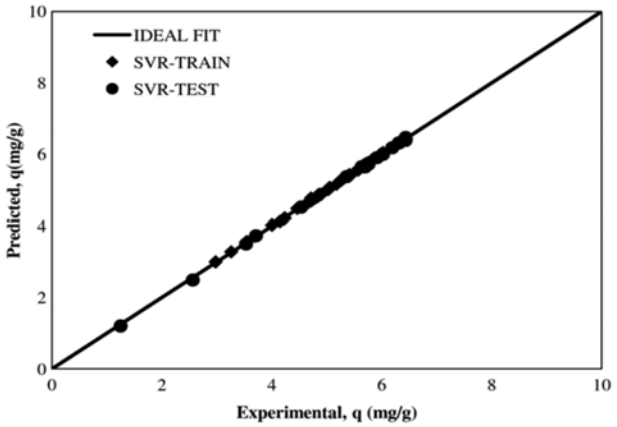

Fig. 5 — SVR simulation for adsorption capacity (q) in cadmium removal process with optimal parameters using training data and test data.

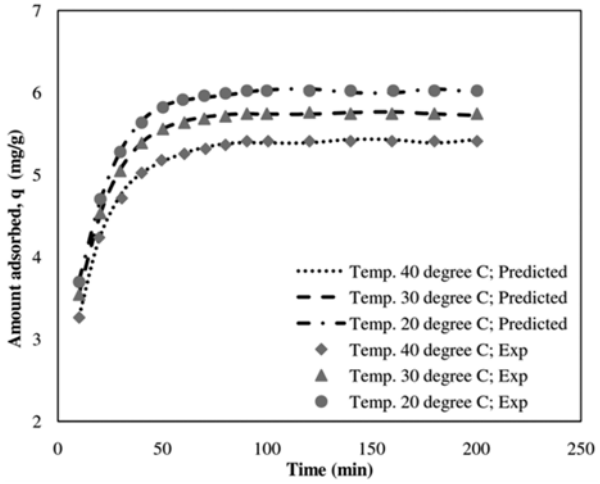

Fig. 6 — Comparison of SVR model with MR model for predicting the adsorption capacity (q) in cadmium removal process using test data

Table 4 — Comparison of SVR model with MR model using test dataset

Model evaluation parameter	SVR-based model	MR model
AARE (%)	0.6784	29.2753
R^2	0.9997	0.2161
RMSE	0.0116	0.1973
SD	1.4634	1.4512
MRE	0.0070	0.0929

from the ideal fit line. From Table 4 and Fig. 6 it can be deduce that SVR-based model has the superior prediction performance with high accuracy and high generalizability. The reason may be attributed to the fact that the conventional regression technique, multiple regression (MR) is based on the empirical risk minimization principle (ERM) which minimizes only the empirical error or training error and does not count the capacity of the machines whereas SVR is

based on structural risk minimization principle (SRM). By SRM principle, the generalization accuracy is optimized over the empirical error and the flatness of the regression function or the capacity of SVM.

Table 5 illustrates the distribution of predicted data points of Cd(II) adsorption capacity by MR and SVR-based model in terms of absolute deviation (AD) for training dataset. It is observed that the MR-based model predicts nearly 2.6% data points within an absolute deviation of less than 0.5 % and a total of 3.9% data points within an absolute deviation of less than 1.0%. And 96.05% data points have an AD of more than 1.0%. Furthermore, this table also shows that the SVR-based model predicts nearly 90.79 % data points within an absolute deviation of less than 0.5 % and a total of 98.68 % data points within an absolute deviation of less than 1.0%. Only 1.3 % data points have an AD of more than 1.0%.

Table 6 shows the distribution of predicted data points of Cd(II) adsorption capacity by MR and SVR-based model in terms of absolute deviation for test dataset. It can be seen that no data points fall within an absolute deviation of less than 0.5 % and 5.26 % data points are within an absolute deviation of less than 1.0%. And most of the data points 94.74 % have an AD of more than 1.0%. While in this table, it is observed that the SVR based model predicts nearly 57.89% data points within an absolute deviation of less than 0.5% and all the data points fall within an absolute deviation of not more than 1.0%. It can be concluded from Table 5 and 6 that the MR model has high deviation between experimental and predicted data points of the adsorption capacity as compared to SVR-based model. Thus, SVR-based model is found to be highly accurate.

Performance of the SVR-based model in the light of existing theory and experimental findings

The effect of the various independent variables such as contact time (min), initial adsorbate concentration (mg/L), pH of the medium, and temperature (°C) on adsorption capacity of Cd(II) predicted by the SVR-based model isbeing discussed in the following section.

Effect of contact time and concentration on the adsorption capacity of Cd(II)

Experiments were performed by Singh *et al*.[32] with different initial adsorbate concentrations, viz., 100, 125 and 150 mg/L at constant temperature of 30°C and *p*H 8.6. The percentage removal of Cd(II) was found to be 96.95, 92.15 and 85.80, respectively. Figure 7 shows that the rate of adsorption increases till the equilibrium is achieved. It can also be

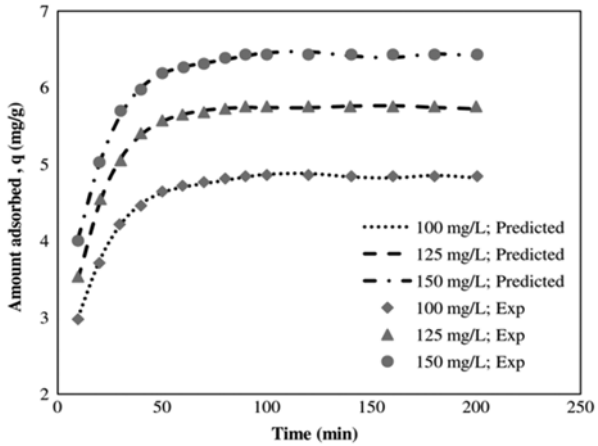

Fig. 7 — Effect of contact time and initial adsorbate concentration on adsorption capacity of Cd(II)-comparison between SVR predicted and experimental values of adsorption capacity of Cd(II) (temperature = 30 °C, pH = 8.6, agitating speed = 125 rpm)

Table 5 — Percentage distribution of predicted data points of the adsorption capacity of Cd(II) by MR model and SVR- based model in terms of absolute deviation (AD) for training data.

Absolute deviation (AD) (%)	% of SVR model predicted values	Cumulative score	% of MR model predicted values	Cumulative score
AD< 0.5	90.79	90.79	2.63	2.63
0.5<AD< 1.0	7.89	98.68	1.32	3.95
AD>1.0	1.32	100	96.05	100
Total	100		100	

Table 6 — Percentage distribution of predicted data points of the adsorption capacity of Cd(II) by MR model and SVR- based model in terms of absolute deviation (AD) for test data

Absolute deviation (AD) (%)	% of SVR model predicted values	Cumulative score	% of MR model predicted values	Cumulative score
AD< 0.5	57.89	57.89	0	0
0.5<AD< 1.0	31.58	89.47	5.26	5.26
AD>1.0	10.53	100	94.74	100
Total	100		100	

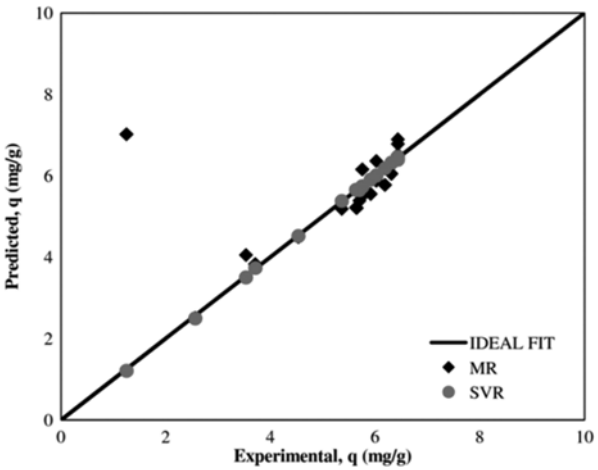

Fig. 8 — Effect of temperature on adsorption capacity of Cd(II) - comparison between SVR predicted and experimental values of adsorption capacity of Cd(II) as a function of time at different temperature (initial adsorbate conc. = 125mgL^{-1}, pH = 8.6, agitating speed = 125 rpm)

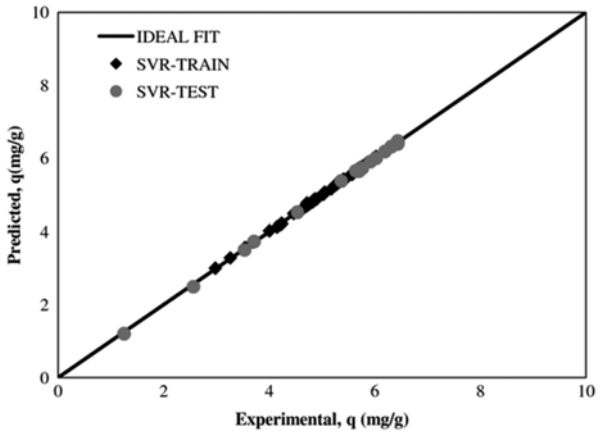

Fig. 9 — Effect of pH on adsorption capacity of Cd(II) - comparison between SVR predicted and experimental values of adsorption capacity of Cd(II) at different pH (initial adsorbate conc. = 125mgL^{-1}, temperature = 30°C, agitating speed = 125 rpm)

concluded that Cd(II) removal was highly dependent on the adsorbate concentration. Furthermore, it is also observed that SVR-based model successfully predicts the effect of contact time and concentration on the adsorption capacity of Cd(II). This is due to the fact that SVR-based model seeks to have low empirical or training error and the complexity of the model.

Effect of temperature on the adsorption capacity of Cd(II)

Experiments were also performed at different temperatures 20, 30 and 40°C keeping the adsorbate concentration of 125 mg/L and a *p*H of 8.6 to a constant value. Figure 8 shows that with the rise in temperature from 20 to 40°C, the percentage of adsorption decreases. The decrease in adsorption happens with the rise in temperature because of the weakening of adsorptive forces between the active sites of the adsorbents and adsorbate and also between the adjacent molecules of the adsorbed phase. Here also, predicted curve shows a close agreement with the experimental data points. Thus, it can be deduced that the SVR model based on superior SRM principle correctly predicts the effect of temperature on the adsorption capacity of Cd(II).

Effect of pH on the adsorption capacity of Cd(II)

Experiments were performed at different *p*H (4.0, 6.4, 7.0, 8.6 and 10.0) to a constant condition of temperature 30°C and concentration 125 mg/L. The percentage removal Cd(II) increased from 40.90 to 92.15 as *p*H increases from 4.0 to 8.6 and thereafter percentage of Cd(II) removal decreased from 92.15 to 19.75 when *p*H increases from 8.6 to 10. The optimum value of *p*H for the removal of Cd(II) on rice polish was obtained 8.6. Figure 9 shows that the adsorption of cadmium was higher in alkaline *p*H. However, the removal was less in acidic range and reached a maximum around *p*H 8.6 and above *p*H 8.6 removal of cadmium started to decrease due to soluble hydroxyl complexes of cadmium hydroxides[32,34]. Figure 9 also shows the predicted values of the amount of Cd(II) adsorbed by SVR-based model. Since, SVR model is based on SRM principle which minimizes the empirical error as well as the complexity of the model. Therefore, the predicted values of the adsorption capacity of Cd(II) are close to experimental values and follows the same trend.

Conclusions

Support vector regression (SVR) has been successfully used for predicting the adsorption capacity of Cd(II). The developed SVR-based model is then compared with the MR model and the obtained results reveal that the SVR model fitted the data better than the commonly used MR model. The effects of various operational parameters (*p*H, initial adsorbate concentration, contact time and temperature) on the adsorption capacity of Cd(II) have been studied and it is observed that the SVR model effectively predicts the behaviour of the complex interaction process under various experimental conditions. SVR model based on structural risk minimization (SRM) principle minimizing the empirical error as well as the complexity of the model is found to be highly accurate and generalized. Good SVR prediction results can be helpful in more efficient design of the adsorption process for the removal of heavy metals.

Acknowledgement

The authors sincerely extend their thanks and gratitude to Messrs K.K. Singh, R. Rastogi, and S.H. Hasan for the availability of their published paper from which the adsorption data has been retrieved for carrying out the present study for model formulation and validation.

NOMENCLATURE

C cost function
$f(x)$ regression function
$K(x_i,x_j)$ kernel function
L Lagrangian function (dual form)
Q^2_{ext} Leave-one-out cross validation on the test set
Q^2_{LOO} Leave-one-out cross validation on the training set
x_i Input vector
y_i Output vector

Greek symbols

σ Width parameter of RBF kernel
ε loss function
γ regularization parameter

α and α* Lagrange multipliers
$\varphi(x_i)$ mapping function to high dimensional feature space for input vector **x**

References

1 Al-Rashdi B, Somerfield C & Hilal N, *Sep Purif Rev*, 40 (2011) 209.
2 Balkaya N & Cesur H,*Chem Eng J*, 140 (2008) 247
3 Qdias H A & Moussa H, *Desalination*, 164 (2004)105
4 Charerntanyarak L,*Water Sci Technol*, 39 (1999) 13
5 Černá M, *Env Monit Assess*, 34 (1995)151
6 Demirbas A, Pehlivan E, Gode F, Altun T& Arslan G, *J Colloid Interface Sci*, 282 (2005) 20
7 Abas S N A, Ismail M H S, Kamal M L & Izhar S, *World Appl Sci J*, 28 (2013) 1518
8 Tripathi A M R R, *J Bioremed Biodeg*, 6 (2015) 1
9 Keng P S, Lee S L, Ha S T, Hung Y T & ST O, *Env Chem Lett*, 12 (2014) 15.
10 Zabochnicka-Świątek M & Krzywonos M, *Pol J Environ Stud*, 23 (2014) 551.
11 Brey T, Jarre-Teichmann A & Borlich O, *Mar Ecol Prog Ser*,140 (1996) 251.
12 Eyupoglu V, Eren B & Dogan E, *Int J Environ Res*, 4 (2010) 463.
13 Vasseghian Y, Ahmadi M, Dolati F & Heydari A, *J Chem Pet Eng*, 48 (2014) 69.
14 Akande K O, Owolabi T O, Twaha S & Olatunji S O, *Journal Comput Eng IOSR-JCE*, 16 (2014) 88.
15 Sheta A F, Ahmed S E M & Faris H, *Int J Adv Res Artif Intell*,4 (2015) 55.
16 Thissen U, Van Brakel R, De Weijer A P, Melssen W J, Buydens L M C, *Chemom Intell Lab Syst*, 69 (2003) 35.
17 Zaidi S, *Chem Eng Sci*, 69 (2012) 514.
18 Zaidi S, *Chem Eng Res Des*, 98 (2015) 44.
19 Peng H & Ling X, *Int J Heat Mass Transf*, 84 (2015) 203.
20 Parveen N, Zaidi S & Danish M, *New Polym Compos Mater Environ Biomed Actuator Fuel Cell Appl*, 5 (2016) 402.
21 Barletta M, Gisario A, Palagi L & Silvestri L, *Powder Technol*, 258 (2014) 85.
22 Cortes C & Vapnik V, *Mach Learn*, 20 (1995) 273.
23 Gunn S, *ISIS Tech Report*, Univ Southampton, Southampton, (1997) 1.
24 Mat Deris A, Mohd Zain A & Sallehuddin R, *Procedia Eng*, 24 (2011) 308.
25 Marti A, *IEEE Intell Syst*,13 (1998)18.
26 Nandi S, Badhe Y, Lonari J, Sridevi U, Rao B S, Tambe S S & Kulkarni B D, *Chem Eng J*, 97 (2004) 115.
27 Vapnik V, Golowich S E & Smola A, *Adv Neural Inform Process Syst*, 9 (1996) 281.
28 Pan Y, Jiang J, Wang R, Cao H & Cui Y, *J Hazard Mater*, 164 (2009) 1242.
29 Smola AJ & Schölkopf B, *Stat Comput*, 14 (2004)199.
30 Shi XZ, Zhou J, Wu BB, Huang D & Wei W, *Trans Nonferrous Met Soc China*, 22 (2012) 432.
31 Lee C & Chern S, *J Mar Sci Technol*, 21 (2013) 318.
32 Singh K K, Rastogi R & Hasan S H, *J Hazard Mater*, 121 (2005) 51.
33 Chang C C, Lin C J, *LIBSVM: A Library for Support Vector Machines*. (Taipei, Taiwan) 2001.
34 Hasan S H, Singh K K, Prakash O, Talat M & Ho YS, *J Hazard Mater*,152 (2008) 356.