

Modeling of Arsenic (III) Removal by Evolutionary Genetic Programming and Least Square Support Vector Machine Models

S. Mandal · S. S. Mahapatra · S. Adhikari · R. K. Patel

Received: 18 August 2014 / Accepted: 27 November 2014 / Published online: 18 December 2014
© Springer International Publishing Switzerland 2014

Abstract In this study, the co-precipitation method was used to synthesize the cerium oxide tetraethylenepentamine (CTEPA) hybrid material with the variation of the molar concentration of the metal oxide. Physicochemical techniques, like FESEM, XRD, FTIR, HR-TEM and TGA-DSC, were used to characterize the hybrid material. The adsorption experiment was carried out to estimate the optimum condition for adsorption with the variation of adsorbent dose, pH of the solution, time and initial concentration of the adsorbate. These variables are further used to develop an approach for the evaluation of As(III) removal from water by using evolutionary genetic programming techniques (GP) and least square support vector models (LS-SVM). GP model was found to be the best performing model for understanding the nonlinear behavior and prediction of As(III) removal (minimum standard error GP_{training} 0.411, GP_{testing} 0.658). The adsorption process in this study followed second order kinetics. The experimental data were best fitted to linearly transformed Langmuir isotherm with an R^2 (correlation coefficient) value of >0.99 and having a maximum adsorption capacity of 124.8 mg/g at 25 ± 2 °C. The type of adsorption process was derived from the Dubinin–Radushkevich isotherm model. A comparison between the model and actual experimental values gave a high correlation coefficient value ($R^2_{\text{training (GP)}}$ 0.988, $R^2_{\text{testing (GP)}}$ 0.977).

Keywords Arsenic · Genetic Programming · HRTEM · Hybrid Materials

Electronic supplementary material The online version of this article (doi:10.1007/s40710-014-0050-6) contains supplementary material, which is available to authorized users.

S. Mandal · R. K. Patel (✉)

Department of Chemistry, National Institute of Technology, Rourkela 769008 Odisha, India
e-mail: rkpatel@nitrkl.ac.in

S. Mandal

e-mail: sandipmandal9@gmail.com

S. S. Mahapatra

Department of Mechanical Engineering, National Institute of Technology, Rourkela, Odisha, India

S. Adhikari

Department of Ceramic Engineering, National Institute of Technology, Rourkela, Odisha, India

1 Introduction

The availability of fresh drinking water is a focus of major concern throughout the globe. The presence of arsenic and its compounds makes water unsuitable for drinking purposes. The contamination of water by arsenic is mainly due to natural or anthropogenic processes such as oxidation, reduction, biochemical methylation, precipitation, metal plating, glass industries, mining and fertilizers (Choong et al. 2007; Shevade and Ford 2004). Arsenic usually exists as arsenic (III) and arsenic (V) in natural water. As(III) species are more harmful than arsenic (V) because of lower stability and high mobility in the water environment (Wang and Mulligan 2006; Vaughan and Reed 2005; Asli et al. 2013). Research reports have shown that arsenic poisoning causes skin, lung and liver cancers besides other effects (Jain and Ali 2000). Due to its detrimental effects on human health, the World Health Organization (WHO) recommends a maximum permissible limit of 10 $\mu\text{g/L}$, beyond which water is undesirable for drinking (WHO 2008).

The removal of arsenic from water by a cost effective method is a challenging task that has gained much attention by researchers. A number of methods and materials have been used for removal of arsenic: among them adsorption is a promising and attractive process when the adsorbent is suitable and versatile (Ma et al. 2011; Chandra et al. 2010; Wan Ngah et al. 2011).

In order to study the effectiveness of the process, modeling and optimization of the removal process is viewed as an important aspect of the study. For example, in the past artificial neural network, response surface methodology, neuro-fuzzy and swarm particle optimization have been used widely in the prediction of contaminant removal from water systems with reasonable accuracy (Kotti et al. 2013; Dora et al. 2013; Aleboyeh et al. 2008; Akratos et al. 2008; Chu 2003; Saha et al. 2010; Elmolla et al. 2010; Khajeh and Modarress 2010); presently some researchers have tried to develop applications of genetic programming (GP) and support vector machine (SVM) tools in prediction and optimization of various aspects, such as prediction of carbon monoxide concentrations by using hybrid Partial Least Square Support Vector Machine (LS-SVM) models (Garg et al. 2014; Yeganeh et al. 2012), and photovoltaic plant output forecasting (Russo et al. 2014), vaporization enthalpy of petroleum fractions and pure hydrocarbon modeling (Parhizgar et al. 2013) and groundwater level prediction and simulation (Fallah-Mehdipour et al. 2013) by using GP techniques.

However, As(III) removal by cerium oxide tetraethylenepentamine and prediction of As(III) removal processes by GP and SVM techniques has not been tested yet. Hence, it is necessary to study it in more detail in order to understand the prediction behavior of both models. The main motivation behind this study was to investigate the possibility of cerium oxide tetraethylenepentamine as a low cost hybrid material for As(III) removal by using GP and LS-SVM techniques. Furthermore, a comparative evaluation was done between the two models based on statistical techniques, sensitive nature of variables and confirmatory experiments.

2 Experimental

2.1 Chemicals

All chemicals used were of AR grade and procured from Merck, Frankfurt, Germany or Sigma Aldrich, St. Louis, United States of America. Standard solutions were prepared according to procedures available in literature (Jain and Ali 2000). The glassware used was obtained from Borosil.

2.2 Synthesis and Characterization of Adsorbent

A series of hybrid materials were synthesized by following the procedure according to our previous publication (Mandal et al. 2011). The required pH 9 of the solutions was maintained by the addition of required amount of 1 M NH_4OH . Three different solutions, i.e., solution-A containing 0.5 M of adipic acid (100 mL), solution-B containing 1 M of cerium oxide (100 mL) and solution-C containing 2 M (100 mL) tetraethylenepentamine, were added drop wise to 500 mL round bottomed flask containing 100 mL of distilled water following the above sequence. The solutions were kept under stirring at a temperature of 50 °C maintaining inert atmosphere. After a predetermined time of 3 h, the white colored product so obtained was filtered. The product obtained was kept undisturbed in the oven at 70 °C for drying. Other hybrid materials were prepared accordingly by changing solution-B (magnesium oxide/zirconium oxychloride/ calcium chloride). The dried material was calcined at 200 °C for 2 h and used as adsorbent. The characterizations were carried out using various sophisticated analytical instruments and techniques as presented in Table S-1 [see Supplemental online material (S-1)].

2.3 Adsorption Experiments

The residual As(III) solution was measured in the solution after separation by filtration after taking 0.3 g of each hybrid material for 30 min in a series of conical flasks by taking 100 mg/L of As(III) in order to know the removal efficiency of each material. The adsorption experiments were carried out according to methods available in literature (Vatutsina et al. 2007) by taking 5 mg/L, 25 mg/L and 50 mg/L As(III) solution without any other ions. Table 1 presents the experimental parameter settings for As(III) removal by batch process.

In order to understand the behavior of competing ions affecting the removal efficiency, interaction of the following anions (CO_3^{2-} , SO_4^{2-} , HCO_3^{2-} , NO_3^{2-} , F^- , Cl^-) were also studied. The As(III) adsorption capacity and % removal were calculated using the following relations:

$$q_e = \left[\frac{(C_o - C_e)V}{m} \right] \quad (1)$$

$$\% \text{removal} = \left(\frac{C_o - C_e}{C_o} \right) \times 100 \quad (2)$$

Table 1 Experimental parameter settings for As(III) removal from water by batch process

Variables studied	Ranges	
	Minimum	Maximum
Dose (x_1)	0.1 g	1.0 g
pH (x_2)	1	12
Time (x_3)	10 min	90 min
Temperature (x_4)	20 °C	80 °C
Concentration (x_5)	5 mg/L	50 mg/L

where q_e is equilibrium adsorption capacity in mg/g, C_o and C_e are the initial and equilibrium concentrations of As(III) in solution (mg/L), m is the adsorbent mass (g) and V is the volume of solution taken.

The isoelectric point (pHzpc) is determined using electrophoretic mobility in a solution of ionic strength of 0.01 M NaNO₃ and 0.01 M NaCl according to procedure available in literature (Ma et al. 2011).

3 Computational Modeling Approach

3.1 Computational Models

3.1.1 Genetic Programming

Genetic programming (GPTIPS) is a biological inspired technology (multi gene) based on Darwinian Theory of evolution and on Herbert Spencer phrase “survival of the fittest” (Garg et al. 2014). In Genetic programming, the fittest population has to be developed from the best performing trees of different sizes randomly by performing tasks based on physical data of mutation and crossover process. The mathematical models generated are represented by the population size. The models are generated by combining the elements randomly from the functional and terminal sets. A model example is presented in Fig. S-1a. The performance of initial population is checked on the training input data based on the fitness function, generally known as root mean square error (RMSE, See Eq. 3).

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{\text{model},i} - Y_{\text{actual},i})^2} \quad (3)$$

where $Y_{\text{model},i}$ is the value predicted of i^{th} data sample by GP model, $Y_{\text{actual},i}$ is the actual value of the i^{th} sample and N is the number of training samples.

Based on the above fact, the genetic algorithm selects model for next genetic operation known as mutation and crossover. The need of genetic operation is to form the best performing genetic population that represent a new generation. The phenomenon of generating population through mutation and crossover continues until the termination criterion is achieved. Termination criterion is the maximum threshold limit and maximum number of generations. The parameter setting of genetic programming provides a significant role in implementing the model efficiently. The selection of parameter settings is done by trial and error method. The population size and number of generations depend upon the involvement of the input data. A study by Garg et al. (2014) suggests that minimum error will be achieved for complex data, if the population size and number of generation values are kept at a higher range.

3.1.2 Support Vector Machines Model

The support vector machines are powerful learning tools for data classification and regression based on statistical theory (Andrés et al. 2012). The model can easily predict both linear and non-linear data series by mapping. Least Square Support Vector Machines is an advanced version of standard SVM, which changes the inequality limits by equality limits in solving the quadratic equations. Input variables in the lower dimensional space are projected into a higher dimensional space H so as to convert the regression problem with non-linearity to the linear regression problem. To assist with such conversion, several transfer function can be used.

Following the literature (Garg et al. 2014), we have formulated the LS-SVM (See Fig. S-1b) for prediction in the present study as follows (Garg et al. 2014),

$$Y = f(x) = \sum_{i=1}^N w_i \varnothing_i(x) + b = w^H \varnothing(x) + b \tag{4}$$

$$W = [W_1 W_2 W_3 \dots W_N]^H \text{ and } \varnothing = [\varnothing_1 \varnothing_2 \varnothing_3 \dots \varnothing_N]^H$$

where x_i are the values of i^{th} input variable and y_i are the values calculated numerically. The function $\varnothing_i(x)$ is the transformed higher dimensional space. The model given by $\varnothing(x)$ is linear in nature and is a converted form of a nonlinear model in higher dimensional space. In this study, there are five input variables and one output variable. In the present LS-SVM model, the kernel function is used for learning which minimizes the regularized risk (L_r). Minimizing the regularized risk is estimated by the weight (w) and bias (b). The L_r equation can be used as studied by Garg et al. (2014). The parameter setting used in the present study is presented in Table 2.

The normalization of input and output variables is done by using the following equation (Garg et al. 2014):

$$X_{norm} = 0.8 \times \left(\frac{X_i - X_{min}}{X_{max} - X_{min}} \right) + 0.1 \tag{5}$$

where X_i is i^{th} input, X_{min} and X_{max} are minimum and maximum values of variable X_i . The problem formulation of As(III) removal by experimental and computational modeling using GP and LS-SVM models is presented in Fig. 1.

3.2 Model Performance

Several statistical methods are used for validation of model performance (Garg et al. 2014; Andrés et al. 2012; Pai et al. 2011). These methods are the: mean square error (MSE), mean absolute percentage error (MAPE), average absolute relative error (AARE), normalized bias

Table 2 Parameter settings for GP and LS-SVM modeling for As(III) removal

Parameters	GP values	Parameters	LS-SVM values
Inputs	5	Inputs	5
Runs	20	Datasets	131
Population size	200	Grain	6
Generations	300	Iteration	3
Method	Tour	Kernel function	RBF Kernel
Tournament size	2	Optimization method	grid search
Max tree depth	6	Data splitting method	cross validate
Function	Times, minus, plus, multiply, tan, sin, cos	Gam	803.507
Max genes	5	Sig2	22.96
Constant range	[-10,10]	F(x)	186.75
Crossover probability rate	0.85		
Reproduction probability rate	0.10		
Mutation probability rate	0.50		
Seed value	2.053		

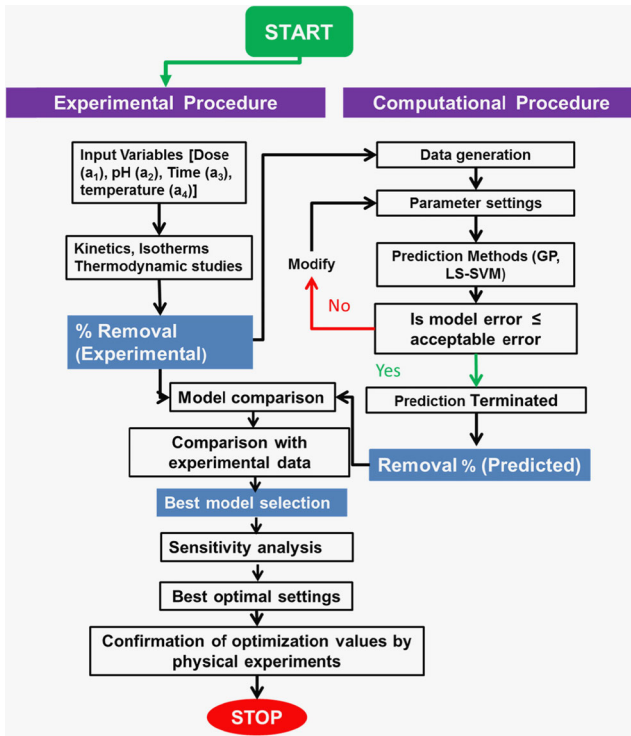


Fig. 1 Problem formulation of modeling of arsenic (III) removal

(NB), standard deviation (σ), chi square (χ^2) and correlation coefficient (R^2), which are described in Eq. (6) to (12):

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_{\text{model},i} - Y_{\text{actual},i})^2 \tag{6}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{Y_{\text{model},i} - Y_{\text{actual},i}}{Y_{\text{actual},i}} \right| \times 100 \tag{7}$$

$$AARE = \frac{1}{N} \sum_{i=1}^N \left| \frac{Y_{\text{model},i} - Y_{\text{actual},i}}{Y_{\text{actual},i}} \right| \tag{8}$$

$$\sigma = \sqrt{\sum_{i=1}^N \frac{1}{N-1} \left[\left| \frac{Y_{\text{model},i} - Y_{\text{actual},i}}{Y_{\text{actual},i}} \right| - AARE \right]^2} \tag{9}$$

$$\chi^2 = \sum_{i=1}^N \frac{(Y_{\text{model},i} - Y_{\text{actual},i})^2}{Y_{\text{actual},i}} \tag{10}$$

$$NB\% = \sum_{i=1}^N \frac{(Y_{\text{model},i} - Y_{\text{actual},i})}{N} / Y_{\text{actual},i} \quad (11)$$

$$R^2 = \frac{\sum_{i=1}^N (Y_{\text{actual},i} - Y_{\text{model,mean}})^2 - \sum_{i=1}^N (Y_{\text{model},i} - Y_{\text{actual},i})^2}{\sum_{i=1}^N (Y_{\text{actual},i} - Y_{\text{model,mean}})^2} \quad (12)$$

where $Y_{\text{actual},i}$ is experimental value of % removal, $Y_{\text{model},i}$ is predicted % removal by the model, $Y_{\text{model,mean}}$ is the average value of model % removal prediction, and N is the number of experiments.

4 Results and Discussion

According to initial adsorption experiments, cerium oxide tetraethylenepentamine (CTEPA) hybrid material exhibited maximum removal of As(III) from water as compared to other materials, and hence the CTEPA was selected for further study. The ion exchange capacity, specific surface area and removal percentage of all hybrid materials synthesized are presented in Table 3.

The thermogravimetric analysis (TGA) of the materials is represented in Fig. 2. From Fig. 2, it is clear that there is a rapid weight loss of 53 % in the temperature range 230–300 °C, and the corresponding endothermic DSC peak is due to the decomposition of organic component present in the material. Another weight loss of 8.1 % is observed in the range of temperature 150–230 °C, which may be due to the loss of free and/or bonded water molecules present in the material. The other weight loss of 15.9 % is observed in the range of 300–350 °C, which is due to the decomposition of residual organic components present in the material. The above observations indicate that the organic component of the material tetraethylenepentamine is an essential component of the material either in bonded or adsorbed form on the cerium oxide gel.

Phase analysis of CTEPA and As(III) adsorbed CTEPA have been carried using the obtained X-ray diffraction pattern, as shown in Fig. 3. A series of sharp and intense peaks can be seen with (111) as the highest intense and sharp peak with clear reflections at low 2 theta values. This characteristic peak is observed for mesoporous cerium oxide (Laha and

Table 3 Selection of best hybrid material towards As(III) removal efficiency

Various hybrid materials	BET surface area (m ² /g)	Particle size (nm)	Zero point charge		Porosity (%) (dry)	Bulk density (g/cm ³)	Percentage removal of As(III) from 100 mg/L As(III) solution
			0.01 M NaNO ₃	0.01 M NaCl			
MgTEPA	144	200–450	3.3	3.7	18 %	2.32	47 %
CaTEPA	178	200–350	6.2	6.1	21 %	1.86	58 %
CTEPA	394	20–100	6.3	6.9	68 %	0.22	81.2 %
ZTEPA	96	300–500	4.7	4.5	8 %	2.87	44 %
MnTEPA	65	400–700	5.2	5.8	2 %	3.14	37 %
LaTEPA	238	100–250	8.7	9.1	34 %	1.22	60.4 %

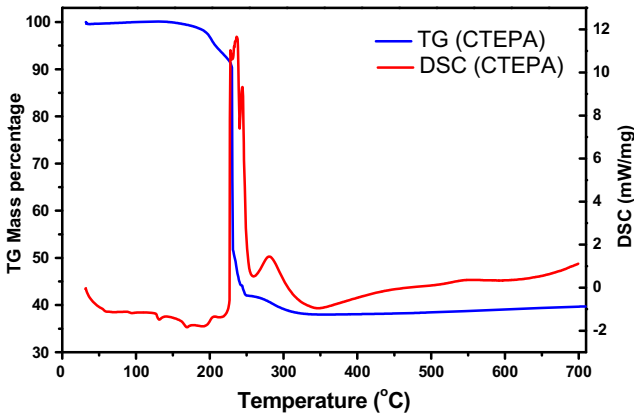


Fig. 2 TGA and DSC of CTEPA hybrid material

Ryoo 2003) that matches well with JCPDS No. 810972. The sharp peaks suggest high crystalline nature of the hybrid material before and after adsorption. Presence of arsenic is seen at lower 2 theta values with small peaks. After adsorption, slight shift of peaks to higher angle is observed suggesting inter-planar rearrangements due to the introduction of arsenic species and formation of complex framework with CTEPA hybrid material, as marked in figure. The peak intensity of (111) increases after arsenic adsorption relating to increase in crystallite size of the particles. The crystallite sizes are calculated using Debye Scherrer equation (Holzwarth and Gibson 2011):

$$D = 0.9\lambda/\beta\cos\theta \quad (13)$$

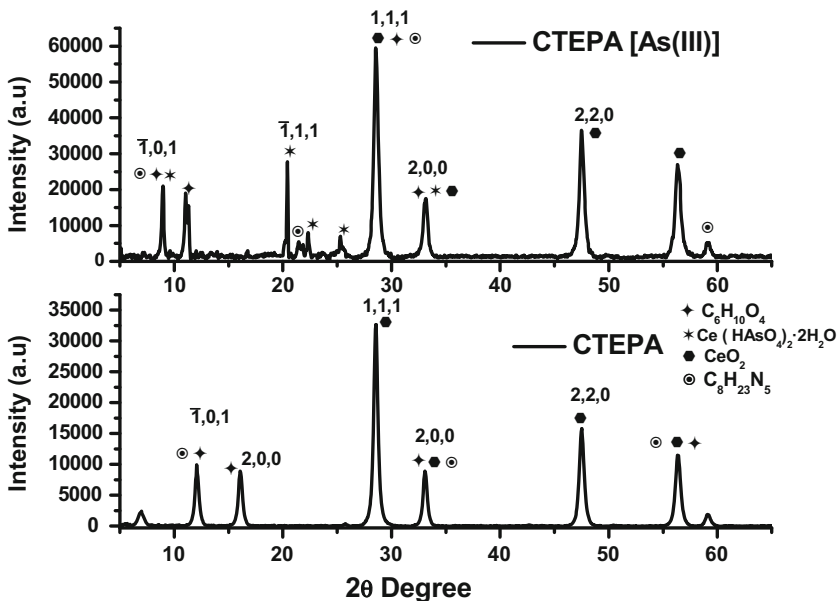


Fig. 3 The powder XRD of fresh CTEPA and CTEPA adsorbed arsenic (III)

where β is the full width half maximum at 2θ value. The average crystallite size before and after adsorption is calculated to be 21 nm and 102 nm, respectively. The above results suggest increase in crystallite size due to adsorption of arsenic on the material sites.

Field emission scanning electron microscopy (Fe-SEM) images of CTEPA hybrid material and As(III) adsorbed material are presented in Fig. 4, along with EDS micrographs. CTEPA is observed to have highly porous spongy structure with non-uniform pore structure. Presence of high amount of cerium and oxygen in EDS micrograph attributes to the absence of any other adsorbed impurities before the adsorption. After adsorption, a fairly high amount of arsenic is found from EDS micrograph supporting the particle flood on CTEPA surface with small fine agglomerated particles indicating possible adsorption of As(III). High Resolution TEM (Fig. 5) has been carried to better understand the morphology of CTEPA hybrid material. CTEPA is seen to be mesoporous in nature with low crystallinity as observed in corresponding selected area electron diffraction (SAED) pattern showing major concentric circles. As(III) adsorption is confirmed from high-resolution transmission electron microscopy (HR-TEM) image showing small particles having hard agglomeration over the surface of CTEPA material with

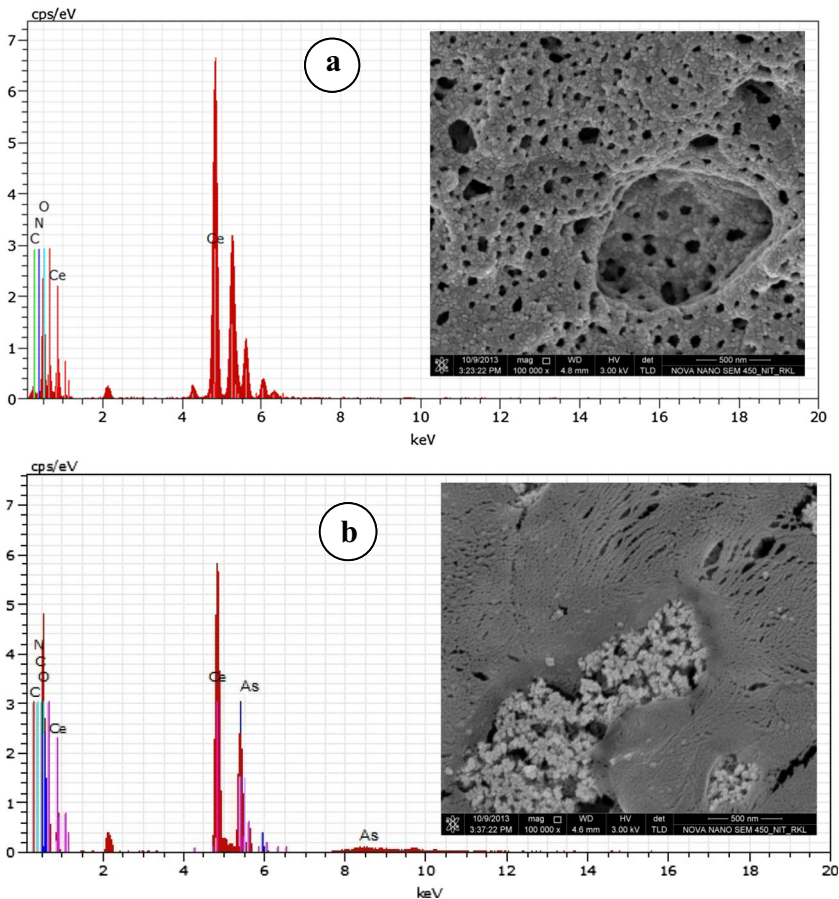


Fig. 4 Fe-SEM and EDS micrographs of CTEPA material **a** before and **b** after adsorption of arsenic (III)

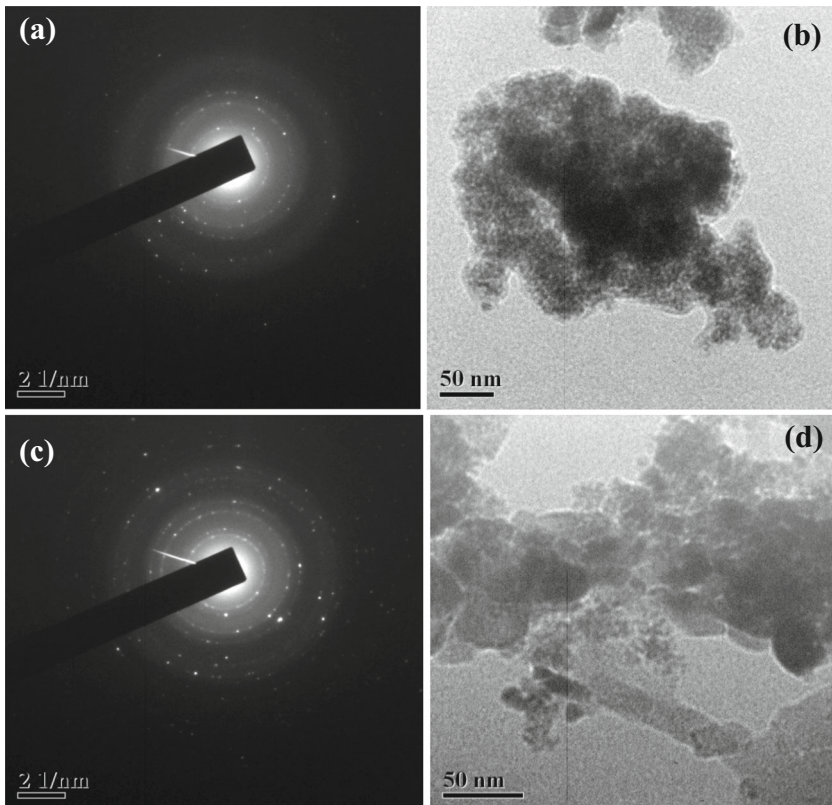


Fig. 5 SAED and TEM pattern of CTEPA hybrid material **a-b** before and **c-d** after adsorption of arsenic (III)

high crystallinity. All the above results support the adsorption of arsenic species on the surface active sites of the adsorbent.

The isotherm t-plot and pore size distribution of the material are represented in Fig. 6. The material exhibits the specific surface area of $394 \text{ m}^2/\text{g}$, micropore volume of $0.522 \text{ cm}^3/\text{g}$ and average micropore diameter of 17.332 \AA . A specific surface area ($24.59 \text{ m}^2/\text{g}$) of the material after adsorption of As(III) is found to be reduced, which indicates that the pores may be closed due to adsorption of As(III). The material shows Type II and Type III isotherm before and after adsorption, indicating indefinite multi-layer formation after completion of the monolayer. A study by Chandra et al. (2010) suggests the possibility of the above observation.

4.1 Batch Adsorption Experiments

4.1.1 Influence of Dose and Solution pH on Adsorption Process

The influence of adsorbent dose on percentage removal of As(III) ions from aqueous solution is graphically represented in Fig. S-2. It is observed from the figure that the removal of As(III) increased from 16 to 95.6 %, 11.8 to 87.1 % and 7.36 to 79.2 % for 0.1–0.9 g/L of the material of initial concentration 5 mg/L, 25 mg/L and 50 mg/L

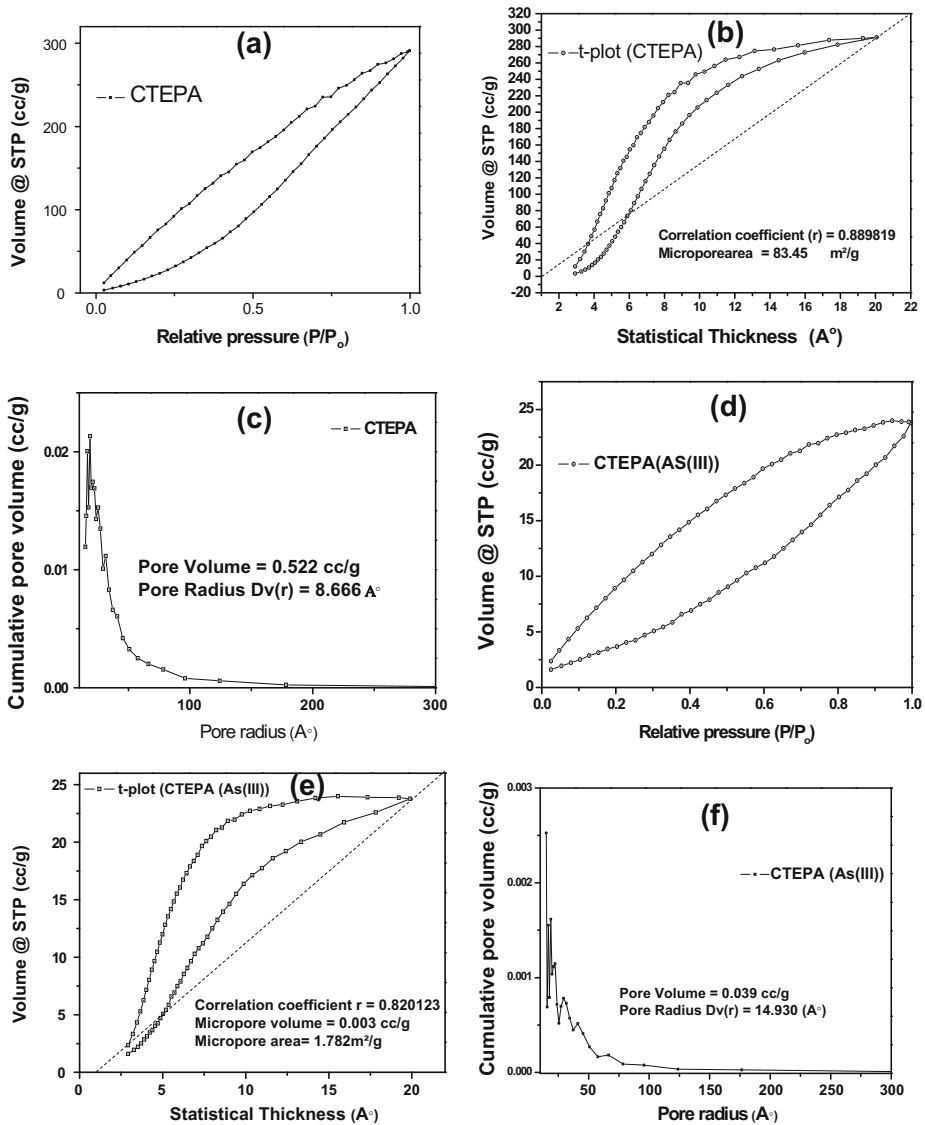


Fig. 6 The typical nitrogen adsorption-desorption isotherms; t-plot and pore size distributions for **a-c** before and **d-f** after adsorption of arsenic (III)

As(III) solution, respectively. The optimum dose is found to be 0.7 g/L, as there is no noticeable change in percentage removal of As(III) beyond this dose. This dose was used for further adsorption studies.

The effect of pH on removal of As(III) ions and zeta potential of the material was studied and presented in Fig. S-3a-b. The isoelectric point of CTEPA is found to be 6.7 and 6.9 for 0.01 M NaCl and 0.05 M NaCl solutions, respectively.

The removal of As(III) was examined with a change in solution pH from 1 to 12. The percentage removal of As(III) rapidly decreases from 97.8 to 12.4 %, 90.6 to 3.5 % and 87.1 to

2.7 % with an increase in pH from 8 to 12, for initial As(III) concentrations of 5 mg/L, 25 mg/L and 50 mg/L, respectively, which indicates that pH plays an important role in the adsorption process. With increase in pH, the OH⁻ ion concentration increases in the solution which may be competing for the adsorption with As(III) species. The suitable pH for removal of As(III) species is between pH 6 and 7, which is in good agreement with the obtained isoelectric (pHzpc) values. The above experimental data for variable pH and adsorbent dose is trained by best fit model, and is found that the best fit model satisfactorily predicts the observation of the experimental data.

4.1.2 Effect of Temperature

The effect of temperature on percentage removal of As(III) with initial concentration of 5 mg/L, 25 mg/L and 50 mg/L onto the material was studied using optimum dose and pH. The results are graphically presented in Fig. S-4, the percentage removal of As(III) increases from 73.2 to 96.8 % (5 mg/L), 46.6 to 96.2 % (25 mg/L) and 47.3 to 97.1 % (50 mg/L), respectively in the temperature range of 20 °C to 70 °C. The above observations are further supported by calculating thermodynamic parameters. The following parameter (free energy (ΔG), change in enthalpy (ΔH) and change in entropy (ΔS)) of adsorption process are calculated using the following equations (Asli et al. 2013):

$$\log K_c = \frac{\Delta s}{2.303R} - \frac{\Delta H}{2.303RT} \quad (14)$$

$$\Delta G = \Delta H - T\Delta S \quad (15)$$

A plot of $\log K_c$ versus $1/T$ for initial As(III) concentration of 5 mg/L, 25 mg/L and 100 mg/L was found to be linear in nature. The value of K_c was calculated using the following equation (Asli et al. 2013):

$$K_c = \frac{C_1}{C_2} \quad (16)$$

where C_1 is the amount of As(III) ion adsorbed per unit mass of CTEPA and C_2 is the concentration of As(III) in aqueous phase. The values obtained are presented in Table 4. For a chemical reaction to occur, bonds must break before new bonds can be formed. The bond breaking and forming is a process of energy absorption

Table 4 Thermodynamic parameters obtained for As(III) removal using CTEPA, As(III) concentration of 5 mg/L, 25 mg/L and 50 mg/L

Initial As(III) concentration	ΔH (kJ/mol)	ΔS (kJ/mol)	ΔG (kJ/mol)								R^2
			20 °C	25 °C	30 °C	35 °C	40 °C	45 °C	50 °C	55 °C	
5 mg/L	-32.14	0.25	-0.38	-0.39	-0.39	-0.40	-0.40	-0.41	-0.41	-0.41	0.994
25 mg/L	-92.27	0.43	-0.80	-0.80	-0.81	-0.82	-0.83	-0.83	-0.84	-0.85	0.991
50 mg/L	-100.03	0.45	-0.85	-0.85	-0.86	-0.87	-0.88	-0.89	-0.90	-0.90	0.991
Best fit model (Predicted)											
5 mg/L	-33.32	0.25	-0.39	-0.39	-0.40	-0.40	-0.41	-0.41	-0.42	-0.42	0.983
25 mg/L	-85.25	0.40	-0.74	-0.75	-0.76	-0.77	-0.77	-0.78	-0.79	-0.79	0.985
50 mg/L	-93.00	0.43	-0.80	-0.81	-0.82	-0.83	-0.84	-0.84	-0.85	-0.86	0.976

(endothermic, $+\Delta H$) and release (exothermic, $-\Delta H$), respectively. So, possibly new bond formation takes place between As(III) and adsorbent resulting to negative enthalpy (exothermic nature) of the ongoing reaction. The positive value of ΔS suggests the structural modification in the system and increases the affinity of adsorbent for As(III) species. The positive value of entropy indicates the increase in randomness in the system. Negative value of ΔG at each temperature indicates the spontaneous nature and feasibility of ongoing adsorption. A decrease in values of ΔG with increase in temperature suggests more impulse of As(III) adsorption at higher temperature.

4.1.3 Effect of Contact Time and Adsorption Kinetics

The effect of contact time on percentage removal of As(III) is graphically presented in Fig. S-5. It is evident from the figure that the removal of As(III) increased from 42.6 to 95.8 %, 47.1 to 90.6 % and 41.5 to 90.8 % for a contact time of 10 to 50 min for initial As(III) concentration of 5 mg/L, 25 mg/L and 50 mg/L, respectively. It is clear that about 40 % removal took place within 10 min, and equilibrium is established after 40 min. The rapid removal during the first 10 min may be due to availability of vacant sites and high concentration gradient. Further, kinetics of adsorption and mass transfer was also computed using different mathematical models such as pseudo first-order, second-order and intraparticle diffusion (Weber-Morris equation) for initial As(III) concentration of 5 mg/L, 10 mg/L and 50 mg/L. The integrated form of the pseudo-first order rate equation can be represented as:

$$\log(q_e - q_t) = \log q_e - K_1 \left(\frac{t}{2.303} \right) \quad (17)$$

$$\frac{t}{q_t} = \left[\left(\frac{1}{K_2 q_e^2} \right) + \frac{1}{q_e} \right] t \quad (18)$$

$$q_t = K_p t^{1/2} + C \quad (19)$$

where q_e and q_t (mg/g) are the amounts of As(III) adsorbed at equilibrium and at time 't', respectively. K_1 , K_2 and K_p are first-order rate constant, second-order rate constant and intraparticle diffusion rate constant, respectively. The values are calculated from their respective slope and intercept (Fig. S-5b-d). The data obtained are presented in Table 5. Due to poor regression coefficient (R^2) values obtained for pseudo-first order and intraparticle diffusion, it may be considered that the adsorption process is best fitted with second-order rate equation. The contact time and kinetic data were trained with the best fitted model and it was found that the model predicts satisfactorily with high correlation values. The trend of observation is presented in respective plots (Fig. S-5b-d).

It is observed from Fig. S-5d that there is a deviation of line from the origin which indicates that the intraparticle transport is not the only rate limiting step, but the transport of adsorbate through the pores of the adsorbent and adsorption on the surface of the material are also responsible. Subsequently, the acceleration of the adsorption process indicates that the diffusion is not consecutive due to pore size (Erhan et al. 2004). At the final phase of adsorption process, the rate of diffusion

Table 5 Kinetic and isotherm parameter for As(III) removal by CTEPA

Initial Concentration	Pseudo First Order	Second Order	Intraparticle diffusion rate constant				
	K_1	q_e (mg/g)	R^2	K_2	R^2	C	R^2
5 mg/L	0.056	32.5	0.883	0.000185	0.998	1.46	0.897
25 mg/L	0.038	75.9	0.742	0.000055	0.940	3.35	0.954
50 mg/L	0.034	136.5	0.912	0.000009	0.982	8.13	0.947
Best fit model (Predicted)							
5 mg/L	0.086	31.1	0.929	0.000207	0.995	1.98	0.971
25 mg/L	0.042	93.6	0.856	0.000057	0.965	3.26	0.929
50 mg/L	0.035	141.0	0.914	0.000009	0.982	7.45	0.957
Temperature	Langmuir	Freundlich	$1/n$	R^2	E (kJ/mol)	q_m	Dubinin-Radushkevich
	q_e (mg/g)	K_f (mg/g)	R^2	R^2	E (kJ/mol)	q_m	K mol ² kJ ⁻²
25 °C	124.8	0.605	0.985	0.626	70.3	1.785	1.0×10^{-07}
50 °C	201.4	0.502	0.994	0.555	91.2	1.794	6.0×10^{-08}
75 °C	358.4	0.581	0.985	0.473	112.0	1.883	3.0×10^{-08}
Best fit model (Predicted)							
25 °C	117.6	0.463	0.988	0.627	66.7	1.787	1.1×10^{-07}
50 °C	188.9	0.517	0.994	0.594	83.5	1.794	7.1×10^{-08}
75 °C	367.2	0.622	0.987	0.507	105.1	1.883	4.5×10^{-08}
							R^2
							0.642
							0.746
							0.699

remains constant due to exhaustion of pores in the material. When the K_p value is higher, the rate of adsorption is increased, but when the C value is higher, the adsorption is better due to better bonding between adsorbate and adsorbent. The concentration gradient decides the flow of particles into the pores of material. With the increase in the concentration of As(III), the K_p value increases, suggesting that the intraparticle diffusion is considered as the concentration diffusion (Erhan et al. 2004; Biyan et al. 2009).

4.1.4 Impact of Adsorbate Concentration and Adsorption Isotherms

The effect of initial concentration on % removal of As(III) from aqueous solution is graphically presented in Fig. 7a, which presents the impact of concentration of adsorbent on the removal efficiency of the material. It is evident from the graph that with increasing initial concentration of As(III) from 1 to 180 mg/L in the solution, the removal efficiency decreases from 92 to 55.2 % due to the lack of active sites at higher concentration of As(III).

To understand the adsorption behaviour, adsorption data obtained from initial concentration was fitted to linear transformed Langmuir (Eq. 20), Freundlich

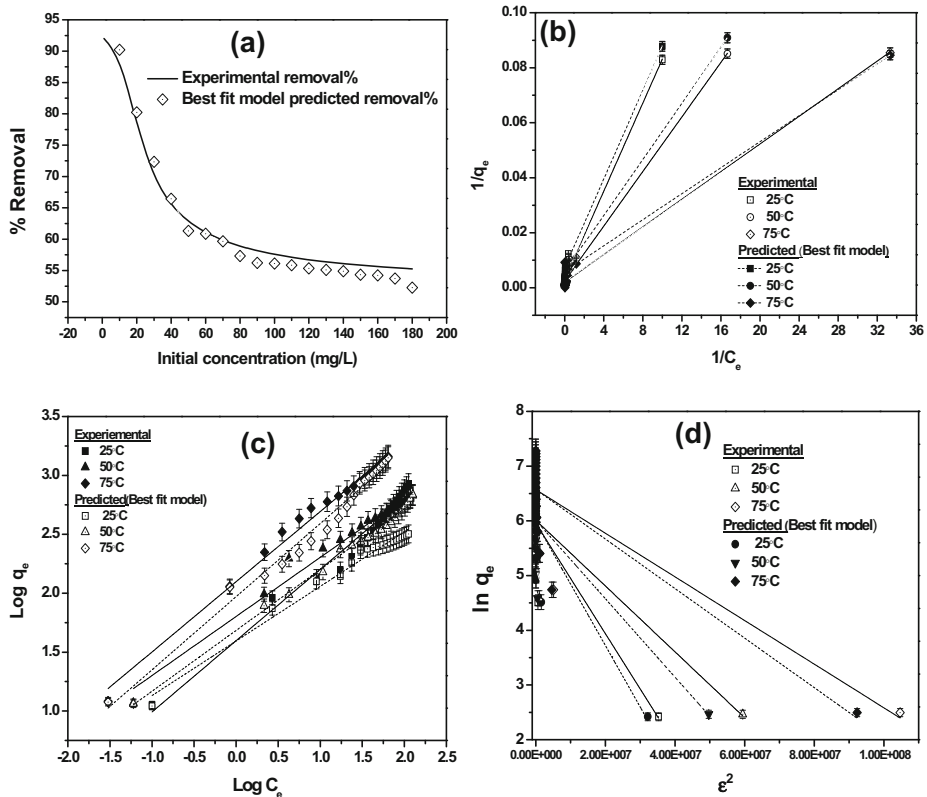


Fig. 7 Experimental and model prediction a effect of initial concentration b Langmuir isotherm plot c Freundlich isotherm plot d D-R isotherm plot

(Eq. 21) and Dubinin-Radushkevich (D-R) (Eq. 22) adsorption isotherms, as presented in the following equations (Foo and Hameed 2010):

$$\frac{1}{q_e} = \frac{1}{q_o b C_e} + \frac{1}{q_o} \quad (20)$$

$$\log q_e = \log K_f + \frac{1}{n} \log C_e \quad (21)$$

$$\ln q_e = \ln q_m - K \varepsilon^2 \quad (22)$$

where q_o is the maximum amount of the As(III) ion adsorbed per unit weight of the material to form a complete monolayer on the surface (adsorption capacity), q_e is the amount of As(III) adsorbed at equilibrium (mg/g), C_e is the equilibrium adsorbate concentration (mg/L), and b is the binding energy constant. K_f and $1/n$ are the constants representing Freundlich adsorption capacity and adsorption intensity, respectively. The Polanyi potential is represented as ε and is equal to $RT \ln(1+1/C_e)$, q_m is the theoretical adsorption capacity, K is the constant and T is the temperature in degrees Kelvin. The plots of Langmuir, Freundlich and D-R isotherms are presented in Fig. 7b-d.

The values of different parameters are calculated from the slope and intercept of the plot and are presented in Table 5. The high correlation coefficient (R^2) value suggests that the adsorption data were best fitted to Langmuir adsorption isotherm with a maximum adsorption capacity of 124.8, 201.4 and 358.4 mg/g for 25, 50 and 75 °C, respectively. The mean free energy of adsorption (E) is calculated from the constant K using the relation (Foo and Hameed 2010)

$$E = (2K)^{-1/2} \quad (23)$$

It is defined as the change in free energy when 1 mol of ion is transferred from the solution to the surface of the solid. The value of E is very useful in predicting the type of adsorption. The different types of adsorption processes are: (a) Physical (< 8 kJ/mol); (b) ion exchange (8 < 16 kJ/mol); and (c) chemisorption (> 16 kJ/mol). As presented in Table 5, the value obtained at different temperatures is higher than 16 kJ/mol, suggesting chemisorption nature of the adsorption process.

The dimensionless equilibrium parameter r is calculated by using the following equation (Foo and Hameed 2010):

$$r = \frac{1}{1 + b C_o} \quad (24)$$

where C_o is initial As(III) concentration in mg/L, where values of $r < 1$ indicate favourable adsorption process. The values of r are found to be 0.340, 0.094 and 0.020 for initial concentrations of 5 mg/L, 25 mg/L and 50 mg/L, respectively, which indicates a favourable adsorption system. A similar trend of observation has been observed with high correlation value when the above isotherm and concentration data are trained by using best fitted model. The observations at prediction are presented in their respective plots (Fig. 7a-d).

4.1.5 Competitive Ions and Regeneration Studies

All the above studies were conducted by synthetic As(III) solution only. However, in real water, several other anions exist, which may compete for the adsorption. In order to understand the effect of anions on adsorption of As(III), a mixture of common anions such as hydroxide, nitrate, chloride, phosphate, sulphate, fluoride, carbonate and bicarbonate were added in known quantities of As(III) solution. The initial concentrations of As(III) were kept fixed at 5 mg/L, 25 mg/L and 50 mg/L, while the concentrations of the other anions are varied from 5 to 100 mg/L. The observation is reported in the form of a plot in Fig. S-6.

It is understood from the figure that presence of these anions highly reduces the As(III) removal efficiency. The anions reduced the As(III) adsorption in the order of hydroxide > bicarbonate > carbonate > phosphate > nitrate > fluoride > sulphate > chloride. This analysis depicts that hydroxides, bicarbonates and carbonates has the most and sulphate has the least effect on the removal of As(III) by the material. It is found that the percentage removal of As(III) is reduced by >71 % in the presence of hydroxide, bicarbonate and carbonate in As(III) solution. This may be due to material having higher interaction and strong attraction towards hydroxide, bicarbonate and carbonate ions as they have a tendency to form correspondingly hydroxide, carbonates and bicarbonates. Presence of sulphate and chloride reduces As(III) removal <15 %. It can be concluded that this material can be used for As(III) removal in the presence of chloride and sulphate at high concentrations; however, presence of hydroxide, bicarbonate and carbonate can affect the removal percentage. So, the alkalinity test must be carried out before using CTEPA for As(III) removal.

A desorption and a regeneration study was carried out in order to understand the reusability of the material. The results obtained are presented in Table S-2.

Generally, poor desorption occurs, when the adsorption process is of chemical nature and this happens because in chemisorption process, adsorbate and adsorbent are associated with stronger bonds (Ma et al. 2011). The data obtained suggest the chemisorption nature of the adsorption process, which is in good agreement with the above data presented in free energy.

Adsorbents are usually regenerated in acidic or alkali medium. In the present study, the material is unstable in extreme acidic conditions (below pH of 3), so the regeneration tests are carried out between pH 8 and 11. It is observed from Table S-2 that regeneration is less than 10 % at different pH values. This suggests poor regeneration of CTEPA material. Studies by Ma et al. (2011) suggest the possibility of above regeneration and desorption studies.

4.1.6 Mechanism of As(III) Adsorption

The FTIR spectra of CTEPA before and after adsorption are shown in Fig. 8. Broad and intense bands at $\sim 3128\text{ cm}^{-1}$ indicate the presence of overlapped O-H and NH_2 stretching vibration groups (Ma et al. 2011). The band at $\sim 1383\text{ cm}^{-1}$ corresponds to carbonate anion which may be due to absorption of atmospheric CO_2 gas. Few lattice vibrations of metal-oxygen bonds M-O are also observed at $\sim 661\text{ cm}^{-1}$ and $\sim 486\text{ cm}^{-1}$ (Hameed et al. 2006). Two new peaks appear at $\sim 1625\text{ cm}^{-1}$ and $\sim 560\text{ cm}^{-1}$ after As(III) adsorption onto the material, possibly indicating the formation of surface complexation (M-O-X) (where $\text{X}=\text{HAsO}_3^{3-}$, H_2AsO_4^- , AsO_2^-). It is well known that As(III) species are adsorbed via electrostatic attraction and complexation (Ma et al. 2011). The presence of M-OH and M-NH_2^+ functional groups plays an important role in the adsorption process according to FTIR spectra. Possibly, there is electrostatic attraction between negatively charged As(III) species and positive surface of the adsorbent (NH_2^+), since the solution pH is below the isoelectric point (pH_{zpc}) (See Section 4.1.1). At low pH, surface hydroxyl groups are protonated and provide the

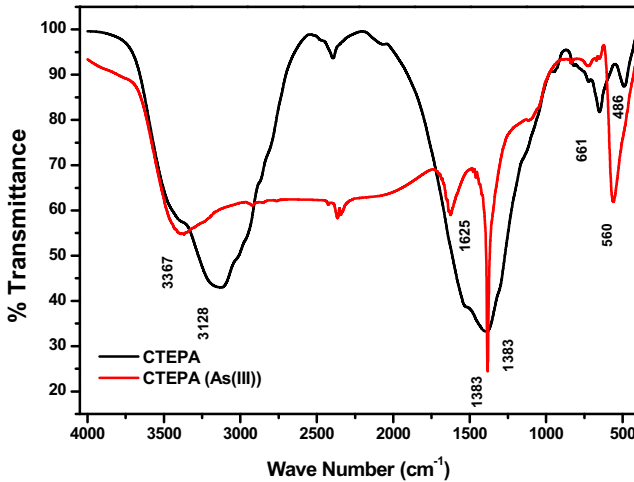


Fig. 8 FTIR comparison spectrum of CTEPA hybrid material before and after adsorption of arsenic (III)

complexation process, because $-\text{OH}^{2+}$ is easier to displace from the metal binding system than OH functional groups (Ma et al. 2011). At the active adsorption sites, As(III) species replaced the hydroxyl groups which are then released to the water solution. The adsorption process could be described in two steps, as illustrated in Fig. 9: (a) electrostatic interaction between positively charged center (nitrogen) and negatively charged arsenite molecule in solution; and (b) complexation between positively charged surface hydroxyl group and arsenite. Table 6 presents the comparison of adsorption capacity of As(III) ions with that of various other hybrid materials reported in the literature, also showing pH range.

4.2 Computational Modeling Approach

4.2.1 GP and LS-SVM Models

In the GP and LS-SVM modeling, input and output variables are normalized first in order to achieve the uniformity level of individual factors. The models are executed on the basis of trial and error method, as discussed in Section 3. Figures 10a and b illustrate the prediction of arsenic removal at training and testing using GP and LS-SVM model. As seen in both figures, there is a good agreement between models and experimental results and the models represent reasonable prediction of the system behaviour. The LS-SVM prediction is terminated after satisfying the optimized parameters ($dF=0.1129$, $dX=0.13198$, $X=6.6667$, $F(X)=174.2688$, Hyper parameter ($\text{gamma}=785.7$, $\text{sig}2=40.24$)) and the best-so-far GP model (see Eq. 25) obtained after satisfying the termination criteria is:

$$\begin{aligned} \text{ypred} = & 15.71 \times (x_1) - 0.5959 \times (x_3) + 0.5959 \times (x_4) - 0.7665 \times (x_5) + \\ & 0.6579 \times \tan(x_3) \times (\cos(x_3) - (x_5) + 0.9605) - 0.03035 \times \tan(x_2) \times (x_5 - \\ & 1.148) + 0.7665 \times \cos(x_1) + 15.71 \times \cos(x_3) + 0.5959 \times \cos(x_5) - 15.71 \times \\ & \sin(x_2) - 0.5959 \times \cos(\sin(x_5) \times \sin(x_1) \times (x_1) - (x_3)) + 0.03035 \times (x_3) \times \\ & (x_5) \times (\sin(x_5) - (x_3) \times \tan(x_1)) + 0.4997 \times (x_1) \times (x_3) \times (x_5) + 10.85 \end{aligned} \quad (25)$$

The equation obtained through genetic programming satisfied the proposed parameter settings, as presented in Table 2. In Eq. (25), x_1 , x_2 , x_3 , x_4 , and x_5 denote normalized values

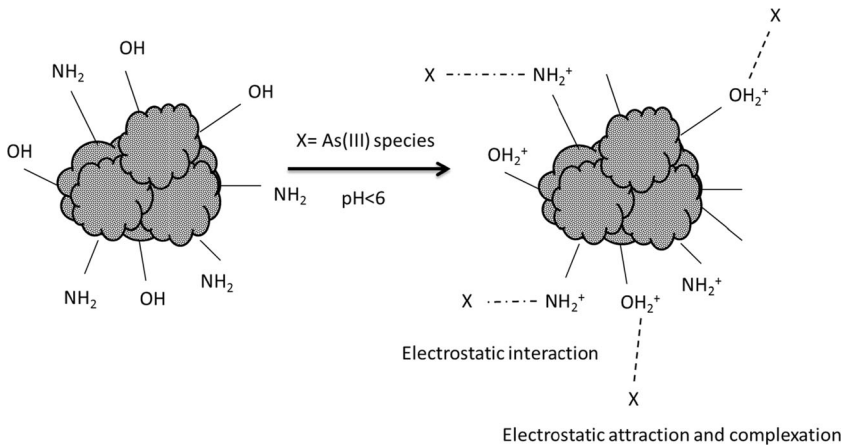


Fig. 9 Mechanism of arsenic (III) removal by CTEPA hybrid material

of dose, pH, time, temperature and concentration, respectively. The execution time of running the GP and LS-SVM algorithm for the present study was approximately 10 min on Dell OPTIPLEX 980 (core i5, RAM 4GB, Windows 7). The number of experimental runs and its prediction behaviour by the GP and LS-SVM model is illustrated in Fig. S-7a-b.

The residuals versus experimental run for both models (GP and LS-SVM) (See Fig. 11a-b) presents uniform nature of residues at training and testing, suggesting that the models were drifting slowly to lower values as the prediction continued. The distribution of points scattered randomly about -10 to $+10$ and -20 to $+40$ for GP and LS-SVM, respectively, regardless of the size of the fitted value, but the residual values may increase with the increase in the size of the fitted value. Having this condition, the residual points appear to be “funnel shaped” as larger end toward larger fitted values. This is due to the fact that with the increase in the value of the response, the residuals have larger and larger scatter. In the present investigation the position of residuals is well within the range of values used for prediction in genetic programming, but the residuals obtained by using LS-SVM was somewhat higher than the GP model because LS-SVM was found unable to capture satisfactorily the relationship between the process variables.

The percentage error distribution at prediction for both models at training and testing is presented in Fig. S-8a-b. The % error distribution lies within the range of $+25$ and -25 for GP and -70 to $+60$ for LS-SVM. In order to understand the reliability and performance of the models, a comparison among the models is discussed in Section 4.2.2.

4.2.2 Comparison and Performance of the Models

A separate comparison at prediction data for thermodynamic parameter, kinetic parameter and isotherm parameters are presented and discussed in Section 4.1. In this section the overall performance of both models in prediction of As(III) removal at batch studies are done by the statistical parameters presented in Section 3.3. The obtained values are presented in Table 7. The values (MSE, RMSE, MAPE, AARE, chi square, NB% and standard deviation) suggest that the GP model has impressively well learned the non-linear relationship between the input and output process variables with high correlation coefficient values (R^2) of 0.988 and 0.977 at the training and testing phases, respectively. Compared to the LS-SVM model, the GP model has shown better performance. A statistical comparison based on box plots of relative error (%)

Table 6 Comparison of present adsorbent material with materials available in literature

Sl.no.	Adsorbents	As(III) removal %	pH range	Adsorption capacity As(III)	Isotherm models Supported	References
1	Magnetic reduced graphene oxide	99.9 %	4–10	13.10 mg/g	Langmuir and freundlich	(Chandra et al. 2010)
2	FIBAN	95 %	3–12	2.9 g/Kg	Langmuir	(Vatusina et al. 2007)
3	Zero-valent iron (ZVI) particles	99 %	4	–	–	(Calo et al. 2012)
4	Hybrid inorganic/organic polymer type material based on hydrated ferric oxide (HFO)	100 %	7–9	5.55 mg/g	–	(Iesana et al. 2008)
5	Zirconium-loaded activated carbon (Zr-AC) with	–	8	7 g/L	Langmuir and Freundlich	(Daus et al. 2004)
6	Iron-chitosan composites	–	7	16.15 mg/g	Langmuir	(Gupta et al. 2009)
7	Al-HDTMA-Sericite (AH)	–	2–10	0.433 mg/g	–	(Tiwari and Lee 2012)
8	Mesoporous hybrid material (Thiol-functionalized silica coated activated alumina)	93 %	3–12	8.41 mg/g	Freundlich	(Hao et al. 2009)
9	Hydrous iron(III) oxide polyacrylamide	–	5–8	43.0 mg/g	–	(Shigetomi et al. 1980)
10	Nano-hydrated zirconium oxide (HZO)/polystyrene hybrid adsorbent	–	7	88.74 mg/g	–	(Pan et al. 2014)

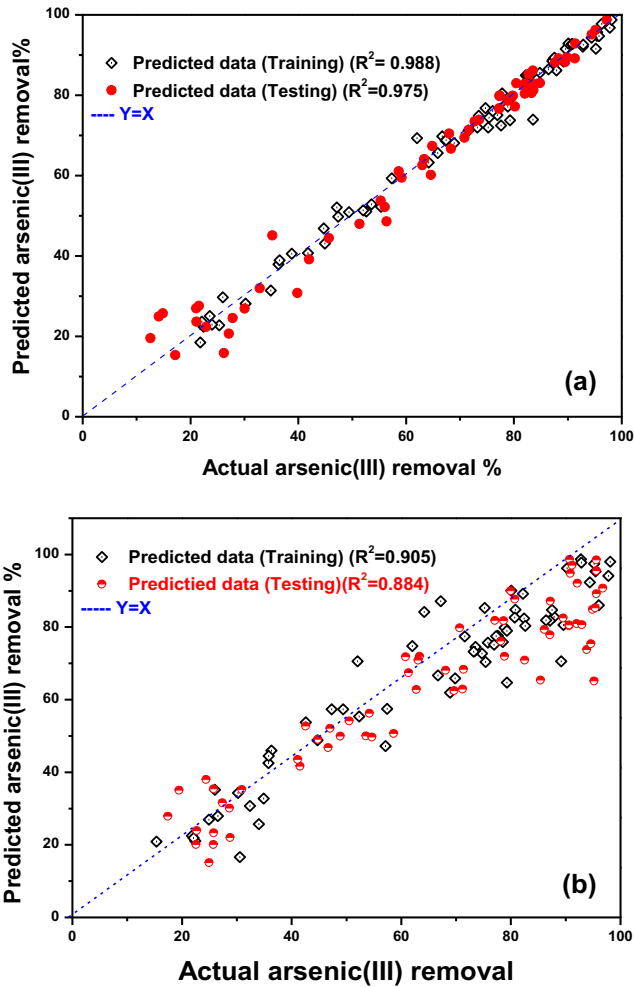


Fig. 10 Correlation prediction plot of arsenic (III) removal at training and testing using **a** GP and **b** LS-SVM models

for the two models at training and testing (Fig. 12) is done by using the following relation (See Eq. 26):

$$Relative\ error\ (\%) = \frac{|\overline{model}_i - \overline{actual}_i|}{actual_i} \times 100 \tag{26}$$

where \overline{model}_i and \overline{actual}_i are the average values of predicted and actual quantities, respectively. A lower mean relative error of 3.48 % and 3.86 % at training and testing, respectively, shows that GP is able to capture the relationship between process variables satisfactorily. The descriptive statistics of the errors are shown in Table 7. The table illustrates mean, standard error, standard deviation, median, maximum and minimum errors. Lower

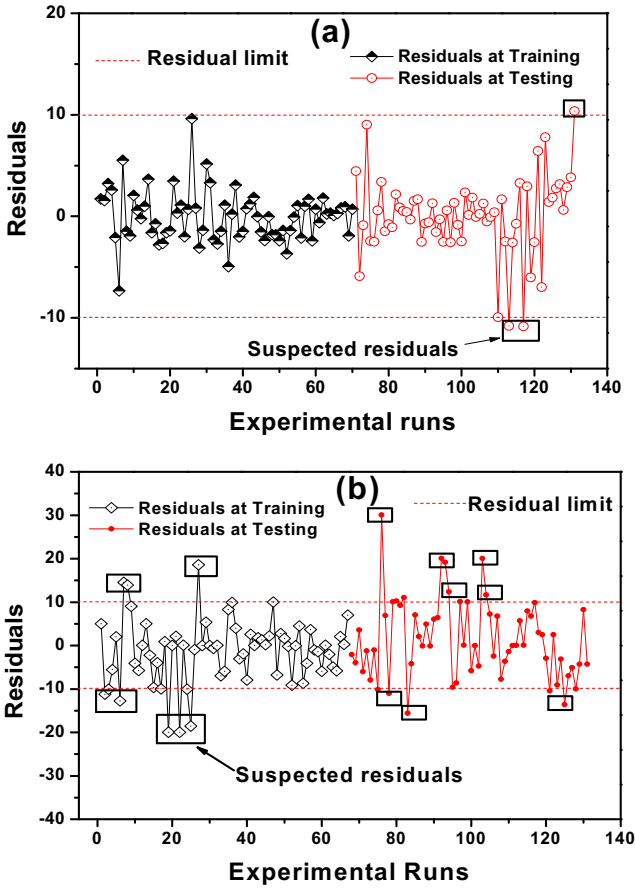


Fig. 11 Experimental runs versus residuals obtained at training and testing by using **a** GP and **b** LS-SVM models

range of confidence intervals at training and testing for the proposed GP model indicates that overall the GP model performance is found to be better than LS-SVM model. Thus, from the statistical comparison presented, it can be concluded that the proposed evolutionary GP model is better compared to LS-SVM in prediction of As(III) removal from water by CTEPA hybrid material.

4.2.3 Sensitivity Analysis

Once the prediction and selection of best model is done by establishing the relationship between input and output, the assessment of impact of each input on output is done by using the following Eq. (27) (Elmolla et al. 2010):

$$SCO = \frac{(Scaled\ Output\ for\ 10\% \ increase\ in\ input) - (Scaled\ Output\ for\ 10\% \ decrease\ in\ input)}{2} \quad (27)$$

The scaled change in output (SCO) is calculated with the current input increased by 10 % and the current input decreased by 10 %. Thus, the results obtained are the

Table 7 Model comparison and performance at training and testing (GP and LS-SVM models)

Models	Genetic programming		LS-SVM	
	Training	Testing	Training	Testing
Error methods				
MSE	0.0998	0.1068	1.423	0.380
RMSE	0.0278	0.0284	0.112	0.047
MAPE	0.0497	0.0632	0.200	0.0950
AARE	0.0003	0.0004	0.002	0.0010
Standard deviation (σ)	0.0491	0.0587	0.145	0.1965
Chi square (χ^2)	1.6159	0.2397	1.648	0.5200
NB%	0.0004	-0.0002	0.028	-0.002
R ²	0.988	0.977	0.905	0.884
Relative error (%)				
Statistical comparison				
GP	Count	Mean	LCI 95 %	Maximum
	70	1.799	2.7	15.162
LS-SVM	Count	Mean	LCI 95 %	Maximum
	61	5.792	2.6	28.222
GP	Standard deviation	Standard error	Median	Minimum
	3.443	0.411	2.195	0.054
LS-SVM	Standard deviation	Standard error	Median	Minimum
	11.06	1.352	5.940	-3.32
GP	UCI 95 %			
	4.3			
LS-SVM	UCI 95 %			
	16.9			

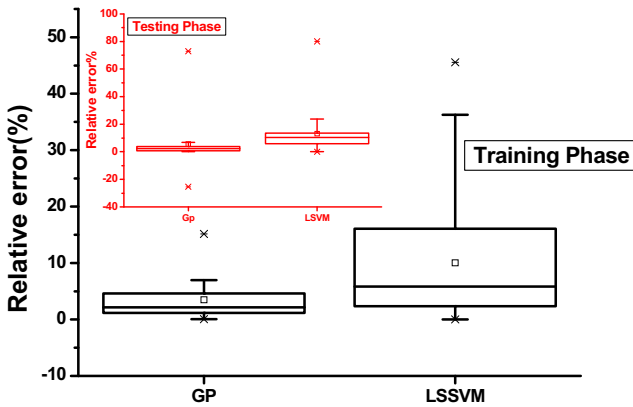


Fig. 12 Statistical comparison based on box plots of relative error (%) for GP and LS-SVM models at training and testing

scaled output change per 10 % change in input. Increase/decrease of an input from its base value results in increase/decrease performance level. Logically, the net effect of change in input results in a positive score for average scaled change in output. The data obtained from sensitivity analysis is presented in Table 8.

The percentage relative contribution of each input variable in removal percentage is calculated using the following equation (Bring 1996):

$$Relative\ contribution\ (\%) = \frac{variance_x}{variance_{x(total)}} \times 100 \tag{28}$$

The relative contribution of input variables in As(III) removal is presented in Fig. 13. The variables such as dose, pH and time are important influencing variables, which effect As(III) removal at highest. However, the sensitivity data suggest that variable concentration is more sensitive than others, but its contribution towards removal percentage is low, which makes it a less significant variable.

Table 8 Net sensitivity at output and relative contribution of input variables affecting removal percentage of arsenic (III)

Variables	Average removal (%) experimental	Average removal (%) with increase and decrease in variable standards		Scaled changed in output	Change in correlation (R ²) value		
		+10 %	-10 %			±10(%)	Experimental
Dose	55.1	53.4	54.8	-0.7	0.991	0.967	0.977
pH	46.8	45.9	46.7	-0.4	0.983	0.978	0.966
Time	82.1	81.6	81.6	0	0.983	0.977	0.970
Temperature	79.2	79.3	79.3	0	0.998	0.999	0.997
Concentration	64.0	60.7	64.1	-1.7	0.971	0.974	0.974

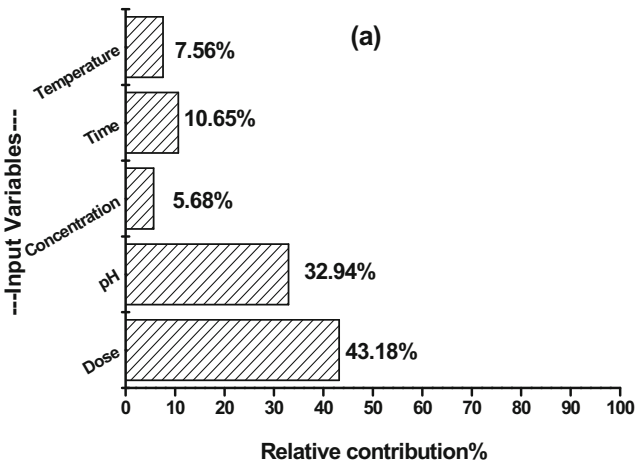


Fig. 13 The relative contribution of input variables in arsenic (III) removal by CTEPA hybrid material

4.2.4 Best Optimal Settings

In the present study, a best optimal setting was obtained by feeding all possible experimental runs of batch studies in GP model (parameters used, as presented in Table 2). The optimum removal of 99.7 % is achieved with a dose=0.7 g/L, pH=6, time=50 min, temperature=65 °C, and initial concentration=25 mg/L. The best-so-far GP model for predicting the optimized value is:

$$\begin{aligned}
 y_{pred} = & 11.79 \times \cos(\sin(x_1) \times (x_4)) + ((\tan(x_3) + (x_5)) - 1.127) - \\
 & (5.952 \times \cos(\cos(x_2) + 0.8083)) - 11.8 \times (\sin(x_3) - 5.221 \times \sin(x_4)) - \\
 & 3.516 \times (x_1) \times ((x_1) - 3.504) + 5.952 \times \sin((x_3) - (x_5)) + 6.278 \times \\
 & \tan(\sin(\sin((x_3) \times (x_4)) \times ((x_2) + 0.8083))) - 0.1955 \times \\
 & \tan(5.221 \times \cos((x_2) + 0.8083) - ((x_1) - 3.504) \times ((x_2) - 0.7015)) - 10.92 \times \\
 & \sin(3.516 \times (x_1) \times ((x_1) - 3.504) + 5.221 \times (x_1) \times (x_4) + 0.7015) - 5.952 \times \\
 & \tan(3.516 \times (x_2)) + 7.442 \times \tan(\sin((x_1) \times (x_4)) - 0.9009 \times (x_2)) - 0.4905 \times \\
 & ((x_1) - 3.516) \times (\cos((x_2) \times (x_4)) - 0.7015) \times ((x_1) - (x_2) + (x_4)) + 1.667 \times \\
 & ((x_1) - 3.504) \times (\cos((x_2) \times (x_4)) - 0.7015) \times \\
 & (\cos((x_2) + 0.8083) + \cos((x_2) \times (x_4)) + 7.352) + 88.72
 \end{aligned}
 \tag{29}$$

The GP model has impressively learned the nonlinear relationship between input and output with correlation coefficient (R^2) of 0.987 at the optimization stage. The above observations suggest that, if optimized input combinations are followed, then it is possible to achieve maximum removal percentage.

4.2.5 Confirmatory Experiments

After finding out the best optimal settings from the modeling, ten numbers of actual experiments are performed with the same conditions in the laboratory to confirm the model adequacy

and predictive relationship. The 95 % confidence interval of confirmation experiments (CI_{CE}) is calculated by using the equation reported by Ross (1988), Roy (1990) and Krishnaiah and Shahabudeen (2012). Confidence interval is estimated as 34.78. Therefore, the predicated confidence interval for the confirmatory experiments is in the range of:

$$34.78 < \mu < 99.7$$

The actual average of ten removal experiments of confirmation test is found to be 96.52. As the overall average values of the confirmatory experimental test fall well within the 95 % CI_{CE} , the confidence interval obtained by using the equation reported by Ross (1988), Roy (1990) and Krishnaiah and Shahabudeen (2012) suggests the accuracy of genetic modeling in prediction of removal percentage.

5 Conclusions

The hybrid material cerium oxide tetraethylenepentamine (CTEPA) was successfully synthesized by co-precipitation method and was characterized using analytical techniques. This study presents an alternate approach using genetic programming and least square support vector machine tools to investigate removal percentage of As(III) ions from water using the hybrid material with a maximum removal percentage (97.2 %). The XRD study confirms the crystalline nature of the material having average crystallite size of 21 and 120 nm before and after adsorption, respectively. The porous nature of the material was confirmed from Fe-SEM studies and As(III) adsorption was confirmed from corresponding EDS. The thermodynamic study indicated that the adsorption process is exothermic, spontaneous and feasible in nature. The kinetic study indicated that the adsorption process is of second order. The Langmuir isotherm was found to best fit the experimental data resulting in high correlation coefficient of $R^2 > 0.98$, with maximum adsorption capacity of 124.8 mg/g at 25 °C. The chemisorption nature of the adsorption process was confirmed by D-R isotherm studies. The FTIR studies suggested the mechanism of electrostatic attraction and complexation for As(III) ions onto the hybrid material. The least values of MSE, RMSE, MAPE, AARE, chi square, NB%, standard deviation and high correlation coefficient values (R^2) (0.988 and 0.977) suggested that the genetic programming (GP) was best fitted to the experimental data and had higher predictive capability than LS-SVM. The variable concentration has the highest sensitivity among other variables, but it has the least contribution to the removal percentage. The study concludes that GP modeling can be used as effective prediction tool in As(III) removal from water. The results also indicate the superiority of genetic programming tools in capturing the nonlinear behaviour of the adsorption system. Based on the above findings, the study can be used as a guide for prediction of As(III) removal from water.

Acknowledgments The authors are thankful to UGC, Govt. of India, New Delhi and Prof. S.K. Sarangi, NIT Rourkela for financial support. The authors are also thankful to esteemed editor and reviewers for useful guidance in improving the content of the manuscript.

References

- Akratos CS, Papaspyros JNE, Tsihrintzis VA (2008) An Artificial neural network model and design equations for BOD and COD Removal Prediction in horizontal subsurface flow constructed Wetlands. *Chem Eng J* 143: 96–110

- Aleboye H, Kasiri MB, Olya ME, Aleboye H (2008) Prediction of azo dye decolorization by UV/H₂O₂ using artificial neural networks. *Dyes Pigments* 77:288–294
- Andrés E, Salcedo-Sanz S, Monge F, PérezBellido AM (2012) Efficient aero-dynamic design through evolutionary programming and support vector regression algorithms. *Expert Syst Appl* 39:10700–10708
- Asli THO, Ercan O, Esra SB, Ulker B (2013) Removal of As (V) from aqueous solution by activated carbon-based hybrid adsorbents: Impact of experimental conditions. *Chem Eng J* 223:116–128
- Biyani J, Fei S, Hu G, Zheng S, Zhang Q, Xu Z (2009) Adsorption of methyl tert-butyl ether (MTBE) from aqueous solution by porous polymeric adsorbent. *J Hazardous Mater* 161(1):81–87
- Bring J (1996) A geometric approach to compare variables in a regression model. *Am Statistician* 50:57–62
- Calo JM, Madhavan L, Kirchner J, Bain EJ (2012) Arsenic removal via ZVI in a hybrid spouted vessel/ fixed bed filter system. *Chem Eng J* 189–190:237–243
- Chandra V, Park J, Chun Y, Lee WJ, Hwang CL, Kim S (2010) Water dispersible magnetic reduced graphene oxide composites for arsenic removal. *ACS Nano* 4(7):3979–3986
- Choong YST, Chuah GT, Robia HY, Koay LFG, Azni I (2007) Arsenic toxicity, health hazards and removal techniques from water: an overview. *Desalination* 217:139–166
- Chu KH (2003) Prediction of two-metal biosorption equilibria using a neural network. *Eur J Miner Process Environ Prot* 3:119–127
- Daus B, Wennrich R, Weiss H (2004) Sorption materials for arsenic removal from water: a comparative study. *Water Res* 12(2948):2954
- Dora TK, Mohanty YK, Roy GK, Sarangi B (2013) Adsorption studies of As(III) from wastewater with a novel adsorbent in a three-phase fluidized bed by using response surface method. *J Environ Chem Eng* 1(3):150–158
- Elmolla ES, Chaudhuri M, Eltoukhy MM (2010) The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. *J Hazard Mater* 179:127–134
- Erhan D, Kobya M, Elif S, Ozkan T (2004) Adsorption kinetics for Cr (VI) removal from aqueous solution on activated carbon prepared from Agro wastes. *Water SA* 30(4):533–541
- Fallah-Mehdipour E, Haddad BO, Mariño MA (2013) Prediction and simulation of monthly groundwater levels by genetic programming. *J Hydro Environ Res* 7(4):253–260
- Foo KY, Hameed BH (2010) Insights into the modeling of adsorption isotherm systems. *Chem Eng J* 156:2–10
- Garg A, Garg A, Tai K, Sreedeeep S (2014) An integrated SRM-multi-gene genetic programming approach for prediction of factor of safety of 3-D soil nailed slopes. *Eng Appl Artif Intell* 30:30–40
- Gupta A, Chauhan VS, Sankararamakrishnan N (2009) Preparation and evaluation of iron–chitosan composites for removal of As(III) and As(V) from arsenic contaminated real life groundwater. *Water Res* 43:3862–3870
- Hameed BH, Din AM, Ahmad AL (2006) Adsorption of methylene blue onto bamboo based activated carbon: kinetics and equilibrium studies. *J Hazard Mater* 137(3):695–699
- Hao J, Han MJ, Wang C, Meng X (2009) Enhanced removal of arsenite from water by a mesoporous hybrid material – Thiol-functionalized silica coated activated alumina. *Microporous Mesoporous Mater* 124:1–7
- Holzwarth U, Gibson N (2011) The Scherrer equation versus the ‘Debye-Scherrer equation’. *Nat Nanotechnol* 6: 534
- Iesana CM, Capat C, Ruta F, Udrea I (2008) Evaluation of a novel hybrid inorganic/organic polymer type material in the Arsenic removal process from drinking water. *Water Res* 42:4327–4333
- Jain CK, Ali I (2000) Arsenic: occurrence, toxicity and speciation techniques. *Water Res* 34:4304–4312
- Khajeh A, Modarress H (2010) Prediction of solubility of gases in polystyrene by adaptive neuro-fuzzy inference system and radial basis function neural network. *Expert Syst Appl* 37:3070–3074
- Kotti IP, Sylaios GK, Tsihrintzis VA (2013) Fuzzy logic models for BOD and COD removal prediction in free-water surface constructed wetlands. *Ecol Eng* 51:66–74
- Krishnaiah K, Shahabudeen P (2012) Applied design of experiments and taguchi methods. PHI Learning, Business & Economics
- Laha SC, Ryou R (2003) Synthesis of thermally stable mesoporous cerium oxide with nano crystalline frameworks using mesoporous silica templates. *Chemical Commun* 17:2138–2139
- Ma Y, Zheng YM, Chen JP (2011) A zirconium based nanoparticle for significantly enhanced adsorption of arsenate: Synthesis, characterization and performance. *J Colloid Interface Sci* 354:785–792
- Mandal S, Padhi T, Patel RK (2011) Studies on the removal of arsenic (III) from water by a novel hybrid material. *J Hazardous Mater* 192:899–908
- Pai TY, Yang PY, Wang SC, Lo MH, Chiang CF, Kuo JL, Chu JH, Su HC, Yu LF, Hu HC, Chang YH (2011) Predicting effluent from the waste water treatment plant of industrial park based on fuzzy network and influent quality. *Appl Math Model* 35:3674–3684
- Pan B, Li Z, Zhang Y, Xu J, Chen L, Dong H, Zhang W (2014) Acid and organic resistant nano-hydrated zirconium oxide (HZO)/polystyrene hybrid adsorbent for arsenic removal from water. *Chem Eng J* 248:290–296
- Parhizgar H, Dehghani MR, Eftekhari A (2013) Modeling of vaporization enthalpies of petroleum fractions and pure hydrocarbons using genetic programming. *J Pet Sci Eng* 112:97–104

- Ross PJ (1988) Taguchi techniques for quality engineering. McGraw-Hill Book Company, New York
- Roy RK (1990) A primer on Taguchi method. Van Nostrand Reinhold, New York
- Russo M, Leotta G, Pugliatti PM, Gigliucci G (2014) Genetic programming for photovoltaic plant output forecasting. *Sol Energy* 105:264–273
- Saha D, Bhowal A, Dutta S (2010) Artificial neural network modelling of fixed bed biosorption using radial basis approach. *Heat Mass Transfer* 46:431–436
- Shevade S, Ford R (2004) Use of synthetic zeolites for arsenate removal from pollutant water. *Water Res* 38: 3197–3204
- Shigetomi Y, Hori Y, Kojima T (1980) The removal of arsenate in waste water with an adsorbent prepared by binding hydrous iron(III) oxide with polyacrylamide. *Bull Chem Soc Jpn* 53:1475–1476
- Tiwari D, Lee SM (2012) Novel hybrid materials in the remediation of ground waters contaminated with As(III) and As(V). *Chem Eng J* 204–206:23–31
- Vatutsina OM, Soldatov VS, Sokolova VI, Johann J, Bissen M, Weissenbacher A (2007) A new hybrid (polymer/inorganic) fibrous sorbent for arsenic removal from drinking water. *React Funct Polymers* 67:184–200
- Vaughan RL, Reed BE (2005) Modelling of As (V) removal by an iron oxide impregnated activated carbon using the surface complexation approach. *Water Res* 39:1005–1014
- Wan Ngah SW, Teong CL, Hanafiah MKAM (2011) Adsorption of dyes and heavy metal ions by chitosan composites a review. *Carbohydr Polym* 83:1446–1456
- Wang S, Mulligan CN (2006) Occurrence of arsenic contamination in Canada: sources, behaviour and distribution. *Sci Total Environ* 366:701–721
- World Health Organization (WHO) (2008) Guidelines for drinking-water quality, Recommendations, WHO (3)
- Yeganeh B, Shafie M, Motlagh P, Rashidi Y, Kamalan H (2012) Prediction of CO concentrations based on a hybrid partial least square and support vector machine model. *Atmos Environ* 55:357–365