Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2018.21930

103 (2018) 182–198
January

A comparative study of hybrid artificial neural network models for predicting Cr(VI) adsorption onto activated carbon

Clint Sutherland[a,*], Beverly S. Chittoo[a], Chintanapalli Venkobachar[b]

[a]*Project Management and Civil Infrastructure Systems, The University of Trinidad and Tobago, Trinidad and Tobago (WI), Tel. 868 497 5744, email: clint.sutherland@utt.edu.tt (C. Sutherland), Tel. 868 491 6879, email: beverly.chittoo@utt.edu.tt (B.S. Chittoo)*
[b]*Formerly with Department of Civil and Environmental Engineering, The University of the West Indies, St. Augustine, Trinidad and Tobago (WI), email: cvenkobachar@gmail.com*

Received 19 July 2017; Accepted 20 January 2018

ABSTRACT

A comparative analysis of the application of the standard, hybrid serial gray-box, and hybrid parallel gray-box artificial neural networks was carried out to predict Cr(VI) adsorption by activated carbon. The dataset was developed through batch kinetic experiments with varying operational parameters. The major reaction transport mechanism was found to be intraparticle diffusion while the major attachment mechanism was chemical bonding. Adsorption kinetics was well represented by the diffusion-chemisorption (DC) model while desorption kinetics followed the pseudo-second order model. It was discovered that the overall DC kinetic rate, K_{DC} was inversely proportional to both the effective diffusion coefficient as well as the diffusivity and proves useful to indirectly assess diffusional effects created by changing operational parameters. The DC kinetic model was subsequently instituted into the hybrid models. The development of the hybrid neural networks was built-in with the joint effect of operating parameters of time and particle size for the standard (SANN) and parallel neural network (PANN), together with the kinetic parameters of the DC model such as overall rate, initial rate and sorption capacity for the serial neural network (DC-ANN). The comparative performance of the networks was subsequently evaluated using error functions and the Bland-Altman plot. The hybrid neural networks produced the best prediction to the target data. Sensitivity analysis revealed a substantial positive influence on the DC-ANN model due to the inclusion of the kinetic equation parameters. Consequently, the DC-ANN model required significantly less computational effort.

Keywords: Activated carbon; Chromium (VI); Artificial neural network; Adsorption; Black-box model; Gray-box model; Kinetic studies

1. Introduction

The use of heavy metals particularly chromium continues to be an indispensable component in industrial applications such as electroplating, textile dying and metal finishing [1]. The most common valency states of chromium on the earth's surface and near-surface environments are trivalent [Cr^{+3} and $Cr(OH)^{2+}$] and hexavalent [$HCrO_4^-$, CrO_4^{2-} or $Cr_2O_7^{2-}$] [2]. According to Patterson as cited in [3], the concentration of Cr(VI) in industrial wastewaters can vary from 0.5 mg/L to 270,000 mg/L. As a result of the high toxicity of Cr(VI), it has been made a priority metal of public health significance. World Health Organization (WHO) recommends that the level of Cr(VI) in wastewater be regulated below 0.05 mg/L [2]. Some of the primary treatment methods adopted for industrial-scaled remediation of such metal ions are ion exchange, chemical precipitation, and adsorption [4]. Adsorption remains one of the most effective and promising techniques for chromium removal [5]. This can be attributed to the low attainable concentration levels [6].

*Corresponding author.

The adsorption potential of a commercial activated carbon derived from bituminous coal was examined in this study. The material possesses a high surface area and pore distribution as well as the presence of functional groups such as the carboxylic groups [7]. Batch adsorption and desorption kinetic studies were used to elucidate and model the sorption process. The fitting of the experimental data to kinetic models gives valuable insights into the mechanisms of adsorption [8]. Thus, kinetic operational parameters were varied to evaluate the optimal adsorption conditions. The development of predictive models can not only save cost and time required for experimental studies but also improve the efficiency in up-scaling and the application of full-scale adsorption systems [9]. However, according to Saraf and Vaidya [10], it is difficult to predict and simulate adsorption processes using the classical single variable optimization method. This is mainly due to the interaction of multiple sorption process variables, and highly non-linear relationships.

Artificial neural networks (ANN) are computational models developed following the functionality of the biological neural system [11]. The first artificial neuron was produced in 1943 by McCulloch and Pits [12]. The first practical application of backpropagation was for estimating a dynamic model to predict nationalism and social communications in 1974 [13]. In recent years, the back propagation artificial neural network (ANN-BP) has gained wide popularity in the application of complex engineering simulations [14]. This is largely because the ANN has the capability of learning highly non-linear and complex relationships, which provides an advantage over traditional methods of optimization [15]. Unfortunately, they (black-box ANNs) often suffer from drawbacks such as inconsistency of the output and the need for a reasonable set of experimental data, which is not always available [16]. Psichogios and Ungar [17] also iterated that standard neural network performs arbitrarily well in approximating a dynamic system if a large data set is available. The authors went on to explain that when the training dataset size is small, then the state space is not sampled sufficiently densely, and a traditional neural network relies heavily on interpolation to approximate the dynamic system. Hussain and Rahman [18] further described a typical neural network as having many internal parameters (the weights and biases), and these could lead to over-fitting of noise as well as the underlying function. This results in poor generalization. In order to achieve a good prediction where training data are sparse or absent, prior information on the behavior of the function should be introduced [19]. These weaknesses have encouraged recent research to focus on combining ANN with other components, such as first principle model, fuzzy logic, and Kalman filtering [17].

In this study, a standard ("black-box") artificial neural network (SANN) was developed and compared to a parallel hybrid neural network (PANN) modeling method (gray-box) and a series hybrid neural network modeling method (gray-box). Numerous authors have successfully demonstrated the use of the SANN to predict the adsorption of Cr(VI) onto various types of activated carbon (Table 1). Consequently, to assess the performance of all competing ANN methods, a batch adsorption system designed for the removal of Cr (VI) from aqueous solution using commercial activated carbon was selected.

In the parallel hybrid model structure, the neural network is used to estimate the difference between mechanistic model predictions (adsorption performance) and the corresponding experimental performance (i.e., the residuals) [24]. The neural network is then built-in with the joint effect of operating parameters viz. time and particle size as input and the residuals as the target. The PANN prediction of residuals are subsequently applied as corrections to the mechanistic model predicted performance. A serial approach was developed which involved instituting the mechanistic model parameters along with the experimental data as input to the ANN structure and adsorption performance was set as the target. This approach for the prediction of batch kinetic systems has not been reported in the literature. However, a similar technique was reported by Faur-Brasquet and Cloirec [25], whereby parameters related to the adsorbate-adsorbent affinity in a batch reactor were introduced in the input layer of the neural network (intraparticular coefficient and Freundlich parameters) together with experimental data to improve the prediction of column breakthrough curves. Consequently, the success of such a serial hybrid approach crux on the selection of a suitable mechanistic model which adequately represents the entire adsorption reaction.

The objectives of this work are: 1) to describe the process of adsorption through batch kinetic and desorption studies; 2) determine and verify the most appropriate kinetic model through mathematical analysis and elucidation of rate-limiting steps in the adsorption process; and 3) through comparative analysis of black and gray-box ANN models develop an efficient neural network for the prediction of batch kinetics under conditions of limited data.

2. Materials and methods

2.1. Preparation of adsorbent

The granular activated carbon used in these experiments was Calgon Filtrasorb 300. According to the manufacturer

Table 1
Application of Artificial Neural Network for the prediction of Cr(VI) adsorption

Adsorbent	ANN Type	Algorithm	Neurons	R^2	Reference
Commercial PAC	FFANN-BP	quasi-Newton backpropagation	10	0.968	[20]
GAC from Coconut Shell	FFANN-BP	–	20	0.982	[21]
Commercial GAC	FFANN-BP-GA	–	10	0.980	[22]
Rice Husk Carbon	FFANN-BP	Levenberg–Marquardt backpropagation	5	0.987	[23]

Calgon Corp. Canada, this carbon has an iodine number of 900 mg/g, an intrinsic pore volume of 0.85 and a BET surface area of 950–1050 m^2/g. The carbon was crushed, sieved, repeatedly washed with distilled water, dried and stored in a desiccator. The average particle size of activated carbon retained on a sieve was calculated as the geometric mean of the diameter openings in two adjacent sieves in the stack. The Geometric Mean Size, (GMS) is expressed as (diameter of upper sieve × diameter of lower sieve)$^{0.5}$ [26].

2.2. Preparation and testing of adsorbate

The chemical stock solution was prepared using potassium chromate. Reaction solutions were prepared using distilled water (prepared by a Corning Mega Pure System MP^{-1}) of pH approximately 7 and conductivity < 5 µm hos/cm. Solution pH was adjusted using appropriate solutions of HCl and NaOH and kept constant throughout the reaction by a 0.01 M acetate buffer for pH 4–5.5 and phosphate buffer for pH 8. Solution pH was measured by a pH meter (Accumet Research-AR10, Fisher Scientific). Chromium (VI) concentrations were determined by the Atomic Adsorption Spectrophotometer (Perkin-Elmer 3030B).

2.3. Experimental procedure

2.3.1. Adsorption kinetic studies

Kinetic studies were conducted using the parallel method according to EPA OPPTS method 835.1230 [27]. The study of metal uptake was carried out in duplicate at room temperature (26 ± 2°C) in a batch reactor with an adsorbent mass 1.0 g/L and spiked with 50 ml of 50 mg/L synthetic metal ion solution. Sorbent masses were accurate to ±0.001 g and solution volumes to ±0.5 ml. Identical reaction mixtures were prepared for each time interval, agitated to maintain complete mixed conditions on a mechanical shaker and removed at predetermined time intervals [28]. The adsorbent was then separated by using Whatman No. 2 qualitative filter paper.

2.3.2. Desorption kinetic studies

Batch desorption studies were carried out after adsorption in a 50 mL reaction mixture which comprised 1.0 g/L absorbent and 50 mg/L adsorbate. The media was separated and added to an equivalent volume (50 mL) of desorbing solution (0.1 M HCl and Distilled water) and agitated until equilibrium was established. The supernatant was subsequently tested for residual Cr(VI) ions.

2.3.3. Adsorption yield and concentration

The adsorption yield or the ratio of adsorbed metal ion concentration to the initial metal ion concentration was calculated from Eq. (1):

$$\%Adsorption = \frac{C_0 - C_t}{C_0} * 100 \quad (1)$$

where C_o (mg/L) is the initial concentration of metal ions in solution, C_t (mg/L) is the concentration of metal ions in solution at any time t.

The concentration of metal ions on activated carbon was determined using the mass balance equation expressed as follows:

$$q_t = \frac{(C_o - C_t)}{m} * V \quad (2)$$

where q_t (mg/g) is mg of adsorbate adsorbed per g of sorbent at any time t, C_o (mg/L) is the initial adsorbate concentration in solution, C_t (mg/L) is the adsorbate concentration in solution at any time t, V (L) is the volume of synthetic adsorbate solution, and m (g) is the mass of the adsorbent.

2.4. Modelling approach

2.4.1. Lagergren model

In 1898, Lagergren as cited in [28], developed a first-order rate equation to describe the kinetic process of oxalic acid and malonic acid onto charcoal. Ho and McKay [29] described the equation as pseudo-first order. The Lagergren equation is given by the derivatives:

$$q_t = q_e(1 - \exp^{-K_{PFO}t}) \quad (3)$$

where K_{PFO} (min^{-1}) is the rate constant of pseudo-first-order adsorption, q_t (mg/g) is mg of adsorbate absorbed per gram of sorbent at any time t, and q_e (mg/g) is the amount of adsorption at equilibrium,

2.4.2. Pseudo-second order model

The pseudo-second order equation was developed for the sorption of divalent metal ions onto peat moss [30]. According to Ho and McKay [31], the model is based on pseudo-second order chemical reaction kinetics. The pseudo-second order equation is as follows:

$$q_t = \frac{K_{PSO} q_e^2 t}{1 + K_2 q_e t} \quad (4)$$

where q_t is mg adsorbate per gram of sorbent at any time, q_e (mg/g) is the amount of adsorption at equilibrium, and K_{PSO} is the pseudo-second-order rate, (g/mg-min).

The initial sorption rate, h, as $t \to 0$ is expressed as:

$$h = (K_{PSO}) q_e^2 \quad (5)$$

2.4.3. Weber and Morris intraparticle diffusion model

Weber and Morris [32] proposed that the rate of intraparticle diffusion varies proportionally with the half power of time and is expressed as Eq. (6). According to Ofomaja [33], when the model is linearised, q_t (mg/g) is the adsorbate uptake at time, t (min), K_{id} (mg/g-t$^{1/2}$) is the rate constant of intraparticle transport and the intercept c (mg/g), is taken to be proportional to the extent of the boundary layer thickness. If the rate limiting step is intraparticle diffusion, a plot of solute adsorbed against the square root of the contact time should yield a straight line passing through the origin [32].

$$q_t = K_{id}\left(t^{1/2}\right) + c \tag{6}$$

The constant K_{id} is related to the intraparticle diffusivity, D by the following equation [34]:

$$K_{id} = \left[\frac{3q_e}{d_p}\right]\sqrt{\frac{D}{\pi}} \tag{7}$$

where d_p (cm) is the diameter of the particle, and q_e (mg/g) is the solid phase concentration at equilibrium.

2.4.4. Boyd model

The Boyd's model also known as Boyd's film-diffusion model was proposed for intraparticle diffusion in a spherical particle. It is frequently used to gain insight related to the mechanisms of adsorption [35]. The general solution of the model is given by Eq. (8):

$$F = 1 - \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp(-n^2 Bt) \tag{8}$$

where $F = q/q_e$, the fractional approach to equilibrium, B is the rate coefficient related to the effective diffusion coefficient, (D_{eff}) and the particle radius, as $B = \pi 2D_{eff}/R^2$. By applying Fourier transform, the relation between the fractional approach to equilibrium F and the rate coefficient B is written as:

$$Bt = \left(\sqrt{\pi} - \sqrt{\pi - \left(\pi^2 F/3\right)}\right)^2 \quad \text{for } F < 0.85 \tag{9a}$$

and $Bt = -0.498 - ln\,(1 - F)$ for $F > 0.85$ (9b)

A plot of Bt vs. time is used to distinguish the film and particle diffusion controlled rates of adsorption. If a straight line passing through the origin is obtained, then particle diffusion mechanism governs the rate of adsorption otherwise it is governed by film diffusion.

2.4.5. Diffusion-chemisorption model

The diffusion-chemisorption kinetic model was developed to simulate biosorption of heavy metals onto heterogeneous media [36]. It is based on the assumption that both diffusion and chemisorption control the adsorption process. To obtain the derivatives a correlation was made where the rate of change of concentration of the solid phase, (q_t, mg/g) is equated as a function of rate of mass transfer of adsorbate from the fluid phase to the adsorption site (K_{DC}, mg/g-t$^{0.5}$); the equilibrium sorption capacity (q_e, mg/g); and time to the power of n^{-1}, where $n = 0.5$. The model is presented as Eq. (10):

$$q_t = \frac{1}{\frac{1}{q_e} + \frac{t^{0.5-1}}{K_{DC}}} \tag{10}$$

where K_{DC} (mg/g-t$^{0.5}$) is the diffusion-chemisorption constant, q_t (mg/g) is the mass of ions adsorbed per gram of sorbent at any time and q_e (mg/g) is the adsorption at equilibrium. The initial slope of the kinetic curve was found to be a function of the diffusion-chemisorption constants K_{DC} and q_e. The following relationship is obtained by assuming a linear region, as $t \to 0$. The initial rate k_i (mg/g-t) is presented as follows:

$$k_i = {K_{DC}}^2 / q_e \tag{11}$$

2.4.6. Theory of artificial neural network

In this study, multi layer feed-forward neural networks trained by backpropagation was adopted using the neural network toolbox from MATLAB (R2012a). The backpropagation training algorithm is one of the most widely used algorithms in multilayer neural networks [2,37]. It has been successfully applied in modeling adsorption systems such as phenol adsorption by activated carbon [38], and cadmium adsorption by inorganic sorbents [39]. The ANN architecture is formed by the number of layers, number of neurons in different layers, transfer function and initial weights which interconnect each layer. Lek and Guegan [40] explained that the non-linear elements (neurons) are arranged in successive layers, and the information flows unidirectional, from input layer to output layer, through the hidden layers. The input values are weighted before entering the hidden layer while the bias units improve convergence by adding a constant term to the weighted sum. The output is based on the sum of the weighted values from the input layer and modified by a transfer function. Table 2 presents the transfer functions at both the hidden and output layer used in this study to optimise the ANN models. After the network's output is compared with the target vector, error values for the hidden units are calculated, and the weights are changed. The backward propagation starts at the output layer and moves backward through the hidden layers until it reaches the input layer [41]. The flow chart of the backpropagation learning algorithm is illustrated in Fig. 1 [42].

Duarte and Saraiva [43] iterated that the modelling of multiple input/output systems with pure black-box tools (Fig. 2a.) permit only limited extrapolation beyond the domain of the data from which they were derived thereby potentially resulting in unreasonable results. Consequently, large amounts of data are required in order to

Table 2
Transfer function

Name of transfer function	Algorithm
Logsig	$f(n) = 1/\left(1 + exp(-n)\right)$
Tansig	$f(n) = \left[2/\left(1 + exp(-2*n)\right)\right] - 1$
Purelin	$f(n) = n$

lead to reliable predictions. An approach that combines the kinetic properties of the reaction with that of the black-box techniques serves to integrate the best of both paradigms. Such combining can be pursued through serial or parallel strategies [44]. Nascimento et al. [45] explained that although the neural network absorbs most of the nonlinearity of the process, it is highly dependent on the quality and the range of the input data. The authors presented a promising alternative by applying the neural network to the parts of the problem where it is not possible to use phenomenological models. Thus, a hybrid model in the form of a serial gray-box model and referred here as DC-ANN model is employed as presented in Fig. 2b. Su et al. [46] proposed a parallel structure to model a chemical reactor system. The parallel approach was also successfully used by Hussain and Rahman [18], to predict the pore formation in foods during drying. In this study, the parallel artificial neural network structure outlined in Fig. 2c. was used to develop a predictive model.

2.5. Error analysis

The goodness of fit of the various kinetic models to the experimental data was evaluated using the linear coefficient of determination (R^2), as well as the chi-square test, root mean square error (RMSE), relative percent error (RPE), normalized mean bias error (NMBE %) and mean square error (MSE) which are presented in Table 3.

3. Results and discussion

3.1. Selection of the most appropriate kinetic model

The analysis of the experimental kinetic data was carried out using four (4) kinetic models: the pseudo-first order model [29]; pseudo-second order model [30]; Weber and Morris intraparticle diffusion model [32]; and the diffusion-chemisorption model [36]. The applicability of

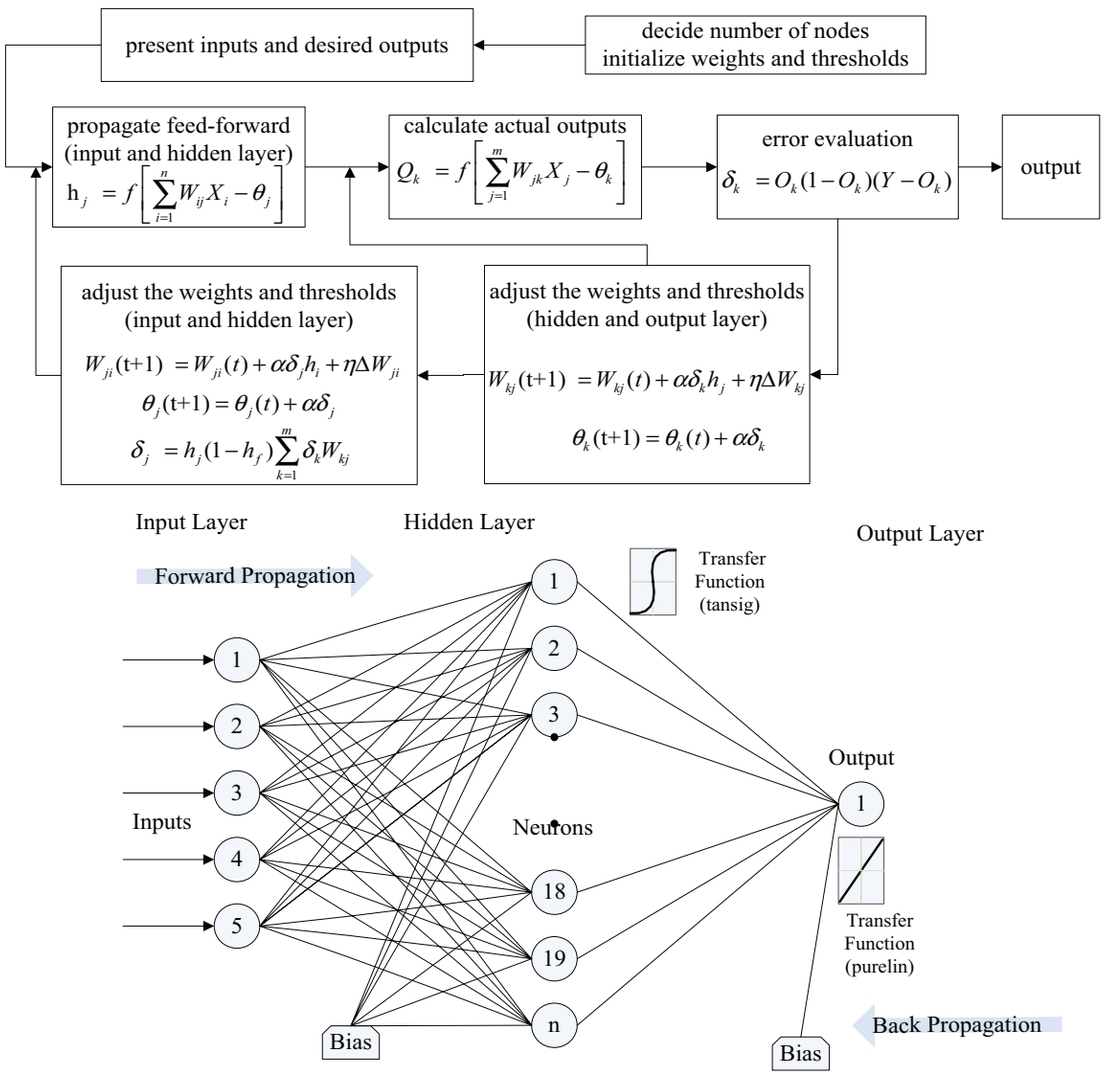

Fig. 1. Structure of backpropagation ANN model.

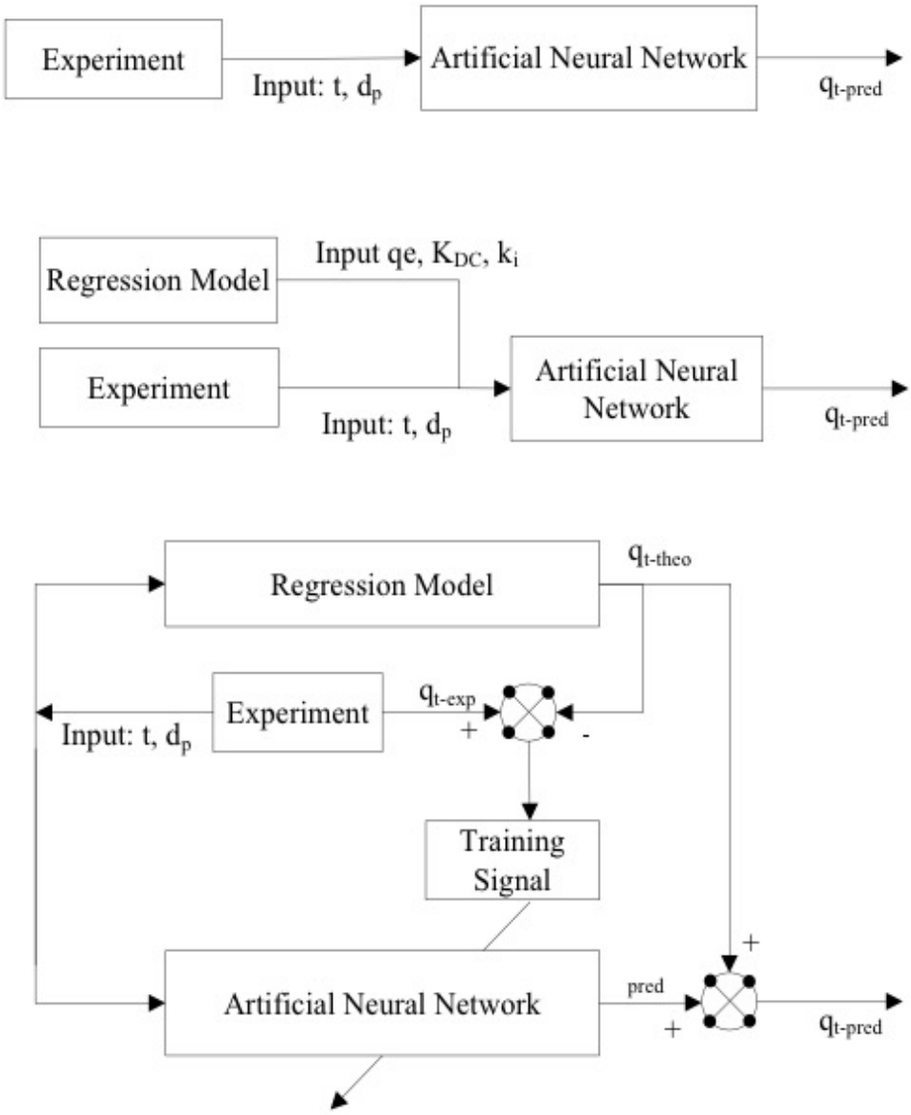

Fig. 2. (a) Schematic diagram of the standard artificial neural network (SANN) (b) Schematic diagram of the gray-box serial artificial neural network (DC-ANN), (c) Schematic diagram of the gray-box parallel artificial neural network (PANN).

these kinetic models to simulate the experimental data was assessed by non-linear regression using the Levenberg–Marquardt algorithm. The resulting theoretical prediction curves were compared to the primary kinetic curves using the RPE, chi-squared test, and RMSE error functions, and are presented in Table 4. The RPE is a matrix which shows how large the error is in relation to the experimental value. The chi-square test is used to determine whether there is a significant difference between the experimental and predicted values. The lower the error value, the greater the conformity. RMSE statistics computes the error by comparing the target or experimental values with the predicted outputs. A lower RMSE indicates good prediction, but this statistic is biased towards high error values [47]. The pseudo-first order model produced the highest error values and as such implies that the reaction is more complex than pseudo-first order kinetics. The intraparticle diffusion model also did not produce a good simulation of the data. However, it is noted that as particle size increased, the conformity of the model also increased. This infers that diffusion may be an operative mechanism as operational parameters vary. The kinetic data was found to conform best to the diffusion-chemisorption model which produced the lowest error values for all particle size studied.

3.2. Verification of the applicability of the diffusion-chemisorption model

3.2.1. Effect of pH on Cr(VI) adsorption

Hexavalent chromium exists in different forms in aqueous solution, and the stability of these forms is mainly

Table 3
Error functions

Error functions	Expression	Equation number
Relative percent error	$RPE\% = \frac{1}{N}\sum_{i=1}^{N}\frac{\left[\left\lvert\left(q_{e_i}\right)_{pred} - \left(q_{e_i}\right)_{\exp}\right\rvert\right]}{\left(q_{e_i}\right)_{\exp}} * 100$ where, N is the number of experimental points	(12)
Chi-square test	$\chi^2 = \sum_{i=1}^{N}\frac{\left(\left(q_{e_i}\right)_{pred} - \left(q_{e_i}\right)_{\exp}\right)^2}{\left(q_{e_i}\right)_{pred}}$ where, N is the number of experimental points.	(13)
Root mean square error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\left(q_{e_i}\right)_{\exp erimental} - \left(q_{e_i}\right)_{predicted}\right)^2}$ where, N is the number of experimental points.	(14)
Coefficient of determination (R^2)	$R^2 = \frac{\sum_{i=1}^{N}\left(\left(q_{e_i}\right)_{\exp} - q_{e_{\exp,mean}}\right)^2 - \sum_{i=1}^{N}\left(\left(q_{e_i}\right)_{\exp} - \left(q_{e_i}\right)_{pred}\right)^2}{\sum_{i=1}^{N}\left(\left(q_{e_i}\right)_{\exp} - \left(q_{e_i}\right)_{pred}\right)^2}$ where, N is the number of experimental points.	(15)
Mean square error (MSE)	$MSE = \left(\frac{1}{N}\sum_{i=1}^{N}\left[\left(q_{e_i}\right)_{\exp} - \left(q_{e_i}\right)_{pred}\right]^2\right)$ where, N is the number of experimental points.	(16)
Normalized mean bias error (NMBE %)	$NMBE\% = \frac{1}{N}\sum_{i=1}^{N}\frac{\left[\left(q_{e_i}\right)_{pred} - \left(q_{e_i}\right)_{\exp}\right]}{q_{e_{\exp,mean}}} * 100$ where, N is the number of experimental points.	(17)

Table 4
Performance comparison of kinetic models using error functions of non-linear regression for chromium uptake by activated carbon

Particle size	Model	RPE	Chi	RMSE
0.21 mm	Pseudo-first order	12.8673	18.0596	2.9059
	Pseudo-second order	10.5336	4.4949	2.1708
	Intraparticle diffusion	23.6147	17.1236	5.1765
	Diffusion-chemisorption	3.2686	0.2130	0.7818
0.65 mm	Pseudo-first order	12.3622	9.6768	1.4542
	Pseudo-second order	11.8954	5.8241	1.3698
	Intraparticle Diffusion	14.1763	2.6454	2.3151
	Diffusion-chemisorption	5.6776	0.4653	1.1257
1.05 mm	Pseudo-first order	12.0908	3.8481	0.8244
	Pseudo-second order	9.7836	2.4902	0.6637
	Intraparticle diffusion	4.3946	0.3158	0.9336
	Diffusion-chemisorption	5.1087	0.2308	0.6573

dependent on the solution pH. The predominant Cr(VI) species are the acid chromate ion $HCrO_4^-$ and the chromate ion CrO_4^{2-}. Below pH 4.0, the $HCrO_4^-$ complex is the major form while at pH above 9.0 the most abundant species is CrO_4^{2-} [3]. The removal efficiency vs. pH is presented in Fig. 3. Adsorption of Cr(VI) was not observed at pH above 6.5, indicating that Cr(VI) was adsorbed as $HCrO_4^-$, the predominant species between pH 1.5 and 4.0.

The surface of activated carbon contains functional groups which can be protonated or deprotonated, depend-

ing on the pH of the aqueous media [48]. As the pH of functional groups is reduced below its pKa value, the surface groups become protonated, and the surface charge approaches zero. Consequently, any further reduction in pH results in surface charge reversal [49]. Fig. 3 shows that as the pH is reduced below 4.5, the removal efficiency of Cr(VI) anions increases significantly. This increase may be accounted for by surface charge reversal causing increased electrostatic interactions between the surface and the Cr(VI) complex. The maximum removal was observed at pH 2.5. A similar observation was reported by Singh and Tiwari [50], for the uptake of Cr(VI) by carbon slurry. As such, it may be surmised that the variation in Cr(VI) adsorption is due in large part to the influence of pH on the surface chemistry of the activated carbon.

3.2.2. Chemical desorption studies

Chromium desorption efficiencies were examined using distilled water and 1.0 M HCl [51] as the desorbing eluents. Complete desorption of Cr(VI) was not attained by any of the eluents. Kinetic analysis of the desorption data was performed using non-linear regression. A comparison of the goodness of fit by each model was assessed using error functions and is presented in Table 5. The data representing desorption by both HCl and distilled water conformed best to the pseudo-second order model. The overall and initial rate of desorption was determined using Eqs. (4) and (5) respectively, and is also presented in Table 5. HCl was the most effective eluent producing a significantly higher rate of Cr(VI) release. The initial release of Cr (VI) ions by HCl was approximately 40 times greater than that of distilled water. The distilled water wash achieved a maximum release of 35% while the HCl wash attained 70% release after 165 min of agitation. The time taken to reach desorption equilibrium was approximately twice that required for adsorption. These results imply that the mechanism of adsorption, for at least 35% of the sorbed anions, is the result of physical adsorption, while sorption of the remaining ions may be attributed to ion exchange or chemical bonding. Similar results were reported by Anandan and Janakiram [51], for the desorption of Cr(VI) from activated carbon made from Albizia lebbeck stem. It is postulated that a higher concentration of HCl, beyond 1.0M, may increase the desorption efficiency.

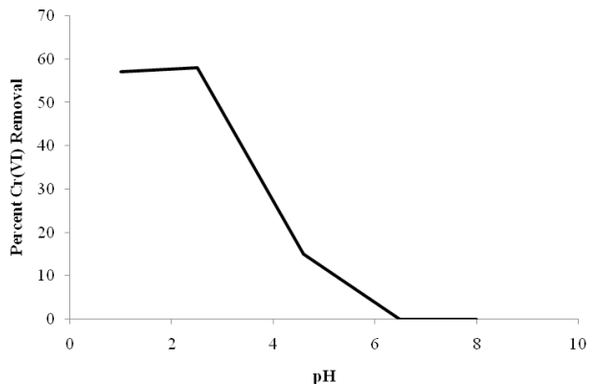

Fig. 3. Effect of pH on Cr(VI) removal efficiency by activated carbon.

3.2.3. Effect of particle size

The influence of activated carbon particle size on sorption kinetics was studied and analysed by non-linear regression using the diffusion-chemisorption model. The overall rate of adsorption, K_{DC}, shown in Table 6, increased as particle size decreased. This phenomenon was expected since a greater surface area of the adsorbent was available for sorption. It was also observed that the time taken to reach equi-

Table 6
Diffusion-chemisorption rate parameters for varying particle sizes

Particle Size, d_p (cm)	k_i (mg/g-s)	K_{DC} (mg/g-s$^{0.5}$)	D (cm^2/s)	D_{eff} (cm^2/s)
0.021	0.0525	0.6504	9.882E-08	9.636E-09
0.065	0.0066	1.5278	2.964E-07	2.852E-08
0.105	0.0013	2.4907	4.953E-07	4.790E-08

Table 5
Desorption kinetic modeling for different eluents

Eluent	Model	Error functions			Pseudo-second order model	
		RPE	Chi	RMSE	Overall rate, K_{PSO} (g/mg-min)	Initial rate, h (mg/g-min)
HCl	Pseudo-first order	6.9136	0.3198	0.6474	0.0865	5.7947
	Pseudo-second order	1.6234	0.1722	0.4551		
	Intraparticle diffusion	37.1867	15.0453	2.4255		
	Diffusion-chemisorption	3.4778	0.0021	0.0464		
H_2O	Pseudo-first order	162.1643	4.0849	1.6024	0.004	0.1271
	Pseudo-second order	3.4959	0.0163	0.0756		
	Intraparticle diffusion	37.6794	0.3071	0.2869		
	Diffusion-chemisorption	46.3148	0.3273	0.2395		

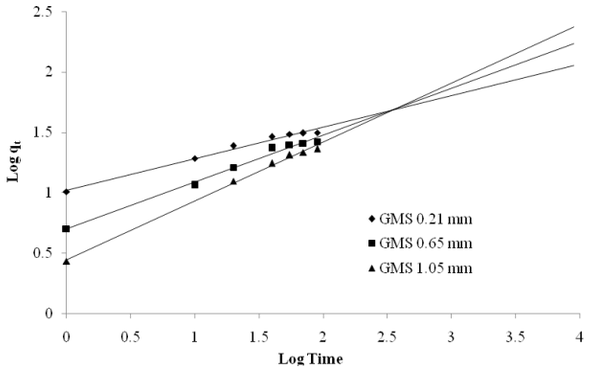

Fig. 4. log-log plot of Cr(VI) uptake by activated carbon for various particle sizes.

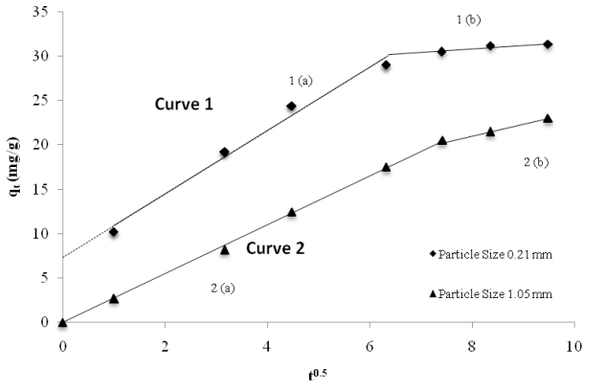

Fig. 5. Plot of intraparticle diffusion model for various particle size.

librium increased as particle size increased. As such, the internal surface area associated with adsorption could have been at least partially occluded. This may be attributed to the difference in diameter of the large $HCrO_4^-$ ion and that of the micro and mesopores or the impact of intraparticle diffusion. To expound this phenomenon a log-log plot of the kinetic data for different particle size is shown in Fig. 4. The plot shows the curves to eventually converge when time is increased, indicating that each particle size will eventually attain the same sorption capacity. Consequently, the variation in equilibrium time with varying particle size verifies the presence and effect of intraparticle diffusion.

3.2.4. Elucidating the mechanisms of adsorption

The phenomenon whereby ions are transferred from a liquid phase to a solid phase usually involves the following steps: (i) boundary layer mass transfer across the liquid film surrounding the particle; (ii) internal diffusion/mass transport within the particle boundary as pore and/or solid diffusion; and (iii) adsorption within the particle and on the external surface [52].

3.2.4.1. Transport mechanisms

A plot of q_t vs. $t^{0.5}$ in accordance with the Weber and Morris model for two sorbent sizes viz. GMS 0.21 mm and 1.05 mm is shown in Fig. 5. It is observed that the plot depicting the largest particle size passed through the origin, indicating the dominance of intraparticle diffusion. Similar observations for the adsorption of cadmium onto various particle sizes of bone char have been reported by Choy et al. [53]. As the particle size decreased (which accompanies an increase in surface area and a reduction in pore length), the plots move further from the origin. Such deviation from the origin infers that intraparticle transport is not the only rate limiting step [54]. Previous studies [33,55,56], have reported that this increase in intercept reveals the growing effect of the boundary layer. The plot of 0.21 GMS (curve 1) shows two distinct slopes. The first slope depicted as 1(a), occurs from 1–40 min and reveals the influence of intraparticle diffusion which may be rate limiting. Some researchers have reported that the final slope depicted as 1(b), corresponds to the slowing of the reaction, possibly due to a reduction in concentration gradient as the reaction approaches equilibrium [57,58]. It is therefore evident that both film and intraparticle diffusion has some measured influence on the reaction rate-limiting step and the magnitude of their influent varies with reaction time.

The corresponding intraparticle diffusivity, *D*, was calculated using Eq. (7), and is shown in Table 6. An increase in diffusivity is observed as particle size increased. A similar trend was also reported by Shanthi et al. [59], for the adsorption of dye on activated carbon. Due to the larger particle size, there is less removal by surface adsorption. Consequently, this increased number of Cr(VI) ions in solution may have created a concentration gradient between the outer regions of the particle and the inner pores and thus a greater driving force was created.

The contribution of boundary layer or film diffusion was further investigated using the Boyd model given by Eq. (9). A plot of *Bt* vs. *t* was carried out (plot not shown) which revealed that the curve representing the largest particle size, GMS 1.05, passed through the origin. As particle size reduced the plots deviated further from the origin confirming the dominance of film diffusion. This observation not only validates the findings of the Weber and Morris model but also underscores the presence of both film and intraparticle diffusion even though they cannot be sharply demarcated. The corresponding effective diffusion coefficient, D_{eff} was also calculated and is presented in Table 6.

3.2.4.2. Attachment mechanism

Walker and Weatherley [60] conveyed that the rate of adsorption and the magnitude of the effective diffusion coefficient are dependent upon the nature of the adsorption process. The authors stated that for physisorption processes, the magnitude of the diffusion coefficient ranges from 10^{-2} to 10^{-5} (cm^2 /s). Diffusion coefficient for chemisorption systems ranges from 10^{-5} to 10^{-13} (cm^2 /s) due to the stronger bonds holding the molecules tighter to the adsorbent pore walls thus lowering the rate of molecular migration. It was found that D_{eff} lies between 10^{-5} to 10^{-13} (cm^2/s) which infers that chemisorption was an influential attachment mechanism in the sorption process. This finding is in agreement with the DC model and the findings from the preceding desorption studies.

3.2.4.3. Mass transfer

The overall rate of adsorption relies on the slowest stage or stages in the adsorption process. Table 6 reveals that the overall rate of adsorption K_{DC} is inversely proportional to particle size. A plot of $1/K_{DC}$ vs. diffusivity (D) and effective diffusion coefficient (D_{eff}) is presented in Fig. 6. The plot reveals a linear relationship in both instances that passes through the origin. This finding shows that the overall diffusion-chemisorption rate, K_{DC} may be used indirectly to identify the influence of changes in process variables on the diffusivity or effective diffusion coefficient. This further verifies the applicability of the DC model.

3.2.4.4. Rate-limiting steps

The preceding discussion revealed the sorption of Cr(VI) to be a complex reaction. The attachments of Cr(VI) as revealed by desorption studies may be attributed to physical forces for at least 35% of the sorbed ions and 65% chemically bonded. The presence of chemical sorption was further verified through mass transfer analysis using the Boyd model. The limiting transport mechanism was found to involve film diffusion as well as intraparticle diffusion. It was evident from mass transfer analysis that the mechanisms that govern the rate-limiting steps as the reaction proceeds changes with time. Consequently, both transport and attachment mechanisms play a significant role in the adsorption process and thus may account for the superior simulation by the diffusion–chemisorption model.

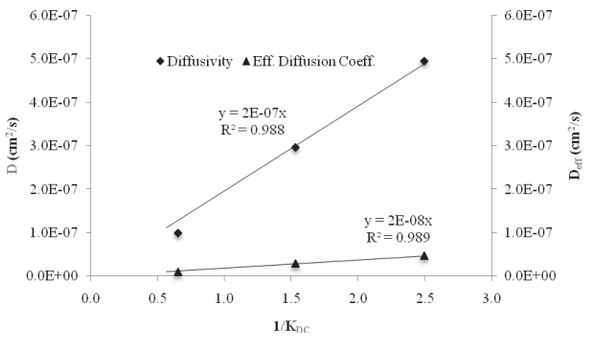

Fig. 6. Plot showing relationship between DC overall kinetic rate (K_{DC}), diffusivity (D) and effective diffusion coefficient (D_{eff}).

3.3. Development of ANN to simulate the process kinetics

3.3.1. Neural network optimization

The neural network models were developed and optimized to simulate and predict the kinetic manner of chromium (VI) adsorption by activated carbon. A three layer feed-forward backpropagation network was adopted with input and output layer comprising neurons as shown in Table 7. The impact of training function on the networks was first examined (Table 8). Using the tansig and purelin transfer functions at the hidden and output layer respectively and 10 neurons at the hidden layer, the BFGS quasi-Newton backpropagation algorithm produced the lowest MSE for the SANN and DC-ANN models while the Levenberg–Marquardt algorithm best suited the PANN model. Using the optimum algorithms, transfer functions from Table 2 were then varied to determine the impact on the networks. The performance of these transfer functions at the hidden and output layers was assessed using the MSE and coefficient of determination. Using the first two training, the optimum functions for all three networks were found to be a tangent sigmoid transfer function (tansig) at the hidden layer and a linear transfer function (purelin) at the output layer (Table 9).

The number of neurons in the hidden layer is critical to the performance of an ANN model. Too few neurons can lead to under-fitting while too many neurons may result in over-fitting [61]. In this protocol, the number of neurons was varied from 2 to 20 and its impact on performance assessed using the MSE as shown in Fig. 7. The figure also reveals fluctuations in MSE as the number of neurons increased. This may have resulted from the network being trapped into the local minima [38]. The least MSE within 10 training runs for the SANN, DC-ANN, and PANN were found to be 0.0013 at neuron 7, 0.0001 at neuron 17 and 5.3×10^{-19} at 15 neurons respectively. The optimized SANN, DC-ANN and PANN networks are presented in Table 10.

3.3.2. Evaluation criteria

According to Dawson et al. [62], there is no universally accepted measure of ANN skill, and therefore the combined use of the performance measures allows comparison and an unbiased estimate of prediction ability. Using these optimised neural networks, the output value was compared to the target value using the coefficient of determination, NMBE, RMSE and the Bland-Atman plot. The experimental dataset was divided into three parts:

Table 7
Range of data for SANN, DC-ANN and PANN models

Layer	Variable	SANN	DC-ANN	PANN
Input Layer	Time of contact, t (min)	0–90	0–90	0–90
	Particle size (GMS), d_p (mm)	0.21–1.05	0.21–1.05	0.21–1.05
	Initial kinetic rate, k_i (mg/g-t)	–	0.078 – 3.15	–
	Overall kinetic rate, K_{DC} (mg/g-$t^{0.5}$)	–	3.11 – 11.91	–
	Relative sorption capacity, q_m (mg/g)	–	45.02 – 123.63	–
Output Layer	Adsorbed concentration q_t (mg/g)	2.7–31.3	2.7–31.3	2.7–31.3

Table 8
Performance of varying training algorithms of ANN models

	SANN		DC-ANN		PANN	
Backpropagation (BP) algorithms	MSE	R^2	MSE	R^2	MSE	R^2
BFGS quasi-Newton backpropagation	0.9008	0.9682	0.1332	0.9933	0.2221	0.3533
Powell–Beale conjugate gradient BP	2.6524	0.9103	1.8535	0.9631	0.2906	0.4808
Fletcher–Reeves conjugate gradient BP	1.6587	0.9308	1.8588	0.9651	0.2716	0.4717
Polak-Ribiere conjugate gradient BP	1.9381	0.9370	4.0071	0.8995	0.2907	0.4808
Gradient descent	270.83	0.0650	96.2580	0.8191	4.4884	0.1244
Gradient descent with momentum	69.6986	0.5288	96.2580	0.8191	4.4884	0.1244
Gradient descent with adaptive learning rate	20.7431	0.6925	2.2059	0.7745	4.4884	0.1244
Gradient descent with momentum & AL	43.6512	0.2664	10.1557	0.7488	4.4884	0.1244
Levenberg–Marquardt backpropagation	1.0754	0.1664	4.8134	0.9564	0.1264	0.8373
One step secant backpropagation	1.2842	0.9674	0.4888	0.8515	0.2118	0.4776
Random weight/Bias	270.83	0.9777	96.2580	0.4647	4.4884	0.5541
Resilient backpropagation	5.8969	0.8627	3.7096	0.7744	0.3027	0.6452
Scaled conjugate gradient backpropagation	1.1567	0.9434	1.6737	0.9693	0.2792	0.4753

Table 9
Performance of varying transfer functions on SANN, DC-ANN, and PANN models

Neural network	Activation function Layer 1	Activation function Layer 2	MSE (first iteration)	MSE (second iteration)	R^2 (first iteration)	R^2 (second iteration)
SANN	Tansig	Purelin	0.9008	0.8633	0.9682	0.9682
DC-ANN	Tansig	Purelin	0.1332	0.0096	0.9933	0.9969
PANN	Tansig	Purelin	0.1264	0.0269	0.8373	0.8373

Table 10
Optimized structure of SANN, DC-ANN and PANN models

Network characteristics	SANN Feed-forward BP	DC-ANN Feed-forward BP	PANN Feed-forward BP
Transfer function (hidden layer)	Tansig	Tansig	Tansig
Transfer function (output layer)	Purelin	Purelin	Purelin
Training function	BFGS quasi-Newton	BFGS quasi-Newton	Levenberg–Marquardt
Performance function	MSE	MSE	MSE
Neurons in input layer	2	5	2
Neurons in hidden layer	7	17	15
Neurons in output layer	1	1	1
Data used for training	70%	70%	70%
Data for cross-validation	15%	15%	15%
Data for testing	15%	15%	15%

70% training data; 15% validation data and 15% testing data. The training dataset is used to build several predictive algorithms. The validation dataset is used to compare and select the most optimal algorithm. The testing dataset presents an opportunity to test and assess the best performing algorithm on un-seen data. Table 11 reveals a comparison of the experimental data and the training dataset, validation dataset, testing dataset and the entire dataset. A positive NMBE would indicate over-prediction while a negative NMBE would indicate under-prediction from a model. The NMBE statistics for PANN produced negative values for validation and testing which reveals under-prediction of the adsorptive performance. However, consistency and an improvement in performance were observed.

Table 11
Statistical performance of SANN, DC-ANN, and PANN models using NMBE and RMSE error functions

Models	NMBE	RMSE
Training		
SANN	0.1167	0.4686
DC-ANN	0.3523	0.2539
PANN	–0.3242	0.4186
Validation		
SANN	0.0248	0.0365
DC-ANN	0.0616	0.0102
PANN	0.0000	0.0000
Testing		
SANN	13.2751	2.3485
DC-ANN	3.212	1.2089
PANN	–0.9997	0.2186
Entire Dataset		
SANN	1.0840	0.9129
DC-ANN	0.7754	0.5046
PANN	–0.3159	0.3287

The DC-ANN showed positive values of NMBE training, validation and testing and thus an overestimation of the adsorptive performance, but also reveals consistency and improved performance by the model. Both models reveal a near optimal prediction accuracy of the adsorption data. The RMSE gives a relatively high weight to large errors [47]. The DC-ANN produced the lowest training RMSE while the PANN had the lowest for validation and testing. When the entire dataset was analysed by this matrix, the PANN produced the lowest RMSE followed by the DC-ANN and the finally the SANN which implies that the SANN is most prone to large errors which are undesirable in adsorption processes.

Correlation is a statistical technique that can show whether and how strong pairs of variables are linearly related [63]. Further, linear regression finds the best line that predicts one variable from the other one then quantifies the goodness of fit using the coefficient of determination. The plots in Figs. 8a to 8d describes the linear relationship (blue lines) between the experimental and predicted values. In all instances, a high correlation of determination is observed (>0.95). However, deviations are observed between the line of equality (red line) which signifies the $y = x$ and the line representing the coefficient of determination, R^2. The R^2 tells us the proportion of variance that the experimental and predicted values have in common, however, Giavarina [63] explained that this does not imply good agreement. In all the plots the data points were clustered close to the regression line which makes it difficult to assess the agreement or differences between the data.

The Bland-Altman method involves plotting the difference in experimental and predicted values against the mean. In this study, the technique was employed to provide a comparative analysis of each neural network's performance in predicting the experimental data. These plots are shown in Figs. 9a–9d for the DC kinetic model, SANN, PANN and the DC-ANN respectively. The plot allows us to assess the relationship between the predicted error and the so called true value [64]. The lack of agreement between the experimental data and the predictions are summerised by calculating the bias. For the prediction of experimental values by the neural networks, the limits of agreement (LOA) are set at mean +/– 1.96 SD [64].

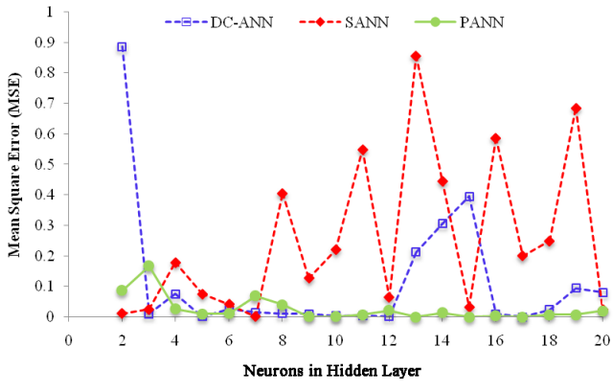

Fig. 7. Relationship between MSE and number of neurons at hidden layer for SANN, DC-ANN, and PANN models.

From Fig. 9b. the predictions by the SANN were on average 0.276 mg/g higher than the experimental values. The DC-ANN model (Fig. 9d) was also higher by 0.153 mg/g while the PANN (Fig. 9c) predicts on average 0.0623 mg/g lower than the experimental values. A significant trend was observed by the DC kinetic model (Fig. 9a), which reveals a prediction of an average of 0.023 mg/g higher than the experimental values and thus a low bias. Consequently, adjusting the reading of the prediction by adding or reducing the values by the bias would make it agree more to the experimental value. Despite the low bias, the LOA estimates the DC kinetic model may produce q_t values within an interval of 1.739 mg/g above or 1.785 mg/g below the experimental value. Further, the LOA estimates an interval for the SANN of 1.725 to –2.277; the DC-ANN of 0.813 mg/g to –1.119 mg/l; and the PANN of 0.781 mg/g to –0.656 mg/g.

The prediction made by the DC kinetic model and the SANN to the experimental data was the least precise among the models tested. The DC-ANN and the PANN both generated a superior prediction with a lower LOA interval. The SANN (black-box model) has the advantage that the model may be applied to predict adsorption removal outside the range of operational parameters. The PANN and ANN-DC are limited to only produce predictions within the range of operational parameters since they rely on input from the DC kinetic model. The PANN requires the error to be first generated followed by analysis using the DC kinetic model to determine the adsorption removal. Although the PANN produced the smallest error, the comparison revealed the DC-ANN produced a sufficiently close simulation and required less computational effort and is thus more practical for up-scale applications.

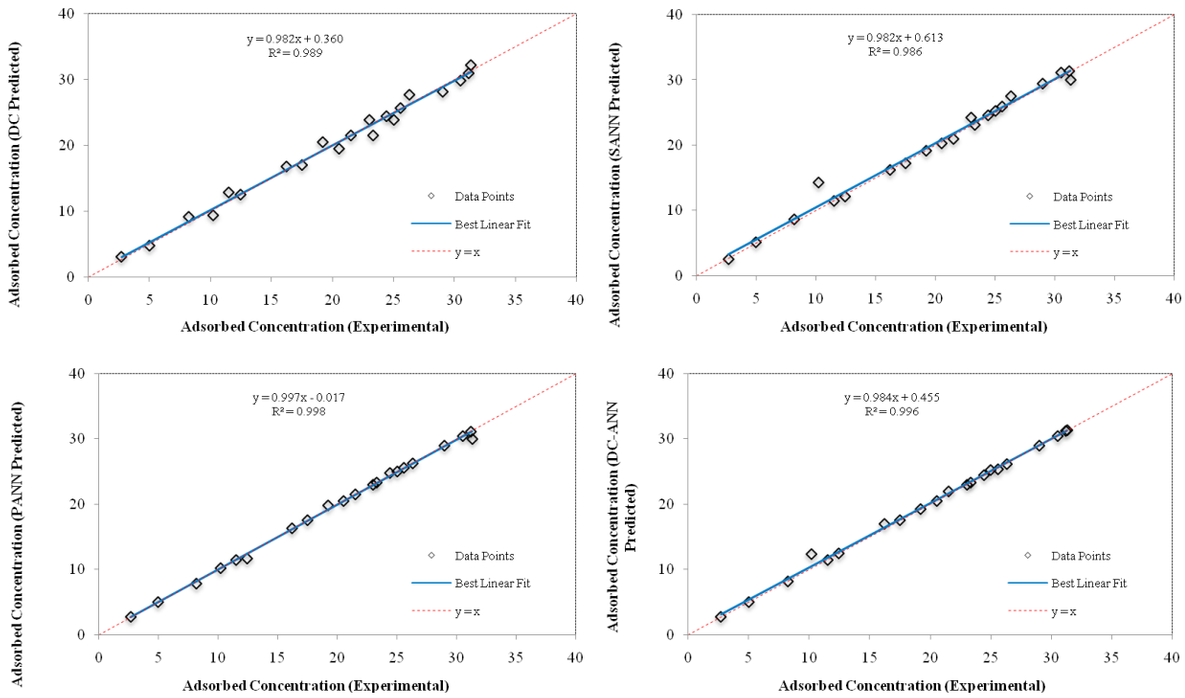

Fig. 8. (a) Plot of DC kinetic model outputs vs. the corresponding experimental targets, (b) Plot of SANN network outputs vs. the corresponding experimental targets, (c) Plot of PANN network outputs vs. the corresponding experimental targets, (d) Plot of DC-ANN network outputs vs. the corresponding experimental targets.

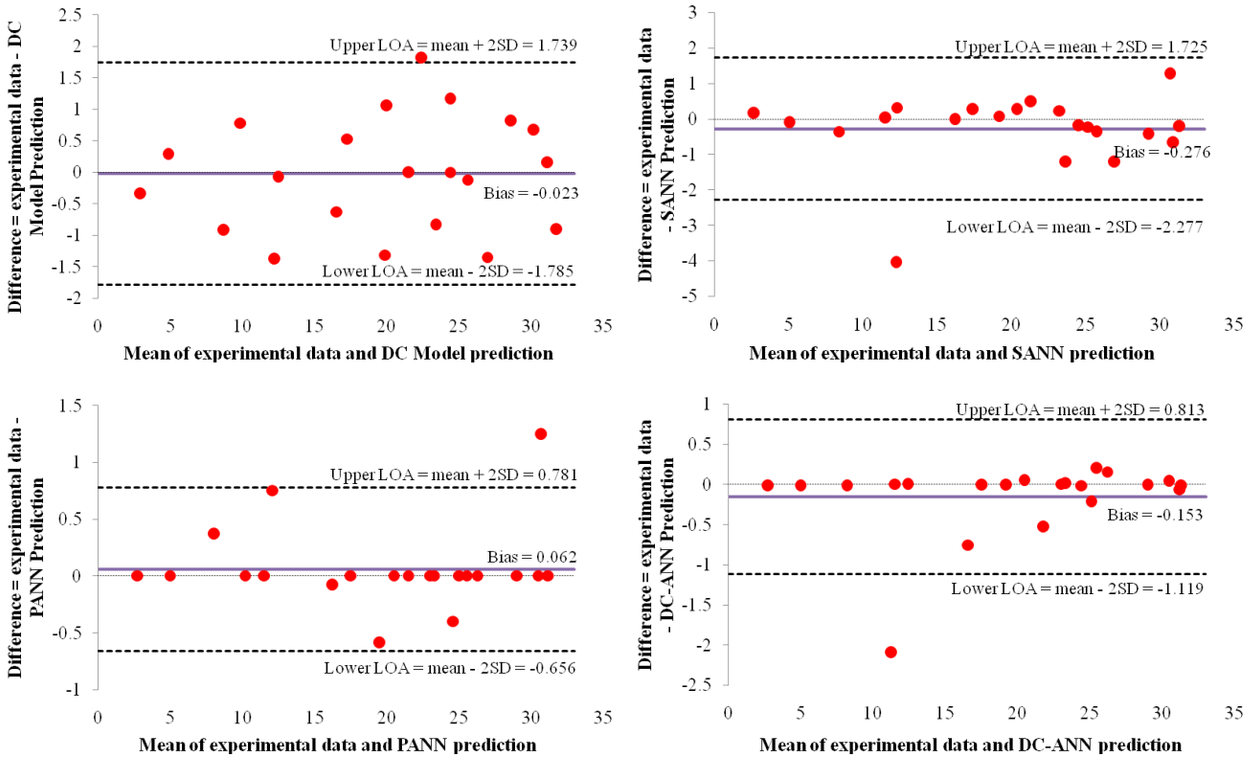

Fig. 9. (a) Bland–Altman plot showing the difference against the average of DC-ANN prediction and experimental data with LOA (broken lines), (b) SANN prediction, (c) PANN prediction, (d) DC kinetic model prediction.

3.3.3. Formulation of empirical equation

An empirical expression was developed to predict the kinetic adsorption process without the need to run the DC-ANN. Using the optimized DC-ANN and Eq. (19) which exhibits the fitness function that correlates the input with the output, the following empirical expression is developed:

$$q_{t,pred} = -0.2375F1 + 0.4387F2 + 0.4285F3 - 0.5440F4 + 0.1897F5 - 0.4676F6 - 0.7580F7 - 0.0410F8 + 0.1234F9 + 0.3883F10 - 0.2581F11 + 0.1364F12 + 0.1115F13 - 0.6453F14 - 0.4122F15 - 0.9162F16 + 0.7232F17 - 0.5957 \quad (18)$$

where F_i is the *tansig* activation function used at the hidden layer and is given as Eq. (19):

$$F_i = \frac{2}{\left[1+\exp(-2*E_i)\right]} - 1 \; ; \; i = 1{:}17 \quad (19)$$

The input data is normalized in the range –1 to 1 using Eq. (20):

$$X_{norm} = 2\left[\frac{X_i - X_{\min}}{X_{\max} - X_{\min}}\right] - 1 \quad (20)$$

where X_i is the input or output variable X, and X_{min} and X_{max} are the minimum and maximum value of variable X. E_i is the weighted sum of the normalized input calculated using values from Table 12 and is defined as follows:

$$E_i = W_{i1} * t + W_{i2} * k_i + W_{i3} * K + W_{i4} * q_m + W_{i5} * d_p + b_i \quad (21)$$

3.3.4. Sensitivity analysis

Based on the analysis of Figs. 8a–8d, it is significant to note that the R^2 for the entire SANN dataset was 0.9892. However, the R^2 increased to 0.9969 upon the inclusion of the DC kinetic model parameters to create the DC-ANN model. A sensitivity analysis was carried out to determine the effect of each variable, particularly the DC kinetic parameters on the performance of the DC-ANN model. To conduct this analysis, performance evaluation of various possible combinations of variables were investigated [65]. The variables were combined to form five groups as presented in Table 13. The input variables were defined as follows: *p1* is time, *p2 is* initial rate, *p3 is* overall rate, *p4* is relative sorption capacity, and *p5* is particle size. The combinations were examined by the optimum DC-ANN model using the BFGS quasi-Newton backpropagation algorithm with 17 neurons in the hidden layer. The table shows *p1* to be the most effective parameter in the group of one variable with an MSE of 8.1159 and R^2 of 0.7447. The minimum value of MSE and highest R^2 in the group of two variables was determined to be 0.0641 and 0.9877 respectively with the interaction of *p1* (time) and *p3* (overall rate). When the interaction of *p1* + *p3* was used with *p5* (particle size), the minimum MSE was observed among the group of three variables. The lowest value of MSE and best R^2 was found to be 0.0321 and 0.9909 respectively. The values of MSE decreased significantly from 0.0321 to 0.0092 when *p2* (initial rate) was added to the interaction. The lowest MSE (0.0001) and highest R^2 *(0.9969)* were observed when all variables were combined. Based on the preceding, it is evident that the greatest improvement in performance occurred with the inclusion of the diffusion-chemisorption kinetic parameters. Consequently, the relative impact of each parameter allows the designer to make a more informed decision regarding process efficiency.

Table 12
Optimized weights and biases for DC-ANN model

	Input 1 (t)	Input 2 (k_i)	Input 3 (K_{DC})	Input 4 (q_m)	Input 5 (d_p)	Bias 1	Output
Node 1	1.0475	1.3636	–1.2657	0.5431	–1.1586	–2.4177	–0.2375
Node 2	1.4876	1.1445	1.2604	0.5271	–0.9572	–2.1103	0.4387
Node 3	–1.0822	1.3822	0.5269	–0.8653	1.2065	1.8860	0.4285
Node 4	1.5389	0.4871	–0.7545	–1.4736	–0.9069	–1.4148	–0.5440
Node 5	0.3620	–1.5491	1.5334	–0.0038	1.0541	–1.2595	0.1897
Node 6	–0.6512	1.0027	–1.2887	1.3257	–0.7055	0.9434	–0.4676
Node 7	–0.6793	1.8681	0.0098	–0.6515	1.4246	0.5667	–0.7580
Node 8	0.4296	1.5543	–1.0696	0.7659	–1.3079	–0.3065	–0.0410
Node 9	1.5248	0.9147	0.9335	–0.9352	–1.0469	–0.0383	0.1234
Node 10	1.3821	0.9950	1.1434	0.9441	–0.9042	0.0038	0.3883
Node 11	–1.5764	–0.3833	–1.2850	–1.3278	0.3490	–0.5106	–0.2581
Node 12	2.2965	0.8613	0.0265	0.0017	–0.1862	0.9855	0.1364
Node 13	1.8416	–1.2765	–0.1842	0.7775	–0.6179	1.2128	0.1115
Node 14	0.0965	0.9994	0.7415	1.5787	1.3193	–1.6459	–0.6453
Node 15	0.9248	–1.5495	0.6831	1.4067	0.2337	1.9876	–0.4122
Node 16	–2.0297	–1.1318	1.2755	0.0535	0.1526	–2.3325	–0.9162
Node 17	–0.2471	–1.4986	–0.7433	–1.1841	1.3720	–2.4540	0.7232
Bias 2							–0.5957

Table 13
Performance evaluation of combinations of input variables

CN	Combination	MSE	R^2
Group of one variable			
1	p1	8.1156	0.7447
2	p2	12.7766	0.1980
3	p3	11.1346	0.1609
4	p4	9.0725	0.1629
5	p5	9.3725	0.2088
Group of two variable			
6	p1+p2	0.4432	0.9701
7	p1+p3	0.0641	0.9877
8	p1+p4	0.1896	0.8217
9	p1+p5	0.2005	0.9744
10	p2+p3	1.4604	0.0176
11	p2+p4	17.5219	0.1811
12	p2+p5	9.3730	0.1440
13	p3+p4	1.7700	0.0804
14	p3+p5	1.0633	0.0804
15	p4+p5	2.4150	0.0031
Group of three variable			
16	p1+p2+p3	0.1304	0.9852
17	p1+p2+p4	0.1203	0.9893
18	p1+p2+p5	0.1224	0.9873
19	p1+p3+p4	0.1333	0.9248
20	p1+p3+p5	0.0321	0.9909
21	p1+p4+p5	0.1453	0.9815
22	p2+p3+p4	21.3603	0.2379
23	p2+p3+p5	23.8005	0.2159
24	p2+p4+p5	13.9080	0.2294
25	p3+p4+p5	19.3844	0.2334
Group of four variable			
26	p1+p2+p3+p4	0.0654	0.9688
27	p1+p2+p3+p5	0.0092	0.9928
28	p1+p2+p4+p5	0.1635	0.9648
29	p1+p3+p4+p5	0.1877	0.9299
30	p2+p3+p4+p5	19.5909	0.1717
Group of five variable			
31	p1+p2+p3+p4+p5	0.0001	0.9969

4. Conclusion

In this paper, the performance of a standard artificial neural network (SANN) was compared to that of a serial gray-box (DC-ANN) and parallel gray-box (PANN) for the prediction of Cr(VI) adsorption by activated carbon. The constructed predictive hybrid models were built-in with the joint effect of operating parameters such as time and particle size (PANN) together with the kinetic parameters of the diffusion-chemisorption (DC) model such as overall rate, initial rate, and sorption capacity (DC-ANN). A fraction of the experimental dataset (70%) was used to train the network. The other data points, not employed in the network training, were used to validate and test the network. The hybrid models produced a significantly higher correlation to the experimental data over that of the DC kinetic model and the SANN. The DC-ANN produced significantly high correlation to the experimental data and required less computational effort when compared to the PANN.

The removal efficiency of Cr(VI) was found to be pH dependent. The optimum pH was determined to be 2.5 whereby the Cr(VI) was adsorbed as $HCrO_4^-$. The major transport mechanism of the reaction was found to be intraparticle diffusion while the major attachment mechanism was chemical bonding. It was also discovered that the DC overall kinetic rate, K_{DC} was inversely proportional to both the effective diffusion coefficient and the diffusivity. This parameter opportunely allows the assessment of diffusional effects.

References

[1] Y. Nakano, K. Takeshita, T. Tsutsumi, Adsorption mechanism of hexavalent chromium by redox within condensed-tannin gel, Water Res., 35 (2001) 496–500.
[2] D. Krishna, R. Sree, Artificial neural network (ANN) approach for modeling chromium (VI) adsorption from aqueous solution using a Borassus flabellifer coir powder, Int. J. Appl. Sci. Eng., 12 (2014) 177–192.
[3] P.C. Mane, A.B. Bhosle, P.D. Deshmukh, C.M. Jangam, Chromium adsorption onto activated carbon derived from Tendu (Diospyros molanoxylon) leaf refuse; Influence of metal/carbon ratio, time and pH, Adv. App. Sci. Res., 1 (2010) 212–221.
[4] C. Sutherland, C. Venkobachar, Equilibrium modeling of Cu (II) biosorption onto untreated and treated forest macro-fungus Fomes fasciatus, Int. J. Pl. An. Env. Sci., 3 (2013) 193–203.
[5] M.R. Gandhi, N. Viswanathan, S. Meenakshi, Adsorption mechanism of hexavalent chromium removal using amberlite IRA 743 resin, Ion Exch. Lett., 3 (2010) 25–35.
[6] T.C. Wang, J.C. Weissman, G. Ramesh, R. Varadarajan, J.R. Benemann, Parameters for removal of toxic heavy metals by water milfoil (Myriophyllum spicatum), Bull. Environ. Contam. Toxicol., 57 (1996) 779–786.
[7] D.W. Hendricks, Water Treatment Unit Processes: physical and Chemical, CRC press, 2006.
[8] B.H. Hameed, D.K. Mahmoud, A.L. Ahmad, Equilibrium modeling and kinetic studies on the adsorption of basic dye by a low-cost adsorbent: Coconut (Cocos nucifera) bunch waste, J. Hazard. Mater., 158 (2008) 65–72.
[9] G. Naja, V. Diniz, B. Volesky, Predicting metal biosorption performance, In: Proceedings of the 16th International Biohydrometallurgy Symposium, STL Harrison, DE Rawlings, J. Peterson, Eds., Compress Co., Cape Town, South Africa, 2005, pp. 553–562.
[10] S. Saraf, V.K. Vaidya, Statistical optimization of biosorption of reactive orange 13 by dead biomass of Rhizopus arrhizus NCIM 997 using response surface methodology, Int. J. Ind. Chem., 6 (2015) 93–104.
[11] P. Kumar, P. Sharma, Artificial neural networks - A study, Int. J. Emerg. Eng. Res. Technol., 2 (2014) 143–148.
[12] W.S. McCulloch, W.A. Pitts, Logical calculus of the ideas immanent in nervous activity, Bull. Math. Biol., 52 (1990) 99–115.
[13] P.J. Werbos, Backpropagation through time: what it does and how to do it,: Proc. IEEE, 1990, pp. 1550–1560.
[14] T.H. Kim, Pattern recognition using artificial neural network: a review, In: Information Security and Assurance, Springer Berlin Heidelberg, 2010, pp. 138–148.

[15] A. Gamal El-Din, D.W. Smith, Modeling a full-scale primary sedimentation tank using artificial neural networks, Environ. Technol., 23 (2002) 479–496.
[16] G. Stuart, R.A.F Machado, A.M.C. Uller, E.L. Lima, J.V. Oliveira, Hybrid artificial neural network applied to modeling SCFE of basil and rosemary oils, Food Sci. Technol., 17 (1997) 501–505.
[17] D.C. Psichogios, L.H. Ungar, A hybrid neural network-first principles approach to process modeling, AIChE J., 38 (1992) 1499–1511.
[18] M.A. Hussain, M.S. Rahman, C.W. Ng, Prediction of pores formation (porosity) in foods during drying: generic models by the use of hybrid neural network, J. Food Eng., 51 (2002) 239–248.
[19] M.A. Kramer, M.L. Thompson, P.M. Bhagat, Embedding theoretical models in neural networks, In: American Control Conference, IEEE, 1992, pp. 475–479.
[20] K. Anupam, S. Dutta, C. Bhattacharjee, S. Datta, Artificial neural network modelling for removal of chromium (VI) from wastewater using physisorption onto powdered activated carbon, Desal. Water Treat., 57 (2016) 3632–3641.
[21] G. Halder, S. Dhawane, P.K. Barai, A. Das, Optimizing chromium (VI) adsorption onto superheated steam activated granular carbon through response surface methodology and artificial neural network, Environ. Prog. Sustain. Energy, 34 (2015) 638–647.
[22] M. Gholipour, H. Hashemipour, M. Mollashahi, Hexavalent chromium removal from aqueous solution via adsorption on granular activated carbon: adsorption, desorption, modeling and simulation studies, J. Eng. Appl. Sci., 6 (2011) 10–18.
[23] T. Khan, M.H. Isa, M.R.U. Mustafa, H. Yeek-Chia, L. Baloo, T.S.B.A. Manan, M.O. Saeed, Cr (VI) adsorption from aqueous solution by an agricultural waste based carbon, RSC Advances, 6 (2016) 56365–56374.
[24] D.S. Lee, C.O. Jeon, J.M. Park, K.S. Chang, Hybrid neural network modeling of a full-scale industrial wastewater treatment process, Biotechnol. Bioeng., 78 (2002) 670–682.
[25] C. Faur-Brasquet, P.L. Cloirec, Neural network modeling of organics removal by activated carbon cloths, J. Environ. Eng., 127 (2001) 889–894.
[26] H. Pfost, V. Headley, Methods of determining and expressing particle size, Feed Manuf. Technol., (1976) 512–517.
[27] U.S. Environmental Protection Agency (USEPA). Fate, transport, and transformation test guidelines. Adsorption/desorption (batch equilibrium), Washington, DC, OPPTS 835.1230, 2008.
[28] W.J. Weber, C.T. Miller, Modeling the sorption of hydrophobic contaminants by aquifer materials—I. Rates and equilibria, Water Res., 22 (1988) 457–464.
[29] Y.S. Ho, G. McKay, A comparison of chemisorption kinetic models applied to pollutant removal on various sorbents, Process Saf. Environ., 76 (1998) 332–340.
[30] Y.S. Ho, G. McKay, Sorption of dye from aqueous solution by peat, Chem. Eng. J., 70 (1998) 115–124.
[31] Y.S. Ho, G. McKay, Pseudo-second order model for sorption processes, Process Biochem., 34 (1999) 451–465.
[32] W.J. Weber, J.C. Morris, Kinetics of adsorption on carbon from solution, J. Sanit. Eng. Div., 89 (1963) 31–60.
[33] A.E. Ofomaja, Intraparticle diffusion process for lead (II) biosorption onto mansonia wood sawdust, Bioresour. Technol., 101 (2010) 5868–5876.
[34] R. Teixeira, V.O. Sousa Neto, J.T. Oliveira, C. Thalles, D.Q. Melo, M.A. Silva, R.F. Nascimento, Study on the use of roasted barley powder for the adsorption of Cu^{2+} ions in batch experiments and in fixed-bed columns, Bioresources, 8 (2013) 3556–3573.
[35] I. Tsibranska, E. Hristova, Comparison of different kinetic models for adsorption of heavy metals onto activated carbon from apricot stones, Bulg. Chem. Commun., 43 (2011) 370–377.
[36] C. Sutherland, C. Venkobachar, A diffusion-chemisorption kinetic model for simulating biosorption using forest macro-fungus, Fomes fasciatus, Int. Res. J. Pl. Sci., 1 (2010) 107–117.
[37] K. Mohd-Yusof, Development, analysis and comparison of connectionist models for real-time optimization, PhD Thesis, University of Waterloo, Ontario Canada, 2001.
[38] Z. Shahryari, A. Sharifi, A. Mohebbi, Artificial neural network (ANN) approach for modeling and formulation of phenol adsorption onto activated carbon, J. Eng. Thermophys., 22 (2013) 322–336.
[39] S. Kuvendziev, M. Marinkovski, K. Lisichkov, P. Paunović, Artificial neural network modeling of Cd (II) ions adsorption on nano-porous inorganic sorbents, In: Nanoscience Advances in CBRN Agents Detection, Information and Energy Security, Springer Netherlands, 2015, pp. 469–476.
[40] S. Lek, J.F. Guegan, Artificial neural networks as a tool in ecological modeling, an introduction, Ecol. Modell., 120 (1999) 65–73.
[41] M. Van Der Bean, C. Jutten, Neural networks in geophysical applications, Geophysics, 65 (2000) 1032–1047.
[42] Y. Babazadeh, S.M. Mousavi, M.R. Akbarzadeh, Multidimensional dynamic modeling of milk ultrafiltration using neuro-fuzzy method and a hybrid physical model, Iranian J. Chem. Eng., 5 (2008) 3–22.
[43] B. Duarte, P.M. Saraiva, C.C. Pantelides, Combined mechanistic and empirical modelling, Int. J. Chem. React. Eng., 2 (2004) 1–19.
[44] M.L. Thompson, M.A. Kramer, Modeling chemical processes using prior knowledge and neural networks, AIChE J., 40 (1994) 1328–1340.
[45] C.A. Nascimento, R. Giudici, N. Scherbakoff, Modeling of industrial nylon-6, 6 polymerization process in a twin-screw extruder reactor. II. Neural networks and hybrid models, J. Appl. Polym. Sci., 72 (1999) 905–912.
[46] H.T. Su, N. Bhat, P.A. Minderman, T.J. McAvoy, Integrating neural networks with first principles models for dynamic modeling, IFAC Proceedings, 25 (1992) 327–332.
[47] T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature, Geosci. Model Dev., 7 (2014) 1247–1250.
[48] P. Chingombe, B. Saha, R.J. Wakeman, Surface modification and characterisation of a coal-based activated carbon, Carbon, 43 (2005) 3132–3143.
[49] E. Romera, F. Gonzalez, A. Ballester, M.L. Blazquez, J.A. Munoz, Biosorption with algae: A statistical review, Crit. Rev. Biotechnol., 26 (2006) 223–235.
[50] V.K. Singh, P.N. Tiwari, Removal and recovery of chromium (VI) from industrial waste water, J. Chem. Technol. Biotechnol., 69 (1997) 376–382.
[51] A. Anandan, T. Janakiram, Adsorption of Cr(VI) from aqueous solutions by low cost activated carbon, J. Chem. Pharm. Res., 4 (2012) 2900–2905.
[52] Y.J.P. Poots, G. McKay, J. Healy, The removal of acid dye from effluent using natural adsorbents - I Peat, Water Res., 10 (1976) 1061–1066.
[53] K.K. Choy, D.C. Ko, C.W. Cheung, J.F. Porter, G. McKay, Film and intraparticle mass transfer during the adsorption of metal ions onto bone char, J. Colloid Interface Sci., 271 (2004) 284–295.
[54] M.A.O. Badmus, T.O.K Audu, B.U. Anyata, Removal of lead ion from industrial wastewaters by activated carbon prepared from periwinkle shells (Typanotonus fuscatus), Turkish J. Eng. Env. Sci., 31 (2007) 251–263.
[55] J. Gao, Q. Zhang, K. Su, R. Chen, Y. Peng, Biosorption of acid yellow 17 from aqueous solution by non-living aerobic granular sludge, J. Hazard. Mater., 174 (2010) 215–225.
[56] S.S. Gupta, K.G. Bhattacharyya, Kinetics of adsorption of metal ions on inorganic materials: a review, Adv. Colloid Interfac., 162 (2011) 39–58.
[57] B.S. Chittoo, C. Sutherland, Phosphate removal and recovery using lime-iron sludge: adsorption, desorption, fractal analysis, modeling and optimization using artificial neural network-genetic algorithm, Desal. Water Treat., 63 (2017) 227–240.
[58] Z. Aksu, E. Kabasakal, Batch adsorption of 2, 4-dichlorophenoxy-acetic acid (2, 4-D) from aqueous solution by granular activated carbon, Sep. Purif. Technol., 35 (2004) 223–240.
[59] P. Shanthi, G. Tamilarasan, K. Anitha, S. Karthikeyan, Film and pore diffusion modeling for adsorption of reactive red-4 onto sterculia quadrifida seed shell waste as activated carbon, Rasayan J. Chem., 7 (2014) 229–240.

[60] G.M. Walker, L.R. Weatherley, Kinetics of acid dye adsorption on GAC, Water Res., 33 (1999) 1895–1899.
[61] K. Yetilmezsoy, Modeling studies for the determination of completely mixed activated sludge reactor volume: Steady-state, empirical and ANN applications, Neural Netw. World, 20 (2010) 559–589.
[62] C.W. Dawson, C. Harpham, R.L. Wilby, Y. Chen, Evaluation of artificial neural network techniques for flow forecasting in the River Yangtze, China, Hydrol. Earth Syst. Sci. Discuss., 6 (2002) 619–626.
[63] D. Giavarina, Understanding bland altman analysis, Biochemia. Medica., 25 (2015) 141–151.
[64] C. Bunce, Correlation, agreement, and Bland–Altman analysis: statistical analysis of method comparison studies, Am. J. Ophthalmol., 148 (2009) 4–6.
[65] K. Yetilmezsoy, S. Demirel, Artificial neural network (ANN) approach for modelling of Pb(II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells, J. Hazard. Mater., 153 (2008) 1288–1300.