Desalination and Water Treatment
www.deswater.com

doi:10.5004/dwt.2018.23030

132 (2018) 42–51
November

Adsorption-diffusion model with neural network-based equilibrium relationship

E.C. Herrera-Hernández[a], R. Ocampo-Perez[b,*], C.G. Aguilar-Madera[c], J.V. Flores-Cano[c]

[a]*CONACYT-Centro de Ingeniería y Desarrollo Industrial, Sede Campeche, Querétaro, 76125, Mexico, email: erik.herrera@cidesi.edu.mx (E.C. Herrera-Hernández)*
[b]*Universidad Autónoma de San Luis Potosí, Facultad de Ciencias Química, 78260 San Luis Potosí, Mexico, email: raul_iqi@yahoo.com.mx (R. Ocampo-Perez)*
[c]*Universidad Autónoma de Nuevo León, Facultad de Ciencias de la Tierra, Linares N.L, México 67700, Mexico, email: carlos_aguilarmadera@hotmail.com (C.G. Aguilar-Madera), jvfc_iqi@yahoo.com.mx (J.V. Flores-Cano)*

Received 5 February2018; Accepted 18 August 2018

ABSTRACT

In this work, we propose a novel approach to predict adsorption equilibrium using artificial neural networks. The equilibrium model (particularly the derivative of such approach) is employed into a surface diffusion model to interpret the concentration decay curves during the adsorption of pyridine onto activated carbon in aqueous solution. Moreover, we estimated the external mass transfer and surface diffusion coefficients through an experimental-based inverse problem formulation. Also, we predict and compare with lab measurements the adsorption equilibrium curve and concentration decay dynamics. We found that results obtained by an artificial neural networks-based equilibrium model coupled with diffusional model agree well with experimental data. Thus, the artificial neural networks capabilities suggest that using a subrogate approach to predict the equilibrium relationship is an appropriate alternative when standard isotherm models fail.

Keywords: Artificial neural networks; Adsorption-diffusion model; Method of lines; Inverse problem

1. Introduction

Adsorption-diffusion phenomenon plays a crucial role in many natural or artificial processes. It has been studied from different point of views; experimentally and/or theoretically. Such phenomenon is complex since it can be governed by different physical or chemical stages like mass transport in the aqueous phase, intra-particle-diffusion, surface diffusion and adsorption/desorption on/from an active site. Due to the uncertainty related to governing mechanisms in experimental adsorption-diffusion systems, care must be taken to select the adequate model. In addition to the mass balances which need to be stated and solved, an equilibrium relationship is required to describe the functionality of the mass adsorbed at equilibrium with respect to the concentration of the solute in the fluid within the pores. Classical non-linear and multi-parametric models (Langmuir, Freundlich and Prausnitz-Radke isotherms) seem to work well for prediction of adsorption equilibrium. These models are useful when only final states need to be modeled, but when dynamical features are desired then other mathematical models need to be considered. Among dynamic models the adsorption-diffusional models offer some advantages as the computation of i) intra-particle mass flux, ii) porous-solution mass flux, and iii) the possibility of using more sophisticated models accounting for the complex microstructure of adsorbent particles. However, adsorption-diffusional models require the evaluation of the derivative of adsorbed mass with respect to the concentration, and this calculation is highly sensitive to small concentration changes. In this context, models based on artificial neural networks (ANN) [1] represent a reliable alternative if enough experimental information is available. ANN-based models have been applied to model adsorption equilibrium since several decades ago, and in this context, Ghaedi and Vafaei [1] have conducted a comprehensive review for adsorption of dyes. Such review includes feed forward neural networks, support vector machine, adaptive neuro-fuzzy inference system and hybrid with genetic algorithm and particle swarm optimization. They found that ANN

*Corresponding author.

models can be used to predict the equilibrium of adsorption of colorants with acceptable precision. ANN has been used by some authors to predict the adsorption of organic and inorganic compounds satisfactorily onto different adsorbent materials such as zeolites, agricultural waste, composites and nanomaterials [2–10]. On the other hand, the adsorption kinetics of organic contaminants on different adsorbent nanomaterials has been studied by some authors by using ANN to model the adsorption phenomenon [10,11]. Alharbi [10] and Agarwal et al. [11] used pseudo-first-order, pseudo-second-order, and Elovich models to analyze the adsorption kinetic of contaminants. However, they do not use diffusional models which consider solute mass transport in the aqueous solution, pore volume diffusion, surface diffusion, or some combination of them [12]. In this sense, it is important to highlight that to use the diffusional model it is necessary to calculate the derivative ($\partial q/\partial C_{AP}$) from a model of two or three parameters that adequately fit the adsorption equilibrium such as the Langmuir, Freundlich or Redlich-Peterson model. Nevertheless, due to the experimental error in the lab measures the estimation of the derivative from these non-linear models may be not accurate, and sometimes they present some drawbacks at low concentrations. Therefore, it is important to mention that ANN capabilities allow capturing non-linear behavior, predicting information outside the measured range, filtering noisy information [13], and the essential quality is the ability to estimate derivative regarding input information [14–18]. Such properties have been exploited to construct models from experimental information in both static and dynamic processes.

With this background, the objective of this study is to use an ANN-based relationship to describe the equilibrium process into a diffusional model. To the best of our knowledge, the coupling of the equilibrium isotherm and its derivative obtained from ANN with an adsorption-diffusion model has not been yet reported. Such a coupling yields a much more complex non-linear problem that needs to be solved to get predictions of concentration decay curves. The usage of ANN to replace the equilibrium model yields smooth predictions of derivative, which is required in the intra-particle mass balance, and thereby problems related to its estimation are avoided since it is well-known its sensitivity to small changes in concentration. To this end, it is necessary that ANN being trained and tested with adsorption equilibrium data. Thus, the non-uniqueness problem is not related to model parameters with physical meaning, but it is related to weights and bias involved in the formulation of the ANN.

2. Materials and methods

2.1. Surface diffusion model (SDM)

The design of an adsorption system requires data about the adsorbate concentration decay curves and the adsorption capacity of the adsorbent. Additionally, the estimation of mass transport parameters involves comparison of experimental kinetic data with those predicted with a mathematical model, which can have an empirical or theoretical framework. Within a deterministic framework, some models can be utilized depending if the adsorption process is controlled by the external mass transfer, intra-particle diffusion or by the surface diffusion/reaction [19–24].

One of the most used mathematical models is the surface diffusion model (SDM), which can be applied when the adsorption capacity is high because the main driving force is the concentration gradient at solid phase [25]. The SDM model considers the following transport stages: external mass transport, surface diffusion and adsorption on an active site inside the pores. The following equations state the initial value problem associated with the mass balances in the physical system [12,26] described by Fig. 1.

The solution mass balance:

$$V\frac{dC_A}{dt} = -mSk_L\left(C_A - C_{AP}\big|_{B_{p-s}}\right) \tag{1}$$

The intra-particle mass balance:

$$\varepsilon_p\frac{\partial C_{AP}}{\partial t} + \rho_p\frac{\partial q}{\partial t} = \nabla\cdot\left(D_s\rho_p\nabla q\right) \tag{2}$$

The initial concentrations:

$$t = 0,\quad C_A = C_{A0},\quad C_{AP} = 0 \tag{3}$$

The adsorbent-solution mass flux:

$$-\mathbf{n}_{p-s}\cdot\left(D_s\rho_p\nabla q\right) = k_L\left(C_A - C_{AP}\big|_{B_{p-s}}\right) \tag{4}$$

Zero mass flux in the center of adsorbent particles:

$$\left.\frac{\partial C_{AP}}{\partial r}\right|_{r=0} = 0 \tag{5}$$

Assuming the mass transport rate on active sites is instantaneous, the local equilibrium takes place between the solute concentration inside the pore, and the solute adsorbed onto the pore surface. Such adsorption equilibrium is a mathematical relationship between q and C_{AP} given by:

$$q = f\left(C_{AP}\right) \tag{6}$$

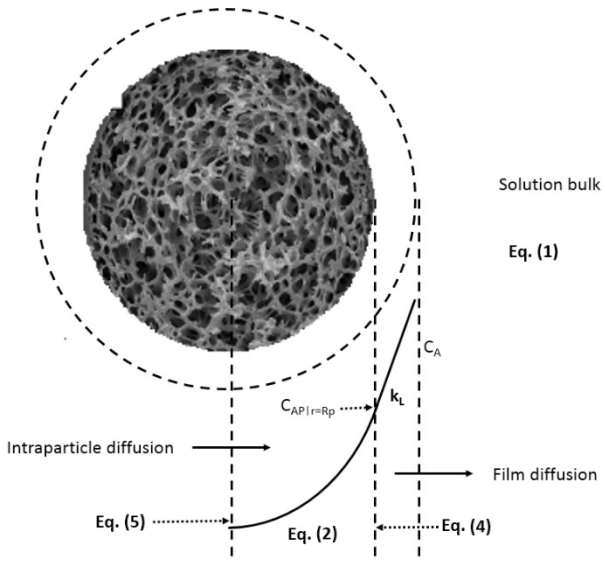


Fig. 1. Adsorption-diffusion system and the domains of governing equations.

Indeed if $f(C_{AP})$ is known, the derivatives of q in (2) can be written according to:

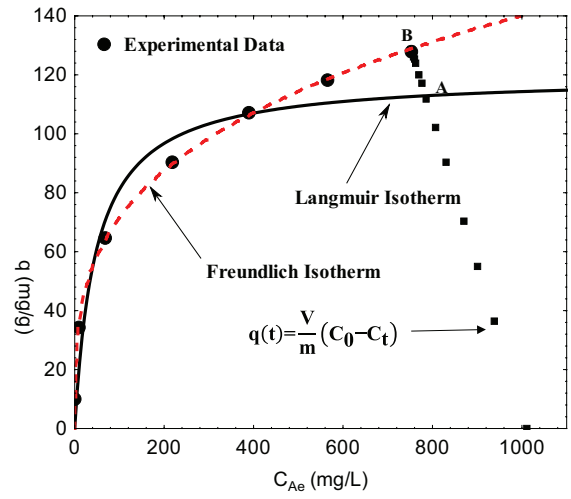
$$\frac{\partial q}{\partial t} = \frac{\partial q}{\partial C_{AP}} \frac{\partial C_{AP}}{\partial t} \tag{7}$$

$$\nabla q = \frac{\partial q}{\partial C_{AP}} \nabla C_{AP} \tag{8}$$

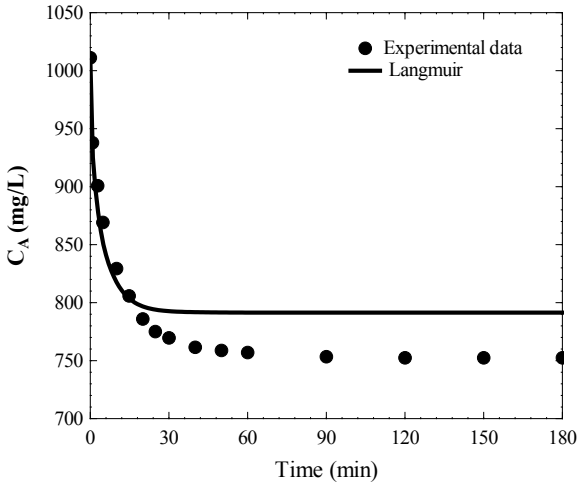
Note that the SDM only requires the derivative $\partial q/\partial C_{AP}$. However, the estimation of this derivative sensibly depends on the proposed model for $f(C_{AP})$. Commonly, isotherm models are fitted with experimental data, but the derivative is left aside.

Eq. (6) plays a crucial role in the prediction of concentration decay curves in all adsorbate-adsorbent systems. To illustrate this fact, Fig. 2a depicts the adsorption equilibrium data of pyridine onto granular activated carbon (GAC) at pH 10 and 25°C [26], whereas Fig. 2b shows the concentration decay curve at an initial concentration of 1011 mg/L. Additionally, in Fig. 2a the black filled squares presents the operation line or mass balance at different times. The operation line matches with the equilibrium data at point B indicating that equilibrium in both experimental conditions is equal.

Fig. 2b deploys predictions of concentration decay curve obtained with the SDM model employing the Langmuir isotherm to calculate $\partial q/\partial C_{AP}$. The mass transport parameters and the physical properties of GAC come from published work [26]. It is evident that the SDM only predicts the experimental data accurately at short times and fails to obtain the equilibrium because theoretically this condition is reached at point A in Fig. 2(a). Thus, the use of Langmuir model to calculate $\partial q/\partial C_{AP}$ results inadequate for application of SDM. On the other hand, the application of Freundlich model, which has a good fit of equilibrium data, results in numerical divergence for low concentrations because its derivative becomes undetermined. These facts illustrate that isotherm models must fulfill not only the prediction of equilibrium but also the prediction of the change of solute adsorbed when the intra-particle concentration changes.



(a)



(b)

Fig. 2. (Color online) a) Comparison between experimental equilibrium data of pyridine on GAC at 298.15 K and pH 10 (bullet), and predictions with Langmuir (black line) and Freundlich (dashed-red line) isotherm models. b) Concentration decay curve obtained with the SDM using Langmuir equilibrium model.

2.2. The artificial neural networks formulation

We propose a surrogate model to represent the equilibrium inside porous medium through ANN, *i.e.*

$$q(C_{AP}) = f_{ANN}(C_{AP}) \tag{9}$$

where $f_{ANN}(C_A)$ is the ANN output after training. The idea is to replace the mathematical model of isotherm, Eq. (6), by a simple static ANN which requires experimental equilibrium data. As a result, it is not required an explicit model to evaluate the amount of solute adsorbed within pores as a function of the pore concentration, since the ANN is the model itself. The ANN topology we used to train the equilibrium data is a multi-layer feed forward neural network composed by three layers; one neuron in the first layer, five neurons in the hidden layer and one neuron in the output layer (see Fig. 3). Hidden layer makes nonlinear processing of information, whereas at the output layer such processing is linear. The configuration of the hidden layer was the result of evaluations and analysis not only of the equilibrium data prediction but also on the derivative behaviors.

As learning rule to adjust the ANN parameters, we employed the simulated annealing method which is a stochastic global search method. The ANN parameters were initialized randomly following a uniform distribution in the interval (–0.25, 0.25), and they were allowed to vary freely.

Prediction after training process can be explicitly derived by evaluating the ANN forwardly according to:

$$f_{ANN}(C_{AP}) = \sum_{j=1}^{M} w_{j3}\left[g\left(w_{j2}C_{AP} + b_{j2}\right)\right] + b_{31} \tag{10}$$

where b_{j2} are the bias of the second layer, b_{31} is the bias of output layer, w_{j2} and w_{j3} are the weights of second and third

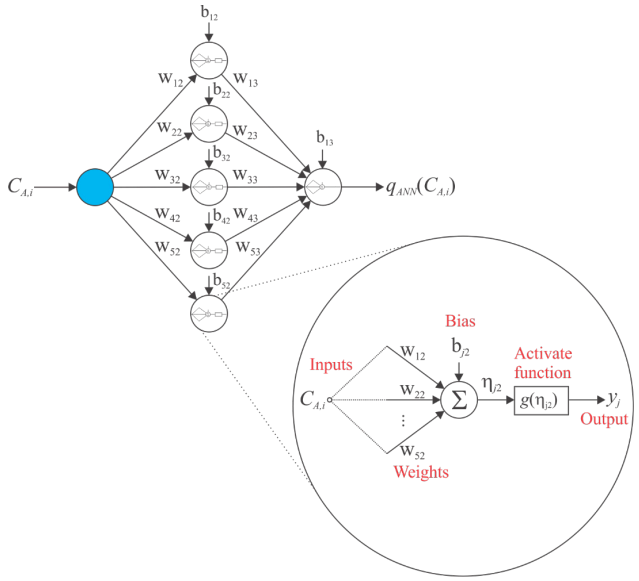


Fig. 3. Artificial neural network components (top) and local processing in a neuron of the hidden layer (bottom).

layers, respectively, and M is the number of neurons in the hidden layer.

Finally, the derivative is predicted by evaluating the ANN backwardly as follows:

$$\frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}} = \sum_{j=1}^{M} w_{j3} w_{j2} \left[g'\left(w_{j2} C_{AP} + b_{j2}\right)\right] \tag{11}$$

In the hidden layer we used the following sigmoidal activation function:

$$g(x) = \frac{2}{1+e^{-2x}} - 1 \tag{12}$$

whose derivative is given by

$$g'(x) = \operatorname{sech}(x)^2 \tag{13}$$

To train the ANN, we define an error function which measures the difference between experimental data and ANN predictions

$$E_{ANN}(\mathbf{W},\mathbf{b}) = \frac{\sum_{i=1}^{N}\left[q^{\exp}\left(C_{Ai}\right) - f_{ANN}\left(C_{Ai};\mathbf{W},\mathbf{b}\right)\right]^2}{\overline{q}^{\exp}\left(C_A\right)\overline{f}_{ANN}\left(C_A;\mathbf{W},\mathbf{b}\right)} \tag{14}$$

where $\mathbf{W}$ and $\mathbf{b}$ are the ANN parameters and N is the number of experimental data. $q^{\exp}(C_{Ai})$ is the experimental equilibrium adsorption value associated to the equilibrium concentration C_{Ai}, while $\overline{q}^{\exp}(C_A)$ and $\overline{f}_{ANN}(C_A;\mathbf{W},\mathbf{b})$ are their respective mean values. The training begins with random values of ANN parameters, and it stops when Eq. (14) satisfies some convergence criterion. This implies finding, through one optimization technique, the optimal set of parameters that minimizes the error function. As mentioned previously, we use the simulated annealing optimization technique for this purpose. This is a stochastic global search method that does not use gradients so that the best solution may be close to the global minimum. The discussion about the non-unicity problem is left aside since it is out of the scope of this work, but we recognize that the set of parameters may be not unique. Once the training procedure finishes, the optimal parameters can be used either in Eq. (10) or in (11) to predict the equilibrium or its derivative, respectively.

2.3. Method of lines to solve the SDM model

Following the assumption of spherical particles and the fact that the adsorption-diffusion phenomenon takes place only in the radial direction, the intra-particle mass balance can be rewritten as follows:

$$\varepsilon_p \frac{\partial C_{AP}}{\partial t} + \rho_p \frac{\partial q}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 D_s \rho_p \frac{\partial q}{\partial r}\right] \tag{15}$$

Furthermore, if we consider a neural network-based equilibrium model given by Eq. (9) to describe the relationship between the intra-particle concentration and the amount of solute adsorbed, then we evaluate the derivative of $q(C_{AP})$ respect to time according to

$$\frac{\partial q}{\partial t} = \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}}\frac{\partial C_{AP}}{\partial t} \tag{16}$$

and for a spherical particle, where relevant changes take place symmetrically in the radial direction, we only require the radial component of the spatial derivative, i.e.

$$\frac{\partial q}{\partial r} = \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}}\frac{\partial C_{AP}}{\partial r} \tag{17}$$

After substituting the last expressions into the mass balance, we get

$$\left(\varepsilon_p + \rho_p \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}}\right)\frac{\partial C_{AP}}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left[D_s \rho_p r^2 \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}}\frac{\partial C_{AP}}{\partial r}\right] \tag{18}$$

If we define the following coefficients:

$$\Phi(r,t) = \varepsilon_p + \rho_p \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}} \tag{19}$$

$$D_T(r,t) = D_s \rho_p r^2 \frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}} \tag{20}$$

the compact form of the mass balance is

$$\Phi(r,t)\frac{\partial C_{AP}}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(D_T(r,t)\frac{\partial C_{AP}}{\partial r}\right) \tag{21}$$

By developing the right-hand side we get

$$\frac{\partial C_{AP}}{\partial t} = \frac{1}{\Phi(r,t)r^2}\left(\frac{\partial D_T(r,t)}{\partial r}\frac{\partial C_{AP}}{\partial r} + D_T(r,t)\frac{\partial^2 C_{AP}}{\partial r^2}\right) \tag{22}$$

where derivative of $D_T(r,t)$ is given by

$$\frac{\partial D_T(r,t)}{\partial r} = D_s \rho_p \left[2r\frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}} + r^2 \frac{\partial}{\partial r}\left(\frac{\partial f_{ANN}\left(C_{AP}\right)}{\partial C_{AP}}\right)\right] \tag{23}$$

From now on, we eliminate the subscript in the intra-particle concentration, thus $C = C_{AP}$.

To apply the method of lines, we discretize the spatial domain by using uniform finite differences and leave the time dependence as continuous. For doing so, firstly we divide the radial domain into M elements of size $r_i = r_0 + (i-1)\Delta r$ for $i = 1, 2, \ldots, M$, where $r_0 = 0$, $r_M = R_p$ and $\Delta r = (r_M - r_0)/M$. Using symmetric finite differences, we discretize the space derivatives of first and second order and get the following expressions

$$\frac{\partial C_i}{\partial r} \approx \frac{C_{i+1}(t) - C_{i-1}(t)}{2\Delta r} \tag{24}$$

$$\frac{\partial^2 C_i}{\partial r^2} \approx \frac{C_{i+1}(t) - 2C_i(t) + C_{i-1}(t)}{\Delta r^2} \tag{25}$$

Space derivative of $D_T(r,t)$ is numerically approximated according to

$$\frac{\partial D_{Ti}(t)}{\partial r} \approx D_s \rho_p \left[2r_i \frac{\partial f_{ANN}(C_i(t))}{\partial C_i(t)} + \frac{r_i^2}{2\Delta r}\left(\frac{\partial f_{ANN}(C_{i+1}(t))}{\partial C_{i+1}(t)} - \frac{\partial f_{ANN}(C_{i-1}(t))}{\partial C_{i-1}(t)} \right) \right] \tag{26}$$

where as $\Phi(r,t)$ as follows

$$\Phi_i(t) = \varepsilon_p + \rho_p \frac{\partial f_{ANN}(C_i(t))}{\partial C_i(t)} \tag{27}$$

Substituting the last discretized equations into the mass balance and grouping some terms we obtain an equation for space nodes, *i.e.*

$$\frac{dC_i(t)}{dt} = \frac{1}{\Phi_i(t) r_i^2}\left(\frac{\partial D_{Ti}(t)}{\partial r} \frac{C_{i+1}(t) - C_{i-1}(t)}{2\Delta r} + D_{Ti}(t) \frac{C_{i-1}(t) - 2C_i(t) + C_{i+1}(t)}{\Delta r^2} \right) \tag{28}$$

$$\frac{dC_i(t)}{dt} = \left[\beta_i(t) - \alpha_i(t)\right] C_{i-1}(t) - 2\beta_i(t) C_i(t) + \left[\beta_i(t) + \alpha_i(t)\right] C_{i+1}(t) \tag{29}$$

where $\alpha_i(t) = \dfrac{1}{2r_i^2 \Phi_i(t) \Delta r} \dfrac{\partial D_{Ti}(t)}{\partial r}$ and $\beta_i(t) = \dfrac{D_{Ti}(t)}{\Phi_i(t) r_i^2 \Delta r^2}$ are space-time dependent coefficients.

This equation is used to generate the corresponding expressions for nodes $2 \le I \le M-1$. For nodes $i = 1$ and $i = M$ a special treatment is necessary as they correspond to the system boundaries. Due to the divergence at the origin in radial coordinates, we used $r_0 = 1.0 \times 10^{-9}$ instead of $r_0 = 0$, and defined that at this point $\partial C_1/\partial r = 0$ according to Eq. (5). In this way, the solute dynamic close to the origin is given by

$$\frac{dC_1(t)}{dt} = (\alpha_1 + \beta_1)\left(C_1(t) - C_2(t)\right) \tag{30}$$

At the interphase solution-particle, we used the equation of continuity of fluxes, Eq. (4), and the dynamic of this point is described by

$$\begin{aligned} \frac{dC_M(t)}{dt} &= \left[\beta_M(t) - \alpha_M(t)\right] C_{M-1}(t) - 2\beta_M(t) C_M(t) \\ &+ \left[\beta_M(t) + \alpha_M(t)\right]\left\{ C_{M-1}(t) + \frac{2k_L \Delta r}{D_{TM}}\left(C_A(t) - C_M(t)\right) \right\} \end{aligned} \tag{31}$$

The solute mass balance in the solution is given by

$$\frac{dC_A(t)}{dt} = -\frac{mSk_L}{V}\left(C_A(t) - C_M(t)\right) \tag{32}$$

As can be seen, we have a system of $M + 1$ ordinary differential equations and the same number of variables, so that to solve this initial value problem we only need the initial condition, which is given in Eq. (3), and a time integrator like those of Runge-Kutta family.

2.4. Scaling parameters

We analyze the adsorption equilibrium data and select a normalization rule to avoid saturating the hidden units of the ANN and trivial solutions in the ANN parameters. In this way, before training the input (equilibrium concentration) and output information (amount of solute adsorbed) were normalized according to

$$\bar{C}_{Ai} = \frac{C_{Ai} - \mu_{input}}{\sigma_{input}} \tag{33}$$

where μ_{input} is the average equilibrium concentration evaluated as follows:

$$\mu_{input} = \frac{1}{N}\sum_{i=1}^{N} C_{Ai} \tag{34}$$

And σ_{input} is the standard deviation, calculated according to

$$\sigma_{input} = \sqrt{\frac{\sum_{i=1}^{N}\left(C_{Ai} - \mu_{input}\right)^2}{N-1}} \tag{35}$$

In the same way, we normalized q_e, it is

$$\bar{q}_{ei} = \frac{q_{ei} - \mu_{output}}{\sigma_{output}} \tag{36}$$

whose average values and standard deviation are given by

$$\mu_{output} = \frac{1}{N}\sum_{i=1}^{N} q_{ei} \tag{37}$$

$$\sigma_{output} = \sqrt{\frac{\sum_{i=1}^{N}\left(q_{ei} - \mu_{output}\right)^2}{N-1}} \tag{38}$$

Once the training stage finishes, we rescale variables as $C_{Ai} = \sigma_{input}\bar{C}_{Ai} + \mu_{input}$ and $q_e = \sigma_{output}\bar{q}_e + \mu_{output}$. When we evaluate the derivative of the output regarding the input, i.e. $\partial f_{ANNi}/\partial C_{Ai}$, regarding scaled variables, we have

$$\frac{\partial f_{ANNi}}{\partial C_{Ai}} = \frac{\sigma_{output}}{\sigma_{input}} \frac{\partial \bar{q}_e}{\partial \bar{C}_{Ai}} \tag{39}$$

2.5. Inverse problem to estimate the mass transport parameters

The mass transport parameters (k_L and D_s) reported in Table 2, were estimated by solving the inverse problem related to each concentration decay curve. To formulate the inverse problem,we define the direct problem as an initial value problem with experimental or guess model parameters besides the initial and boundary conditions, Eqs. (1)–(5). By using the method of lines as described in Sec. 2.3 and a fourth-order Runge-Kutta time integrator, we solved the

direct problem which allows us to obtain the model predictions. Thus, we defined a dimensionless error function based on the difference between the concentration decay from experimental and model prediction data, *i.e.*

$$E_{IP}(k_L, D_s) = \frac{\sum_{i=1}^{N}\left[C_A^{Exp}(t_i) - C_A^{Model}(t_i)\right]^2}{\bar{C}_A^{Exp}(t) * \bar{C}_A^{Model}(t)} \quad (40)$$

where $C_A^{Exp}(t_i)$ is the experimental datum and $C_A^{Model}(t_i)$ is the numerical solution to the SDM model for some guessed values of the mass transport parameters, both at a time t_i $\bar{C}_A^{Exp}(t)$ and $\bar{C}_A^{Model}(t)$ are the corresponding mean values. The resulting non-linear inverse problem is

$$\min_{k_L, D_s} E_{IP} \quad subjet\ to: \quad k_L > 0 \ and \ D_s > 0 \quad (41)$$

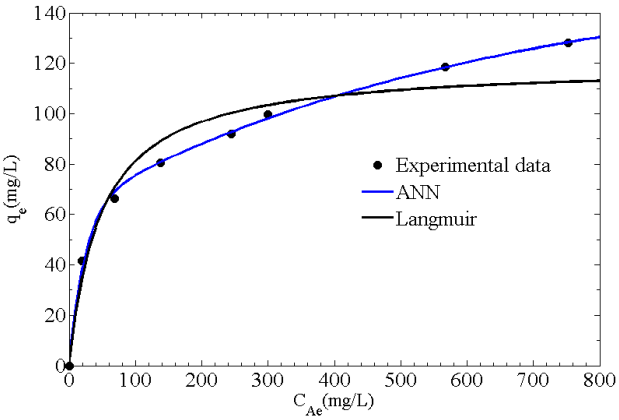
Note that each value of the error function implies a solution to the direct problem. The error function was systematically minimized through the genetic algorithms optimization method. We used reported values and random values as the starting point in the optimization process, and the result did not change. This fact is a signal that such mass transport values are optimal.

2.6. Adsorption equilibrium and rate adsorption data of pyridine onto granular activated carbon

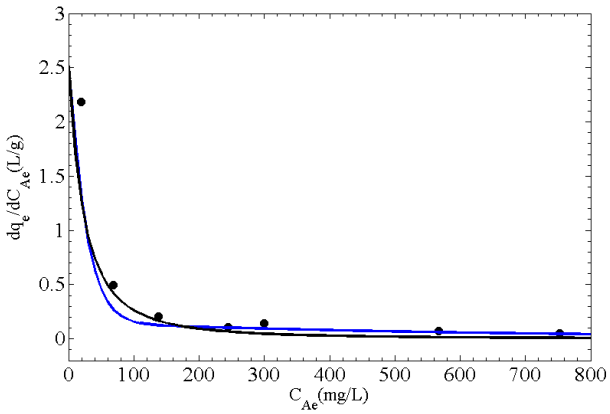
Adsorption equilibrium and rate adsorption data of pyridine on GAC used in this work were obtained from Ocampo-Perez et al. [26]. The adsorption equilibrium was obtaining by contacting 1 g of GAC and 480 mL of a solution with a known initial concentration of pyridine at pH = 10 in a batch adsorber. The initial pyridine concentration was varied from 20 to 800 mg/L. The pyridine solution remained in contact with the GAC particles until equilibrium was reached. On the other hand, a rotating basket batch adsorber was used to obtain the experimental concentration decay curves for the pyridine adsorption on GAC. This adsorber was composed of a 1 L three-neck reaction flask and an impeller with its blades replaced with stainless steel baskets. A pyridine solution of initial concentration between 100 to 1000 mg/L was poured in to the adsorber and a mass of 2 g of GAC particles were placed in the stainless steel mesh baskets, which were attached to a shaft connected to a variable speed motor (200 rpm). The adsorber was partially immersed in a constant temperature water bath controlled by a recirculator. The solution was periodically sampled (5 mL) and analyzed to determine the pyridine concentration. Sampling times were 0, 1, 3, 5, 10, 15, 20, 25, 30, 40, 50, 60, 90, 120, 150, and 180 min. The concentration of pyridine in solution vs. time, were plotted to obtain the concentration decay curve.

3. Results and discussions

Fig. 4a shows the experimental data of the pyridine adsorption equilibrium on activated carbon, where it is observed that the adsorption capacity reached was 124 mg/g at an equilibrium concentration of 800 mg/L. Furthermore, it is evident that experimental data showed the L type of Giles classification [27], indicating both that the aromatic ring of this molecule was adsorbed parallel to the carbon surface and there is no major competition between contaminant molecules and water molecules for the active adsorption centres on the carbon surface.



(a)



(b)

Fig. 4. (Color online) (a) Comparison between experimental data (black filled circles), predictions with Langmuir isotherm (continuous black line) and the ANN-Surface Diffusional Model (continuous blue line) for (a) equilibrium isotherm, and (b) derivative of the equilibrium isotherm.

The adsorption equilibrium data were interpreted with the Langmuir adsorption isotherm model. This model can be mathematically represented by the following equation:

$$q = \frac{q_m K C_{Ap}}{1 + K C_{Ap}} \quad (42)$$

The adsorption constants were evaluated by using the Rosenbrock and quasi-Newton estimation method. The values of the isotherm constants were q_m = 119.76 mg/g and K = 0.021 L/mg.

The adsorption equilibrium experimental data were employed to train ANN. Prior to this, we normalize the experimental data (inputs and outputs) with average val-

Table 1
Scaling parameters derived from the lab adsorption equilibrium data

Parameter	Value
σ_{output}	42.001
μ_{output}	78.289
σ_{input}	271.369
μ_{input}	261.328

ues, μ, and their standard deviations, σ, as presented in Table 1. We used a three-layered feed forward ANN to learn the equilibrium behavior, and for doing so, we minimize the error function, Eq. (16), by fitting the ANN parameters. Since the available experimental equilibrium data are few, we generate the final training set by interpolating the available information. Once we get the optimal ANN parameters, we used them to predict equilibrium and its derivative.

Fig. 4a ilustrates the prediction of experimental data with Langmuir isotherm model highlighting a good approach for low equilibrium concentration values (C_{Ae}<100 mg/L) whereas for higher values the prediction worsen; between 100<C_{Ae}<400 mg/L an overestimation of q_{Ae} is observed, whereas for C_{Ae}<400 mg/L underestimation is predicted. Note that Langmuir's estimation of asymptotic behavior does not conform to experimental data. This behavior remains for prediction of derivative of solute adsorbed. Fig. 4b also shows that derivative decreases sharply for low concentrations and then it monotonically decreases until it gets a constant value. Even though equilibrium is not well predicted by Langmuir isotherm, its derivative present reasonable estimations for C_{Ae}<150 mg/L. For larger values, predictions differ from the main trend as almost no changes in the derivative are observed. As a consequence, predictions at low concentrations will be better than at higher values of C_{A0}.

On the other hand, the ANN prediction agrees well with experimental data for all values of equilibrium concentrations; however, when predicting derivatives, some discrepancies emerge mainly at low equilibrium concentrations (C_{Ae}<140 mg/L) where underestimations are observed. This fact results in differences in the dynamic of concentration decay curves, but not in the equilibrium as can be seen in Fig. 4(b) for C_{Ae}<100 mg/L. We stress that a reasonable prediction of equilibrium does not necessarily guarantee accuracy in the prediction in its derivative. Nevertheless, prediction of concentration decay strongly depends on the combination of both quantities.

Once we get the optimal ANN parameters, we used them to predict equilibrium and its derivative. In this way, the equilibrium relationship derivative given by the ANN, Eq. (11), was substituted into Eqs. (16) and (17) to proceed with the numerical solution of the SDM. The ANN-based SDM, Eqs. (1)–(5), was solved numerically, and this permitted us to obtain the concentration decay dynamic curves. We used the method of lines, Eqs. (29)–(32), to get the time evolution of solute concentration not only at the solution but also inside adsorbent particles.

Turning our attention to the mass transport parameters reported in Table 2, we remark that their estimations depend on the solution to the appropriated inverse problem formulated for each concentration decay curve. In this sense, firstly we solved the direct problem as an initial-value problem given all model parameters. Thus, the inverse problem consists in constructing an error function based on the difference between the concentration decay from experimental data and model prediction. In this case, we used the genetic algorithms optimization method to minimize such error function systematically.

Table 2
Mass transport parameters and initial and equilibrium concentrations for pyridine adsorption onto GAC at pH = 10, particle radius = 0.51 mm, RPM = 200 and T = 25°C

C_{A0} (mg/L)	q_e (mg/g)	Ocampo-Perez et al. [26]		This Work	
		$k_L\times10^3$ (cm/s)	$D_s\times10^7$ (cm²/s)	$k_L\times10^3$ (cm/s)	$D_s\times10^7$ (cm²/s)
102.0	41.5	20.4	0.55	6.08	1.36
201.0	66.3	8.7	1.54	6.50	1.47
300.0	80.8	8.4	1.85	5.77	2.02
500.0	99.9	14.7	2.00	5.23	2.89
1011.0	129.4	10.4	3.77	5.25	6.03

Figs. 5a–d show the concentration decay curves for experiments showed in Table 2 and the prediction of SDM coupled with the ANN. If we define three regimes in the decay curves; short- middle- and long-term, we can analyze results in term of them. Such regimes correspond to the dynamic behavior of the solute concentration in solution; short-term can be related to the concentration dynamic where a constant slope is observed and the mass transfer in the solution is dominant. Middle-term can be associated to the concentration dynamic where the slope changes continuously what means that both solution mass transfer and surface diffusion compete and it is not clear which of them is dominant. Finally, long-term correspond to a condition where zero-slope is observed, and the whole process is dominated by the equilibrium taking place at the porous surface. It is interesting to observe an increment in the slope when the initial concentration increases as well. This short-term behavior is inversely associated with the mass transfer coefficient, just as shown in Table 2, which decreases accordingly. The middle-term behavior is difficult to predict even with an ANN-based isotherm model, see Figs. 5a and 5d, since this regime demands to know with high precision the contribution of the neglected stages in the adsorption-diffusion process. As can be remembered, in our formulation we do not consider thin-film diffusion either the intra-particle diffusion, and it is possible that in some cases their contributions may be relevant in the transitory regime. The long-term behavior is well-captured in all cases what means that prediction of the isotherm derivative through an ANN is a trustable option when classical adsorption equilibrium models do not interpret lab data accurately. In general, it is appreciated a good agreement between the prediction of model and lab data at both short and long times.

Complementary, values of mass transport parameters estimated with the SDM model are summarized in Table 2, where it can be noted that the mass transfer coefficient k_L remained constant regardless the operating

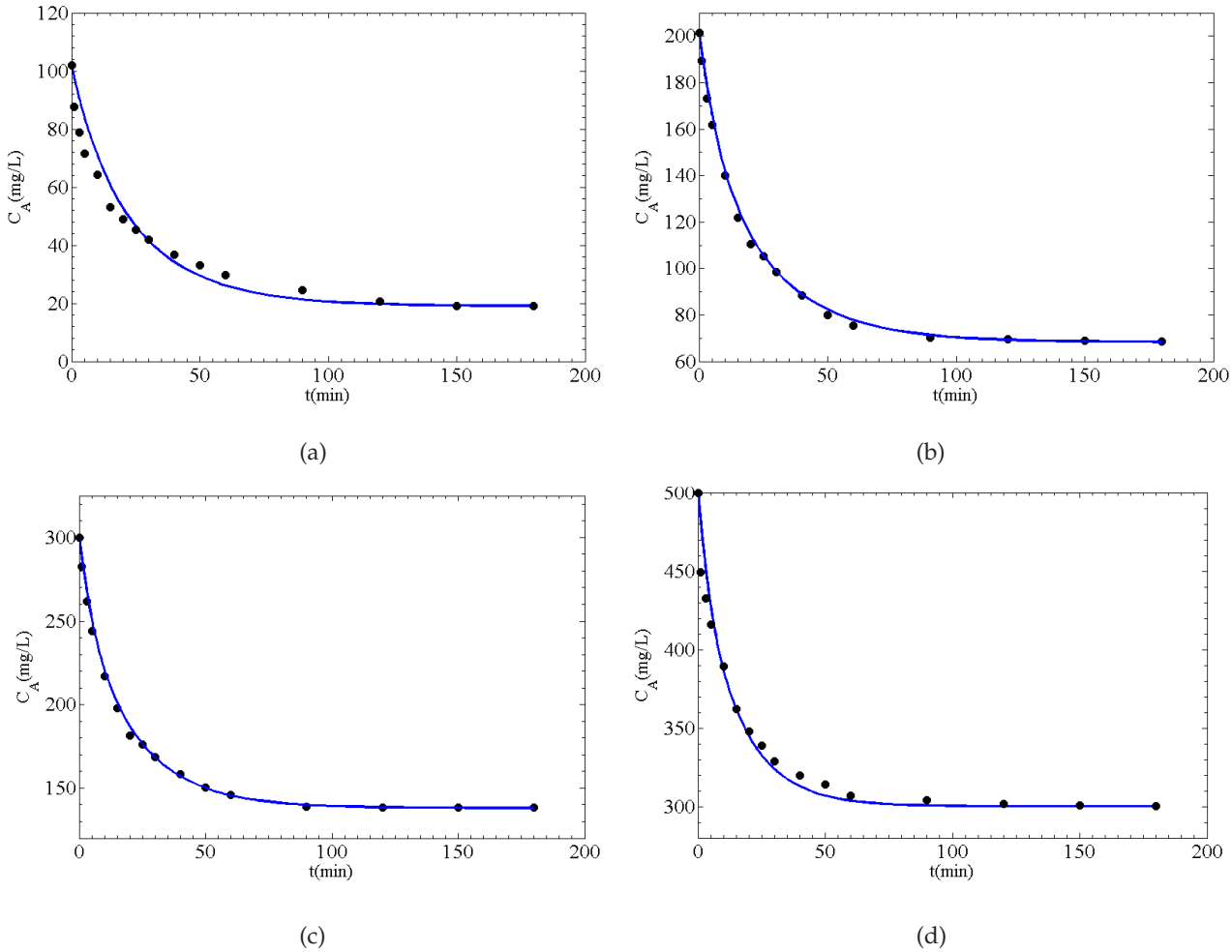


Fig. 5. (a) Comparison between experimental data (black filled circles) and predictions with the ANN-Surface Diffusional Model for concentration decay for the cases: (a) $C_{A0} = 100$, (b) $C_{A0} = 200$, (c) $C_{A0} = 300$, and (d) $C_{A0} = 500$ *mg/L*.

conditions. On the other hand, the surface effective diffusion coefficient D_s increases as the adsorbed pyridine does it. This phenomenon has already been reported in the literature and it indicates that larger adsorbed mass promotes the mobility of adsorbed molecules toward energetically less favorable sites [26]. Finally, coefficients k_L and D_s were compared with those reported in the literature for the pyridine-GAC system. The values reported in Table 1 have the same order of magnitude indicating that our approach represents an accurate option to surrogate classical models when available experimental data are scarce. Another interesting and valuable resulting fact is that with the ANN-based equilibrium relationship the equilibrium values for all presented cases were well-predicted, the observed regularity in the mass transfer coefficient k_L and the clearly trend for D_s.

4. Conclusions

In this work, a novel approach to predict the equilibrium relationship with an artificial neural network is proposed. We take advantage of the capabilities of the artificial neural networks to construct a function from a reduced set of equilibrium adsorption data. With the ANN, we found that both equilibrium and its derivative can be well-predicted with a precision depending on the amount of available experimental data. The derivative of equilibrium relationship was employed in a surface diffusion model to determine concentration decay curves in the liquid solution. Good agreement of model predictions with experimental data was found for short- and long-term. Some drawbacks were observed for the middle-term predictions, what suggest that additional information should be included in the surface diffusion model. Also, through a lab-based data inverse problem formulation, we determine the mass transport parameters for each experimental case following consistent tendencies with those reported in previous work. For instance, the mass transfer coefficient k_L almost remains constant (as expected) while the surface effective diffusion coefficient D_s follows an increasing trend with solute concentration which agrees with published result.

Both the equilibrium relationship and its derivative are relevant to predict the dynamics of solute concentra-

tion decay curves in the short- and long-term. Short-term predictions can be associated with the performance of the derivative whereas the long-term behavior depends on the capability of equilibrium model to predict the experimental isotherm data. Our results suggest that ANN as a surrogate equilibrium model is a valuable tool which could be used to design equipment for industrial applications and fundamental research.

Acknowledgments

EC-HH acknowledges the support from CONACYT trough cátedras program at CIDESI.

Symbols

C_A — Pyridine concentration in the solution, *mg* L^{-1}
C_{A0} — Initial concentration of pyridine in the solution, *mg* L^{-1}
C_{AP} — Pyridine concentration within adsorbent particles, *mg* L^{-1}
C_p — Effective intra-particle diffusivity, cm^2 s^{-1}
D_S — Effective surface diffusivity, cm^2 s^{-1}
D_T (r,t) — Global coefficient defined by ()
E_{ANN} — Error function in the artificial neural network
E_{IP} — Error function in the inverse problem
k_L — Solution mass transfer coefficient, *cm* s^{-1}
M — Number of experimental data
m — Mass of GAC, *g*
N — Number of neurons in the hidden layer
q_e — Mass of pyridine at equilibrium per mass of GAC, mg g^{-1}
R_p — Particle radius, *cm*
S — External surface area of GAC per unit mass, cm^2 g^{-1}
V — Volume of the pyridine solution, *L*
$\Phi(r,t)$ — Global coefficient defined by ()

Greek

ε_P — Void fraction of GAC
μ_{input} — Average equilibrium concentration, *mg* L^{-1}
μ_{output} — Average mass of pyridine at equilibrium per mass of GAC, *mg* g^{-1}
σ_{input} — Standard deviation of the average equilibrium concentrarion, *mg* L^{-1}
σ_{output} — Standard deviation of mass of pyridine at equilibrium per mass of GAC, *mg* g^{-1}
ρ_P — Mass density of GAC, *g* cm^{-3}

References

[1] A.M. Ghaedi, A. Vafaei, Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review, Adv. Colloid Interface Sci., 245 (2017) 20–39.
[2] K.M. Chu, Prediction of two-metal biosorption equilibria using a neural network, Eur. J. Mineral Proc. Environ. Protect., 3(1) (2003) 119–127.
[3] K. Snigdha, Modeling phenol adsorption in water environment using artificial neural network, Int. Res. J. Environ. Sci., 2(7) (2013) 39–43.
[4] R. Gomez-Gonzalez, F.J. Cerino-Córdova, A.M. Garcia-León, E. Soto-Regalado, N.E. Davila-Guzman, J.J. Salazar-Rabago, Lead biosorption onto coffee grounds: Comparative analysis of several optimization techniques using equilibrium adsorption models and ANN, J. Taiwan Inst. Chem. Eng. 68 (2016) 201–210.
[5] B.G. Saucedo-Delgado, D.A. De Haro-Del Rio, L.M. González-Rodríguez, H.E. Reynel-Ávila, D.I. Mendoza-Castillo, A. Bonilla-Petriciolet, J. Rivera de la Rosa, Fluoride adsorption from aqueous solution using a protonated clinoptilolite and its modeling with artificial neural network-based equations, J. Fluor. Chem., 204 (2017) 98–106.
[6] P.S. Ghosal, K.V. Kattil, M.K. Yadav, A.K. Gupta, Adsorptive removal of arsenic by novel iron/olivine composite: Insights into preparation and adsorption process, J. Environ. Manage., 209 (2018) 176–187.
[7] I. Ali, O.M.L. Alharbi, Z.A. Alothman, A.Y. Badjah, A. AAlwarthan, A.A. Basheer, Artificial neural network modelling of amido black dye sorption on iron composite nanomaterial: Kinetics and thermodynamics studies, J. Mol. Liq., 250 (2018) 1–8.
[8] S.K. Ashan, N. Ziaeifar, R. Khalilnezha, Artificial neural network modelling of Cr(VI) surface adsorption with NiO nanoparticles using the results obtained from optimization of response surface methodology, Neural Comput. Appl., 29 (2018) 969–979.
[9] M. Pazouki, M. Zabihi, J. Shayegan, M.H. Fatehi, Mercury ion adsorption on AC@Fe_3O_4-NH_2-COOH from saline solutions: Experimental studies and artificial neural network modeling, Korean J. Chem. Eng., 35(3) (2018) 671–683.
[10] O.M.L. Alharbi, Sorption, kinetic, thermodynamics and artificial neural network modelling of phenol and 3-amino-phenol in water on composite iron nano-adsorbent, J. Mol. Liq., 260 (2018) 261–269.
[11] S. Agarwal, I. Tyagi, V.K. Gupta, M. Ghaedi, M. Masoomzade, A.M. Ghaedi, B. Mirtamizdoust, Kinetics and thermodynamics of methyl orange adsorption from aqueous solutions—artificial neural network-particle swarm optimization modeling, J. Mol. Liq., 218 (2016) 354–362.
[12.] R. Leyva-Ramos, C.J. Geankoplis, Diffusion in liquid-filled pores of activated carbon. I. Pore volume diffusion, Can. J. Chem. Eng., 72 (1994) 262–271.
[13] K. Hornik, M. Stinchcombe, H. White, Multilayer feed forward networks are universal aproximators, Neural Netw., 2 (1989) 359–366.
[14] P. Cardaliaguet, G. Euvrard, Approximation of a function and its derivative with a neural network, Neural Networks, 5 (1992) 207–220.
[15] Z. Zainuddin, O. Pauline, Function approximation using artificial neural networks, Int. J. Syst. Appl. Eng. Develop., 1(4) (2007) 173–178.
[16] S. Yang T.O. Ting, K.L. Man, S. SU. Guan, Investigation of neural networks for function approximation, Procedia Comp. Sci., 17 (2013) 586–594.
[17] S. Ferrary, R.F. Stengel, Smooth function approximation using neural networks, Trans. Neural Networks, 16(1) (2005) 24–38.
[18] T. Nguyen-Thien, T. Tran-Cong, Approximation of functions and their derivatives: A neural network implementation with applications, Appl. Math. Model., 23 (1999) 687–704.
[19] R. Ocampo-Pérez, R. Leyva-Ramos, J. Rivera-Utrilla, J.V. Flores-Cano, M. Sánchez-Polo, Modeling adsorption rate of tetracyclines on activated carbons from aqueous phase, Chem. Eng. Res. Des., 104 (2015) 579–588.
[20] S. Tripathi, R.F. Tabor, Modeling two-rate adsorption kinetics: Two-site, two-species, bilayer and rearrangement adsorption processes, J. Colloid Interf. Sci., 476 (2016) 119–131.
[21] M. Schwaab, E. Steffani, E. Barbosa-Coutinho, J.B. Severo Júnior, Critical analysis of adsorption/diffusion modelling as a function of time square root, Chem. Eng. Sci., 173 (2017) 179–186.
[22] A. Muthukkumaran, K. Aravamudan, Combined Homogeneous Surface Diffusion Model – Design of experiments

approach to optimize dye adsorption considering both equilibrium and kinetic aspects, J. Environ. Manage., 204 (2017) 424–435.
[23] S. Eris, S. Azizian, Extension of classical adsorption rate equations using mass of adsorbent: A graphical analysis, Sep. Purif. Technol., 179 (2017) 304–308.
[24] G. Marbán, L.A. Ramírez–Montoya, H. García, J.A. Menéndez, A.M.A. Montes–Morán, Load–dependent surface diffusion model for analyzing the kinetics of protein adsorption onto mesoporous materials, J. Colloid. Interf. Sci., 511 (2017) 27–38.
[25] R. Ocampo-Perez, R. Leyva-Ramos, M.Sanchez-Polo, J. Rivera-Urtilla, Role of pore volumen and surface diffusion in the adsorption of aromatic compounds on activated carbon, Adsorption, 19(1) (2013) 945–957.
[26] R. Ocampo-Perez, R. Leyva-Ramos, P. Alonso-Davila, J. Rivera-Utrilla, M. Sanchez-Polo, Modeling adsorption rate of pyridine onto granular activated carbon, Chem. Eng. J., 165 (2010) 133–141.
[27] C.H. Giles, D. Smith, A. Huitson, A general treatment and classification of the solute adsorption isotherm. I. Theoretical, J. Colloid Interf. Sci., 47(3) (1974) 755.