Chemical Engineering Journal xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Chemical Engineering Journal

journal homepage: www.elsevier.com/locate/cej

Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks

C. Faur[a,*], A. Cougnaud[a], G. Dreyfus[b], P. Le Cloirec[c]

[a] UMR Cirad 016 Génie des Procédés Eaux – Bioproduits, Université de Montpellier 2, place Eugène Bataillon, 34095 Montpellier Cedex 5, France
[b] Ecole Supérieure de Physique et de Chimie Industrielles (ESPCI-ParisTech), Laboratoire d'Electronique (CNRS UMR 7084), 10 rue Vauquelin, 75231 Paris Cedex 05, France
[c] Ecole Nationale Supérieure de Chimie de Rennes (ENSCR), UMR CNRS 6226 "Sciences Chimiques de Rennes", Campus de Beaulieu, 263 Avenue du Général Leclerc, 35700 Rennes, France

ARTICLE INFO

Article history:
Received 6 December 2007
Received in revised form 6 February 2008
Accepted 20 February 2008

Keywords:
Feed-forward neural networks
Recurrent neural networks
Water treatment
Adsorption
Activated carbon
Pesticides

ABSTRACT

A black-box approach is performed to model the breakthrough of activated carbon filters by pesticides present in surface waters with a recurrent neural network (in input–output form) and, as a baseline, by a feed-forward neural network, which includes time as an input variable.

In a first part, isotherm experimental runs are performed in static reactors, using five activated carbons and three pesticides, under different operating conditions. The influence of adsorbent and adsorbate properties on adsorption performance in highlighted for pure and natural waters. The modelling of competitive adsorption isotherms by the equivalent background compound (EBC) model enables to determine the Freundlich parameters of the EBC which is the part of natural organic matter in competition with the pesticide.

In a second part, experimental breakthrough curves of pesticide in a surface water are assessed in fixed-beds and modelled using neural network approaches. The selection of data is based on physical and statistical approaches, equilibrium parameters assessed in static reactors being considered as influential variables to take into account the competitive adsorption phenomenon. Static and recurrent neural networks provide both high determination coefficients ($R^2 > 0.990$) and low root mean square modelling errors (RMSE < 0.035 while standard deviation of data is equal to 2.9%) for the prediction of the global breakthrough curves. To model the breakthrough zone ($C/C_0 < 0.1$), the recurrent neural network, with a smaller number of parameters, is however more accurate than the feed-forward one, since the process to be modelled is dynamic.

1. Introduction

In the last decade, the scientific community has shown a growing interest in compounds interfering with the normal endocrine function of wildlife and possibly humans. Pesticides belong to these endocrine disrupters active at low concentration and their removal from surface waters is thus essential [1]. Adsorption onto activated carbon filters is the most effective and widely used method for removing pesticides. Traditionally, the breakthrough time, which indicates the time for termination of the operation and replacement (or regeneration) of the contaminated adsorbent, is estimated by the process knowledge of managers of drinking water supply systems, or by measurements of the outlet concentrations. However, these measurements are expensive and time-consuming.

To overcome the above limitations, models have been developed to predict the concentration profiles at the outlet of the filter ("breakthrough curves"). At present, these models are knowledge-based models based on transient material balance, liquid and intra-particle mass transfer, and adsorption equilibrium equations, usually solved numerically using finite element or finite difference methods [2]. They differ mainly in the assumptions made on the transfer mechanism(s) within the particles, which can be diffusion in the liquid-filled pores (pore diffusion), diffusion in the adsorbed phase (surface diffusion), and both mechanisms in parallel. Among mathematical models of current use, the *linear driving force model* [3], the *pore diffusion model* [4] or the *homogeneous surface diffusion model* [5] are notable. Such models were used for describing the breakthrough curves of monocomponent solutions of metal ions [6], organic micropollutants [7] or dyes [8] onto activated carbon. Some studies have showed that those approaches are

Abbreviations: A, atrazin; AC, activated carbon; ADE, atrazin-desethyl; DOC, dissolved organic compound; EBC, equivalent background compound; NN, neural network; NOM, natural organic matter; TFSM, triflusulfuron-methyl; TOC, total organic carbon.

* Corresponding author. Tel.: +33 4 67 14 38 70.
E-mail address: catherine.faur@univ-montp2.fr (C. Faur).

1385-8947/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cej.2008.02.015

Nomenclature

C_i	concentration of compound i [M L^{-3}]
C_0	inlet concentration in the fixed-bed [M L^{-3}]
$C_{0,i}$	initial concentration of compound i [M L^{-3}]
d_p	particle diameter [L]
E	relative error (defined by Eq. (12))
E_p	leave-one-out score (defined by Eq. (5))
f	activation function of the neural network
J	least square function
k	positive constant which describes discrete time, $t = kT$
K_i	Freundlich parameter of compound i [M M^{-1} M$^{3/n}$ L$^{-3/n}$]
m	active carbon weight [M]
M	molecular weight [M mol^{-1}]
n_i	Freundlich parameter of compound i
N	number of experimental data
q_i	adsorption capacity of compound i [M M^{-1}]
$RMSE_T$	root mean square error on the test set (defined by Eq. (6))
$R_k^{(-k)}$	prediction error on example k when the latter is withdrawn from the training set
R_T^2	determination coefficient on the test set (defined by Eq. (7))
S	solubility [M L^{-3}]
t	time [T]
T	period [T]
$\mathbf{u}$	m-vector of external variables
u_j	input neuron j
u_0	velocity in the activated carbon bed [L T^{-1}]
V	volume of the solution [L^3]
V_{meso}	mesopore volume [L^3 M^{-1}]
$V_{microsec}$	volume of secondary micropores [L^3 M^{-1}]
w_{ij}	parameters of the neural network, connecting neurons i and j
$\mathbf{x}$	n-vector of the state variables
y_i	output of neuron i
y_k	prediction of the output by the model for observation k
y_p^k	measured value of the process output for observation k

Greek symbols

ψ	difference between experimental and modelled values (defined by Eq. (11))
$\emptyset_{med\ \mu pores}$	median micropores width [L]

satisfactory for modelling binary dynamic adsorption, but they are not appropriate in the case of complex multi-solute solutions [9]. In addition, despite most of them rely on mass transfer parameters which are related to pore characteristics known to influence strongly the adsorption of microorganics alone in solution or in the presence of natural organic matter [10,11], they do not take them explicitly into account.

In recent years, neural networks (NN) have been successfully applied to a wide variety of domains; in the field of water treatment, they have been used, e.g. for predicting the behaviour of wastewater treatment plants [12], for simulating a combined humic substances coagulation and membrane filtration process [13] or for estimating H_2O_2 addition critical point in a decoloration process [14]. Neural networks are parameterised non-linear models whose parameters are estimated by training from examples. They can be advantageously used when no satisfactory prior knowledge is available about the physical process, and about the mechanisms that it involves. Furthermore, they require a smaller number of parameters than alternative parameterised non-linear regression models such as polynomials: the number of parameters of neural networks varies linearly with the number of variables of the process, whereas it varies exponentially for polynomials [15,16].

A neural network (see for instance [17,18]) is a combination of parameterised functions called neurons. The most popular neural network (termed multi-layer perceptron) features "hidden" neurons, each of which computes a non-linear function of a weighted sum of its variables, and an "output neuron", which computes a linear combination of the outputs of the hidden neurons. Therefore, a multi-layer perceptron is essentially a static model, since it involves merely a combination of functions of the variables of the process to be modelled. In order to model dynamic processes, recurrent neural networks are appropriate. In their simplest form (*input–output form*), the model output is fed back to its input with one (or more) unit time delay(s) and is thus a non-linear function of exogenous inputs and of its past output(s). Input–output recurrent neural networks are a restricted form of the more general *state-space* recurrent neural network [19]. Feed-forward neural networks were applied successfully to the modelling of adsorption processes in a static reactor [20,21] or to the prediction of breakthrough parameters of volatile organic compounds onto a granular activated carbon [22]. However, despite the widespread use of recurrent neural networks for modelling non-linear dynamic phenomena (see for instance [23–25]), they have never been applied to the prediction of pollutant removal by adsorption onto activated carbon.

The present paper reports an investigation of the adsorption of three pesticides onto activated carbons (AC) with different characteristics in terms of shapes, dimensions and pore properties. In a first part, experimental isotherm curves of pesticides in synthetic and natural waters are modelled to get equilibrium parameters of monocomponent and competitive adsorption. In the second part of the paper, the breakthrough curves of activated carbon filter are modelled by a recurrent neural network, and a feed-forward neural network as a baseline. Equilibrium parameters assessed in the first part are considered as influent factors of the model.

2. Theory of modelling by neural networks

By contrast to a knowledge-based approach, a black-box model relies on measurements only: a model is sought within a family of parameterised functions, and its parameters are estimated by minimizing the modelling error on a set of available measurements. Both static and dynamic black-box models can be designed.

In the present work, *static* models were sought in the family of feed-forward neural networks with one layer of hidden neurons (shown graphically in Fig. 1a), where the output y_i of neuron i is a non-linear function f of the weighted sum of its variables u_j:

$$y_i = f\left(\sum_j w_{ij} u_j\right) \quad (1)$$

w_{ij} are the parameters of the neuron, and f is its activation function. Just as in affine or polynomial models, one of the variables is a constant equal to 1 (called "bias"). When designing black-box models, the activation function is generally taken as an s-shaped ("sigmoid") function, such as a tanh function, which guarantees the universal approximation property [16]. The output neuron simply performs a weighted sum of its inputs (without non-linearity) so that the model output is a weighted sum of the outputs of the hidden neurons.

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

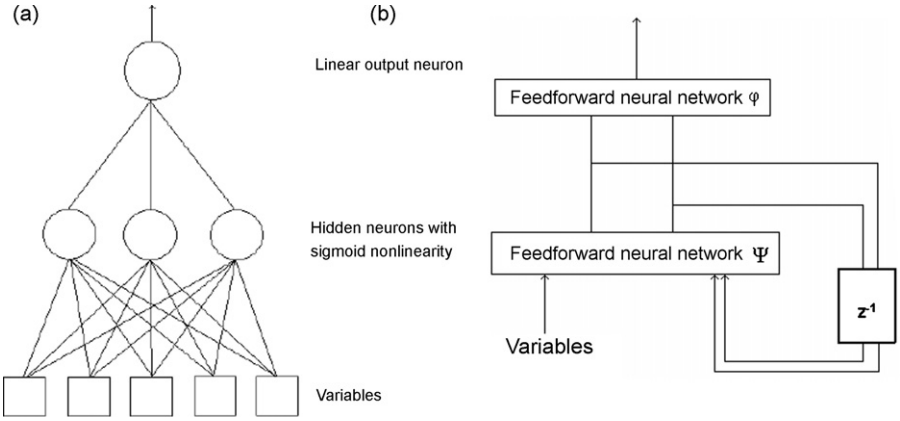

Fig. 1. Pictorial representations of (a) a feed-forward neural network and (b) a state-space recurrent neural network of order 2. z^{-1} is the conventional symbol for a unit delay.

The "training" of the neural network consists in estimating the values of the parameters that minimize the least squares cost function:

$$J(\mathbf{w}) = \sum_k (y_p^k - y_k(\mathbf{w}))^2 \quad (2)$$

where the summation runs on all elements of an available set of observations (or "examples") called the "training set", y_p^k is the measured value of the process output for observation k, y_k the prediction of the model for observation k, and $\mathbf{w}$ is the vector of parameters of the model. Any conventional gradient-based optimization algorithms can perform the minimization of that cost function with respect to the parameters of the model.

The above description is relevant to *static* black-box neural models, which are considered as a baseline in this study. When a *dynamic* black-box model is necessary in order to provide an appropriate model of the process, the most general form of the model is the state-space form, which can be written, in discrete time form, as

$$\begin{aligned} \mathbf{x}(k+1) &= \Psi(\mathbf{x}(k), \mathbf{u}(k)), \quad \text{state equation} \\ y(k+1) &= \varphi(\mathbf{x}(k+1)), \quad \text{observation equation} \end{aligned} \quad (3)$$

where $\mathbf{x}$ is the n-vector of the state variables, $\mathbf{u}$ the m-vector of external variables, y the model output (a single-output model is considered here), and k a positive integer that describes discrete time ($t = kT$, where T is the sampling period). The dimension n of vector $\mathbf{x}$ is the *order* of the dynamic model.

If the state vector is the vector of the n most recent outputs of the model $[\, y(k) \quad y(k-1) \quad \cdots \quad y(k-n+1)\,]^T$, the model can be written as

$$y(k+1) = \Phi(y(k), y(k-1), \ldots, y(k-n+1), \mathbf{u}(k)) \quad (4)$$

which is called the input–output form of a dynamic model. It can be shown that state-space forms are sometimes more difficult to train but may be more general and more parsimonious than input–output forms. Functions φ, ψ and the resulting Φ are computed by feed-forward neural networks. A state-space neural network is shown pictorially in Fig. 1b.

The design of a static black-box model requires: (i) selecting the relevant variables, whose influence on the output is larger than the influence of noise, (ii) training models of increasing complexity, (iii) selecting the model that generalizes best, i.e. provides the most accurate predictions in situations that are not present in the training set, (iv) estimating the performance of the selected model on fresh data, which was used neither for training nor for model selection.

In the present investigation, variable selection was performed with the random probe method, as described in [26].

Networks of increasing complexity were subsequently trained. As mentioned above, the training of neural networks consists in minimizing the least squares cost function (2). The gradient of the cost function is first computed by an algorithm (specific to neural networks) called "backpropagation" [18]; that gradient is provided to a general-purpose optimization algorithm, which, in the present study, was the Levenberg–Marquardt algorithm.

Model selection was performed by estimating the generalization ability of the models trained as described, using the "leave-one-out score" E_p:

$$E_p = \sqrt{\sum_{k=1}^{n} [R_k^{(-k)}]^2} \quad (5)$$

where $R_k^{(-k)}$ is the prediction error on the example k when the latter has been withdrawn from the training set and the model has been trained with all other examples. The leave-one-out score E_p is known to be an unbiased estimate of the generalization error of the model. Since the computation of the leave-one-out score is computer-intensive, approximations of the leave-one-out errors $R_k^{(-k)}$ were computed by the "virtual leave-one-out" method, described in [27].

In addition, a test set, made of examples that were used neither for training nor for model selection, was used for estimating the performance of the model selected by the procedure described above. The performance was estimated through the root mean square error RMSE_T and determination coefficients R_T^2 between predicted and observed responses on the test set:

$$\mathrm{RMSE}_T = \sqrt{\frac{\sum_{k=1}^{n} (y_p^k - y_k)^2}{n}} \quad (6)$$

$$R_T^2 = \frac{\sum_{k=1}^{n} (y_k - \bar{y}_p^k)^2}{\sum_{k=1}^{n} (y_p^k - \bar{y}_p^k)^2} \quad (7)$$

3. Materials and methods

3.1. Adsorbents

Five commercial activated carbons were selected for this study. They were produced from different raw materials and provided under different forms (pellets, extruded, fibres), yielding a large range of pore characteristics, which are known to have an influence

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

Table 1
Main characteristics of activated carbon materials

Commercial name	NC-60	Picabiol	Cal 12-40	Row 0.8	WWP3
Manufacturer	Pica	Pica	Chemviron	Norit	Actitex
Precursor	Coconut	Wood	Coal	Peat	RAYON
Form	Pellets	Pellets	Pellets	Extruded	Fibers
Particle diameter, d_p (mm)	1.23	1.16	1.04	0.92	0.01
Specific surface area ($m^2\,g^{-1}$)	1232	1575	1103	913	1148
Total pore volume ($cm^3\,g^{-1}$)	0.590	0.992	0.646	0.647	0.608
Micropore volume ($cm^3\,g^{-1}$)	0.513	0.625	0.466	0.388	0.467
Microporosity (vol.%)	87.0	63.1	72.2	59.9	76.8
Primary microporosity (%)	81.5	46.8	59.8	54.6	36.6
Secondary microporosity (%)	5.5	16.3	12.4	5.3	40.2
Median micropores width (Å)	5.97	6.82	5.83	5.52	5.45
Macropore volume ($cm^3\,g^{-1}$)	0.242	1.069	0.646	0.375	0.287

on pesticide adsorption [28]. Prior to their use, AC were washed with deionised water and dried at 100 °C overnight.

Their main properties are described in Table 1. Particle diameters were measured from scanning electron microscopic observations for fibres [29] and are mean diameters determined by sifting of granules and extruded particles. Mercury porosimeter Micromeritics Autopore IV 9500 allowed an assessment of particles macroporosity.

Pore parameters were measured by nitrogen adsorption at 77.7 K with a Micromeritics ASAP 2010 apparatus. Micropore volume was assessed by the Howarth and Kawazoe theory [30] whereas specific surface area was determined according to the Brunauer–Emmett–Teller method [31]. The percentages of primary micropores (<0.8 nm) and secondary micropores (0.8–2 nm) were also calculated. According to micropore percentage and median micropore width, mesopore concentrations may be ranked as: Picabiol > Norit Row 0.8 > Chemviron Cal 12-40 > Pica NC-60 > WWP3.

3.2. Adsorbates

Three pesticides were used as the target adsorbates: atrazin (A, $M = 216\,g\,mol^{-1}$, solubility $S = 35\,g\,L^{-1}$, $V_m = 180\,cm^3\,mol^{-1}$, $pK_a = 1.68$), atrazin-desethyl (ADE, $M = 187\,g\,mol^{-1}$, solubility $S = 3200\,g\,L^{-1}$, $V_m = 157\,cm^3\,mol^{-1}$, $pK_a = 1.00$) and triflusulfuron-methyl (TFSM, $M = 492\,g\,mol^{-1}$, solubility $S = 43\,g\,L^{-1}$, $V_m = 339\,cm^3\,mol^{-1}$, $pK_a = 4.40$), all supplied by Riedel-de Haën (France). Atrazin was selected because it was the most frequent pesticide until its prohibition in 2001 in many European countries [32], and its degradation by microorganisms is very slow [33]. ADE, the most common atrazin metabolite, is phytotoxic-like atrazin [34]. Finally, TFSM has been recommended by the French regulation in force [35] to replace atrazin. Concerning the surface water (Erdre river, France), its characteristics are given in Table 2.

In the range of low concentrations (<1 $mg\,L^{-1}$), all solutes were analysed using a Waters 600 HPLC provided with a UV Waters 486 detector and a Novapack C18 non-polar column coupling with solid phase extraction when the concentration is below 5 $\mu g\,L^{-1}$. A Shimadzu UV 1601 spectrophotometer was used in the range of high concentrations (>1 $mg\,L^{-1}$).

Table 2
Characteristics of the surface water

	Surface water
T (°C)	20–22
pH	7.8–8.2
Turbidity (NTU)	0.65–0.80
Conductivity (μS)	600–700
$Abs_{254\,nm}$	0.11–0.14
Oxidation $KMnO_4$ ($mg\,O_2\,L^{-1}$)	0.7–5.0
Dissolved organic carbon ($mg\,L^{-1}$)	2.0–5.0

3.3. Adsorption data collection

Isotherm experiments were performed at 20 ± 1 °C in static reactors to get equilibrium parameters for synthetic and natural waters. Two hundred and fifty milliliters of distilled or natural water containing an initial concentration of pesticide ranging from 0.005 to 22 $mg\,L^{-1}$ were stirred with a given weight (20 mg) of active carbon at 300 rpm for 48 h, previous works having demonstrated that this time enables to reach adsorption equilibrium [36]. pH ranged from 5.5 to 5.6 for synthetic waters and from 7.8 to 8.2 for natural waters.

The dynamic adsorption onto activated carbon was performed at 20 ± 1 °C in water-jacketed glass XK 26/20 columns from Amersham Biosciences Company (France), with an inside diameter of 2.5 cm and a length of 10 cm. A Masterflex peristaltic pump enables natural water (Erdre water, Nantes) containing a pesticide to flow through an activated carbon bed with a velocity u_0 ranging from 5 to 20 $m\,h^{-1}$. Each column was loaded randomly with 10 g of activated carbon. At each column output, sampling is performed as a function of time t to determine breakthrough curves. The value of the initial concentration C_0 was 1 or 15 $mg\,L^{-1}$, for a sampling period of 2 h.

4. Results and discussion

4.1. Modelling competitive adsorption of pesticide and NOM onto activated carbon in a static reactor

Monocomponent adsorption of ADE in synthetic water is presented in Fig. 2. The equilibrium of pesticides onto activated carbon in synthetic waters was modelled by the Freundlich equation [37]:

$$q_1 = K_1 c_1^{1/n_1} \quad (8)$$

where K_1 [$(mg\,g^{-1})(L\,mg^{-1})^{1/n}$] and $1/n_1$ are Freundlich constants that depend on temperature and on the given adsorbent–adsorbate couple and q_1 and c_1 are equilibrium adsorption capacity of pesticide on active carbon ($mg\,g^{-1}$) and equilibrium concentration of pesticide in synthetic water ($mg\,L^{-1}$), respectively. Parameters K_1 and $1/n_1$ are given in Table 3 for the three pesticides and the five active carbons.

The comparison of adsorption isotherm curves and Freundlich parameters shows, for low molecular weight pesticides A and ADE, a preferential adsorption of active carbons which combine a high micropore volume with a narrow pore size distribution (NC-60, WWP3). The low molecular diameter of these pesticides allows them to penetrate into the micropore network of active carbons. For TFSM which presents higher molecular weight and volume, active carbons with a larger mesopore content lead to higher adsorption

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

Table 3
Summary of external Freundlich parameters (K_i and $1/n_i$) for pesticide ($i = 1$) and NOM ($i = 2$)

Parameter	Adsorbate(s)	Activated carbon				
		NC-60	Picabiol	Cal 12-40	Row 0.8	WWP3
Freundlich parameters of pesticides in monocomponent solution						
K_1 ($(mg\,g^{-1})(L\,mg^{-1})^{1/n}$)	A	38.4	22.6	28.7	21.4	27.3
	ADE	26.4	23.1	16.5	21.3	5.2
	TFSM	8.7	19.7	32.1	18.5	8.6
$1/n_1$	A	0.454	0.587	0.642	0.563	0.489
	ADE	0.811	0.729	0.911	0.747	0.523
	TFSM	0.511	0.529	0.732	0.669	0.339
R^2	A	0.950	0.970	0.940	0.950	0.990
	ADE	0.960	0.990	0.970	0.930	0.950
	TFSM	0.950	0.970	0.930	0.960	0.960
Initial concentration of NOM						
$C_{0,2}$ ($mg\,L^{-1}$)	A	0.034	0.030	0.064	0.041	0.107
	ADE	0.023	0.012	0.054	0.039	0.070
	TFSM	0.058	0.047	0.038	0.041	0.087
Freundlich parameters of NOM in presence of a pesticide (EBC model)						
K_2 ($(mg\,g^{-1})(L\,mg^{-1})^{1/n}$)	A + NOM	279.3	1.10	2.54	1.08	132.6
	ADE + NOM	1.11	1.04	1.87	2.09	1.35
	TFSM + NOM	117.7	119.9	1.47	106.9	111.2
$1/n_2$	A + NOM	0.0090	0.0044	0.0063	0.0026	0.0026
	ADE + NOM	0.0036	0.0043	0.0028	0.0060	0.0031
	TFSM + NOM	0.0098	0.0072	0.0030	0.0086	0.0070
Statistical parameters of EBC model						
ψ	A + NOM	0.0008	0.0006	0.0008	0.0330	0.0002
	ADE + NOM	0.0018	0.0025	0.0004	0.0009	0.0005
	TFSM + NOM	0.0026	0.0026	0.0009	0.0027	0.0611
E (%)	A + NOM	1.15	1.00	1.15	7.42	0.58
	ADE + NOM	1.73	2.04	0.82	1.22	0.91
	TFSM + NOM	2.08	2.08	1.22	2.12	10.09

Determination coefficient (R^2), difference (ψ) and relative error (E) between observed and calculated points (C_i, q_i) of the isotherm curve.

capacities as demonstrated by higher values of K_1 parameters with Picabiol or Cal 12-40.

In the case of competitive adsorption between pesticides and natural organic matter (NOM), NOM is ranked as one compound, called equivalent background compound (EBC), according to the Crittenden approach [38] modified by Najm [39]. The EBC model can be used to obtain the adsorption isotherm of competing compound when the adsorption of the target compound (pesticide) is known in the presence and absence of NOM. The model enables to determine the EBC adsorption parameters (Freundlich K_2 and $1/n_2$ and initial EBC concentration $C_{0,2}$). Applying the ideal adsorbed solution theory, see [2] to the binary system (pesticide 1/NOM 2) gives:

$$\begin{aligned} C_{1,0} - \frac{m}{V}q_1 - \left(\frac{q_1}{q_1+q_2}\right)\left[\frac{n_1q_1+n_2q_2}{n_1K_1}\right]^{n_1} &= 0 \\ C_{2,0} - \frac{m}{V}q_2 - \left(\frac{q_2}{q_1+q_2}\right)\left[\frac{n_1q_1+n_2q_2}{n_2K_2}\right]^{n_2} &= 0 \end{aligned} \quad (9)$$

where $C_{1,0}$ and $C_{2,0}$ are the initial concentrations ($mg\,L^{-1}$) of pesticide and NOM, respectively, m the mass of activated carbon introduced in the batch reactor (g), V the volume of solution (L), q_1 and q_2 is the adsorption capacity ($mg\,g^{-1}$) of pesticide and NOM, respectively. The single solute isotherm for pesticide is described by the Freundlich equation (with parameters K_1, n_1) applied to the adsorption isotherm of pesticide in synthetic water.

Adsorption isotherms of pesticide in natural water enable to estimate the equilibrium adsorption capacities and concentrations of pesticide (C_1 and q_1). The initial concentrations of pesticides are then obtained from a mass balance onto the target compound:

$$C_{1,0} = C_1 + \frac{m}{V}q_1 \quad (10)$$

This system has four unknown parameters: Freundlich parameters of NOM (K_2, n_2) and initial concentration of NOM ($C_{2,0}$) we want to determine and adsorption capacity of NOM in binary solution (q_2).

The parameters (C_1, q_1 and q_2) are determined for three experimental points and the parameters ($C_{2,0}$, K_2 and n_2) are thus determined by solving the system of six equations, these parameters being calculated so as to minimize the difference ψ between experimental points and computed points of the isotherm curve

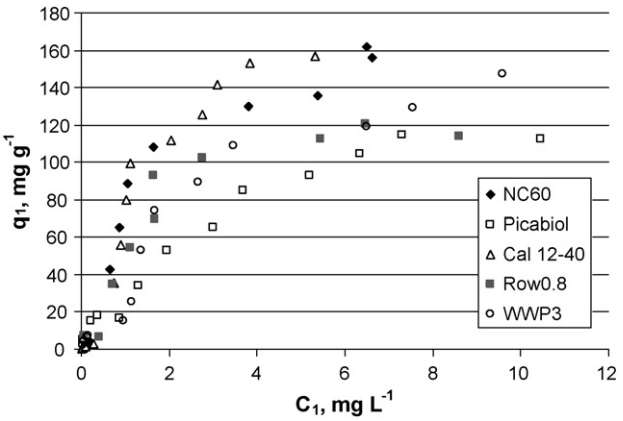

Fig. 2. Monocomponent adsorption isotherms of ADE onto activated carbons. $V = 250$ mL, AC weight = 20 mg, $C_0 = 0.005$–22 $mg\,L^{-1}$, $T = 20 \pm 1$ °C, stirring = 300 rpm for 48 h.

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

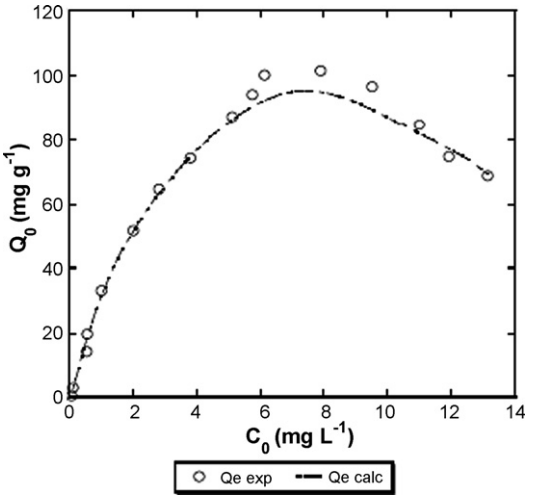

Fig. 3. Modelling of the adsorption isotherm of atrazin onto activated carbon WWP3 in the presence of NOM. V = 250 mL, AC weight = 20 mg, C_0 atrazin = 0.005–22 mg L^{-1}, $T = 20 \pm 1$ °C, stirring = 300 rpm for 48 h.

[40]:

$$\psi = \sum_{i=1}^{3} \left(\frac{q_{1,\mathrm{mod}} - q_{1,\mathrm{exp}}}{q_{1,\mathrm{mod}}} \right)^2 + \sum_{i=1}^{3} \left(\frac{C_{1,\mathrm{mod}} - C_{1,\mathrm{exp}}}{C_{1,\mathrm{mod}}} \right)^2 \quad (11)$$

From ψ values, the relative error between experimental and modelled data may be calculated from the following equation where N is the number of experimental data [40]:

$$E = \sqrt{\frac{\psi}{2N}} \quad (12)$$

This solving method is presented in details by [39]. Solving was performed with Matlab Simulink software. Fig. 3 illustrates modelling of DEA adsorption on Picabiol activated carbon in the presence of NOM by the EBC model. Values of EBC initial concentration $C_{0,2}$, and K_2 and $1/n_2$ Freundlich parameters are given in Table 3. In all cases, the ψ parameter was lower than 0.061 corresponding to a relative error E lower than 10%. Initial EBC concentrations $C_{0,2}$ (mg L^{-1}) are close to values obtained by [41] with the competitive adsorption of NOM and 2-methylisoborneol in different waters. These values are lower than initial DOC measured for surface waters, because the EBC is not considered to be the entire NOM present in natural waters as only an unknown portion of the NOM will compete with pesticides [42]. K_2 values vary in a large range, from 1.04 to 279.3 (mg g^{-1})(L mg^{-1})$^{1/n}$, but are in agreement with values obtained by [40,43] for the adsorption of synthetic organic compounds in the presence of NOM. Low $1/n_2$ values are in agreement with a favorable adsorption of NOM.

4.2. Variable selection and data pre-processing for NN models

A large number of factors may have an influence on the saturation of activated carbon filters, such as the adsorbent pore properties, operating conditions, adsorbate properties or variables related to the presence of natural organic matter. From a literature survey, 15 candidate variables were thus identified as potentially influential for the dynamic adsorption of pesticides from surface water: primary micropore volume $V_{\mathrm{microprim}}$, secondary micropore volume V_{microsec}, median micropore width $\emptyset_{\mathrm{med\,\mu pores}}$, mesopore volume V_{meso}, particle diameter d_{p}, molecular weight M, solubility S, inlet concentration C_0, flow velocity u_0, initial concentration of natural organic matter in terms of total organic carbon TOC_0, Freundlich parameters of pesticides (K_1 and $1/n_1$) and natural organic matter (K_2 and $1/n_2$). Recurrent neural nets have an additional candidate variable, namely, the value of (C/C_0) predicted by the model at the previous time step $(C/C_0)_{(k-1)T}$. The fact that the only previous value of (C/C_0) present in the candidate variables is $(C/C_0)_{(k-1)T}$ reflects the underlying assumption that the recurrent model should be of order 1; that is justified by the fact that knowledge-based models used to describe dynamic adsorption are first order with respect to time [2].

Therefore, the variables of both feed-forward and recurrent neural networks were selected among those 16 parameters by a statistical method, the random probe method [27], in two steps. (1) First, the variables are ranked in order of decreasing relevance by Gram-Schmidt orthogonalization. (2) The second step consists in computing the rank below which variables should be discarded, by generating a "probe" variable (i.e. a random variable that is unrelated to the quantity to be modelled) and ranking it like other candidate variables as described above. The candidate variables that are ranked below the probe variable should be discarded. For the feed-forward neural network, two candidate variables were rejected (with a 10% risk of false positive, i.e. of keeping a variable although it is irrelevant): V_{microsec} and $\emptyset_{\mathrm{med\,\mu pores}}$. For the recurrent model and the same risk level, only eight most influential variables were selected: $(C/C_0)_{(k-1)T} > K_2 > S > 1/n_1 > V_{\mathrm{microsec}} > M > \mathrm{TOC}_0 > C_0$.

The data set featured 9749 values related to the 5 AC and the 3 pesticides with different operating conditions. All data were normalized to have zero mean and standard deviation equal to 1; the experimental breakthrough curves were discretized with the same time step (100 min).

4.3. Model design and network training

The design of the neural network models was performed using the NeuroOne™ v.5 software (Netral S.A., France). For the feed-forward neural network, the 9749 measurements were randomly divided into a training database of 6499 values for training and model selection, and a test database of 3250 values for the final assessment of the generalization performance of the model. In the case of the recurrent networks, the sequential nature of the data is fundamental and must be preserved in the training and test sets; therefore, the breakthrough curves were randomly partitioned into 20 breakthrough curves for training and 10 for testing, resulting in a training database of 6452 values while the test database contained 3297 measurements. In the latter case, data are organized in sequences, each sequence corresponding to one experimental breakthrough curve. The number of epochs was equal to 100.

Because the state-space form did not provide any improvement upon the input–output form, the latter was selected; training was performed under the output-noise assumption, so that a semi-directed algorithm was used: during training, the value of the state input was the value *predicted by the model* at the previous time step.

The central problem in black-box model design is known as the bias-variance dilemma [44]: a model with too few parameters is unable to learn the training data, whereas a model with too many parameters learns both the data *and the noise*, hence generalizes poorly. The goal of black-box model design is to find a model for which the estimate of the root mean square generalization error is on the order of the standard deviation of the noise present in the training data. To that effect, models of increasing complexity (i.e. increasing number of hidden neurons) were trained, and the virtual leave-one-out score E_{p} of each model was computed. The root mean square error on the training set ($\mathrm{RMSE}_{\mathrm{tr}}$) was also computed; those quantities are reported in Table 4. As expected for the recurrent model, the leave-one-out score decreases as the number of hidden neurons increases and starts increasing when the number of parameters is large enough for over-fitting to occur (number of hidden neurons > 2). This is in contrast to the behaviour

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

Table 4
Optimization of the number of hidden neurons for the neural networks

Number of hidden neurons	Recurrent model		Feed-forward model	
	$RMSE_{tr}$	E_p	$RMSE_{tr}$	E_p
1	0.0037	0.086	0.0045	0.095
2	0.0019	0.064	0.0014	0.052
3	0.0001	0.003	0.0008	0.041
4	0.0001	0.002	0.0005	0.032
5	0.0001	0.001	0.0005	0.031
6	0.0001	0.002	0.0004	0.025
7	0.0001	0.003	0.0003	0.024
8	0.0001	0.003	0.0002	0.022
9	0.0001	0.004	0.0002	0.021

E_p: leave-one-out score (see Eq. (5)); $RMSE_{tr}$: root mean squared error on the training set.

of the $RMSE_{tr}$ on the training set, which decreases as the number of hidden neurons increases. For the feed-forward model (see also Table 4), the generalization error E_p does not increase significantly with the number of hidden neurons, in the investigated range. In order to minimize the number of parameters, six hidden neurons were selected for the static model. The final optimized architectures of both NN are shown in Fig. 4, corresponding to 31 and 91 parameters for the recurrent and static models, respectively.

4.4. Generalization performance of the black-box models

Fig. 5 shows the prediction of the breakthrough curve of TFSM on Picabiol by the feed-forward neural network and the recurrent neural network. The agreement between experimental and predicted data for this particular experiment is typical of the results obtained

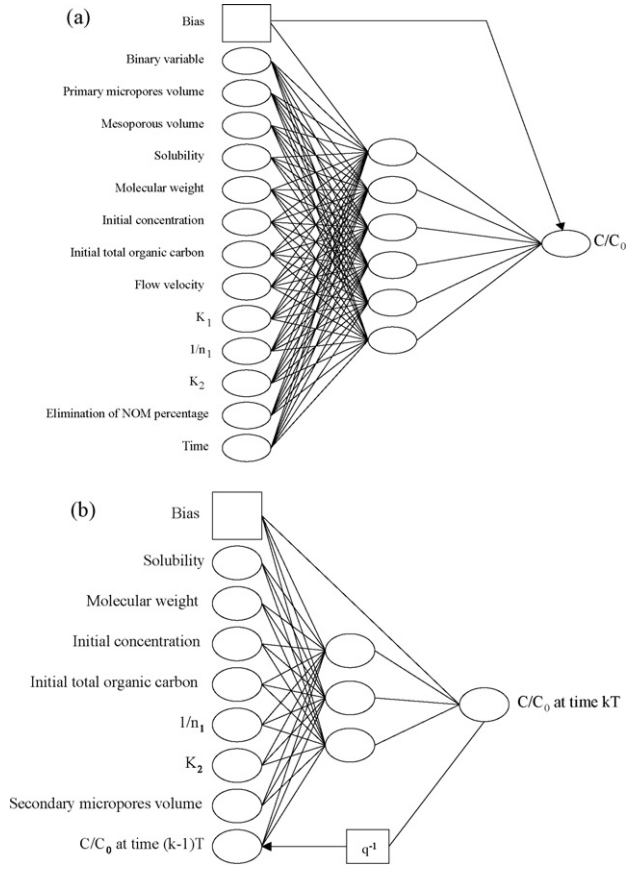

Fig. 4. Optimized architectures of the (a) feed-forward and (b) recurrent neural network. The recurrent network is a first-order model in input–output form.

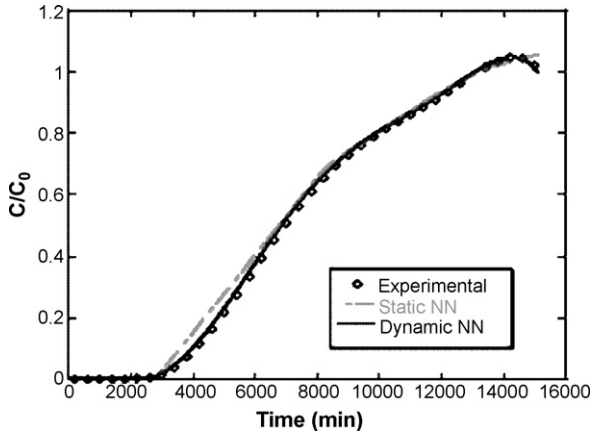

Fig. 5. Prediction of a breakthrough curve by the feed-forward and the feedback NN, for the breakthrough of TFSM in a surface water onto Picabiol ($C_0 = 1\ mg\ L^{-1}$, $u_0 = 20\ m\ h^{-1}$).

for all breakthrough curves. $RMSE_T$ and R_T^2 for the feed-forward neural network were computed using the same observations as those randomly selected for inclusion into the test database of the recurrent neural network. For all breakthrough curves of the test database, observed values as a function of predicted ones are shown in Fig. 6 for static and recurrent neural networks. Computations were performed for specific zones of the breakthrough curves (Table 5): breakthrough zone ($C/C_0 < 0.10$), saturation zone ($C/C_0 > 0.9$), and the global breakthrough curve ($0 < C/C_0 < 1$).

Neural networks provide quite satisfactory predictions for pesticide dynamic adsorption onto activated carbon from a global point of view: R^2 is larger than 0.981 while RMSE is lower than 0.035. However, the recurrent NN generalizes better than the static one, especially for the breakthrough and the saturation zones, despite a lower number of parameters (31 for the recurrent NN and 91 for the feed-forward NN). That is not unexpected, since the dynamic character of the process is taken into account in the recurrent model only.

In all cases, operating conditions and pesticide properties seem to have a great influence on the pesticide adsorption in natural water because these parameters are selected as relevant variables of both neural networks, while adsorbent characteristics have a lower impact, except the porous distributions of active carbons, especially secondary micropore proportions, which is found to be relevant for the predictions made by both networks. It confirms the influence of secondary micropores on the direct competition phenomenon between pesticides and natural organic matter. However, the meso-

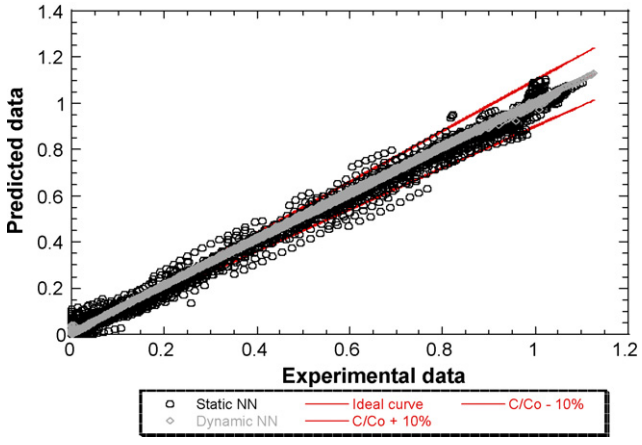

Fig. 6. Parity diagram for all the breakthrough curves of the test database.

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

Table 5
Performances of both models (recurrent and feed-forward neural networks) to predict specific zones of breakthrough curves

Model	$0 < C/C_0 < 0.1$		$0.1 < C/C_0 < 0.9$		$0.9 < C/C_0 < 1$		$0 < C/C_0 < 1$	
	$RMSE_T$	R^2	$RMSE_T$	R^2	$RMSE_T$	R^2	$RMSE_T$	R^2
Recurrent	0.007	0.995	0.013	0.999	0.007	0.981	0.011	0.999
Feed-forward	0.029	0.538	0.034	0.982	0.034	0.695	0.032	0.994

porous volume which, contributes to reduce the pore blockage effect, is only identified as a relevant variable in the feed-forward NN.

5. Conclusion

In the present paper, the feasibility of the prediction of the breakthrough of activated carbon filters was investigated with black-box models.

In a first part, adsorption isotherms of pesticides were performed in a static reactor for synthetic and natural waters. Modelling of experimental curves by the equivalent background compound model enabled to assess design parameters related to the competitive adsorption of pesticides with natural organic matter.

In a second part, these quantities were variables of static and recurrent neural networks designed to model breakthrough curves of the same pesticides in a natural water. Variables related to adsorbents properties and operating conditions were also taken into account as influential factors. Comparison of modelling abilities of both models for specific parts of the section curves shows that both NN provide a satisfactory modelling of the global breakthrough curves. For the breakthrough zone specifically ($0 < C/C_0 < 0.1$), the dynamic character of the process leads to higher performance of the recurrent model despite a lower number of parameters.

Recurrent neural networks are now to be applied to the prediction of breakthrough of activated carbon filters used to treat natural waters containing pesticides in a case of pilot unit located in an industrial site. Variables related to the activated carbon filter (height or volume for example) and the contact time may thus be included in the candidate variables, in order to account for the competition effects that may happen with pesticides.

Acknowledgments

The authors acknowledge Pica Company (France), Actitex Company (France), Chemviron Company (USA) and Norit Company (USA) for providing the activated carbons. They are very grateful to Laurence Le Coq, professor at the Ecole des Mines de Nantes, for help in Matlab Simulink use.

References

[1] J.W. Birkett, J.N. Lester, Endocrine Disrupters in Wastewater and Sludge Treatment Processes, IWA Publishing, Lewis Publishers, London, 2003.
[2] C. Tien, Adsorption Calculations and Modelling: Series in Chemical Engineering, Butterworth–Heinemann, Newton, MA, 1994.
[3] E. Glueckauf, J.I. Coates, Theory of chromatography. Part IV. The influence of incomplete equilibrium on the front boundary of chromatograms and on the effectiveness of separation, J. Chem. Soc. (1947) 1315.
[4] V. Brauch, E.U. Schlünder, The scale-up of activated carbon columns for water purification based on results from batch tests. II. Theoretical and experimental determination of breakthrough curves in activated carbon columns, Chem. Eng. Sci. 30 (5/6) (1975) 539–548.
[5] J.C. Crittenden, W.J. Weber Jr., Predictive model for design of fixed-beds adsorbers; parameters estimation and model development, J. Environ. Eng. Div. (Am. Soc. Civil. Eng.) 104 (1978) 185.
[6] G.H. Xiu, P. Li, Prediction of breakthrough curves for adsorption of lead (II) on activated carbon fibers in a fixed bed, Carbon 37 (2000) 975–981.
[7] G.M. Walker, L.R. Weatherley, Fixed bed adsorption of acid dyes onto activated carbon, Environ. Pollut. 99 (1998) 133–136.
[8] G.U. de Souza, M.A. Selene, L.C. Peruzzo, A.A. Ulson de Souza, Numerical study of the adsorption of dyes in textile effluents, Appl. Math. Modell. (2007), doi:10.1016/j.apm.2007.06.007.
[9] A. Wolborska, P. Pustelnik, A simplified method for determination of the breakthrough time of an adsorbent layer, Water Res. 30 (11) (1996) 2643–2650.
[10] Q. Li, V.L. Snoeyink, B.J. Mariaas, C. Campos, Pore blockage effect of NOM on atrazine adsorption kinetics of PAC: the roles of PAC pore size distribution and NOM molecular weight, Water Res. 37 (2003) 4863–4872.
[11] A. Cougnaud, C. Faur-Brasquet, P. Le Cloirec, Quantitative relationship between activated carbon characteristics and adsorption properties: applications to the removal of pesticides from aqueous solution, Environ. Technol. 26 (8) (2005) 857–866.
[12] L.A. Belanche, J.J. Valdes, J. Comas, I.R. Roda, M. Poch, Towards a model of input–output behaviour of wastewater treatment plants using soft computing techniques, Environ. Model. Software 14 (5) (1999) 409–419.
[13] M. Al-Abri, N. Hilal, Artificial neural network simulation of combined humic substance coagulation and membrane filtration, Chem. Eng. J. (2007), doi:10.1016/j.cej.2007.10.005.
[14] O.L.C. Guimaraes, D.N.Va. Filho, A.F. Siquerra, H.J.L. Filho, M.B. Silva, Optimization of the azo dyes decoloration process through neural networks: determination of the H_2O_2 critical point, Chem. Eng. J. (2007), doi:10.1016/j.cej.2007.10.017.
[15] A. Barron, Universal approximation bounds for superposition of a sigmoidal function, IEEE Trans. Inform. Theory 39 (1993) 930–994.
[16] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.
[17] C. Bishop, Neural Networks for Pattern Recognition, Clarendon, Oxford, 1995.
[18] G. Dreyfus, J.-M. Martinez, M. Samuelides, M.B. Gordon, F. Badran, S. Thiria, L. Hérault, Neural Networks, Methodology and Applications, Springer, 2005.
[19] G. Dreyfus, Y. Idan, The canonical form of non-linear discrete-time models, Neural Comput. 10 (1998) 133–164.
[20] M. Carsky, D.D. Do, Neural network modelling of adsorption of binary vapour mixtures, Adsorption 5 (3) (1999) 183–192.
[21] S.K. Jha, G. Madras, Neural network modeling of adsorption of equilibria of mixtures in supercritical fluids, Ind. Eng. Chem. Res. 44 (7) (2005) 7038–7041.
[22] I.A. Basheer, Y.M. Najjar, M.N. Hajmeer, Neuronet modeling of VOC adsorption by GAC, Environ. Technol. 17 (1996) 795–806.
[23] O. Nerrand, D. Urbani, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, Training recurrent neural networks: why and how? An illustration in process modeling, IEEE Trans. Neural Networks 5 (1994) 178–184.
[24] R. Zbikowski, K.J. Hunt, Neural Adaptive Control Technology, World Scientific, 1995.
[25] R.F. Fu, S.F. Kang, S.I. Liaw, M.C. Chen, Application of an artificial neural network to control the coagulant dosing in a water treatment plant, Water Sci. Technol. 42 (3/4) (2000) 403–408.
[26] H. Stoppiglia, G. Dreyfus, R. Dubois, Y. Oussar, Ranking a random feature for variable and feature selection, J. Mach. Learn. Res. (2003) 1399–1414.
[27] G. Monari, G. Dreyfus, Local overfitting control via leverages, Neural Comput. 14 (2002) 1481–1506.
[28] H. Pignon, C. Brasquet, P. Le Cloirec, Treatment of complex aqueous solutions by the coupling of ultrafiltration and adsorption onto activated carbon cloths, Environ. Technol. 21 (2000) 1261–1270.
[29] C. Brasquet, B. Rousseau, H. Estrade-Szwarckopf, P. Le Cloirec, Observation of activated carbon fibres with SEM and AFM correlation with adsorption data in aqueous solution, Carbon 30 (2000) 407–422.
[30] G. Howarth, K.J. Kawazoe, Method for the calculation of effective pore size distribution in molecular sieve carbon, J. Chem. Eng. Jpn. 16 (1983) 470–475.
[31] S. Brunauer, P.H. Emmett, E. Teller, Adsorption of gases in multimolecular layers, J. Am. Chem. Soc. 60 (1938) 309–319.
[32] S. Dousset, C. Mouvet, M. Schiavon, Leaching of atrazine and some of its metabolites in undisturbed field lysimeters of three soil types, Chemosphere 30 (3) (1995) 511–524.
[33] N. Shapir, R.T. Mandelbaum, C.S. Jacobsen, Rapid atrazine mineralization under denitrifying conditions by *Pseudomonas* sp. strain ADP in aquifer sediments, Environ. Sci. Technol. 32 (1998) 3789–3792.
[34] D.A. Winkelman, S.J. Klaine, Degradation and bound residues formation of four atrazine metabolites, deethylatrazine, deisopropylatrazine, dealkylatrazine and hydroxyatrazine in a western Tennessee soil, Environ. Toxicol. Chem. 10 (1991) 347–354.
[35] French decree no. 2001-1220 of 20 December 2001 relating to water intended for human consumption with the exception of natural mineral waters.
[36] H. Metivier-Pignon, C. Faur-Brasquet, P. Le Cloirec, Adsorption of dyes onto activated carbon cloths: approach of adsorption mechanisms and coupling of ACCs with ultrafiltration to treat coloured wastewaters, Sep. Purif. Technol. 31 (2003) 3–11.

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015

[37] H. Freundlich, W. Heller, The adsorption of *cis*- and *trans*-azobenzene, J. Am. Chem. Soc. 61 (1939) 2228–2230.
[38] J.C. Crittenden, P. Luff, D.W. Hand, L. Oravitz, S.W. Loper, M. Arl, Prediction of multicomponent adsorption, equilibria using ideal adsorbed solution theory, Environ. Sci. Technol. 33 (1985) 2929–2933.
[39] I.N. Najm, V.L. Snoeyink, Y. Richard, Effect of initial concentration of a SOC in natural water on its adsorption by activated carbon, J. Am. Water Works Assoc. Res. Technol. (1991) 57–63.
[40] L. Gicquel, Remobal by Adsorption onto Active Carbon of Atrazine Present in Surface Waters: Study and Modelling of Competition with Dissolved Organic and Mineral Matter, Ph.D. Thesis, Université de Rennes, France, 1996.
[41] M.R. Graham, R.S. Summers, M.R. Simpson, B.W. MacLeod, Modelling equilibrium adsorption of 2-methylisoborneol and geosmin in natural waters, Water Res. 34 (8) (2000) 2291–2300.
[42] G. Newcombe, J. Morrison, C. Hepplewhite, D.R.U. Knappe, Simultaneous adsorption of MIB and NOM onto activated carbon. II. Competitive effects, Carbon 40 (2002) 2147–2156.
[43] S. Qi, L. Schideman, B.J. Marinas, V.L. Snoeyink, C. Campos, Simplification of the IAST for activated carbon adsorption of trace organic compounds from natural waters, Water Res. 41 (2007) 440–448.
[44] S. Geman, E. Bienenstock, R. Doursat, Neural networks and the bias/variance dilemma, Neural Comput. 4 (1992) 1–58.

Please cite this article in press as: C. Faur, et al., Modelling the breakthrough of activated carbon filters by pesticides in surface waters with static and recurrent neural networks, Chem. Eng. J. (2008), doi:10.1016/j.cej.2008.02.015