Experimental and neural model analysis of styrene removal from polluted air in a biofilter

Eldon R. Rene, María C. Veiga, Christian Kennes

Abstract

BACKGROUND: Biofilters are efficient systems for treating malodorous emissions. The mechanism involved during pollutant transfer and subsequent biotransformation within a biofilm is a complex process. The use of artificial neural networks to model the performance of biofilters using easily measurable state variables appears to be an effective alternative to conventional phenomenological modelling.

RESULTS: An artificial neural network model was used to predict the extent of styrene removal in a perlite-biofilter inoculated with a mixed microbial culture. After a 43 day biofilter acclimation period, styrene removal experiments were carried out by subjecting the bioreactor to different flow rates (0.15–0.9 m3 h−1) and concentrations (0.5–17.2 g m−3), that correspond to inlet loading rates up to 1390 g m−3 h−1. During the different phases of continuous biofilter operation, greater than 92% styrene removal was achievable for loading rates up to 250 g m−3 h−1. A back propagation neural network algorithm was applied to model and predict the removal efficiency (%) of this process using inlet concentration (g m−3) and unit flow (h−1) as input variables. The data points were divided into training (115 × 3) and testing set (42 × 3). The most reliable condition for the network was selected by a trial and error approach and by estimating the determination coefficient (R2) value (0.98) achieved during prediction of the testing set.

CONCLUSION: The results showed that a simple neural network based model with a topology of 2–4–1 was able to efficiently predict the styrene removal performance in the biofilter. Through sensitivity analysis, the most influential input parameter affecting styrene removal was ascertained to be the flow rate.

INTRODUCTION

For many years, volatile organic compounds (VOCs) in the air have been shown to have detrimental effects on human health and the environment. Among the 188 VOCs classified under Title III of the Clean Air Act Amendments of 1990 (CAAA–90), styrene is one of the most frequently encountered solvents in many chemical industries. It is an important chemical feedstock, which is used commonly as a raw material for the synthesis of plastics, synthetic resins, butadiene–styrene latex, styrene co-polymers and unsaturated polyester resins.[1] Due to improper practices and treatment, a substantial amount of vapours containing styrene are emitted to the atmosphere from process industries. Exposure to even low concentrations of styrene could cause contact-based skin inflammation, irritation of the eyes, nose and respiratory tract, and may induce narcotism.[2] Thanks to increased attention from regulatory authorities, legislation controlling emissions of air pollutants has proliferated and treatment systems for successful operation and control of VOCs has emerged in recent years.[3]

Biodegradation is a promising method for complete mineralization of VOCs into innocuous end products. In recent years, bioreactor configurations such as biofilters, bioscrubbers, biotrickling filters, rotating drum bioreactors, stirred tank bioreactors and membrane bioreactors have been studied and applied successfully to decontaminate VOC emissions.[4] The simplicity in the operation of biofilters has often resulted in its emergence as the most practical treatment option.[4, 5] Biofilters have proven to be effective in treating large volumes of VOCs at relatively low pollutant concentrations.[4–8] Several studies have shown that biofiltration is a suitable technology to treat air polluted with styrene and other related benzene compounds,[1, 6, 8–10] but hardly any work has been published on artificial neural network (ANN) modelling of such air-pollution treatment systems.

The extent of biodegradation of pollutants in the filter bed depends primarily on waste gas concentration and composition, moisture content, microbial ecology and distribution, nutrient and oxygen availability, temperature and pH. A considerable amount of modelling work has been done with biofilters explaining the micro- and macro-kinetics of the degradation process.[4, 5] These knowledge-driven models explain the underlying phenomenon of the biofilter system with a prior knowledge of sensitive physico-chemical and biological parameters such as microbial growth rate, biofilm thickness, overall mass transfer coefficient, partition, distribution and diffusion coefficients, Henry's rate constant, porosity, O_2 consumption rate and biomass yield coefficient. However the accurate measurement of these variables in a biofilter involves elaborate experimentation that is not only complex, but also difficult. Hence under such circumstances, based on an assumption of the state variables or by conducting batch studies and with the aid of previously reported literature values, biofiltration researchers have been able to successfully develop appropriate models to describe the performance of biofilters. However, biokinetic constants (for example, μ_{max}, K_s) reported in the literature may vary significantly, depending on the experimental conditions used to determinate those parameters.[4] The most proficient way to overcome such practical

limitation is to quantify or predict the performance of the system under consideration using knowledge gained from experimental trials and techniques capable of indicating the non-linearity of the process.[11] These techniques, also called data-driven modelling strategies (borrowed heavily from Artificial Intelligence (AI) techniques), are based on a limited knowledge of the modelling process and can rely on the data describing the input and output functions. The concept of ANN modelling has widespread applications in the field of health and medicine, marketing, applied sciences and engineering.[12] For instance, an ANN-based model was developed to simulate different types of biomass for a gasification process and demonstrated that the model-predicted profiles matched closely the experimental values.[13] A model based on wavelet packet decomposition, entropy and neural networks was used to predict the long-term performance of a wastewater treatment plant.[14] A three-layered neural network with the standard back propagation algorithm was used in their study and they found that the model was able to predict plant performance better. ANNs have been used to model a wide variety of environmental processes such as adsorption, odour monitoring, optimizing media constituents for lipase production, prediction of pollutant concentrations in atmospheric air and, only recently, biofilters.[3, 15–18] Cluster analysis coupled with genetic algorithm was recently used to analyze the initial data subsets, and later, a multi-layered perception (MLP) with the following topology, 2–2–1 and hyperbolic transfer function (tanh) was able to predict the removal efficiency of an inorganic pollutant, i.e. H2S, in a biofilter packed with natural packing material; pig manure and sawdust.[3]

The aforementioned literature and the strategies adopted suggest that neural network based models could efficiently make an abstraction and generalization of the process behaviour and play a complementary role to knowledge-driven models. The present report is the first one applying an ANN to a biofilter packed with a synthetic carrier material and degrading a volatile organic compound, i.e. styrene, from polluted air. The input and output parameters for building the neural network model were chosen based on suggestions given in the literature,[3] where two measured variables, i.e. inlet concentration (g m−3) and unit flow ($\frac{Q}{V}$, h−1), were used as the inputs and removal efficiency (%) as the output parameter for the model.

LIST OF SYMBOLS

Q	Gas flow rate (m3 h−1)
V	Volume of the filter bed (m3)
Cgi and Cgo	Inlet and outlet styrene concentrations, respectively (g m−3)
NI	Number of neurons in the input layer
NH	Number of neurons in the hidden layer
NO	Number of neurons in the output layer
Od	Desired output vector
Op	Predicted output vector from neural model
$\hat{X}$	Normalized value for the variable
Xmin	Minimum value of the variable
Xmax	Maximum value of the variable
NTr	Number of training vectors
NTe	Number of testing vectors
η	Learning rate
E	Global error function
ε	Number of training parameters presented to the network
α	Momentum term
Wij	Connection weights
Tc	Training count
R2	Determination coefficient
Ymodeli	Predictions made by the model
Yobservedi	Observed true values from experiments
N	Number of cases analyzed
Ȳ	Average value
SY	Standard deviations
Ski(p)	Sensitivity of a trained output
p	Number of training patterns

NEURAL MODELLING STRATEGY

The ANN consists of a set of highly interconnected, yet simple processing units commonly referred to as neurons or nodes that are responsible for executing the prescribed rudimentary computation. These neurons are arranged in a series of layers, i.e. input, hidden and output layer of the ANN architecture. The neurons in the input layer represent the set of parameters that are likely to influence the predicting variable, namely removal efficiency (RE %) in this study. The number of hidden layers can be varied depending on the complexity of the relations between the input and output variable. It has been shown that a network architecture with one hidden layer can approximate any continuous function.[3, 19] The optimum number of neurons in the hidden layer is normally estimated by trial and error. The output layer corresponds to the output from the network and the number of neurons in this layer is equal to the number of variables to be predicted. The architecture of a three-layered ANN as employed in this study is shown in Fig. 1. The input layer consists of two neurons (NI = 2) that are connected to the neurons in the hidden layer (NH) and subsequently from the hidden layer to the output layer (NO = 1). Each of these input data (χ_i) are transformed by multiplying with a connection weight (W_{ij}) and a threshold value (θ_j) to generate data to the hidden layer. Each neuron in the hidden layer receives the weighted outputs from the preceding layer, and then feeds it through a transfer function to generate the neuron input, I_j and then the desired output. This is given by
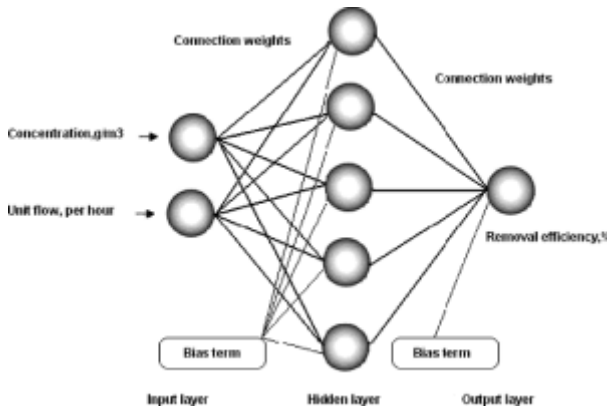
$$I_j = \sum w_{ij} \chi_i + \theta_j \quad (1)$$

Figure 1.

Schematic of a multilayer perceptron used for predicting removal efficiency of biofilter.

INFLUENCE OF INTERNAL NETWORK PARAMETERS

The process of developing the best network architecture involves a number of network specifications to be optimized; the number of neurons in the hidden layer (NH), the learning rate (η), epoch size, momentum (μ), the processing element activation function, training count (Tc) and the training algorithm.

The minimum number of neurons (nodes) in the hidden layer can be equal to or greater than the number of inputs to the network. However, the optimum number of neurons is generally estimated by trial and error. In certain cases, general guidelines such as those given by Hecht–Nielsen[20] below can be followed to choose the upper limit for the number of neurons in the hidden layer.

$$N_H \leq 2N_I + 1 \quad (2)$$

A lot of effort has been made in recent years to develop procedures that automatically adjust the number of neurons during network training, so as to determine the smallest network that is capable of capturing the relationships between different parameters.[21] The rate at which the ANN learns the given pattern depends on several controllable parameters, among which the learning rate is very crucial. The amount a particular weight connection changes is proportional to the learning rate, η, which in turn affects the size of steps taken in weight space.[22] If η is too small, the algorithm will take a long time to converge, and conversely, if η is too large, the network would go through large oscillations during training and never converge. Most learning functions have some provision to change the learning rate and this value remains fixed during the training process.[23] This term is usually positive and varies between 0 and 1. If this value is more than 1, it is easier for the learning algorithm to overshoot in correcting the weights and one would notice more oscillations in the error surface.[24] Hence sufficient care should be taken when choosing this parameter.

The momentum term μ accelerates the convergence of the error during the learning process and is the most popular extension of the back propagation algorithm. This term simply adds a fraction of the previous weight update to the current one and is often related to the learning rate. High η and μ values can accelerate convergence towards a local minimum with huge steps, whereas small η values with high μ can lead to divergent behaviour during training.[22, 24] This value ranges from 0–1 and is normally determined by trial and error for a given network.[25]

The activation function, denoted by f(x), defines the output of a neuron in terms of the induced local field x. Three basic types of activation functions are normally envisioned during the development of networks for prediction purposes; linear, hyperbolic tangent and sigmoid.[26] The most commonly used activation function within the neurons for complex biological systems is the logistic sigmoid function, which takes the form

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

The standard back propagation neural network (BPNN) developed by Rummelhart et al.[27] is the most popular method of optimizing the connection weights during network training. This process involves adjusting the weights of the connections in the network so as to minimize the difference between the output vector predicted by the trained network and the desired (experimental) output vector. Hence when a training vector

sample is presented to the network, the global error function (E) is calculated according to

$$E = \frac{1}{2} \Sigma (O_d - O_p)^2 \quad (4)$$

In the BPNN, the error of the output layer in the network propagates backward to the input layer through the hidden layer to obtain the final desired output. During the learning process, the connection weights are adjusted using a form of the generalized delta learning rule in order to minimize the error function, E and obtain the desired output from a given set of inputs. The connection weights (W_{ij}) are updated using the gradient descent algorithm,[24] which can be represented in the form

$$\Delta W(t) = \sum_{s=1}^{\varepsilon} -\eta \frac{\partial E}{\partial W} + \mu \Delta W(t-1) \quad (5)$$

The value of ε was kept constant at 100 during the training process. The closeness between experimental and model-predicted outputs was evaluated by computing the coefficient of determination values[3] as below:

$$R^2 = \left[\frac{\sum_{i=1}^{N} (Y_{model_i} - \overline{Y_{model}})(Y_{observed_i} - \overline{Y_{observed}})}{(N-1) \, S_{Y_{model}} \, S_{Y_{observed}}} \right]^2 \quad (6)$$

The experimental data collected from the biofilter were also normalized and scaled to the range 0 to 1, so as to suit the transfer function in the hidden (sigmoid) and output layer (linear):

$$\hat{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (7)$$

MATERIALS AND METHODS

Microbial seed

A mixed microbial culture obtained from petrochemical refinery sludge was used to inoculate the biofilter. This was done by filling the biofilter with the sludge and draining it after 12 h. The procedure was repeated several times until visible biomass was noticed on the surface of perlite.

Biofilter construction and experiments

The biofilter was made of glass having a diameter (ID) of 10 cm and 70 cm in height. The biofilter was packed with sieved perlite particles (4–6 mm) to give a total working volume of 5 L. A perforated plate at the bottom provided support for the packing while another plate at the top acted as a distributor for gas flow and mineral salt media addition. Gas sampling ports sealed with airtight rubber septa were provided at equal intervals along the biofilter height. The bed moisture was maintained constant by

periodic, bi-weekly, addition of fresh mineral salt medium from the top.6 During the operation, the relative humidity of the air stream was maintained at 95 ± 3%. Experiments were carried out by varying the flow rates of the styrene vapors and humidified air independently to get different initial concentrations and residence times in the biofilter. The biofilter was operated in a downflow mode. Gas samples were collected from different ports and analyzed for residual styrene and CO2 concentrations.

Analytical methods

Styrene concentration in gas samples was measured by gas chromatography on an HP 5890 (Agilent Technologies, Spain) gas chromatograph (GC), using a 50 m TRACER column and a flame ionization detector (FID). The flow rates were 30 mL min−1 for H2 and 300 mL min−1 for air. Helium was used as the carrier gas at a flow rate of 2 mL min−1. The temperatures at the GC injection, oven and detection ports were 150 °C in all three cases. CO2 was analyzed with a HP 5890 gas chromatograph equipped with a thermal conductivity detector (TCD), using a Porapack Q column (1/8″ diameter × 2 m). The injection and oven temperatures were 90 and 25 °C, respectively, with the TCD set at 100 °C. Relative air humidity was measured using a Lufft thermo-hygrometer, model C-210 fitted with a flexible sampling probe (G. Lufft Mess- und Regeltechnik, GmbH, Germany).

Biofilter performance parameters

Inlet loading rate, removal efficiency and elimination capacity were used to evaluate biofilter performance, using the equations defined below:

$$\text{Inlet loading rate, ILR (g m}^{-3}\text{ h}^{-1}) = \frac{Q\,C_{gi}}{V} \tag{8}$$

$$\text{Elimination capacity, EC (g m}^{-3}\text{ h}^{-1}) = \frac{Q\,(C_{gi} - C_{go})}{V} \tag{9}$$

$$\text{Removal efficiency, RE (\%)} = \frac{C_{gi} - C_{go}}{C_{gi}} \times 100 \tag{10}$$

Software used

ANN-based predictive modelling was carried out using the shareware version of the neural network and multivariable statistical modelling software, NNMODEL (Version 1.4, Neural Fusion, NY). The same software was also used to generate contour plots and carry out sensitivity analysis.

RESULTS AND DISCUSSIONS

Effect of operating conditions on biofilter performance

The biofilter was initially acclimated to styrene by passing low pollutant concentrations and low gas flow rates (0.15 m3 h−1) for several days, which allowed enough biomass to gradually colonize the packing material. Scanning electron microscopy was used to

visualize the perlite particles colonized with microbial populations under steady-state conditions. These images (Fig. 2) clearly show the dominant presence of fungi as potential styrene oxidizers. The degree of acclimatization depends primarily on the adaptive capability of the microorganisms inoculated on the perlite, substrate concentrations, nutrient concentration and its availability and other necessary environmental conditions.
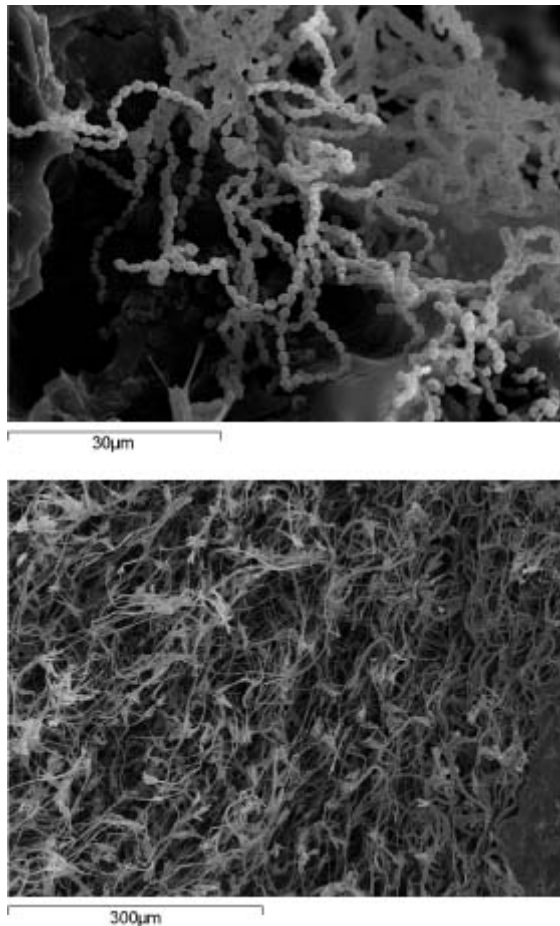

Figure 2.

SEM micrographs of biomass grown on the surface of the perlite packing material, showing the highly dominant presence of fungi.

The combined effect of styrene inlet concentration and gas flow rate was then investigated in different phases of continuous biofilter operation, that corresponded to residence times varying between 20 s and 120 s. Thus, after day 43, once enough biomass was attached to the carrier material, different empty bed residence times (EBRT) were used, while modifying the feed concentration in each case, in order to evaluate the effect of those parameters on the elimination capacity and maximum performance of the reactor. The EC, which reflects the capacity of the biofilter to remove pollutants, is plotted for the different EBRT, i.e. different gas flow rates, as a function of the inlet styrene concentration in Fig. 3. Though there were fluctuations in the EC values during start–up, under steady state conditions, a linear relation between

the two variables was observed with a maximum EC of 260 g m−3 h−1 (RE = 100%). The results from this study are relatively high compared with other studies reported in the literature using biofilters for styrene biodegradation, which could be due to the dominant presence of fungi, as shown in Fig. 2. Indeed, it has been reported that fungal-dominant biofilters allow a better performance to be reached than is usual for hydrophobic VOCs.28
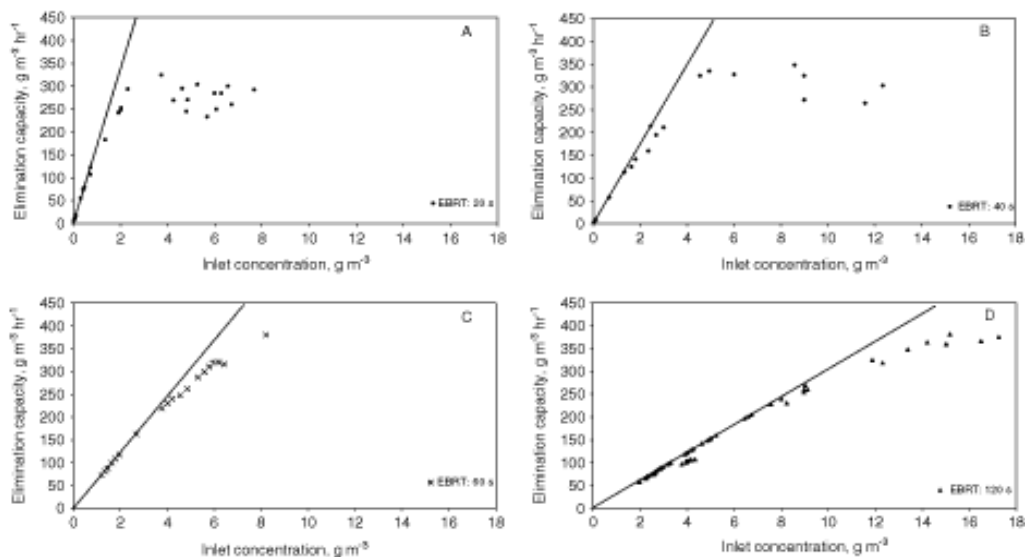

Figure 3.

Elimination capacity profiles versus inlet pollutant concentrations, at different EBRT (i.e. gas flow rates): (A) 20 s, (B) 40 s, (C) 60 s, (D) 120 s. The dotted straight line denotes100% removal efficiency.

In order to understand the styrene elimination mechanism within the reactor, the concentration profile at different bed heights was measured at a constant loading rate. The results indicated that styrene removal was more efficient in the upper section of the filter bed than in the lower section. Nearly 40% of the styrene was removed in the first section followed by 30% in each of the other two sections. This may be due to a higher concentration of microbial populations and higher moisture content in the upper section of the filter bed (data not shown).

Neural modelling

A neural network based predictive model was developed for the biofilter with inlet concentration and unit flow as the model inputs and RE as the sole model output. The entire experimental data were divided into training (NTr—75%) and test data (NTe—25%), leaving behind the quasi-steady-state data during initial biofilter acclimatization (days 1–43). The test data was set aside during network training and was used only for evaluating the predictive potentiality of the trained network. The basic statistics of the variables for the training and test matrix are shown in Table 1.

Table 1. Basic statistics of the data matrix used for ANN model development (a) training and (b) test data

Basic statistics of the training data

(a) Variable	N	Mean	Std Dev	Minimum	Maximum	Sum Sq
Concentration, g m−3	115	4.19	3.83	0.034	17.25	482.64
Unit flow, h−1	115	71.73	57.26	30	180	8250
Removal efficiency, %	115	86.61	20.38	20.99	100	9961.02

Basic statistics of the test data

(b) Variable	N	Mean	Std Dev	Minimum	Maximum	Sum Sq
Concentration, g m−3	42	3.644	2.42	0.52	11.58	153.08
Unit flow, h−1	42	64.28	51.14	30	180	2700
Removal efficiency, %	42	82.45	25.53	25.23	100	3463.16

The number of neurons in the input layer (NI = 2) and output layer (N0 = 1) were chosen based on the number of input and output variables to the network. A detailed study on the effect of internal network parameters on the performance of back propagation networks[19] and the procedure involved in selecting the best network topology has been described elsewhere.[22] However, in most instances, literature suggests the use of a trial and error approach where the performance goal is set by the user.[29] The best network architecture was chosen (Table 2) based on the maximum predictability of the network for the test data by analyzing values of the coefficient of determination.[3]

Table 2. Best values of network parameters used for training the network

Training parameters	Value
Training count	22 000
Number of neurons in input layer	2
Number of neurons in hidden layer	4
Number of neurons in output layer	1
Learning rate	0.8
Momentum term	0.8
Error tolerance	0.0001
Training algorithm	Gradient descent algorithm
Number of training data set	115
Number of test data set	42

The model was trained using different combinations of these network parameters so as to achieve maximum coefficient of determination (R^2) values (target value = 1, i.e. 100% correlation between measured and predicted variables). This was achieved by a vigorous trial and error approach by keeping some training parameters constant and by slowly moving the other parameters over a wide range of values. The R^2 value was 0.973 during training (sum square error, SSE—0.000786), while for the test data it was slightly better at 0.98 (SSE—0.000691). Hence, only about 2% of the total deviations could not be mapped by the model for RE profiles. From Figs 4 and 5, it can be seen that data points before 20 show some differences between measured and predicted RE profiles. During this step of continuous operation in the biofilter (days 53–63), the inlet concentration to the biofilter was suddenly increased from 0.5 to 4 g m−3 after the acclimatization step, where a sudden deterioration in the performance was ascertained. However, in the later steps, the biofilter was able to re–acclimatize to this sudden shock load within a few days and restored its performance completely. This sudden and unexpected decrease in removal profiles would have caused an impact in the neural networks learning/generalization pattern while predicting the performance parameter. However, at the highest flow rate tested in this study (days 163–190, 0.9 m3 h−1) and when the concentration was increased from <0.7 to 4.8 g m−3, the loading rates were sufficiently high (877 g m−3 h−1) that just 30% of the incoming styrene was removed. For this case, the prediction power of the neural network was good without any discrepancies, showing its ability to learn how to do tasks based on the data given for training or initial experience gained from successive training counts. Overall, the ANN model developed was able to demonstrate high accuracy in predicting the various

operating conditions of the biofilter remarkably well. This implies that the degree of complexity of the data for the given network is not high, so that a simple MLP is sufficient to map a relation between the given inputs and the output. Figure 6 depicts several contours of RE (%) generated from the ANN model predictions due to changes in unit flow and inlet concentration to the styrene degrading biofilter. It can be seen that, to achieve high removals, it would be necessary to operate the biofilter under the following conditions: low concentrations (<2 g m−3), irrespective of the range of unit flows (30–180 h−1) and high concentrations (>5.8 but <13 g m−3) and unit flow <40 h−1 (Q—0.2 m3 h−1). Practically, this would be feasible because, as envisaged earlier, the biofilter was able to perform well with 100% RE under different flow conditions (for EBRTs varying between 20 s and 120 s) and at the highest EBRT tested in this study (120 s) the critical load of the biofilter was 260 g m−3 h−1. The connection weights and bias term obtained for the interconnections between different neurons in different layers of the MLP are shown in Table 3.
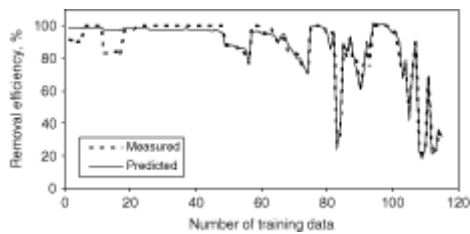

Figure 4.

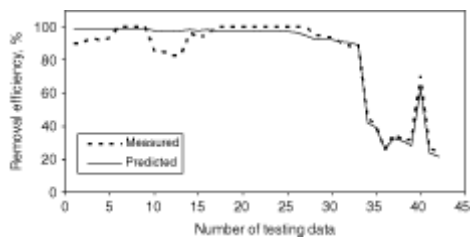Prediction of removal efficiency after network training.

Figure 5.

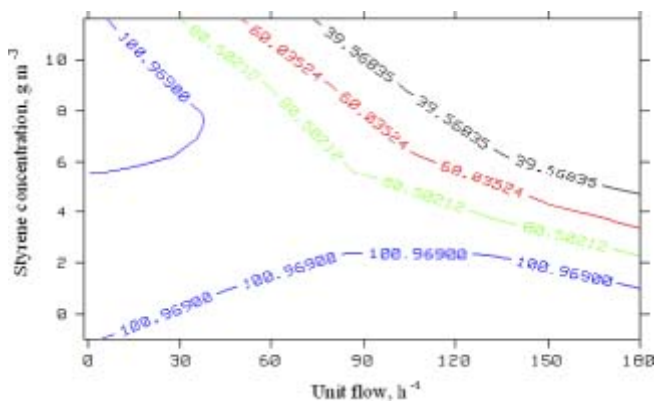Prediction of removal efficiency after network testing.

Figure 6. Contour plot showing the effect of inlet concentration and unit flow on the predicted removal efficiency after ANN training.

Table 3. Hidden later connection weights and bias term after network training

Input to hidden layer weights				
Variable	HID001	HID002	HID003	HID004
V1	9.535	− 5.379	− 13.806	9.912
V2	14.910	− 5.997	8.523	− 15.415
Bias	− 11.661	− 0.484	− 3.600	− 2.731

Hidden to output layer weights	
	V3
HID001	− 2.265
HID002	− 1.030
HID003	4.433	V1-Inlet concentration
HID004	0.946	V2-Unit flow
Bias	1.337	V3-Removal efficiency

By estimating the absolute average sensitivity (AAS), a sensitivity analysis was carried out using the software NNMODEL to identify the most influential parameter affecting the reactor performance. The sensitivity is calculated by summing the changes in the output variables caused by moving the input variables by a small amount over the entire training set.30 The absolute value average sensitivity matrix, Ski, abs can be defined as follows:

$$S_{ki,abs} = \frac{\sum_{p=1}^{p} \left| s_{ki}^{(p)} \right|}{p} \quad (11)$$

These AAS values were 0.5250 and 0.4249 for unit flow and inlet concentration, respectively. The higher values for unit flow than concentration suggest that unit flow change has a significant and greater influence on the RE of the biofilter than concentration. This result also matches the sensitivity analysis results carried out by Elias et al.3 and corroborates the observed decline in RE at higher flow rates in this study. For waste gas treatment systems, it has been reported that at higher gas flow rates (high turbulence), shear force effects causes a decrease in the biofilm thickness. Hence resistance decreased and mass transfer was enhanced.31 The critical load for 100% removal in the biofilter at an EBRT of 20 s (flow rate 0.9 m3 h−1) was just 200 g m−3

h−1, where for high styrene loading rates at this EBRT, the RE dropped significantly to values as low as 25%.

In general, the predictive ability of the proposed model using the concepts of artificial intelligence and the back propagation algorithm was high at the 95% confidence interval and meaningfully significant, as ascertained from the R2 value between the measured and predicted outputs in the training and test data. This work could enable the intensification of research into neural networks for evaluating pilot scale biofilters, besides helping to optimize their state variables.

CONCLUSIONS

The following conclusions are based on the results of the data collected from the perlite biofilter and its performance modelling using neural networks.

- 1.Continuous experiments conducted in a perlite biofilter showed promising results for the treatment of high concentrations of off gas emissions containing styrene. The inlet loads were varied between 60 and 1380 g m−3 h−1, reaching a high maximum EC of 382 g m−3 h−1, which could be due to the dominant presence of fungi.

- 2.Under the process conditions tested, for loading rates less than 260 g m−3 h−1 RE as high as 100% were achieved.

- 3.The results from ANN-based data-driven modelling shows that the biofilter performance, in terms of RE, can be predicted with high confidence level (95%) using easily measurable operational parameters, namely, inlet concentration and unit flow. The model was adequately trained with the lab scale biofilter data and tested with a separate data set.

- 4.A three-layered MLP was sufficient to describe the process behavior using connection weights and bias terms (thresholds). The suitable network architecture of the model was determined through a vigorous trial and error approach. The computed determination coefficient (R2) values for the test data set show high correlation between the predicted and measured performance values. The best topology was found to be a simple three-layered network, 2–4–1, achieved at the following settings of internal network parameters: learning rate (0.8), momentum (0.8) and a training count (epoch size) of 22 000.

- 5.The robustness of the ANN was further explored with a sensitivity analysis of the input parameters, which showed that unit flow was a more critical factor affecting biofilter performance, than the inlet concentration. The results from this study could be used to assess, monitor and design control strategies for continuously operated biofilters treating high concentrations of VOCs.

Acknowledgements

REFERENCES

☐ Jorio H, Bibeau L and Heitz M, Biofiltration of air contaminated by styrene: effect of nitrogen supply, gas flow rate and inlet concentration. Environ Sci Technol 34: 1764–1771 (2000).

☐ 2Fielder RJ and Lowing R, Toxicity Review 1: Styrene. Health and Safety Executive, London (1981).

☐ 3Elias A, Ibarra–Berastegi G, Arias R and Barona A, Neural networks as a tool for control and management of a biological reactor for treating hydrogen sulphide. Bioprocess Biosyst Eng 29: 129–136 (2006).

☐ 4Kennes C and Veiga MC, Bioreactors for Waste Gas Treatment. Kluwer Academic Publishers, Dordrecht, The Netherlands (2001).

☐ 5Devinny JS, Deshusses MA and Webster TS, Biofiltration for Air Pollution Control. Lewis Publisher, Boca Raton (1999).

☐ 6Kennes C, Cox HHJ, Doddema HJ and Harder W, Design and performance of biofilters for the removal of alkylbenzene vapours. J Chem Technol Biotechnol 66: 300–304 (1996).

☐ 7Kennes C and Thalasso F, Waste gas biotreatment technology. J Chem Technol Biotechnol 72: 303–319 (1998).

☐ 8Mohammad BT, Veiga MC and Kennes C, Mesophilic and thermophilic biotreatment of BTEX polluted air in reactors. Biotechnol Bioeng 97: 1423–1438 (2007).

☐ 9Cox HHJ, Moerman RE, van Baalen S, van Heiningen WNM, Doddema HJ and Harder W, Performance of styrene degrading biofilter containing the yeast Exophiala jeanselmei. Biotechnol Bioeng 53: 259–266 (1997).

☐ 10Veiga MC and Kennes C, Parameters affecting performance and modelling of biofilters treating alkylbenzene polluted air. Appl Microbiol Biotechnol 55: 254–258 (2001).

☐ 11Glassey J, Montague GA, Ward AC and Kara BV, Artificial neural network based experimental design procedures for enhancing fermentation development. Biotechnol Bioeng 44: 397–405 (1994)

☐ 12Paliwal M and Kumar UA, Neural networks and statistical techniques: a review of applications. Expert Syst Appl 36: 2–17 (2009).

☐ 13Guo B, Li D, Cheng C, Lu Z and Shen Y, Simulation of biomass gasification with a hybrid neural network model. Bioresource Technol 76: 77–83 (2001).

☐ 14Hanbay D, Turkoglu I and Demir Y, Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. Expert Syst Appl 34: 1038–1043 (2008).

☐ 15Faur–Brasquet C and Le Cloirec P, Neural network modelling of organics removal by activated carbon cloths. J Environ Eng 127: 889–893 (2001).

☐ 16Bockreis A and Jager J, Odour monitoring by the combination of sensors and neural networks. Environ Model Software 14: 421–426 (1999).

☐ 17Haider MA, Pakshirajan K, Singh A and Chaudhry S, Artificial neural network—genetic algorithm approach to optimize media constituents for enhancing lipase production by a soil microorganism. Appl Biochem Biotechnol 144: 225–235 (2008).

☐ 18Mckendry IG, Evaluation of artificial neural networks for fine particulate pollution (PM10 and PM2.5) forecasting. J Air Waste Manage Assoc 52: 1096–1101 (2002).

☐ 19Hornik K, Stinchcombe M and White H, Multilayer feed forward networks as universal approximators. Neural Netw 2: 359–356 (1989).

☐ 20Hecht–Nielsen R, Kolmogorov's mapping neural network theorem. in Proceedings of the First IEEE International Joint Conference on Neural Networks. New York, IEEE Press, pp 11–14 (1987).

☐ 21Verma B, Novel network architecture and learning algorithm for the classification of mass abnormalities in digitized mammograms. Artif Intel Med 42: 67–79 (2008).

☐ 22Maier HR and Dandy GC, The effects of internal parameters and geometry on the performance of back propagation neural networks: an empirical study. Environ Model Software 13: 193–209 (1998).

☐ 23Warner B and Misra M, Understanding neural networks as statistical tools. Am Statistician 50: 284–293 (1996).

☐ 24Maier HR and Dandy GC, Neural network based modelling of environmental variables: a systematic approach. Math Comput Model 33: 669–682 (2001).

☐ 25Dai HC and Macbeth C, Effects of learning parameters on learning procedure and performance of a BPNN. Neural Netw 10: 1505–1521 (1997).

☐ 26Haykin S, Neural Networks: A Comprehensive Foundation. Macmillan College Co, New York (1994).

☐ 27Rummelhart DE, Hinton GE and Williams RJ, Learning representations by back–propagation errors. Nature 323: 533–536 (1986).

☐ 28Kennes C and Veiga MC, Fungal biocatalysts in the biofiltration of VOC polluted air. J Biotechnol 113: 305–319 (2004).

☐ 29Silva RG, Pinotti LM, Cruz AJG, Giordano RLC and Giordano RC, Artificial neural networks to infer biomass and product concentration during the production of penicillin G acylase from Bacillus megaterium. J Chem Technol Biotechnol 83: 739–749 (2008).

☐ 30Zurada JM, Malinowski A and Cloete I, Sensitivity analysis for minimization of input data dimensions for feed forward neural network. In IEEE International Symposium on Circuits and Systems, London, pp 447–450 (1994).

☐ 31Jin Y, Veiga MC and Kennes C, Development of a novel monolith bioreactor for the treatment of VOC polluted air. Environ Technol 27: 1271–1277 (2006).