Membrane Water Treatment, *Vol. 9, No. 2 (2018) 115-121*
DOI: https://doi.org/10.12989/mwt.2018.9.2.115

Modelling of starch industry wastewater microfiltration parameters by neural network

Aleksandar I. Jokić*[1], Laslo L. Šereš[2], Nemanja R. Milović[1], Zita I. Šereš[1], Nikola R. Maravić[1], Žana Šaranović[3] and Ljubica P. Dokić[1]

[1]University of Novi Sad, Faculty of Technology, Bul. cara Lazara 1, 21000 Novi Sad, Serbia
[2]University of Novi Sad, Faculty of Economics in Subotica, Segedinski put 9-11, 24000 Subotica, Serbia
[3]Institute of Economics ad, Ulica Kralja Milana 16, 11000 Belgrade, Serbia, AD Fidelinka, Fidelinka – Starch, Čantavirski put 1, 24000 Subotica, Serbia

(Received March 10, 2017, Revised June 7, 2017, Accepted October 21, 2017)

Abstract. Artificial neural network (ANN) simulation is used to predict the dynamic change of permeate flux during wheat starch industry wastewater microfiltration with and without static turbulence promoter. The experimental program spans range of a sedimentation times from 2 to 4 h, for feed flow rates 50 to 150 L/h, at transmembrane pressures covering the range of 1×10^5 to 3×10^5 Pa. ANN predictions of the wastewater microfiltration are compared with experimental results obtained using two different set of microfiltration experiments, with and without static turbulence promoter. The effects of the training algorithm, neural network architectures on the ANN performance are discussed. For the most of the cases considered, the ANN proved to be an adequate interpolation tool, where an excellent prediction was obtained using automated Bayesian regularization as training algorithm. The optimal ANN architecture was determined as 4-10-1 with hyperbolic tangent sigmoid transfer function transfer function for hidden and output layers. The error distributions of data revealed that experimental results are in very good agreement with computed ones with only 2% data points had absolute relative error greater than 20% for the microfiltration without static turbulence promoter whereas for the microfiltration with static turbulence promoter it was 1%. The contribution of filtration time variable to flux values provided by ANNs was determined in an important level at the range of 52-66% due to increased membrane fouling by the time. In the case of microfiltration with static turbulence promoter, relative importance of transmembrane pressure and feed flow rate increased for about 30%.

Keywords: wheat starch industry wastewater; microfiltration; artificial neural networks

1. Introduction

Nowadays, wastewater treatment and intensification of the separation process is one of the most objectives in environmental protection. Adequate treatment of wastewater can reduce the environmental impact of specific industrial plant as well the consumption of clean water used in the production process. For starch and gluten production in wheat wet-processing, so called Martin process, water is used in the process of flour hydration, starch and gluten separation, starch refinement, as well as for cleaning of the equipment and the work space. The largest amount of the wastewater is created during the starch separation process. The obtained wastewater settles in sedimentation tanks, after which it is decanted and then it undergoes a purification procedure (Šaranović *et al.* 2011). The potential of cross-flow microfiltration as a separation method for the purification of starch industry wastewater is significant.

Membrane separation method is a physical separation technology which includes microfiltration (MF), ultrafiltration (UF), nanofiltration (NF) and reverse osmosis (RO). It is a powerful approach for separating diverse types of particles and molecules with different molecular weights. Microfiltration is the very popular membrane technology to remove dispersed particles during wastewater treatment processes. It is pressure driven separation process with membrane pore diameter in range of about 0.1 to 10 μm. Microfiltration performance is generally expressed in terms of the filtrate flux i.e. the volume of filtrate that passes through unit membrane area in unit time, as it is the case for all filtration processes. During the microfiltration permeate flux declines under the influence of many phenomena. One of them is the formation of a cake layer on the membrane surface that is usually the cause of the steep decline in filtrate flux; however other phenomena such as the plugging of the membrane pores by particles, infiltration of fines into the filter cake or membrane fouling by macrosolutes can also contribute to flux decline. These phenomena are referred to as fouling, where fouling is any phenomenon, other than pure cake formation, that contributes to flux decline (Tanaka *et al*. 1994). The flux decline is then the result of superposition of numerous mechanisms of membrane fouling. Flux decline is a limiting factor in the industrialization of membrane processes (Nourbakhsh *et al.* 2014). It is number of techniques to reduce cake formation as well as the membrane fouling such as backflushing, gas sparging, turbulence promoters or static mixers and many others. The use of turbulence promoters or inserts in the

*Corresponding author, Professor
E-mail: jokic@uns.ac.rs

http://www.techno-press.org/?journal=mwt&subpage=7
ISSN: 2005-8624 (Print), 2092-7037 (Online)

tubular membrane is one of the technique applying hydrodynamic methods in reducing permeate flux decrease i.e., controlling membrane fouling. Turbulence promoters or inserts have many shapes and sizes. There are static rods, Kenics static turbulence promoters, metal grills, spiral wire, cone shape inserts, disc and doughnut shape inserts (Jokić *et al*. 2010, Popović *et al*. 2013).

Number of different models have been proposed in the literature to predict permeate flux decline based on the theory using physical, chemical, and hydrodynamic modelling parameters such as particle size, zeta potential, temperature, solution pH, ionic strength, transmembrane pressure, shear rate, etc. (Hermia 1985, Bowen and Jenner 1995). The disadvantage theoretical modelling is in their accuracy due to the fact of insufficient knowledge regarding the complexity of the microscale phenomena occurring during the filtration processes. This lack of knowledge is avoided by using assumptions; resulting in models which are only valid under certain conditions. On the other side, empirical models are built upon specific experimental observations and then used to predict the filtration system performance and can be used as alternative to the theoretical models. Artificial neural network (ANN) approach is one such method. ANN develops a model based on raw data and uses them to generate an initial network (training). The network is tested with new data but can continually input data to refine the model. Compared to theoretical models, ANNs are a black box technique.

Researchers have applied ANN to diverse fields such as spray drying process (Aghbashlo *et al.* 2012), soil water retention curve (Bayat *et al*. 2013), red plum juice permeate flux (Nourbakhsh *et al.* 2014), optimizing control system for a seawater-desalination solar-powered membrane distillation unit (Porrazzo *et al*. 2013), etc. Applications of ANN for the prediction of membrane separation processes have been reported in a number of papers. ANNs have attracted a lot of interests for the past decade in certain membrane processes. In most of previous studies, the ANN models have been proved to perform better than the conventional modelling methods (Nourbakhsh *et al.* 2014). Dornier *et al.* (1995) presented a possibility for dynamic modelling cross-flow microfiltration of a raw cane sugar suspension using neural networks. Results of the study showed a good convergence (97%) was obtained with 5 neurones in the first hidden layer and 3 neurones in the second hidden layer.

Al-Abri and Hilal (2008) implemented artificial neural network model to predict combined humic substance coagulation and membrane filtration. The difference between predicted and experimental data was lower than 5%. Flux declines versus time during cross-flow microfiltration of a mixture that contains phosphate and fly ash was modelled and compared by using an artificial neural network (Aydiner *et al.* 2005). The experiments were carried out by varying conditions of fly ash dosages, phosphate concentrations, transmembrane pressures and two membrane types. It was shown that all of the experimental conditions can be modelled as a whole or separately, and the model results obtained for one experiment can be used for others at the same conditions with an acceptable correlation level by ANNs.

Chellam (2005) reported results of highly accurate simulations using artificial neural network in prediction of time-variant specific fluxes for several feed suspensions under a wide range of hydrodynamic parameters. It was found that only one hidden layer was to be sufficient; whilst with increasing number of hidden neurons the error initially decreased due to increasing number of weights and later increased probably due to overfitting. The higher number of hidden neurons necessary for was necessary for some suspensions microfiltration simulations suggests that their underlying mechanisms contributing to flux decline were more complicated than others taking into account feed suspensions particle size distributions.

Artificial neural network model can be developed for turbulence promoter-assisted cross-flow microfiltration of particulate suspensions (Liu *et al.* 2014). The inlet velocity, transmembrane pressure and feed concentration were taken as inputs, and the flux improvement efficiency by turbulence promoter was taken as output. Results of the study suggest that an ANN model with one single hidden layer is sufficient to accurately represent experimental data. The trial-and-error method was implemented to determine the number of neurons in the hidden layer; and it was found that 12 hidden neurons is optimal configuration to avoid overfitting.

Although, ANNs can be considered as black box models, some researchers reported an important contribution that demonstrated that ANNs need not be used simply as black boxes but can be used in combination with the connection weight partitioning methodology to determinate cause-effect information that can be quantitatively extracted from network connection weights (Chellam 2005, Liu *et al.* 2014). This method was initially proposed by Garson (1991).

The main objective of the present study was to develop an ANN based model to simulate flux reduction during microfiltration of starch industry wastewater generated during the process of wet-production of wheat starch and vital wheat gluten, in single channel ceramic membrane with and without static turbulence promoter.

2. Materıal and methods

Experiments were conducted on the samples of decanted wastewater after sedimentation in predefined time periods in range 2 to 4 h. The dry matter content of the wastewater varied between 1.2% and 0.85%; the chemical oxygen demand ranged from 23 000 to 20 000 mgO_2/L, and the suspended matter was between 8000 and 4000 mg, for sedimentation times of 2 and 4 h, respectively.

2.1 Microfiltration experiments

The microfiltration experiments were done according to full 2^3 factorial experimental design with added central point, making in total 10 experiments. Lower levels of experimental design were transmembrane pressure 1×10^5 Pa, sedimentation time 2 h and feed flow rate 50 L/h. Upper level values were $3{\times}10^5$ Pa, 4 h and feed flow rate 150 L/h, for transmembrane pressure, sedimentation and feed flow rate, respectively. Values of mean level were $2{\times}10^5$ Pa, 3 h and feed flow rate 100 L/h. All flux measurements i.e.,

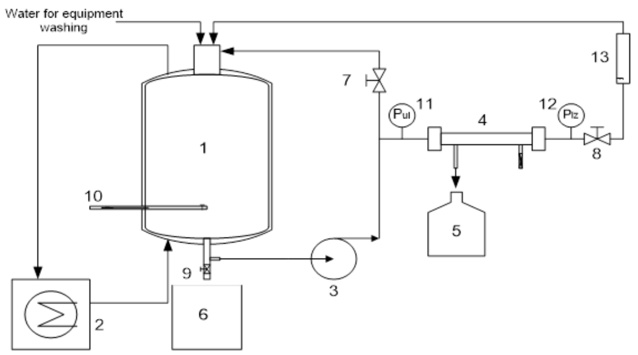

Fig. 1 Laboratory setup for cross-flow microfiltration: 1-feed tank, 2-thermostat, 3-pump (0.25 kW), 4-module with membrane, 5-vessel for permeate, 6-vessel for retentate, 7, 8, 9-valves, 10-thermometer, 11, 12-manometer, 13-rotameter

experiments in this study were carried out in triplicate and the results averaged. The reproducibility of these measurements were good, the deviation between parallel experiments were in the range of ±6.2%. All experiments were carried out at the room temperature (25°C). The laboratory apparatus for microfiltration is shown in Fig. 1.

The single channel ceramic aluminium silicate and titanium oxide membrane (GEA, Germany) had a nominal pore size 200 nm with the length of 250 mm and inner/external diameter of 7/10 mm. The useful membrane surface was 5.00×10^{-3} m^2. After the working parameters have been set, the filtration of wastewater was initiated. Over a predetermined time periods the permeate volume was recorded, based on which the permeate flux is calculated.

The static turbulence promoter used throughout experiments was the Kenics static turbulence promoter (MX8124-AC, Omega, USA). The main geometrical characteristics of the static mixer are: the diameter of the static turbulence promoter is 6.35 mm, its length equals 190.5 mm and the number of elements is 30. The static turbulence promoter was inserted inside the whole membrane tube and was fixed appropriately to avoid any movement due to the fluid flow. It consists of a series of helical mixing elements made from thin, flat strips, twisted through 180° to form helices. Helices are turned around their main axis by 90° against the next element. Its characteristic geometric design produces the unique patterns of flow division and radial mixing simultaneously. Additionally, the Kenics static turbulence promoter has “streamlined” shape which presents minimal surface area in the plane normal to the tube axis and prevents the creation of stagnation regions where impurities may collect and eventually foul the membrane. These features strongly favoured the Kenics static turbulence promoter in respect to other commercial static turbulence promoter for cross-flow filtration applications (Krstić *et al.* 2004).

2.2 Data compilation

Reliable experimental data for training of the ANN models is of utmost importance. Total data for changing flux during microfiltration experiments are divided into two groups: first one microfiltration without static mixer and the second group data of experiments done with static turbulence promoter inserted into membrane channel. In the case of microfiltration without static turbulence promoter data consisted of 646 data points while in the case of microfiltration with static turbulence promoter there is 550 data points. Prior to the artificial network training data sets were normalized; scaling or normalized; that brings all data within a specific range. Min–Max normalization was used i.e., subtracts the minimum value of an attribute from each value of the attribute and then divides the difference by the range of the attribute. It has the advantage of preserving exactly all relationship in the data, without adding any bias. The values predicted by the ANN during the simulation must be de-normalized to generate data in the original range.

Operation conditions including feed flow rate, transmembrane pressure and wastewater sedimentation time were taken as the input variables of ANN model in this study. Additional input variable was microfiltration time. As output, permeate flux values for specific filtration times were selected.

2.3 Artificial neural network modelling

Feed-forward neural networks are the most popular and most widely used models in many practical applications and in this study it was used with backpropagation training algorithm. Feed-forward neural networks are made of many computing elements, called neurons, which are connected by weights, which are allowed to be adapted through a learning process. The hyperbolic tangent sigmoid function is the transfer function employed in the hidden and output layers. Among the many backpropagation training methods the Levenberg-Marquardt training algorithm and Bayesian regularization are selected. MATLAB R2012b neural network toolbox was used to implement ANN modelling. The standard statistical indicators are employed in order to determine the best ANN architecture the mean square error (MSE) and the coefficient of determination (R^2) between the experimental and calculated values.

The multi-layer feed-forward network was used in this study. Internal ANN factors such as number of hidden layers, number of neurons in each layer, epoch size, momentum factor, learning rate, transfer functions, and initial weight distribution have great impact on model building. Default values were selected for some of these factors (momentum factor and learning rate), since they only affect the training time (Bayar *et al.* 2009). In our study, the maximum number of epochs, target error goal MSE, and minimum performance gradient are set to 1000, 0, and 10^{-10}, respectively. Training stops when the maximum number of epochs is reached or when either the MSE or performance gradient is minimized to arrive at the pre-determined goal. Since the neural network is highly dependent upon the initial weight values and in order to achieve the best results, the neural networks were run ten times and the average values of statistical indicators, the mean square error (MSE) and the coefficient of determination (R^2), are used for comparing of network performances (Ghandehari *et al.* 2011, Saghatoleslami *et al.* 2011).

2.3.1 Selection of learning algorithm

The Lavenberg-Marquardt and Bayesian regularization are selected as training algorithms. The Lavenberg-Marquardt algorithm uses an early stopping criterion to improve network training speed and efficiency. To determine the criterion, all the data for both microfiltration modes (without or with static turbulence promoter) are divided into three sets. The first set is the training set for determining the weights and biases of the network and it consists of 80% of all data. The second set is the validation set (10% of data) for evaluating the weights and biases and for deciding when to stop training. The validation error generally decreases at the beginning of the training process, but when the network starts to over-fit the data, the validation error begins to increase. The training is stopped when the validation error begins to increase and the weights and biases will then be derived at the minimum error. A maximum validation failure is set to default value of five. The last data set is for testing (10% of data), the weights and biases are used to verify the capability of the stopping criterion and to estimate the expected network operation on new data sets.

Another approach to solve overfitting issue is Bayesian regularization algorithm (trainbr) which is adjusting the optimal regularization parameters in an automated fashion (Demuth and Beale, 1998). Bayesian regularization is a modification of the Levenberg-Marquardt training algorithm to improve the model's generalization. Over-fitting problem or reduced generalization ability occurs when an artificial neural network over learns during a training period. Due to its lack of generalization ability such a too well trained model may not perform well on unseen data set (new data presented to ANN). This approach involves modifying the performance function, which is normally chosen to be the sum of squares of the network errors on the training set (MSE). It is assumed that the weights and biases of the network are random variables following specified distributions and the parameters are related to the unknown variances associated with these distributions. Then, these parameters can be estimated using statistical techniques (Demuth and Beale, 2004). Using this performance function will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to over-fitting. In comparison, the accuracy of the trainbr algorithm is approximately five times more than early stopping method (Demuth and Beale, 1998). The complete data sets for both microfiltration modes were divided into two sets according to software default algorithm values; for training 80% and the remaining 20% for testing.

2.3.2 The number of hidden layer neurons

The optimal number of hidden layer neurons is case dependent and there is no straightforward method to adjust it (Ghafari-Nazari and Mozafari 2012). Increase in the number of the hidden layer neurons leads to enhancement of the approximation ability of the neural networks. Nevertheless, when that number exceeds an optimum, the overfitting problem may occur. In these circumstances, although the statistical indicators of the network are very suitable for training data, its predictions deteriorate for points not included for training.

3. Results and discussion

3.1 Effect of learning algorithm

The network was trained firstly through feed forward of the input training pattern and then the associated error was calculated (back propagation); finally, the weights and biases of network were adjusted. The ANN outputs corresponding to its own input were compared with the target values and the weights and biases were adjusted to reduce the mean square errors (MSE) The Levenberg–Marquardt and Bayesian regularization are selected as training algorithms. The neural network implemented in this work has four inputs and one output. In order to select the best training algorithm neural network with three hidden layer neurons is investigated, so the architecture of the network is 4-3-1.

The Levenberg-Marquardt training algorithm (trainlm) in case of microfiltration without static turbulence promoter yields MSE value 5.82×10^{-3}. For the same data sets coefficient of determination was for training data 0.952 and for testing data 0.927. For the experimental data for microfiltration with static turbulence promoter MSE value is 2.63×10^{-3}, whilst coefficient of determination are 0.968 and 0.942 are for training and testing data sets, respectively. Bayesian regularization showed better results of network performance (MSE) compared to Levenberg-Marquardt training algorithm. In the case of microfiltration without presence of static turbulence promoter the value of MSE is 2.23×10^{-3}, whilst in case of microfiltration enhanced by use of static mixer as turbulence promoter that value is 1.54×10^{-3}. For data sets of microfiltration without static mixer coefficient of determination for training data is 0.968 and for testing data 0.936. The experimental data for microfiltration with static turbulence promoter, ANN model yielded coefficient of determination of 0.989 and 0.978, for training and testing data sets, respectively.

In this study for further ANN model development automated Bayesian regularization is selected, as better generalized neural network model for this data.

3.2 Number of hidden layer neurons

A trial and error based method was selected for defining the number of neurons in the hidden layer. Fig. 2 shows the variation of squared correlation coefficient and MSE versus the number of neurons in the hidden layer for microfiltration experiments without static turbulence promoter (NSM mode).

In this figure, the horizontal axis displays the number of neurons in the hidden layer and the vertical axis denotes squared correlation coefficient for training and testing data sets (right) and the second (left) performance of the network i.e., MSE. It is obvious that increasing the number of hidden neurons from 1 to 10 increased the coefficient of

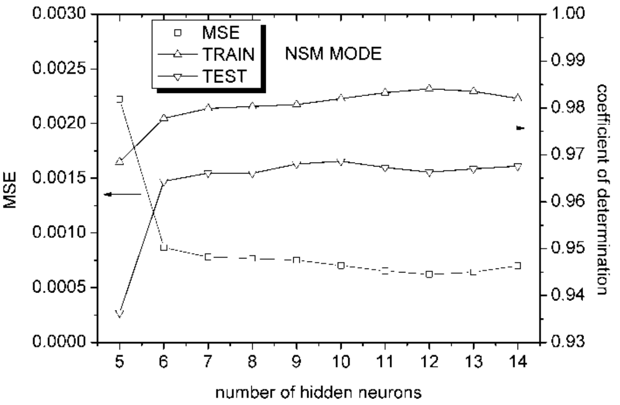

Fig. 2 The variation of squared correlation coefficient and MSE versus the number of neurons in the hidden layer for microfiltration experiments without static turbulence promoter(NSM mode)

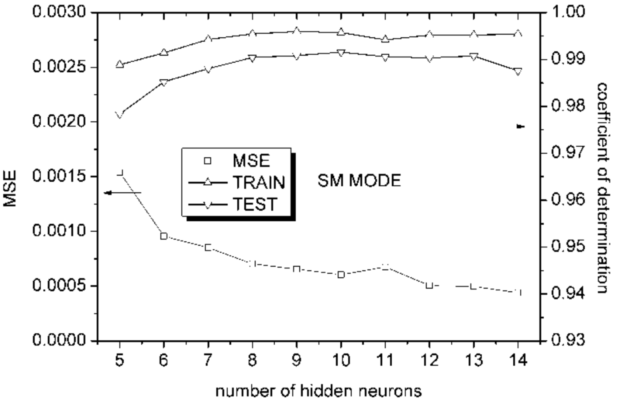

Fig. 3 The variation of squared correlation coefficient and MSE versus the number of neurons in the hidden layer for microfiltration experiments with static turbulence promoter(SM mode)

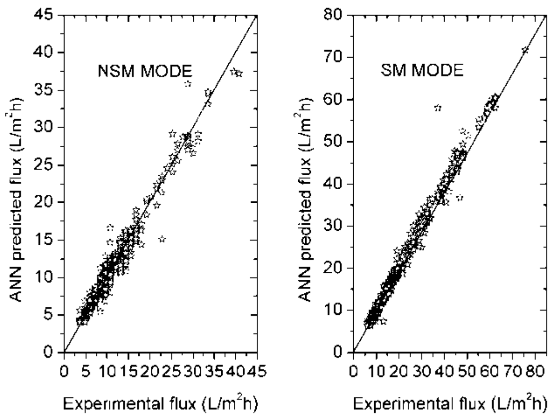

Fig. 4 The scatter plots of predicted and actual output values for all data points

determination; for both data sets, training as well as testing. Further increase in hidden layer neurons resulted in coefficient of determination rise but only for the training data, while for testing data decrease of coefficient values can be seen. In the same time MSE decreased for the all selected hidden neuron numbers up to 12. The reason for this may be found in the fact that although network

Table 1 Absolute relative error distribution of ANN predictions for all experiments

Absolute relative error		<1	<5	<10	<15	<20	>20	Total
NSM mode	Number of observations	70	250	183	90	40	13	646
	%	11	39	28	14	6	2	100
SM mode	Number of observations	72	280	138	41	11	8	550
	%	13	51	25	8	2	1	100

performance is improving, over-fitting problem or reduced generalization ability occurs because an artificial neural network over learns during a training period and in testing data lower value of coefficient of determination is achieved. So for data regarding the microfiltration without static turbulence promoter, the optimal number of hidden layer neurons is ten as the network has coefficient of determination 0.982 and 0.969, for training and testing data, respectively. In the same time MSE is 7.03×10^{-4}. This results show that the network, besides having grate fitting capabilities, has generalization ability.

In Fig. 3, the variation of squared correlation coefficient and MSE versus the number of neurons in the hidden layer for microfiltration experiments with static turbulence promoter (SM mode) is shown. Same as previous figure the horizontal axis displays the number of neurons in the hidden layer and the vertical axis denotes squared correlation coefficient for training and testing data sets (right) and the second (left) performance of the network i.e., MSE.

Once again it can be seen that coefficient of determination is increasing with the rise in number of hidden layer neurons. Nevertheless, in the case of experimental data obtained for microfiltration of starch industry wastewater, coefficient of determination values are slightly higher compared to the data retrieved from the microfiltration experiments without static turbulence promoter. In the SM mode similar ANN results are obtained as for NSM mode. In this case ten hidden neurons are also sufficient for the ANN simulation. The network has coefficient of determination 0.996 and 0.992, for training and testing data, respectively. In the same time MSE is 6.06×10^{-4}.

The experimental results versus neural network predictions are depicted in Fig. 4. It can be found that there are very close agreement between the experimental data and predicted curves for the great majority of cases. Error analysis was carried out for better comparison and understanding of each model results. As summarized in Table 1, ANNs model developed herein were able to predict the great majority of observations with <15% absolute relative error for NSM data. For the data collected during microfiltration without static turbulence promoter 92% of observations were predicted with the absolute relative error less than 15% in the same time almost 50% of predicted data had absolute relative error less than 5%. In the same time only 13 (2%) data points had absolute relative error greater than 20%.

In the case of microfiltration with static turbulence promoter, SM mode, even more accurate predictions were

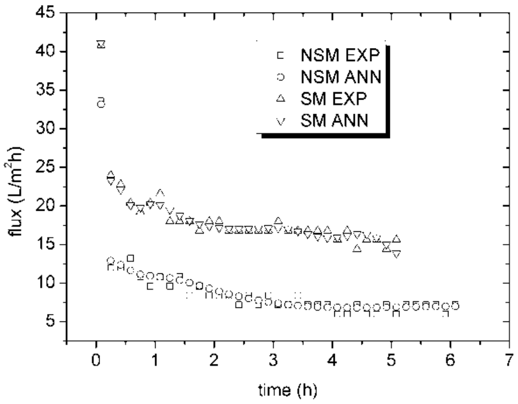

Fig. 5 Flux values for verification data sets using optimized ANN architecture

achieved, 97% of predicted data points had absolute relative error less than 15%; whilst absolute relative error less than 5% was achieved in 64% of observations.

According to absolute relative error of the ANN simulation results (Table 1), network with selected architecture demonstrated the advantage in highly accurate predictions of instantaneous permeate flux under an experimental range of feed rates, sedimentation times and transmembrane pressures.

These flux predictions indicates that ANNs can potentially reduce the number of expensive, pilot-scale tests often conducted in support of membrane system design.

3.3 Model verification

To evaluate the generalization capacity of ANN, new experimental data were presented to the neural networks. It is good to emphasize that these data were unknown to the ANN. The input variables used for the verification of the developed model were transmembrane pressure 1.5×10^5 Pa, sedimentation time 2.5 h and feed flow rate 120 L/h. The microfiltration experiments were done in two modes, NSM and SM. Results of model verification are presented in Fig. 5.

A good correlation among the input and output data could be observed for the additional test set. As shown in Fig. 5, membrane fouling occurred at the initial stage of starch wastewater, consequently flux decline rapidly followed by slow decline of flux values till the end of microfiltration run. At the simulated curves, this problem can be detected. It can be observed that static turbulence promoter increased permeate flux from around 6 L/m^{2}h in NSM mode to around 15 L/m^{2}h in SM mode. Dynamic behaviour of permeate flux during starch industry wastewater has been successfully predicted with good accuracy by developed ANN.

3.4 Relative importance of the input variables

ANNs need not be used simply as black boxes and cause-effect information can be quantitatively extracted from network connection weights to assist in model development and experimental design (Chellam 2005). The method for ranking the relative importance of input variables was first proposed by Garson (1991). This method involves partitioning the connection weights absolute values (the hidden output connection weights) of each hidden neuron into components associated with each input to the neural network. Relative importance of input variable is calculated according to Eq. (1)

$$v = \frac{\sum_{j=1}^{n_H}\left[\left(i_{vj} / \sum_{k=1}^{n_v} i_{kj}\right) O_j\right]}{\sum_{i=1}^{n_v}\left[\sum_{j=1}^{n_H}\left(\left(i_{vj} / \sum_{k=1}^{n_v} i_{kj}\right) O_j\right)\right]} \tag{1}$$

where n_v is the number of input neurons, n_H the number of hidden neurons, i_j the absolute value of connection weights between the input and hidden layers, and $O_j$is the absolute value of connection weights between the hidden and output layers.

Table 2 Relative importance of each input variable to instantaneous specific flux

Input	NSM mode		SM mode	
	Importance (%)	Rank	Importance (%)	Rank
Transmembrane pressure (Pa)	11	4	16.8	3
Feed flow rate (L/h)	12.6	3	18.3	2
Sedimentation time (h)	11.4	2	13.1	4
Filtration time (h)	66	1	51.8	1
Total	100		100	

As seen in Table 2, in all cases time played an important role in determining permeates flux (66.0% in NSM mode and 51.8% in SM mode, decrease for 27%). Similar results were reported by Aydiner *et al.* (2005); the contribution of filtration time as input variable to flux values provided by ANNs during cross-flow microfiltration of a mixture that contains phosphate and fly ash, was determined in an important level at the range of 40-50% due to increasing in membrane fouling by the time.

The other system parameters (transmembrane pressure, feed flow rate and sedimentation time) investigated in the study have a contribution of about 11-12.6% in NSM mode, whilst in SM mode contribution range is 13.1-18.3%. After filtration time, feed flow rate has the greatest contribution to ANN model of dynamic flux behaviour in both modes of microfiltration. Influence of feed flow rate increased for around 30% when static turbulence promoter is placed into the membrane channel. The reason can be found in the fact that flow pattern in the membrane channel is changed so the membrane fouling is reduced (Krstić *et al*. 2004). The same explanation may be applied also to the increase (around 35%) in relative importance of transmembrane pressure in SM mode compared to NSM mode.

In the case of sedimentation time the increase in relative importance in SM mode is about 13% compared to NSM mode. During sedimentation, the largest particles are removed, with the only small particles remaining in suspension, these particles can settle on the membrane surface or within the pores of the membrane. The larger particles still in solution concentrate near the membrane and function as a prefilter, preventing the obstruction of the membrane pores by smaller particles (Cancino-Madariaga and Aguirre 2011). For this reason in the case of microfiltration with static turbulence promoter with changed feed flow patterns hindered influence of sedimentation time compared to the transmembrane pressure influence.

4. Conclusions

The present work was focused on the application of ANN on the modelling microfiltration of wheat starch industry wastewaters. The dynamics of the rate of specific flux decline during cross-flow microfiltration with and without static turbulence promoter of wheat starch industry wastewaters was captured accurately by ANNs. Bayesian regularization showed better results of network performance (MSE) compared to Levenberg-Marquardt training algorithm. The optimal number of hidden layer neurons is ten, the networks have coefficient of determination for testing data 0.969 and 0.992, for the NSM as well as SM mode, respectively. Absolute relative error analysis showed satisfactory predictions of permeate flux with more than 90% of data were predicted with error less than 15%. The relative importance of input variables was investigated by applying the Garson equation. The model finding revealed that filtration time has the most significant effect on permeate flux. In SM mode the relative importance of time was lower by 27%, when compared to NSM mode. As for other input variables the increase in their importance was 35, 31 and 13% for transmembrane pressure, feed flow rate and sedimentation time, respectively.

Acknowledgements

The authors would like to thank to office GEA, Westfalia Separator GmbH in Belgrade for donated ceramic membrane. This study was part of national project of Ministry of Science and Technological Development of the Republic of Serbia.

References

Aghbashlo, M., Mobli, H., Rafiee, S. and Madadlou, A. (2012), “The use of artificial neural network to predict exergetic performance of spray drying process: A preliminary study”, *Comput. Electron. Agric.*, **88**, 32-43.

Al-Abri, M. and Hilal, N. (2008), “Artificial neural network simulation of combined humic substance coagulation and membrane filtration”, *Chem. Eng. J.*, **141**(1-3), 27-34.

Aydiner, C., Demir, I. and Yildiz E. (2005), “Modeling of flux decline in crossflow microfiltration using neural networks: The case of phosphate removal”, *J. Membr. Sci.*, **248**(1-2), 53-62.

Bayar, S., Demir, I. and Engin, GO. (2009), “Modeling leaching behavior of solidified wastes using back-propagation neural networks”, *Ecotoxicol. Environ. Saf.*, **72**(3), 843-850.

Bayat, H., Neyshaburi, M.R., Mohammadi, K., Nariman-Zadeh, N., Irannejad, M. and Gregory, AS. (2013), “Combination of artificial neural networks and fractal theory to predict soil water retention curve”, *Comput. Electron. Agric.*, **92**, 92-103.

Bowen, W.R. and Jenner, F. (1995), “Theoretical descriptions of membrane filtration of colloids and fine particles: An assessment and review”, *Adv. Colloid Interfac. Sci.*, **56**, 141-200.

Cancino-Madariaga, B. and Aguirre, J. (2011), “Combination treatment of corn starch wastewater by sedimentation, microfiltration and reverse osmosis”, *Desalination*, **279**(1-3), 285-90.

Chellam, S. (2005), “Artificial neural network model for transient crossflow microfiltration of polydispersed suspensions”, *J. Membr. Sci.*, **258**(1-2), 35-42.

Demuth, H. and Beale, M. (1998), *Neural Network Toolbox for Use with MATLAB*, TheMathWorks Inc., Massachusetts, U.S.A.

Demuth, H. and Beale, M. (2004), *Neural Network Toolbox for Use with MATLAB (Version 4.0)*, The MathWorks, Inc., Massachusetts, U.S.A.

Dornier, M., Decloux, M., Trystram, G. and Lebert, A. (1995), “Dynamic modeling of crossflow microfiltration using neural networks”, *J. Membr. Sci.*, **98**(3), 263-73.

Garson, GD. (1991), “Interpreting neural-network connection weights”, *AI Expert*, **6**(4), 46-51.

Ghafari-Nazari, A. and Mozafari, M. (2012), “Simulation of structural features on mechanochemical synthesis of Al2O3-TiB2 nanocomposite by optimized artificial neural network”, *Adv. Powder Technol.*, **23**(2), 220-227.

Ghandehari, S., Montazer-Rahmati, M.M. and Asghari, M. (2011), “A comparison between semi-theoretical and empirical modeling of cross-flow microfiltration using ANN”, *Desalination*, **277**(1-3), 348-355.

Hermia, J. (1985), *Blocking Filtration. Application to Non-Newtonian Fluids*, in *Mathematical Models and Design Methods in Solid-Liquid Separation*, Springer, Dordrecht, The Netherlands.

Jokić, A., Zavargo, Z., Šereš, Z. and Tekić, M. (2010), “The effect of turbulence promoter on cross-flow microfiltration of yeast suspensions: A response surface methodology approach”, *J. Membr. Sci.*, **350**, 269-278.

Krstić, DM.,Tekić, MN., Carić, MD. and Milanović, S.D. (2004), “Static turbulence promoter in cross-flow microfiltration of skim milk”, *Desalination*, **163**, 297-309.

Liu, Y., He, G., Tan, M., Nie, F. and Li, B. (2014), “Artificial neural network model for turbulence promoter-assisted crossflow microfiltration of particulate suspensions”, *Desalination*, **338**, 57-64.

Nourbakhsh, H., Emam-Djomeh, Z., Omid, M., Mirsaeedghazi, H. and Moini, S. (2014), “Prediction of red plum juice permeate flux during membrane processing with ANN optimized using RSM”, *Comput. Electron. Agric.*, **102**, 1-9.

Popović, S., Jovičević, D., Muhadinović, M., Milanović, S. and Tekić, M.N. (2013), “Intensification of microfiltration using a blade-type turbulence promoter”, *J. Membr. Sci.*, **425**, 113-120.

Porrazzo, R., Cipollina, A., Galluzzo, M. and Micale, G. (2013), “A neural network-based optimizing control system for a seawater-desalination solar-powered membrane distillation unit”, *Comput. Chem. Eng.*, **54**, 79-96.

Saghatoleslami, N., Vatankhah, GH., Karimi, H. and Noie, S.H. (2011), “Prediction of the overall sieve tray efficiency for a group of hydrocarbons, an artificial neural network approach”, *J. Nat. Gas Sci. Eng.*, **3**(1), 319-25.

Šaranović, Ž., Šereš, Z., Jokić, A., Pajin, B., Dokić, Lj., Gyura, J., Dalmacija, B. and Simović, D.Š. (2011), “Reduction of solid content in starch industry wastewater by microfiltration”, *Starch Stärke*, **63**(2), 64-74.

Tanaka, T., Kamimura, R., Fujiwara, R. and Nakanishi, K. (1994), “Crossflow filtration of yeast cultivated in molasses”, *Biotechnol. Bioeng.*, **43**(11), 1094-1101.

ED