Desalination and Water Treatment
www.deswater.com

14 (2010) 30–36
February

1944-3994/1944-3986
doi: 10.5004/dwt.2010.1008

Short-range forecast of permeate flux in detergent waste water ultrafiltration

Malgorzata Kabsch-Korbutowicz, Malgorzata Kutylowska*

Institute of Environment Protection Engineering, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, -370 Wroclaw, Poland Tel. +48713203639; Fax.: +48713282980; email: malgorzata.kutylowska@pwr.wroc.pl

Received 11 June 2009; accepted 5 November 2009

ABSTRACT

Waste water purification is very crucial nowadays and new techniques like low-pressure driven membrane processes are developed to obtain the sufficient productivity and high efficiency of cleaning. However, selection of membrane requires long-lasting tests. The new methods like artificial neural networks (ANN) are essential to avoid time consuming laboratory measurements. In the article the permeate flow rate of detergent waste water was forecasted using ANN. In order to predict the flow rate, the network used short-range time series regression. During two days of continuous measurements (every 5 min) the following parameters were measured: temperature of waste water, recirculation flow rate, inlet pressure and outlet pressure. Those parameters were considered as inputs to the network while permeate flow rate as output value. The best created network (with the highest value of Pearson correlation coefficient and the lowest value of training error) was characterised by the following parameters: hyperbolic tangent and sine function of activation, Quasi-Newton method of learning, 4 input neurons, 8 hidden neurons and 1 output neuron.

Keywords: Flow rate; Detergent waste water; Ultrafiltration; Short-range forecast; Time series; Artificial neural network

1. Introduction

Nowadays it is needed to apply efficient processes in waste water treatment in order to achieve the assumed level of purification. Ultrafiltration became more effective recently from environmental engineering point of view. Despite substantial progress done in last years in membrane science, we have still many problems in controlling and predicting membrane fouling and membrane selectivity.

At present membranes are often used in domestic and industrial wastewater treatment to remove different pollutants, e.g. dye effluents [1] or detergents [2]. The environmental risk of detergent effluents related to manufacture, use and disposal of these chemicals is of a great interest [2]. These effluents are present in high concentration in the reactor and are mainly responsible for the residual products, which are needed to be washed away because of using the same facility for the manufacture of other products. Detergents are also present in domestic and industrial waste water. The main active ingredients of detergents are called surfactants. Surfactants, which are surface-active amphiphilic agents containing both hydrophilic and hydrophobic components, are classified into four groups depending on the charge of the hydrophilic moiety: anionic, cationic, nonionic or zwitterionic [3]. Detergent effluents influence significantly environment because surfactants and other detergent components can be toxic to aquatic life. Low

*Corresponding author

Presented at the Fourth Membrane Science and Technology Conference of Visegrad Countries (PERMEA 2009), Prague, 7-11 June, 2009.

pressure-driven membrane separation process may be useful in treatment of waste water containing detergents. The membrane filtration process, due to the selectivity of the membrane, make it possible to recover detergents and process water as well as reduce high organic load of disposed waste water.

New mathematical tools can be used in order to solve complex engineering problems [4]. There are known many mathematical models widely used for predicting membrane parameters, but some unconventional methods are also considered as promising implementation for solving complicated issues. New method of mathematical modelling using artificial intelligence is lately very popular. Many chemical and environmental engineering tasks can be modelled using artificial neural network (ANN).

1.1. Principles of artificial neural networks

In artificial neural networks the method of information transferring is imitating the way of human nervous system performance. Natural neurons (Fig. 1), the main elements of nervous system, are responsible for transferring information.

ANN consists of neurons which are data processors. Each neuron is responsible for summarizing input signals. The first model of artificial neuron was created by McCulloch – Pitts in 1943. According to McCulloch – Pitts [5] the output signals are expressed as:

$$y_i = f\left(\sum_{j=1}^{N} w_{ij} \cdot x_j + b_i\right), \quad i,\ j = 1, 2, ..., N \qquad (1)$$

The function $f(^*)$, as shown in the Eq. (1), is called activation function that stimulates the information transmission. Above mentioned model is quite simple and since 1943 ANN has been developed and improved to be sufficient for modelling a lot of dynamic processes. ANN just computes output values from input values. The sum of transferred information is weighted. The sum values are transmitted to the next network layer. Neural networks enable to model non-linear and complex problems. These days the most popular type of the network is multilayer perceptrone with one input layer, one or more hidden layers and one output layer. Generally, it should be remembered that ANN modelling is like ''black box'' approach and that is why it is impossible to penetrate deeply inside the way of forming the network structure. The modelling using ANN may be quite good manner of predicting membrane parameters. The prediction by ANN approach is dynamic and efficient which is important because of changing parameters during operating time.

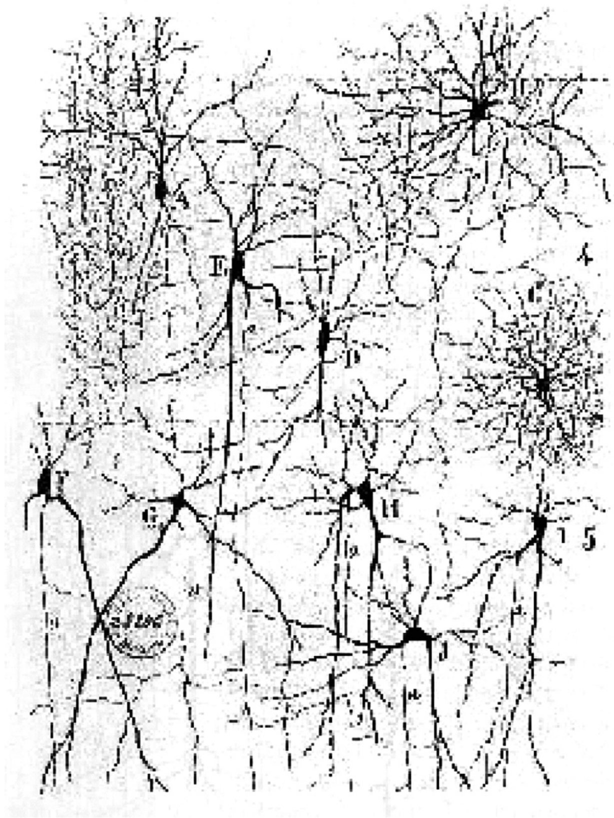

Fig. 1. The image of the natural neurons.

1.2. Previous investigations in modelling membrane parameters using ANN

Up to now ANN were mainly used for prediction of transmembrane pressure, retention coefficient and permeate flux. There are a lot of investigations concerning modelling membrane parameters using neural network approach. Hilal et al. [6] have shown how neural network approach can be used to forecast the dynamic rate of membrane fouling, particle deposition and improvement of flux sustainability and membrane efficiency as a function of pressure, temperature and crossflow velocity. Curcio et al. [7] have developed the hybrid system to model the behavior of membrane units operating in pulsating conditions. In two previous papers [8,9] the authors have already shown how complex models based on ANN can be applied in prediction the behavior of unsteady-state membrane processes. Chen and Kim [10] applied ANN, based on radial basis function, to calculate the permeate flux in crossflow membrane filtration as a function of transmembrane pressure, ionic strength, solution pH, particle size, and elapsed filtration time. Some research were carried out in order to predict the process parameters in membrane treatment of effluent from meat industry [11] or purification of municipal wastewater using membrane bioreactors [12].

There is no information in the literature about short-range forecasting of membranes flux in detergent waste water ultrafiltration using time series regression. The analysis of time series is based on the statement that successive values in the data set represent consecutive measurements taken at equally spaced time intervals. There are two major goals of time series analysis: (i) identifying the character of the phenomenon represented by the sequence of observations, and (ii) forecasting – predicting future values of the time series variable. These goals require that the pattern of experimental time series data is identified and more or less formally described. When the pattern is established, it can be interpreted and integrated with other data. Short-range forecast means the prediction in one day time horizon. Short-range forecast of flow rate in detergent waste water by using time series regression in ANN will be a helpful approach due to the complexity of the detergent effluent modelling.

The aim of this article was to predict permeate flow of ultrafiltration membranes used in the treatment waste water from detergent production. Membrane process was used in order to concentrate and recover detergents.

2. Experimental

2.1. Ultrafiltration of detergent waste water

Experiments and measurements were done at one commercial industrial plant. In the experiments 5 kDa hollow fibre membranes made from polyethersulphone were applied. System was equipped with 30 Romicon type modules produced by Koch. The inner diameter of the fibre amounted to 1.1 mm and the length of one fibre equalled to 1.0 m. One module consisted of approximately 1,755 fibres. The membrane area in the module amounted to approximately 6.13 m^2. The operating data of the module were established as follow: maximal inlet pressure – 270 kPa, maximal transmembrane pressure – 240 kPa, maximal backwash pressure – 140 kPa, maximal temperature at pH equal to 6–333.15 K, pH range 1.5÷13.0. Permeate pressure was maintained at the constant level of 10 kPa. Initial COD varied in the range 70÷1,00,000 $gO_2{\cdot}m^{-3}$ and the measurements were discrete, while the COD in concentrate was changing from 250 to 3,00,000 $gO_2{\cdot}m^{-3}$. Experiments were run continuously for two days. Every 5 min such parameters as: waste water temperature T (K), recirculation flow F_r ($m^3{\cdot}h^{-1}$), inlet pressure P_1 (Pa), outlet pressure P_2 (Pa) and permeate flow rate Q ($m^3 \cdot h^{-1}$) were measured. The installation has been working for about 0.5–2 h and then the backwash was performed. The aim of the ultrafiltration was to concentrate the detergent solution. Results of the experiments carried out in the industrial plant were taken into consideration and used for creating, learning and verification of artificial neural network. 288 results from one day were used to learn ANN and also 288 from another day to verify created network.

2.2. Assumptions of ANN structure for short-range forecast of permeate flux

During the creation of ANN using time series the choice of proper inputs and time horizon as well as the number of late value is very important. In the created network the time horizon (look ahead) was equal to 0 and the number of late value (step) amounted to 1. Also the number of input variables is quite significant matter. The more variables at the beginning, the more accurately output parameter is predicted. Such approach creates function with better fitting. The realization of artificial neural network approach was performed in the program Statistica 8.0. Using time series (short-range), there were created artificial neural networks for permeate flow rate forecasting. Although it is quantitative parameter, it can be argued that the quality parameter of the feed solution like waste water temperature has an influence on the value of the permeate flow rate. This approach was justified because with the temperature variation, the conditions of the ultrafiltration process were also changing. Various models of neural networks were created having as inputs: waste water temperature (T), recirculation flow (F_r), inlet pressure (P_1) and outlet (P_2). Permeate flow rate (Q) was established as output value. During learning of the network input parameters varied in the range of: 290.55÷306.65 K for T, 0.0÷406.0 $m^3{\cdot}h^{-1}$ for F_r, 0÷270 kPa for P_1 and 0÷50 kPa for P_2. Output parameter varied in the range of: 0÷5.5 $m^3{\cdot}h^{-1}$ for experimental Q and −0.05÷5.3 $m^3{\cdot}h^{-1}$ for Q predicted by ANN. During verification of the network input parameters varied in the range of: 291.15÷306.15 K for T, 0.0÷390.0 $m^3{\cdot}h^{-1}$ for F_r, 0÷270 kPa for P_1 and 10÷50 kPa for P_2. Output parameter varied in the range of: 0÷4.3 $m^3{\cdot}h^{-1}$ for experimental Q and −0.02÷4.9 $m^3{\cdot}h^{-1}$ for Q predicted by ANN.

During the analysis of the network structure, there were investigated:

a) choice of subsets:

Table 1
Activation functions and their definitions

Function	Formula	Range
Linear	$f(x) = x$	$(-\infty, +\infty)$
Logistic	$f(x) = \frac{1}{1-e^{-x}}$	(0, +1)
Hyperbolic tangent	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	(−1, +1)
Exponential	$f(x) = e^{-x}$	$(0, +\infty)$
Sine	$f(x) = \sin(x)$	[0, +1]

A subset sampling method was used. Training set was called ''u'' and consisted of 117 (36 values from 0:01 to 2:56 h, 34 values from 10:41 to 13:26 h and 47 values from 18:21 to 22:11 h) selected variables from all 288 data, testing set was called ''t'' and consisted of 89 (52 values from 3:01 to 7:16 h, 16 values from 13:31 to 14:46 h and 21 values from 22:16 to 23:56 h) selected variables from all 288 data, validation set was called ''w'' and consisted of 82 (40 values from 7:21 to 10:36 h, 42 values from 14:51 to 18:16 h) selected variables from all 288 data. This kind of selection was conditioned by the time changes of each variable in 24 h. In verification the best created network used 100% of other 288 values from another day;

b) type of network:

Only multilayer perceptrone (MLP *I-H-O*) with *I* number of input neurons, H number of hidden neurons and O number of output neurons was examined. It is the most popular network architecture used nowadays;

c) minimum number of hidden neurons:

To have reliable boundary conditions there was decided to establish the lowest number of the hidden neurons at the level of 2;

d) maximum number of hidden neurons:

To have reliable boundary conditions there was decided to establish the highest number of the hidden neurons at the level of 10. More neurons lead to network overfitting;

e) activation functions in hidden and output layer:

Typical and simple activation functions were chosen: linear, logistic, hyperbolic tangent, exponential, sine;

f) methods of learning:

For training the following algorithms were investigated: BFGS (Broyden-Fletcher-Goldfarb-Shanno) which is quasi-Newton approach, fastest decline, conjugated gradient descent.

3. Results and discussion

Taking into consideration mentioned in the paragraph 2.2 assumptions, 5 models of neural networks were created. Table 1 sets up the description of used activation functions. From engineering point of view it is very important that mathematical description of the model should be quite simple. Investigated networks used functions that mathematical formulas are not very complicated.

For created networks the learning errors and quality of learning were determined. In Table 2 learning error and main structure parameters for all 5 network models are presented. It was decided to compare training errors in learning step, because such approach gave reasonable base for choosing the best (in verification giving the best convergence) structure of the network. For all 5 models of ANN the method of training based on quasi-Newton approach was chosen as the best in comparison to fastest decline and conjugated gradient descent algorithms. Nowadays it is told that network training based on BFGS algorithm is the most effective. It requires a small number of iterations to train a neural network given their fast convergence rate and more intelligent search criterion. It is the recommended technique for most networks with a small number of weights (less than a couple of hundred).

In time series problems, the aim is to forecast ahead the value of a variable that varies in time, using previous values of that and other variables. In the discussed problem the predicted variable (permeate flow rate) is continuous, so time series prediction is a specialized form of regression. There are no rules about the choice of the most suitable number of hidden neurons. The number of the hidden neurons depends on the assumed level of agreement and computational time. We have to remember that the more hidden neurons, the more complicated is the structure of the network and the number of iterations needed to achieve assumed level of training error. More neurons lead sometimes to overtraining the network. On the other hand, if there is to small number of hidden neurons, the level of prediction could be unsatisfactory. It is necessary to find the ''gold solution'' using often the scientific intuition. It should be remembered that not too large but also not too small number of hidden neurons is required for obtaining good convergence between measured and predicted values.

Table 2
The main parameters of network models

No.	Type of network	Activation function – hidden layer	Activation function – output layer	Training error in learning step
1.	MLP 4-9-1	Exponential	Linear	0.0007
2.	MLP 4-4-1	Logistic	Linear	0.0008
3.	MLP 4-2-1	Linear	Linear	0.0007
4.	MLP 4-8-1	Hyperbolic tangent	Sine	0.0006
5.	MLP 4-10-1	Exponential	Hyperbolic tangent	0.0010

Due to different times of backwashing in analysed days, parameters of feed stream and permeate were random and divergent. Above mentioned problems caused the differences between convergence with experimental and forecasted values in learning and verifying processes. During backwashing permeate flow rate values were equal to zero. In this case created network generated sometimes the negative values of Q which were treated as an error and replaced by real experimental values equal to zero.

Training errors in all 5 models were very small. Even the highest error value (network number 5 in Table 2) gave really great convergence, because Pearson correlation coefficient was equal to 0.9826. However, the network MLP 4-8-1 (in bolds in Table 2) with the lowest value of training error (0.0006) was chosen as the best and the verifying step was performed on the different data set. In learning process the agreement between experimental and predicted value was also almost ideal, because Pearson correlation coefficient was equal to 0.9851. As it is seen the coefficients in the best and the worse networks were established at the same level in spite of quite different values of training errors. It was found that permeate flow rate values were forecasted correctly, but for the sake of the zero values during backwashing the convergence was not so ideal as it should have been required from such new and innovative tool as artificial neural network.

The chosen network consisted of 4 input, 8 hidden and 1 output neurons. Its structure is presented in Fig. 2. Hidden neurons were activated by hyperbolic tangent and one output neuron corresponding to permeate flow rate value was activated by sine function. Sine function performed as periodical function which in the discussed problem was very important because permeate flow rate changed periodically in analysed 24 h. Interesting combination of activation functions in hidden and output layer is found. Quite rarely the sine function is chosen to solve prediction problems, but in time series regression (short-range forecast) problems it could be very useful. Also from engineering point of view it is important that these functions are characterized by quite simple mathematical form. The hyperbolic function is a symmetric S-shaped (sigmoid) function. The output lies in the range (-1, $+1$). Sine function often gives better solutions than the logistic sigmoid function because of its symmetry.

The network provided the best fit outcomes, because the convergence between experimental and forecasted flow rate was actually equal to 98%.

Chosen network MLP 4-8-1 was verified and as a consequence of that the following results were obtained:

- Pearson correlation coefficient equalled 0.9222
- Average measured permeate flow rate (Q_{measured}) at the level of 2.41 $m^3 \cdot h^{-1}$
- Average predicted permeate flow rate ($Q_{\text{predicted}}$) at the level of 2.47 $m^3 \cdot h^{-1}$

Because of the backwash (verification data set), measured flow rate was equal to zero in the following hours: 1:36–2:56, 5:16–7:01, 9:16–10:11, 13:06–13:26, 14:26–15:06, 17:31–18:31, 20:11–20:31 and 21:31–22:11. The network did not understand properly zero values of flow rate in spite of being taught that such values could have occurred and generated negative results or close to zero but not exactly equal to zero. Probably it was caused by

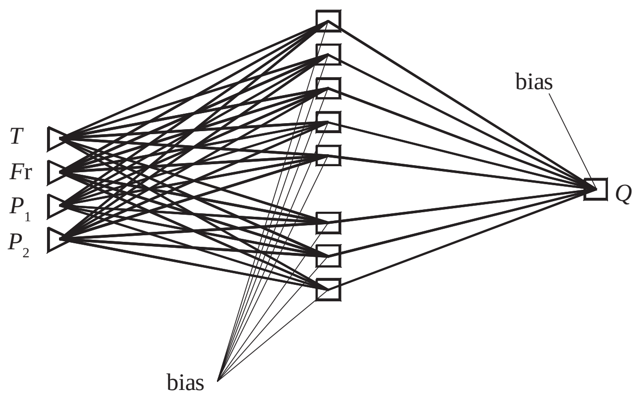

Fig. 2. The structure of the best created network.

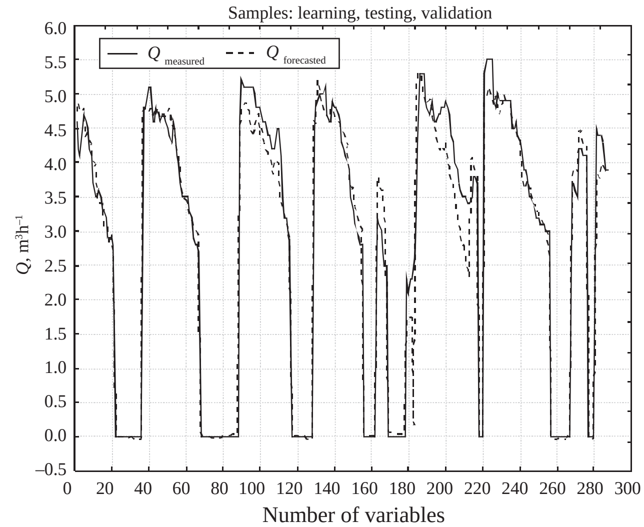

Fig. 3. The relationship between experimental and forecasted permeate flow rate values in learning process.

the fact that the learning parameters were not sufficient and the time duration of backwash in two data sets (learning and verification) was not the same. One could have concluded that the number of hidden neurons or training iterations should have been higher, but the network number 1 or 5 with higher number of neurons (Table 2) also forecasted improper values of permeate flow rate during the backwashing. Another point was: the accuracy of the data collected. There was only one solution after the verification step (of course only for the time being): to replace the predicted negative values by real one equal to zero. As a result of that 93 values of predicted permeate flow rate were turned into zero measured values. But on the other hand the rest of permeate flow rate data were predicted properly. Replacing did not cause the change in the Pearson correlation coefficient value. After replacing the Pearson correlation coefficient was equal to 0.9225, which was almost the same as previously. But in the future, such problems as data replacement should be omitted and it is necessary to find another solution.

Analyzing obtained results it should be remembered that neurones and their behaviour played the great role in the whole modelling process. It is obvious that during the network learning the neurons covered various activation positions. Some neurons were activated more than others. The changes in activity of the neurons were important because showed the behaviour of the network and the different values of the connections weights. The negative values corresponded to the less active behaviour of the neuron. The connection between outlet pressure and hidden neuron number 6 was characterized by the lowest weight value equalled to −1.05. The higher was the value of the weight, the stronger influence and impact on the neuron. In the created network the bias value was very important relating to the training process. Bias is a neuron in which activation function is permanently set to 1. Just as like other neurons, a bias connects to the neurons in the layer above. The weight is often called threshold. The highest weight value (0.85) described the connection between bias and hidden neurone number 8. It shows that thresholds played important role in the whole learning process. The interaction strength provided very significant information about the training process in which some connections were useless. The best network (MLP 4-8-1) provided specified in the Fig. 3 results of permeate flow rate prediction. Fig. 3 shows the relationship between predicted and experimental permeate flow values in learning process. Graph was constructed as time series.

Results shown in the Fig. 3 indicate quite good convergence between measured and predicted values. Only relatively slight deviation between measured and predicted permeate flow rate was observed. It means that the forecasted and measured values were correctly correlated. The predicted vector of permeate flow rate matched closely the pattern vector. This correlation was observed because the network was learnt (using BFGS approach) correctly. Apart from the mentioned earlier problems with zero values, the rest of prediction was from engineering point of view sufficient, because the average relative error was equal to 7.46%. Activation of each single neuron by hyperbolic tangent and sine function respectively in hidden and output layer and also the specific training method turned out the good solutions for the flow rate prediction.

Slightly different situation was observed during verification of the best created network. In verifying process there was put to the network different data set with 288 values of each (T, F_r, P_1 and P_2 as inputs and Q was forecasted as output signal) variable. It should be remembered that in verification step network did not know previously used data. The network just used the knowledge from the learning process and on the basis of this information predicted values of the permeate flow rate. But still the correlation, as shown in the Fig. 4, was satisfactory. Forecasted and measured values of permeate flow rates were in good agreement but apart from the mentioned earlier problems with zero values, the rest of prediction was from engineering point of view sufficient, because the average relative error was equal to 19.27%. Concerning that sometimes the measuring device error reach 10%, obtained error lower than 20% is rather good result. Such convergent results proved that ANN approach may be useful for predicting permeate flow rate in detergent waste water

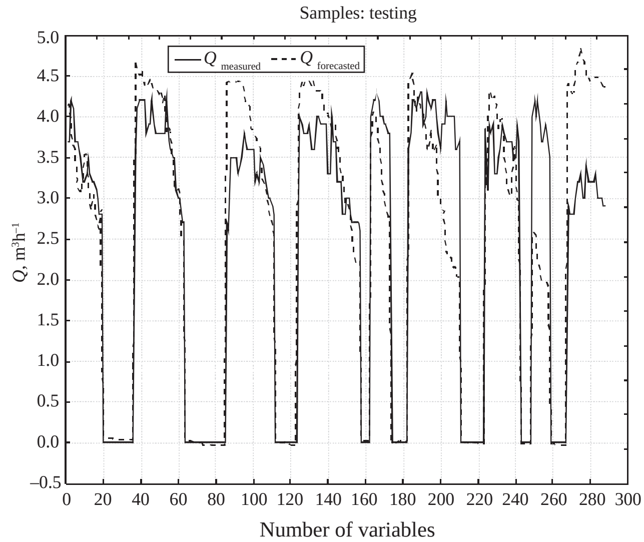

Fig. 4. The relationship between experimental and forecasted permeate flow rate values in verification process.

ultrafiltration. Analysing the effects one can conclude that the results obtained in verifying process were sufficient from engineering point of view and that modelling using ANN gave proper results when the conditions of the membrane process and feed solutions were similar. Created network can be used to predict values of permeate flow rate for other days or other installations equipped with the same kind of membrane.

4. Conclusions

Based on obtained results it might be stated that:

- the prediction of permeate flow rate in detergent waste water ultrafiltration using artificial neural network (prediction by time series regression in short-range forecast) is useful and could be recommended for forecasting hydraulic efficiency of membrane system;
- the best created network (with the highest value of Pearson correlation coefficient and the lowest value of training error) was characterised by the following parameters: hyperbolic tangent and sine function of activation, Quasi-Newton method of learning, 4 input neurons, 8 hidden neurons and 1 output neuron;
- agreement in learning and verifying process between measured and predicted values of the permeate flow rate was established at the good level and from engineering point of view could be acceptable;
- created network could be used many times for short-range forecast of permeate flow rate when the conditions of the membrane processes are almost similar;
- unfortunately, the network did not understand the zero values of permeate flow rate (during backwashing) and that is why there were generated negative results, which from engineering point of view was senseless.

Acknowledgements

The authors would like to thank company, PP-EKO Sp. z o.o.'' for giving the possibility of using their experimental data for artificial neural network modelling.

Symbols

ANN	artificial neural network
b_i	threshold values
COD	chemical oxygen demand, $gO_2 \cdot m^{-3}$
F_r	recirculation flow, $m^3 \cdot h^{-1}$
H	number of hidden neurons
I	number of input neurons
MLP	multilayer perceptrone
O	number of output neurons
P_1	inlet pressure, Pa
P_2	outlet pressure, Pa
Q	permeate flow rate, $m^3 \cdot h^{-1}$
t	name of testing set
T	waste water temperature, K
U	name of training set
W	name of validation set
w_{ij}	weights between node i and j
x_i	input signals
y_i	output signals

References

[1] Z. Badani, C. Cabassud and H. Ait-Amar, DWT, 9 (2009) 105–111.
[2] I. Kowalska, K. Majewska-Nowak and M. Kabsch-Korbutowicz, Desalination, 200 (2006) 274–276.
[3] Sz. Kertész, Zs. László, Zs. Horváth and C. Hodúr, Desalination, 221 (2008) 303–311.
[4] S. Hasani, M. Shakaib, M. Mahmood, DWT, 9 (2009) 211–220.
[5] W. McCulloch, W. Pitts, Bull. Math. Biophys., 5 (1943) 115–133.
[6] N. Hilal, O. Ogunbiyi and M. Al-Abri, Desalination, 228 (2008) 175–182.
[7] S. Curcio, V. Calabro and G. Iorio, Desalination, 236 (2009) 234–243.
[8] S. Curcio, G. Scilingo, V. Calabro and G. Iorio, J. Membr. Sci., 246 (2005) 235–247.
[9] S. Curcio, V. Calabro and G. Iorio, J. Membr. Sci., 286 (2006) 125-132.
[10] H. Chen and A.S. Kim, Desalination, 192 (2006) 415–428.
[11] E. Sroka, W. Kamiński and J. Bohdziewicz, Desalination, 162 (2004) 85–91.
[12] S. Geissler, T. Wintgens, T. Melin, K. Vossenkaul and C. Kullmann, Desalination, 178 (2005) 125–134.