Dynamic modelling of crossflow ultrafiltration of milk using neural networks

Seyed M. A. Razavi[*,#]

[#]Department of Food Science and Technology, University of Ferdowsi,
Mashad P.O.Box: 91775-1163, Iran.

Abstract

Artificial neural networks (ANNs) have been used to dynamically model cross flow ultrafiltration of milk. It aims to predict permeate flux, total hydraulic resistance and the milk components rejection (protein, fat, lactose, ash and total solids) as a function of transmembrane pressure and processing time. In this work, emphasis has been focused on intelligent selection of training data, using few training data points and small network. Also it has been tried to test the ANN ability to predict new data that not be originally available. Two neural network models were constructed to predict the flux/total resistance and rejection during ultrafiltration of milk. The results showed that there is an excellent agreement between actual and modelled data, with average errors less than 1%. Also the trained networks are able to accurately capture the non-linear dynamics of milk ultrafiltration even for a new condition that has not been used in the training process.

Keywords: Ultrafiltration; milk; neural network; flux; resistance; rejection.

1. Introduction

Ultrafiltration is an important process in the food industry, particularly for dairy applications. The efficiency and cost of membrane processing is dependent on flux and rejection, which is a function of different factors. Therefore, permeation flux and solute rejection data are necessary for the design of a specified or new membrane separation process [1]. A precise estimation of the membrane area can be made by using the equations describing the dependence of permeate flux and rejection on the process variables [1, 2]. In addition to the complexity of mathematical equations involved, each of these models has a number of limitations. Hence, modelling methods based on direct analysis of experimental data appear to be good alternative to the theoretical models. One of these methods is Artificial Neural Networks (ANNs). While the effects of process conditions on the dynamic behaviour of membrane performance (flux, rejection and fouling) are non-linear, so neural networks should allow one to account for the non-linearity of many phenomena that take place during membrane processing.

The network theory has been used to dynamically model membrane fouling during crossflow microfiltration of raw cane sugar syrup [3]. The results obtained satisfactorily modelled the effects of both constant and variable transmembrane pressure and crossflow velocity. The neural network model has also been investigated for obtaining an estimation of the permeate flux and rejection during ultrafiltration of the pulp bleach plant effluent [4]. The results showed that using neural networks a membrane process could be simulated with sufficient accuracy for process design. ANNs with a single hidden layer have been applied to predict the dynamic ultrafiltration of bovine serum albumin (BSA) as a function of pH and ionic strength [5]. Testing of the neural network approach showed that it could give excellent agreement with experimental results. The neural network approach has been used to predict the total hydraulic resistance during ultrafiltration of drink water using some parameters concerning

[*] Corresponding author. Tel.: +98 511 8410863; fax: +98 511 8415845; e-mail: Razavi@ferdowsi.um.ac.ir

water quality and operating conditions [6]. Different network structures have been evaluated and some of them allowed a prediction of resistance with very good accuracy. The development of mathematical models based on ANN has been investigated to predict the flux during ultrafiltration of wastewater [7]. The trained networks were able to accurately capture the non-linear dynamics for initial fluxes.
Although there are many papers about ultrafiltration application for milk processing [1, 8, 9, 10, 11, 12], most of these have been focused on the effects of different conditions on process performance and/or modelling of process by theoretical approaches. The aim of the present paper is to develop neural network models in order to the dynamic prediction of permeate flux, total hydraulic resistance and components rejection (proteins, fat, lactose, ash and total solids) during crossflow ultrafiltration of skim milk.

2. Neural networks modelling

The objective of a neural network is to compute output values from input data by some internal calculations. Output from a network is determined by transforming its input using a suitable transfer function. Sigmoidal function is commonly employed for non-linear relationship that it has been used in this study [3, 4, 5, 6, 7, 8]. The most popular ANN that has been used in this work is the multi-layer feed forward network, where the neurons are arranged into layers of input, hidden and output. The training process consists of determining the weights that produce from the inputs the best fit of the predicted outputs over the entire training data set. The difference between the computed output and the target is used to determine the weights using an optimisation procedure in order to minimise a suitable error function. This form of training is termed the supervised training. Several iterative methods have been proposed to minimise the error function. The most widely method which has been used in this paper is the back propagation algorithm [3,4,5]. There are several variables such as the learning rate (η), the momentum coefficient (α), number of hidden layers (L) and the number of hidden neurons (H) that have an effect on the ANN training. Also the predictions can be judged by a combination of some parameters such as the maximum, average and minimum absolute errors (E_{max}, E_{ave} and E_{min} respectively) and number of training cycles or epoch (C). To find the best set of these variables and parameters, all of those varied and the best combination chose. The software used for the ANNs modelling was EasyNN version 8.01(Stephen Wolstenholme Company, USA).
In this paper, 675 experimental data were divided into three sets for developing ANNs model, 84 data for training, 321 data for validation and 270 data for querying. The training data was used for learning the ANN and the validation data was used to test the ANN generalisability, whereas the querying data was used to test the neural networks predictability using data not used in the training/validation process. There was an emphasis on three points in this research, (i) the use of small amounts of data for training, (ii) the sensible selection of training points, and (iii) the use of small networks. The first point is related to the fact that in practical applications, process data might not be widely available. The second point is related to complexity of dynamic behaviour of flux and fouling, so an intelligent selection of training data can improve the network prediction. The third point is also related to the complexity of ANNs. Reducing ANN complexity by using few hidden neurons or layers can reduce the training time and the risk of over-fitting [5].
The training data was divided into two training sets. The first training set contained 54 data for modelling of flux and total hydraulic resistance. The second training set contained 30 data for modelling of components rejection (Table 1). Each of the data consists of two inputs: transmembrane pressure (TMP) and time (t), but the network in the first training set had two outputs: permeate flux (J_P) and total hydraulic resistance (R_T) and in the second training set had five outputs: the rejections of protein (R_P), fat (R_F), lactose (R_L), ash (R_A) and total solids

(R_{TS}). As it is shown in Table 1, the first and second training sets included 54 and 34 data from experiments at TMP 51, 152 and 253 kPa, respectively. The remaining experimental data used for two training sets were used for validation (Table 1). Choosing training and validation data from the same experiments ensures that the network will capture these data with maximum generalisability, which gives better predictions on the unseen (querying) data. The remaining data related to experiments at TMP 101.33 and 203 kPa were used as a querying set to test the ANN ability to predict new data that might not be available originally (Table 1).

Table 1. Number of data used for training, validation and querying in the ANN analysis.

TMP (kPa)	Training Set 1	Training Set 2	Validation Set 1	Validation Set 2	Querying Set 1	Querying Set 2
50	18	10	102	5		
100					120	15
150	18	10	102	5		
200					120	15
250	18	10	102	5		
Total*	54	30	306	15	240	30
Percent (%)	9	40	51	20	40	40

*It should be noted that the total number of data in set (1) and set (2) was 600 and 75, respectively.

3.Materials and methods

Ultrafiltration of milk was carried out using the pilot plant UF-MF membrane system (Biocon company, Moscow, Russia). The membrane was composed of Polysulfone amide with MWCO 20 KD. The average composition of skim milk samples was 2.86% protein, 0.09% fat, 4.73% lactose, 0.77% ash and 8.44% total solids. The effect of varying TMP (51, 101.33, 152, 203, 253 kPa) on flux, total hydraulic resistance and components rejection (Protein, fat, lactose, ash and total solids) were studied. Experiments were carried out at constant temperature and flow rate (40^oC and 15 l/min, respectively). For each set of processing conditions, the feed tank was first filled with distilled water at 40^oC to warm up the system and evaluate the water flux. Then the water was replaced with reconstituted skim milk at 40^oC. The permeate flux was measured and recorded every 30 second. After 30 min operation, the membrane unit was flushed according to manufacturer's instructions.

Protein, lactose, fat, ash and total solids contents of skim milk, permeate and retentate samples were measured using a Lactostar instrument (Funke Gerber Ltd., Germany) after 3, 15 and 30 min operation in each run. All measurements were carried out at least twice.

4. Total hydraulic resistance and rejection

By assuming that the osmotic pressure is very small, the total hydraulic resistance (R_T) can be expressed by Darcy's law (1):

$$R_T = \frac{TMP}{\mu_p \times J_P} \quad (1)$$

Where μ_p is the permeate viscosity, J_p is permeate flux, and TMP is the transmembrane pressure. The observed rejection (R_{obs}) of each component (protein, fat, lactose, ash and total solids) was calculated according to the following equation:

$$R_{obs} = 1 - \frac{C_p}{C_r} \quad (2)$$

Where C_p and C_r are the concentrations of each component in the permeate and retentate, respectively.

5. Results and discussion

5.1. Permeate flux and total hydraulic resistance

The experimental data and the results of modelling using ANNs for J_p and R_T at five TMP are shown in Fig. 1(a) and 1(b), respectively. It can be seen that the complex behaviour (non-linearity) of J_p or R_T is well reproduced by the ANNs. As shown in Fig. 1., there is excellent agreement between the predictions (solid lines) and the experimental data (dots) of full time-dependent J_p/R_T profiles at TMP 50, 150 and 250 kPa and also at intermediate TMP (100 & 200 kPa), except for R_T profile at TMP=100 kPa, where all R_T modelled data were lower than the actual data (under-estimate). The ability to predict J_p and R_T at intermediate TMP (100 and 200 kPa) could significantly reduce the computation time and the amount of practical work required before designing a new membrane process.

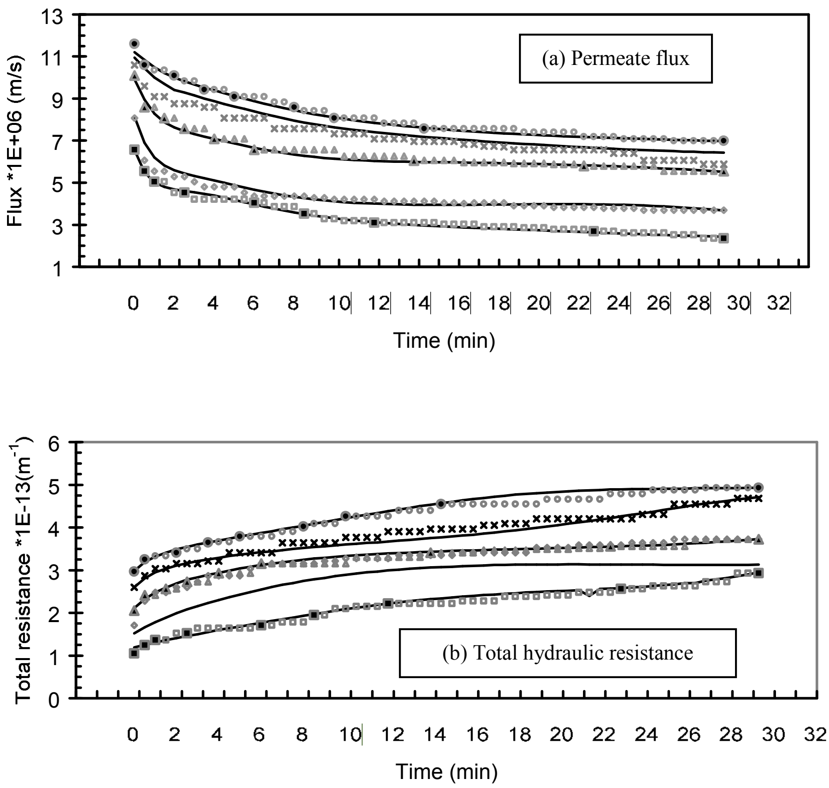

Fig. 1. Dynamic predictions of permeate flux and total hydraulic resistance during ultrafiltration of milk using training set (1). ANN used: 2/10/2. Training points are as solid symbols. Training points / Validation points / Querying points: 54/306/240. (□, 50 kPa; ◇, 100 kPa; △, 150 kPa; ×, 200 kPa; ○, 250 kPa; —, ANN)

Table 2 shows the ANNs parameters (α, η, L and H) giving the best fits for each training data set. As it is seen, for data set (1), the model with one hidden layer and ten hidden neurons produced the best model performance in terms of the E_{min} (0.09%), E_{ave} (1.0%) and E_{max} (3.35%). The table also shows that the amount of all errors is less than 3.35%, while for process design purposes a 10% error margin is usually acceptable. Furthermore, the accuracy of predictions is better than those obtained with theoretical models for milk membrane processing [1, 8, 9, 10, 11, 12].

Table 2. Best ANN results for each training data set.

Data set	α	η	L	H	$E_{min.}$	$E_{ave.}$	$E_{max.}$	C
1	0.76	0.58	1	10	0.09	1.0	3.35	47416
2	0.80	0.60	1	7	0.07	0.54	1.0	5100

Fig. 1 clearly shows that the magnitude of both J_p and R_T varies significantly with TMP and time. The permeate flux at each value of TMP decreased greatly with time during the initial

10 min. processing, whereas the total hydraulic resistance increased with time at the corresponding TMP. Also increasing TMP led to an increase in both J_p and R_T. Therefore, these results show the behaviour of J_p and R_T with TMP and time, but they cannot explain the causes of such complex behaviour.

5.2. Components rejection

The objectives were to study the ability of ANNs to model the milk solutes rejection under different TMP and operating time and also to specify the role of each component on flux decline and fouling development during the process. The results of this part are shown in Figs.2 (a-e). It can be seen that there is excellent agreement between the experimental data and the predictions, so that the simulated curves (solid lines) followed the same trend as the experimental ones (dots) for each TMP and component, although for R_L, R_A and R_{TS} at TMP= 200 kPa, the ANN model slightly underestimated the values (Fig. 2 (d)). These results show how the ANNs approach can be used to predict process variables at other conditions [5].

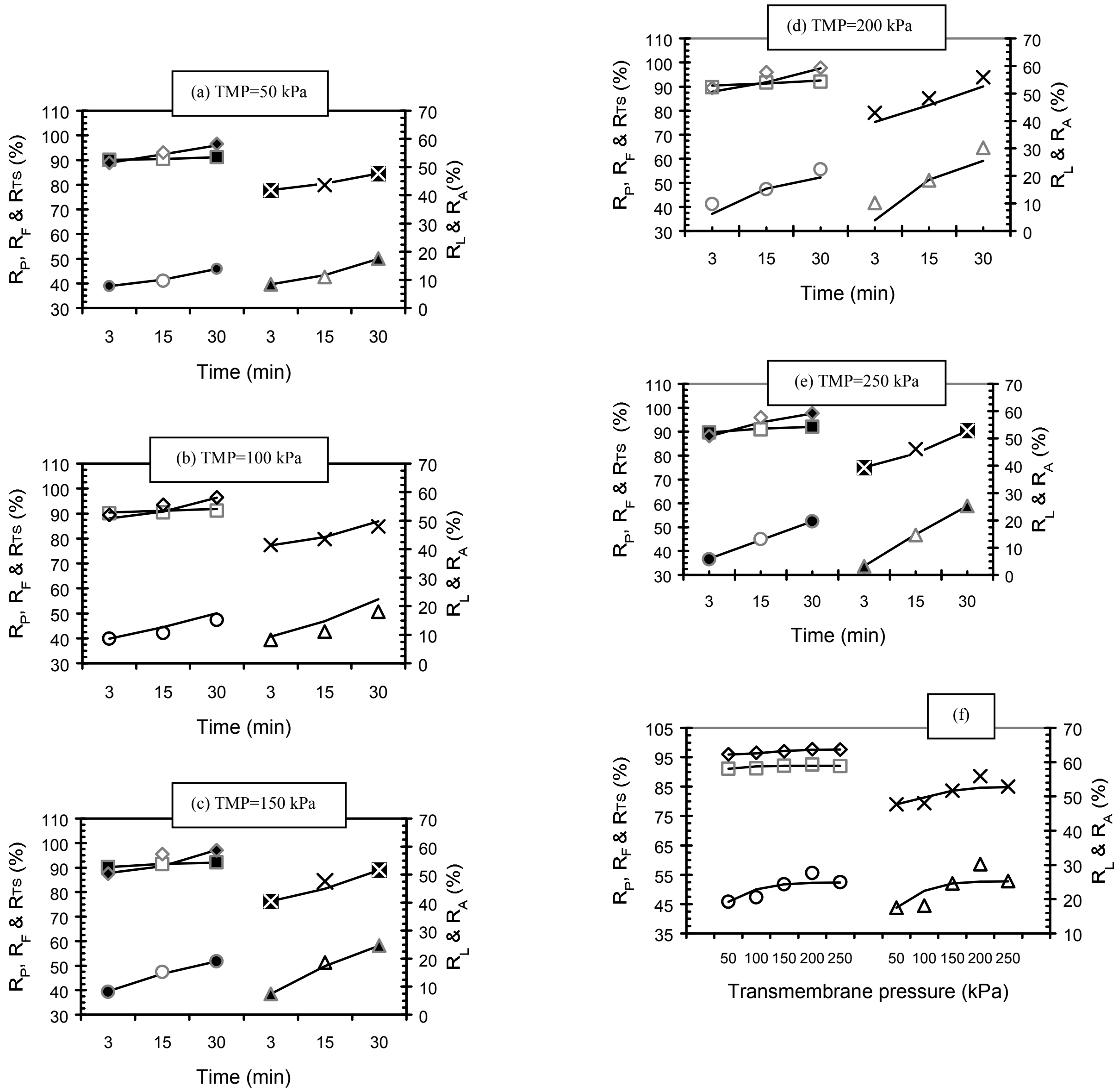

Fig 2. Dynamic predictions of rejection for each component [protein, fat, lactose, ash and total solids (T.S.)] during ultrafiltration of milk using training set (2). ANN used 2/7/5. Training points are as solids symbols Training points/ Validation points / Querying points = 30/15/30. (□, Protein; ◇, Fat; △, Lactose; ×, Ash; ○, Total solids; —, ANN).

From Table 2, it is found that for data set (2), the model with one hidden layer and seven hidden neurons produced the best model performance in terms of the E_{min} (0.07%), E_{ave} (0.54%) and E_{max} (1.0%), thus the amount of all errors is less than 1%. Fig. 2 clearly shows that the protein rejection (R_P) at each value of TMP almost constant with time, but the rejection of other components (such as R_F, R_L, R_A and R_{TS}) have increased significantly with time at each TMP. Thus, it can be concluded that flux decline and increasing of total resistance with time is probably due to decreasing transmission of small soluble compounds (such as lactose, salts) and fat through the membrane and then adsorption of them onto the membrane surface. Fig. 2(f) shows the effect of different TMP on components rejection at the end of each run (30 min.). It can be found that there is very good agreement between the actual values (dots) and the predicted values (solid lines) of components rejection. In addition, it shows that increasing TMP up to 250 kPa results in a small increase in the rejection of each component. Despite flux increasing with TMP (Fig. 1), the rejection of each component has not changed so much (Fig. 2f). This may be due to an increased rate of compounds migration toward membrane surface by convective transport (permeate flux), deposition of them onto membrane and the reduction of effective pore size [1, 8, 9, 10]. Thus it is seen that the fouling or total hydraulic resistance has been increased with the TMP (Fig.1).

5. Conclusions

The possibility of artificial neural network approach was investigated to model permeate flux, total hydraulic resistance and rejection of milk components as a function of transmembrane pressure and operating time. Due to the complexity of milk ultrafiltration prediction using conventional methods, these alternative models allow a unified approach that can be used for analysis of process and design of a new application.

The modelling results showed that the time profiles of the milk ultrafiltration performance can be predicted using single hidden layer network, small number of training data and sensible selection of training points with excellent accuracy, so that very small values of the relative errors obtained after training of data sets. Also, the validity of these models has been tested with very good results for unseen data related to intermediate pressures. As a result, it is unnecessary to carry out extensive pilot plant testing for collection of data and will be able to interpolate the process variables at other conditions of interest, with potentially great savings in time and cost. The experimental results also showed that the permeate flux at each TMP decreased significantly with time, while total resistance increased significantly with time at the corresponding TMP. As TMP increased, both flux and total resistance increased significantly. The rejection of each component increased greatly with time at each TMP, except for protein that was almost constant with time. Meanwhile the increasing of TMP led to a small increase of rejection of each component.

References

[1] M. Cheryan, 1998, *Ultrafiltration and microfiltration handbook*, Technomic Pub.Ltd.
[2] G. Samelsson, et al., 1997, J. Membr. Sci. 129, 277-281.
[3] M. Dornier, et al., 1995, J. Membr. Sci. 98, 263 273.
[4] H. Niemi, et al., 1995, J. Membr. Sci. 102, 185-191.
[5] W. R. Bowen, 1998b, Chem. Eng. Sci. 53, 3793-3802.
[6] N. Delgrange, et al., 1998, Desalination 118, 213-227.
[7] C. Teodosiu, et al., 2000, Wat. Res. 34, 4371-4380.
[8] R. H. G. Ramachandar, 2002, Desalination 144, 319-324.
[9] S. M. A. Razavi, et al., 2002, J. Agic. Sci. Tech., 8, 25-33.
[10] A. S. Grandison, et al., 2000, Lait 80, 165-174.
[11] T. E. Clarke and C. A. Heath, 1997, J. Food Eng. 33, 373-383.
[12] B. H. Chiang and M. Cheryan, 1987, J. Food Eng. 6, 241-255.