Note: Application of neural network modelling for the control of dewatering and impregnation soaking process (osmotic dehydration)

Ioan-Cristian Trelea, A.L. Raoult-Wack, Gilles Trystram

Food Science and Technology International (1997) **3**, 459–465

Note: Application of neural network modelling for the control of dewatering and impregnation soaking process (osmotic dehydration)

Nota: Aplicación del sistema de simulación de redes neurales para el control de la deshidratación osmótica

I.C. Trelea[1], A.L. Raoult-Wack[2]* and G. Trystram[1]

[1]*ENSIA-INRA, 91305 Massy, France*
[2]*CIRAD-SAR, BP 5035, 34032 Montpellier, France*

The aim of this work was to elaborate a predictive model of the mass transfer (water loss and solute gain) that occurs during dewatering and soaking by using neural network modelling. Two separate feedforward networks with one hidden layer were used (for water loss and solute gain respectively). Model validation was carried out on results obtained previously, which dealt with agar gel soaked in sucrose solution over a wide experimental range (temperature, 30–70 °C; solution concentration, 30–70 g sucrose/100 g solution; time 0–500 min; agar concentration, 2–8%). The best results were obtained with three hidden neurons, which made it possible to predict mass transfer, with an accuracy at least as good as the experimental error, over the whole experimental range. The technological interest of such a model is related to a rapidity in simulation comparable to that of a traditional transfer function, a limited number of parameters and experimental data, and the fact that no preliminary assumption on the underlying mechanisms was needed.

Keyword: modelling, neural network, soaking, osmotic dehydration

En este trabajo se ha elaborado un modelo de simulación de transferencia de masa (pérdida de agua y ganancia de sólidos) durante la deshidratación osmótica mediante un sistema de red neural (*neural network*). Se emplearon dos redes por separado,una para la pérdida de agua y otra para la ganancia de sólidos. La validez del modelo se comprobó con datos publicados previamente de gel de agar sumergido en una disolución de sacarosa con un amplio intervalo de condiciones experimentales (temperatura de 30 a 70 °C, concentración de sacarosa de 30 a 70 g/100, tiempo de 0 a 500 min y contenido de agar en el gel de 2 a 8%). Los mejores resultados se obtuvieron con 3 neuronas ocultas (*hidden neurons*), con lo que se consiguió predecir una transferencia de masa con una precisión al menos tan buena como el error experimental para todo el intervalo de condiciones estudiadas. El interés tecnológico del modelo está vinculado a la rapidez en la simulación, comparable a una función de transferencia tradicional, limitado número de parámetros y datos experimentales y del hecho de que no se necesitan consideraciones preliminares sobre los mecanismos propuestos.

Palabras clave: simulación, *neural network*, deshidratación osmótica

INTRODUCTION

The dewatering and impregnation soaking process (DIS process, osmotic dehydration) is a technological pre-treatment, which consists of soaking water-rich

*To whom correspondence should be sent.
Received 18 July 1996; revised 13 January 1997.

solid food materials in concentrated aqueous solutions (mainly of sugars or salts). This operation leads to simultaneous dewatering and direct formulation of the food material (through impregnation and leaching), and can be introduced as a pre-treatment in any conventional processing chain to improve quality and save energy. Recent advances in the field were recently reviewed by Le Maguer (1988), Torreggiani (1993), and Raoult-Wack (1994), and existing models do not permit adequate control of the DIS process in industrial applications, mainly because they cannot take into account the process complexity (Raoult-Wack, 1994).

Neural networks are recognized as good tools for dynamic modelling, and have been extensively studied since the publication of the perceptron identification method (Rumelhart and Zipner, 1985). The interest of such models includes modelling without any assumptions about the nature of underlying mechanisms, and their ability to take into account non-linearities and interactions between variables (Bishop, 1994). Recent results establish that it is always possible to identify a neural model based on the perceptron structure, with only one hidden layer, for either steady-state or dynamic operations (with recurrent models) (Hornik *et al.*, 1989; Hornik 1993). An outstanding feature of neural networks is the ability to learn the solution of the problem from a set of examples and to provide a smooth and reasonable interpolation for new data. In the field of food process engineering, it is a good alternative to conventional empirical modelling based on polynomial and linear regressions. For food processes, the interest in application of neural computing keeps on growing. Applications concern fermentation (Latrille, 1994), extrusion (Linko *et al.*, 1982), filtration (Dornier *et al.*, 1995), and drying (Huang and Mujumdar, 1993).

The aim of the present work was to test the interest and efficiency of neural networks to model and predict the behaviour of food material subjected to a DIS process.

METHODS

Model

Network structure

An artificial neural network is an association of elementary cells or 'neurons' grouped into distinct layers and interconnected according to a given architecture. The standard network structure for function approximation is the multilayer perceptron (or feedforward network) that we use in this work. Figure 1 shows the network for the solute gain model; the water loss model is similar. The input layer consists of $m = 4$ identity fan-out units, one for each input, with no coefficient associated with them. The neurons in the single hidden layer compute a weighted sum of their inputs, and apply a squashing transfer function to the result. The number $r = 3$ units was selected after some tests. The coefficients associated with the hidden layer are grouped into the matrices A1 (weights) and B1 (biases). The output layer contains $p = 1$ neuron in our case (because there is only one output), which computes the weighted sum of the signals provided by the hidden layer. The associated coefficients are grouped into matrices A2 and B2. The matrices have the following dimensions: $A1 \in \Re^{r\times m}$, $B1 \in \Re^{r\times 1}$, $A2 \in \Re^{p\times r}$, $B2 \in \Re^{p\times 1}$,and the total number of network coefficients is:

$$n = r\cdot(m+1) + p\cdot(r+1) = 19 \quad (1)$$

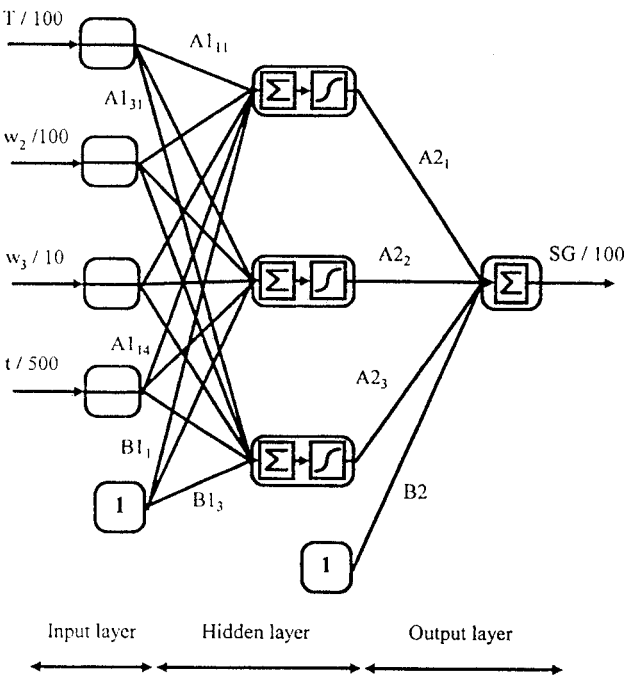

Figure 1. Schematic representation of the single hidden layer, feedforward neural network architecture used for the solute gain model. The network used for the water loss model is similar.

Figura 1. Representación esquemática de una capa oculta sencilla de la arquitectura de red neural utilizada para el modelo de ganancia de solutos. La red para la pérdida de agua es similar.

Using the matrix notation, the network output can be computed from Equation (2) :

$$SG = 100 \cdot [A2 \cdot \tanh(A1 \cdot u + B1) + B2] \quad (2)$$

For further details, the reader is referred to standard textbooks on neural networks (Freeman and Skapura, 1992). Taking into account the normalization of the inputs and outputs, as well as the fact that hyperbolic tangent calculation (if suitably arranged) requires four basic floating point operations plus an exponential function calculation, the number N_f of the elementary floating point operation needed for a neural model simulation is $N_f = 2r \cdot m + 2p \cdot r + m + p + 4r = 47$, and the number of exponential function calculation is $N_e = r = 4$. These numbers are very small compared with those needed for physical models. The simulation of a model based on differential equations, for example, typically requires many thousands of floating point operations. On a standard Pentium-based PC computer at 90 MHz, a basic floating point operation takes 0.11 μs, and an exponential calculation 2.2 μs (figures obtained in Matlab, after several tests on large matrices). A neural model simulation thus takes less than 15 μs, which is comparable to traditional linear transfer function and state space models, with the additional advantage of taking into account non-linear phenomena.

Identification: learning algorithm and model quality assessment

The learning process consists of adjusting the network coefficients (the weight A*i* and B*i*) in order to minimize an error function (usually a quadratic one) between the network outputs for a given set of inputs and the known correct outputs. If smooth non-linearities are used, the gradient of the error function can be easily computed by the classical backpropagation procedure (Rumelhart *et al.*, 1986). Early learning algorithms used this gradient directly in a steepest descent optimization, but recent results and our own experience showed that second order methods are far more effective. In this work, the Levenberg–Marquardt optimization procedure in the Neural Network Toolbox (Demuth and Beale, 1993) was used. Despite the fact that the computations involved in each iteration are more complex than in the steepest descent case, the convergence is faster, typically by a factor of 100. The learning process is much more computationally intensive than the simulation. It typically takes 5–10 min on the PC computer mentioned above. The root mean square error (RMSE) between the experimental values and network predictions was used as a criterion of model adequacy.

DATABASE PREPARATION

In this work, experimental data provided by Raoult-Wack *et al.* (1991) consisted of mass transfer kinetics (water loss and solute gain) measured on agar gel cubes (initial side dimension 9×10^{-3} m) soaked in aqueous sucrose solutions, under constant external conditions, over 480 min. Agar gel consisted of agar, sucrose and water. The concentration of the occluded solution within the agar network was fixed at 10 g sucrose/100 g solution, which is a representative level of sugar concentration in most fruits and vegetables. Mass transfer measurements were water loss and solute gain, denoted WL and SG respectively, expressed in g/100 g initial gel.

Four process factors were varied: initial agar weight fraction in the gel (denoted w_3, in %), processing temperature (denoted T, in °C), sucrose weight fraction in the soaking solution (denoted w_2, in %), and processing time duration (t, in min). Table 1 gives experimental conditions for each of the 17 situations studied. Corresponding files are denoted F1–F17. For each situation, WL and SG measurements were performed at time $t = 5$, 10, 20, 30, 45, 60, 90, 120, 150, 180, 240, 300, 360, 420, 480 min, resulting in 34 experimental kinetics. Figure 2 presents the distribution of various experimental files in a (T, w_2, w_3) space.

Table 1. Experimental conditions studied

Tabla 1. Condiciones experimentales estudiadas.

File number	T (°C)	w_1 (%)	w_3 (%)	t (min)
F1	50	60	2	0–480
F2	50	60	3	0–480
F3	50	60	3.5	0–480
F4	50	60	5	0–480
F5	50	60	6	0–480
F6	50	60	7	0–480
F7	50	60	8	0–480
F8	50	20	4	0–480
F9	50	30	4	0–480
F10	50	40	4	0–480
F11	50	50	4	0–480
F12	50	70	4	0–480
F13	30	60	4	0–480
F14	40	60	4	0–480
F15	50	60	4	0–480
F16	60	60	4	0–480
F17	70	60	4	0–480

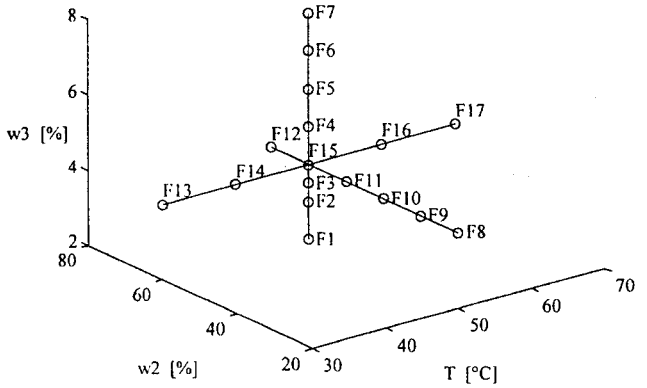

Figure 2. Distribution of the experimental files in a (T, w_2, w_3) space.

Figura 2. Distribución de los ficheros experimentales en un espacio (T, w_2, w_3).

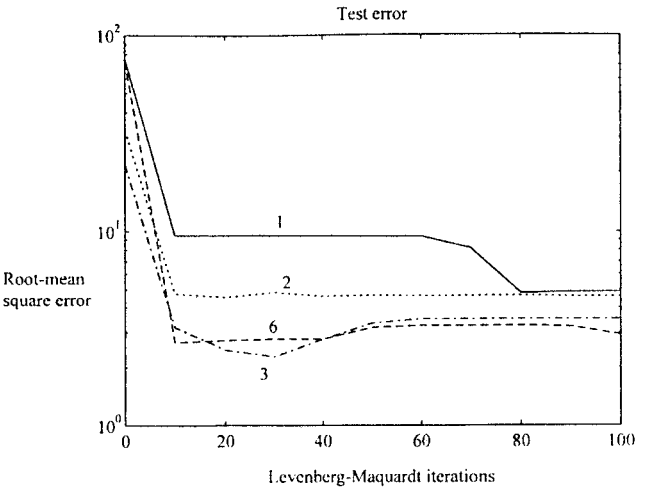

Figure 3. Test root mean square error (RMSE) versus iteration number, in the case of solute gain, for various numbers of hidden neurones.

Figura 3. Test de la raíz del error de mínimos cuadrados (RMSE) frente al número de iteración, en el caso de ganancia de solutos, para varios números de neuronas ocultas.

The selected experimental range is wide, and covers two types of situation, as defined by Raoult-Wack *et al.* (1991): so-called 'dewatering situations' when water loss is higher than sugar gain, and 'impregnation situations' in which sugar gain is higher than water loss. Experimental files were split into learning and test bases to obtain a good representation of the situation diversity ('dewatering' and 'impregnation' situations) in each base: 'learning base', F1, F2, F5, F7, F8, F10, F11, F12, F13, F15, F17; 'test base', F3, F4, F6, F9, F14, F16.

The numerical behaviour of the learning routine is significantly improved if all network inputs and outputs are normalized to similar ranges. Hence T, w_2, WL and SG were divided by 100, w_3 by 10 and t by 500. Hence in Equation (2) u was written as given by Equation (3):

$$u = \begin{cases} T/100 \\ w_2/100 \\ w_3/10 \\ t/500 \end{cases} \quad (3)$$

RESULTS AND DISCUSSION

Our model involved 19 weights for 11 kinetics (as calculated from Equation 1), and hence 38 weights for 22 kinetics because there were two separate networks.

Figure 3 gives the test RSME respectively, against iteration number in the case of SG, for one to six neurons in the hidden layer. Results showed that the learning error typically decreased when the number of neurons in the hidden layer increased, but for more than six neurons in the hidden layer, an additional increase in structure complexity did not decrease the learning RSME further. Figure 3 showed a slight increase in test error, accounting for over-fitting, for more than 30 iterations and three neurons in the hidden layer. Let us point out that the increase in test error may be enhanced in other cases. As for water loss, similar behaviour to that previously described for sugar gain was obtained, hence the results are not presented here.

Tables 2 and 3 give the final choice for the number of neurons in the hidden layer and iterations, as well as weight values and learning and test RSME, for WL and SG respectively. As a whole, the learning and test RSME were similar, which accounts for a good generalization capability of the neural network. The learning and test RSME are close (even inferior) to experimental error of the original database (Raoult-Wack *et al.*, 1991) and also to that encountered in most studies dealing with the behaviour of real raw materials subjected to DIS processes (Saurel *et al.*, 1994).

When simulated against experimental data for test and learning bases, WL and SG showed that in all cases the prediction was correct whatever the level of water loss and sugar gain (Figures 4 and 5).

Table 2. Characteristics of the best neural network in the case of water loss.

Tabla 2. Características de la mejor red neural para la pérdida de agua.

3 hidden neurons, 100 iterations			
RSME learn = 2.83 RSME test = 2.36			
A1 =			
−0.8790	−0.4366	−0.0908	7.5406
−0.4725	−4.0081	−0.6674	0.1077
−0.6339	−0.5002	−0.0941	6.7092
B1 =			
1.1828			
2.5521			
1.6694			
A2 =			
−6.2804	−0.3914	17.8413	
B2 =			
−11.1687			

Table 3. Characteristics of the best neural network in the case of solute gain.

Tabla 3. Características de la mejor red neural para la ganancia de solutos.

3 hidden neurons, 30 iterations			
RSME learn = 2.74 RSME test = 2.60			
A1 =			
−4.3709	4.6844	0.5906	−0.1983
−0.4234	0.4326	−1.4352	−7.6583
−4.5282	5.2997	2.0249	−0.4885
B1 =			
0.0000			
−1.0218			
−0.9562			
A2 =			
0.7720	−2.0429	−0.6476	
B2 =			
−1.9532			

Figure 6 gives simulated and experimental WL and SG kinetics obtained in situation F4 (test base). According to Raoult-Wack *et al.* (1991), general time-course evolution could be divided into two phases, which is typical behaviour in the DIS process, whatever the product nature or experimental conditions. In the first phase (0–180 min), WL and SG were very rapid. Then fluxes progressively decreased, and WL stagnated whereas SG continued to increase slightly (phase 2). Simulated results showed that the neural network model provided accurate simulation of WL and SG evolution in both phases.

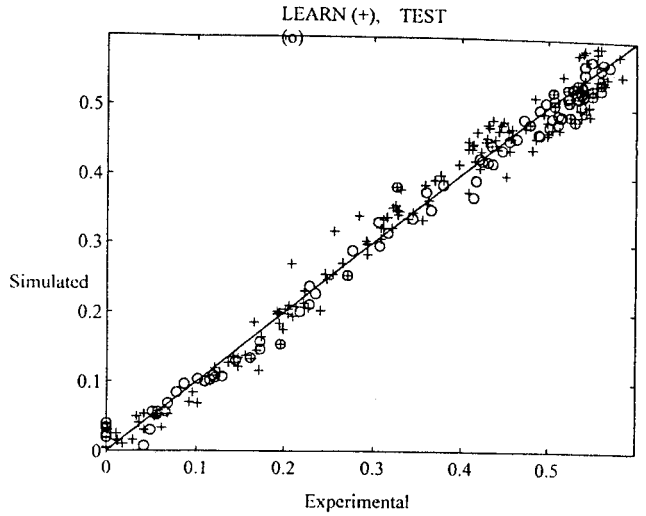

Figure 4. Simulated versus experimental data for test and learning bases, in the case of water loss.

Figura 4. Datos experimentales frente a los simulados en el caso de la pérdida de agua.

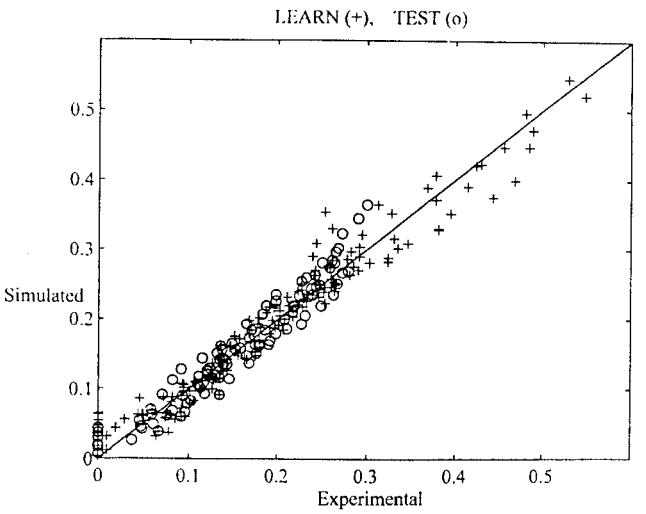

Figure 5. Simulated versus experimental data for test and learning bases, in the case of solute gain.

Figura 5. Datos experimentales frente a los simulados para la ganancia de solutos.

Figure 7 gives simultaneous simulated and experimental kinetics obtained in situation F9 (test base). In contrast to situation F4, which is a dewatering situation, F9 corresponds to an impregnation situation (SG higher than WL). Results showed that the model satisfactorily described the kinetics in each case.

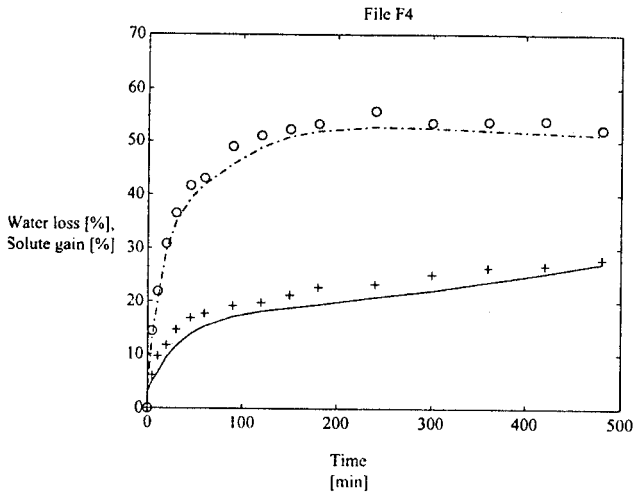

Figure 6. Simulated and experimental kinetics for the water loss (o) and solute gain (+) in the dewatering situation F4.

Figura 6. Cinéticas experimental y simulada para la pérdida de agua (o) y la ganancia de solutos (+) en el caso de la deshidratación osmótica F4.

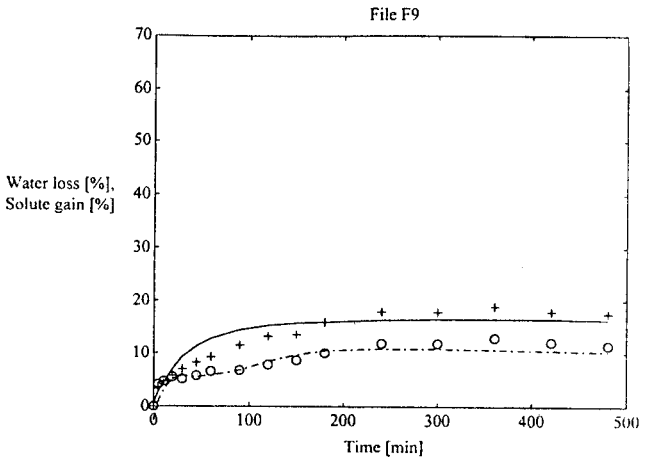

Figure 7. Simulated and experimental kinetics for the water loss (o) and solute gain (+) in the impregnation situation F9.

Figura 7. Cinéticas experimental y simulada para la pérdida de agua (o) y la ganancia de solutos (+) en el caso de la impregnación F9.

CONCLUSION

This study shows that neural network modelling makes it possible to obtain accurate simulation of water loss and sugar gain kinetics during the DIS process, over a wide experimental range, for either dewatering or impregnation situations, with a limited number of parameters and experimental data. The technological interest of this kind of modelling has to be related to the fact that it is elaborated without any preliminary assumptions on the underlying mechanisms, and also to its implementation facilities and speed of simulation (of the order of 15 μs on a standard PC computer). This neural network modelling was validated on results obtained with model foods. Further modelling should be carried out on real raw materials, and for variable operating conditions. Other architectures (in particular recurrent ones) should also be studied. Generally, the neural model should provide a promising tool for process control in DIS processes. It could also be used for data filtering, which could be of greatest interest in phenomenological modelling of the process, or typology of the behaviour of materials subjected to soaking from available experimental data, which are claimed to be the main shortcomings of research in this field to date.

REFERENCES

Bishop C.M. (1994). Neural networks and their applications. *Reviews on Scientific Instrumentation* **65** (6): 1803–1832.

Dornier M., Rocha Meir T., Bardot I., Decloux M. and Trystram G. (1993). Applications of neural computation for dynamic modelling of food processes; drying and microfiltration. In: *Artificial Intelligence for Agriculture and food*. Paris:EC2 ed. pp. 223–240.

Demuth H. and Beale M.(1993). *Neural Network Toolbox for Matlab – User's Guide*. Natrick, MA: The MathWorks Inc.

Hornik K. (1993). Some new results on Neural Network Approximation. *Neural Network* **6**: 1069–1072.

Hornik K., Stinchcombe M. and White H. (1989). Multilayer feedforward networks are universal approximators. *Neural Network* **2**: 359–366.

Huang B. and Mujumdar A. (1993). Use of neural networks to predict industrial dryers performances. *Drying Technology* **11** : 525–541.

Freeman J.A. and Skapura D.M. (1992). *Neural Networks*. New York: Addison Wesley.

Latrille E., Corrieu G. and Thibault J. (1993). pH prediction and final fermentation time determination in lactic acid batch fermentations. *Escape 2, Computer Chemical Engineering* **17**: 423–428.

Le Maguer M. (1988). Osmotic dehydration: review and future directions. In: *Proceedings of Symposium on Progress in Food Preservation Processes*. Vol 1. Brussels. pp. 283–309.

Linko P. and Zhu Y.H. (1992). Neural networks for real time variable estimation and prediction in the control of glucoamylase fermentation, *Progress in Biochemistry* **27**: 275–283.

Narendra K.S. and Parthasaraty K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks* **1** (1): 4–27.

Raoult-Wack A.L. (1994). Recent advances in the osmotic dehydration of foods. *Trends in Food Science and Technology* **5** (8): 255–260.

Raoult-Wack A.L., Guilbert S., Le Maguer M. and Rios G. (1991). Simultaneous water and solute transport in shrinking media. Part 1: application to dewatering and impregnation soaking analysis (osmotic dehydration). *Drying Technology* **9** (3): 589–612.

Rosenblatt F. (1958). The perceptron: a probabilistic model for information storage and organisation in the brain. *Psychology Review* **65**: 386–408.

Rumelhart D. and Zipner D. (1985). Feature discovering by competitive learning. *Cognitive Science* **9**: 75–112.

Saurel R., Raoult-Wack A.L., Guilbert S. and Rios G. (1994). Mass transfer during osmotic dehydration of apples. part 1 : fresh plant tissue. *International Journal of Food Science and Technology* **29**: 531–542.

Torreggianni D. (1993). Osmotic dehydration in fruit and vegetable processing. *Food Research International* **26**: 59–68.