Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
International Communications in Heat and Mass Transfer xx (2005) xxx–xxx
International Communications in HEAT and MASS TRANSFER
www.elsevier.com/locate/ichmt

A novel predictive architecture for microwave-assisted drying processes based on neural networks ☆

J.L. Pedreño-Molina [a,*], J. Monzó-Cabrera [a], A. Toledo-Moreo [b], D. Sánchez-Hernández [a]

[a] *Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Cartagena, Campus Muralla del Mar, E-30202 Cartagena, Spain*
[b] *Departamento de Tecnología Electrónica, Universidad Politécnica de Cartagena, Campus Muralla del Mar, E-30202 Cartagena, Spain*

Abstract

In this contribution, a novel learning architecture based on the interconnection of two different learning-based neural networks has been used to both predict temperature and drying curves and solve inverse modelling equations in microwave-assisted drying processes. In this way, a neural model that combines the accuracy of neural networks based on Radial Basis Functions (RBF) and the algebraic capabilities of the matrix polynomial structures is presented and validated. The architecture has been trained by temperature ($T_c(t)$) and moisture content ($X_t(t)$) curves, which have been generated by a previously validated drying model. The results show that the neural model is able to very accurately predict both kind of curves for any combination of the considered input variables (electric field and air temperature) provided that an appropriate training process is performed. The proposed configuration also permits the solution of the inverse problem in the drying process by finding the optimal value for the electric field. This provides $T_c(t)$ or $X_t(t)$ curves that fit to a desired drying condition in a specific time slot.

Keywords: Predictive system; Neural network modelling; Microwave-heating applications

☆ Communicated by J.W. Rose and A. Briggs.
* Corresponding author.
E-mail address: Juan.Pmolina@upct.es (J.L. Pedreño-Molina).

0735-1933/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.icheatmasstransfer.2005.05.001

1. Introduction

The conventional analysis of heat and mass transfer processes is mainly based on the physical modelling of these phenomena, by finding and solving the differential equations that describe them. In this way, many authors have published very different mathematical models which produce good results under several constraints as described in [1]. Yet, the key shortcomings for this kind of models are the need of an accurate characterization of the model coefficients, particularly temperature and/or moisture content dependency [1,2]. In contrast, the neural-network-based architectures give solutions to the appearance of equations restraints when the operating environment is unknown. Since this kind of structures is learning based, the rest of the process or environment parameters are no longer necessary when the mapping between the input variables of the process and the output functions is known. Although the employment of neural networks allows for the determination of parameter optimization problems [4], they are commonplace for the identification and prediction of dynamic processes. Thus, a fuzzy neural algorithm is employed in [5] to control wood drying processes and a back-propagation neural structure is developed for the precise modelling of rice drying in [6]. On the other hand, most neural network structures applied to drying processes involve solutions in which the output of the network is a unique solution [7] for an input set of one-dimensional parameters, rather than a time-dependent function.

In this work, however, the proposed model is able to learn and predict complete drying curves, such as sample temperature $T_c(t)$ or moisture content $X_t(t)$ evolution, from only two numerical input parameters. The proposed neural structure, based on Radial Basis Functions (RBF) neural networks and polynomial learning structures, has been trained from several temperature and drying curves which have been obtained from a previously validated microwave-assisted drying model [2,3].

2. Theoretical study

2.1. Neural architecture design

From previous contributions concerning microwave-assisted drying processes [2,3,8], one can conclude that $X_t(t)$ varies smoothly along the different drying stages while $T_c(t)$ shows steeper changes during dehydration. Additionally, both curves are highly dependent on the electric field strength (E) and the air-flow temperature (T_{air}). It is for this reason that both E and T_{air} have been chosen as the input parameters for the neural network model.

In order to design a neural architecture which is able to predict both $X_t(t)$ or $T_c(t)$ from E and T_{air}, it is necessary to consider neural network models based on learning algorithms mainly focused on the solution of non-linear problems. In this case, the same neural architecture will be used to learn and predict $X_t(t)$ or $T_c(t)$ regardless of their time evolution characteristics, as illustrated in Fig. 1. Two different levels can be distinguished in this architecture, where w_{jk} are the neuron weights of the first network for the kth trial, $\mathbf{WW}$ is the matrix associated to the second network and $y(t)$ is a general purpose time-dependent curve that represents either $X_t(t)$ or $T_c(t)$. In the first level of the proposed architecture, the output function length, $t=[0, T-1]$, is divided into M time intervals in order to project the T points of $y(t)$ onto M neurons ($M<T$) whose weights have been properly learned. If we assume that Z_i ($i=[1, 2]$) are the input parameters of the drying process (E and T_{air} respectively), the relationship

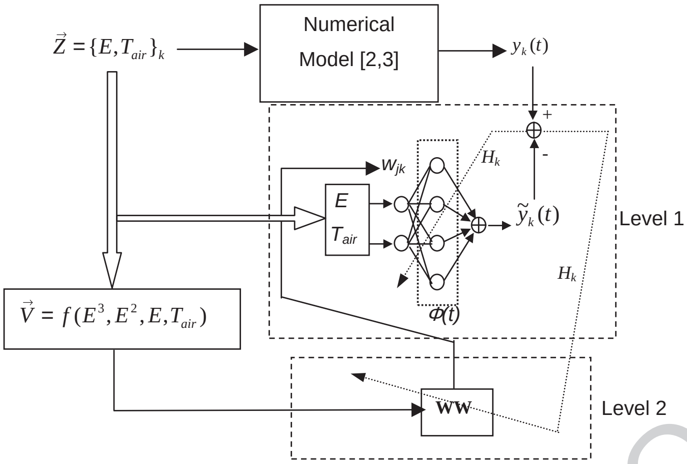

Fig. 1. General scheme of the proposed neural model.

between Z_i and $y(t)$ within each time interval will be learned by the RBF neural network during the training stage and then, it is later assigned to a neuron weight w_j ($j=[1, M]$), where M is the time interval number. The second architecture level projects the M weights of the RBF neurons onto E and T_{air}. This projection is carried out by other network whose structure depends basically on the relationship between each w_k and Z_i. In this case, a third degree polynomial model has been employed, which provides sufficient final precision for the $M \rightarrow 2$ projection in each interval as discussed further on. This second level permits the generation of appropriate RBF neuron weights from whatever value for E and T_{air}.

2.2. RBF neural networks for drying processes identification

RBFs are supervised neural networks [9] whose structure provides a solution to the local interpolation of non-linear functions. This is the case of the generic function $y(t)$ considered in this work since $X_t(t)$ or $T_c(t)$ do not have a linear behaviour. Although the neurons activation in the RBF model is carried out by radial functions, the RBF model has a linear expression for the estimation ($\tilde{y}(t)$) of $y(t)$. Therefore, for each pair of input variables $[E, T_{\text{air}}]_k$ corresponding to the kth trial, the estimation of $\tilde{y}_k(t)$ is given by Eqs. (1) and (2),

$$\tilde{y}_k(t) = \sum_{j=1}^{M} w_{jk} \phi_j(t) \tag{1}$$

$$\phi_j(t) = \exp\left(\frac{-(t-\mu_j)^2}{\sigma_j^2} \right) \tag{2}$$

where $\phi_j(t)$ is the jth Gaussian radial function, μ_j and σ_j are the centre and standard deviation of $\phi_j(t)$, w_{jk} is the weight value associated to $\phi_j(t)$ for the kth trial ($k=[1, N]$), and N is the number of trials during the learning period.

The estimation of w_{jk} is carried out by using the gradient descent algorithm [10] to minimize the cost function described in Eq. (3)

$$H_k = \sum_{t=0}^{T} (y_k(t) - \tilde{y}_k(t))^2 \tag{3}$$

As a conclusion, the application of the RBF neural network to the estimation of $\tilde{y}_k(t)$ permits to obtain, after the training stage, the optimal values for w_{jk} from curves generated by the validated drying model [2,3]. In the prediction stage, Eq. (1) supplies the approximation of $y_k(t)$ for $T_c(t)$ or $X_t(t)$. Transforming Eq. (1) into a matrix notation,

$$\tilde{y}_k(t) = \mathbf{W_k} \times \boldsymbol{\Phi}(\mathbf{t})^T \tag{4}$$

with $\mathbf{W_k}$ being the $1 \times M$ dimension vector containing the RBF neuron weights for the kth trial, and $\boldsymbol{\Phi}(\mathbf{t})^{\mathbf{T}}$ the vector whose elements are the M Gaussian functions, which are independent of k, as described in Eq. (2).

2.3. Interpolation network for curves prediction

In order to obtain information about the dependence order of each RBF neuron with respect to the input variables of the process and to design the second level network, the $\mathbf{W_k}$ obtained in each trial have been projected onto $[E, T_{air}]_k$, which are the input variables used to generate $y_k(t)$. Taking into account all the trials used in the learning stage, this projection results in the **WW** matrix. This matrix contains the mapping between E and T_{air} and the contribution of each Gaussian function to the generation of $\tilde{y}(t)$. In order to establish the dependence of each neuron versus E and T_{air} variations, a two-dimensional projection of all the M neuron weights has been generated. Fig. 2 shows, for both $X_t(t)$ and $T_c(t)$, the projection for the fourth neuron of the weight matrix. The rest of the $M-1$ neuron weights present very similar behaviours with respect to E and T_{air}. From the surface analyses in Fig. 2, one can conclude that there is a third-order dependence of the weights in the RBF network versus E in both $X_t(t)$ and $T_c(t)$

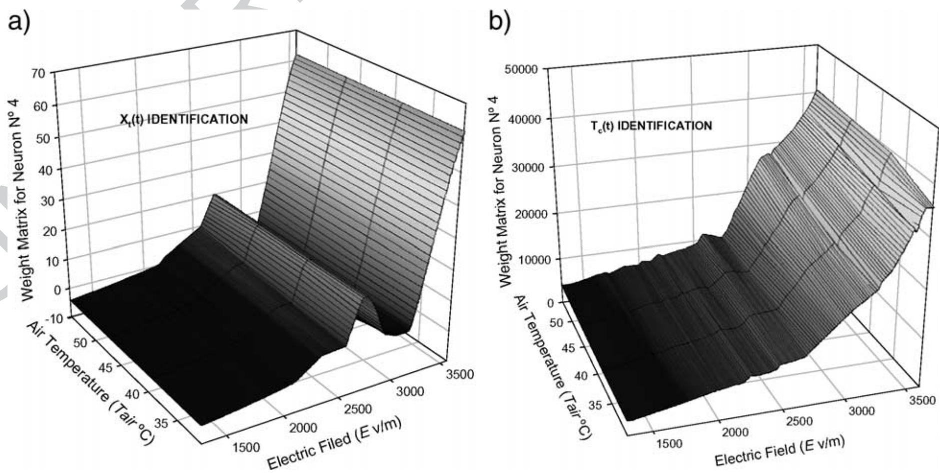

Fig. 2. 3D projection of the 4th RBF weight over E and T_{air} for estimated (a) $X_t(t)$ and (b) $T_c(t)$.

curves. At the same time, these surfaces also show a linear dependence with respect to T_{air}, which justifies the selection of the polynomial structure of the second architecture level.

By following the scheme in Fig. 1, applying the polynomial model to the RBF neuron weights and considering the matrix expression in Eq. (4), $\tilde{y}(t)$ can be generated, for each pair of inputs [E, T_{air}], by means of

$$y(t,T_{\mathrm{air}},E)=w_j\cdot\Phi(t,\mu_i,\sigma)^T=\vec{V}(E,T_{\mathrm{air}})\cdot\mathbf{WW}\cdot\Phi(t,\mu_i,\sigma)^T \tag{5}$$

where the $\vec{V}$ components are represented in Fig. 1, and the matrix **WW** is learned by following the linear regression algorithm for the **R** and **S** matrixes shown in Eqs. (6) and (7).

$$\mathbf{S}=\begin{bmatrix} T_{\mathrm{air}1} & E_1 \\ T_{\mathrm{air}2} & E_2 \\ \vdots & \vdots \\ T_{\mathrm{air}N} & E_N \end{bmatrix} \qquad \mathbf{R}=\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1M} \\ w_{21} & w_{22} & \cdots & w_{2M} \\ \vdots & \vdots & & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NM} \end{bmatrix} \tag{6}$$

$$\mathbf{WW}=(\mathbf{S}^T\cdot\mathbf{S})^{-1}\cdot\mathbf{S}^T\cdot\mathbf{R} \tag{7}$$

2.4. Inverse calculation for electric field intensity estimation

One of the main advantages of using the proposed neural configuration, based on interconnected RBF neural networks and polynomial matrix equations, is the possibility to readily solve the inverse problem. In this case, this implies the estimation of the optimal value of one of the input variables of the process from desired output results in terms of temperature or moisture content. This is particularly important for microwave-assisted drying processes since it provides the initial drying configuration that ensures a final drying level to be reached in a previously established time. In this work, we have fixed T_{air} and a target value (K) for the sample temperature (T_{c0}) or the moisture content (X_{t0}), which has to be reached within t_0 seconds. With these conditions, the proposed model is able to estimate the optimal value for E that generates the $T_c(t)$ or $X_t(t)$ curves that fit to the desired points [T_{c0}; t_0] or [X_{t0}; t_0]. By particularizing expression (5) for a specific t_0 and T_{air}, Eq. (8) is obtained.

$$y(t_0,E)=\vec{V}(E)\cdot\mathbf{WW}\cdot\Phi(t_0)^T=K \tag{8}$$

Solving Eq. (8) for the variable E, a solution for the inverse problem can be reached provided that the desired target point [K, t_0] belongs to the learned range for E and T_{air}. Thus, the solution for the optimal electric field, considering the form of the vector $\vec{V}$ from Fig. 1, will be obtained by solving the third degree equation given by,

$$aE^3+bE^2+cE+d=K \tag{9}$$

$$\left.\begin{aligned} a&=A_6(t_0)\\ b&=A_3(t_0)+T_{\mathrm{air}}A_5(t_0)\\ c&=A_1(t_0)+T_{\mathrm{air}}A_4(t_0)\\ d&=A_7(t_0)+T_{\mathrm{air}}A_2(t_0)\end{aligned}\right\} \tag{10}$$

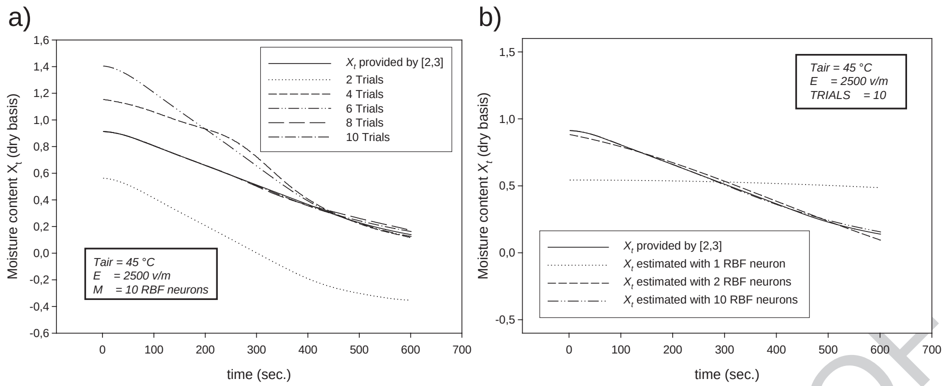

Fig. 3. Predictive behaviour of the proposed model for estimation of $X_t(t)$ from E and T_{air} inputs.

where the coefficients $A_k(t_0)$ are calculated as:

$$A_k(t_0) = \sum_{j=1}^{M} \mathbf{WW}_{kj}\mathbf{dU_k}(\mathbf{t_0}) \tag{11}$$

3. Results and discussion

The diagram in Fig. 1 and Eqs. (1)–(11) have been used both for curves' learning and prediction and to solve the inverse problem in microwave-assisted drying processes. The temperature and drying curves for the learning and validation procedures have been provided by the drying model described in [2,3]. During the learning stage, random values for the input parameters E and T_{air} have been generated and different curves were obtained. A moisture content of 0.912 (dry basis) and sample temperature of 26 °C have been used as initial boundary conditions for all the simulations. For other initial simulation parameters, the readers should refer to [2,3]. The time interval for training and testing the neural model was set to $T = 600$ s. For all the Gaussian functions of the RBF neural network $\mu_i = T \cdot i/M$ ($i \in [1, M]$) and $\sigma = 3T/M$. The learning intervals for each input parameter (E and T_{air}) are [1246.6; 3630.8] and [30.30; 69.84], respectively. Moreover, it must be pointed out that the proposed neural model has been trained separately for the identification of $X_t(t)$ or $T_c(t)$ curves.

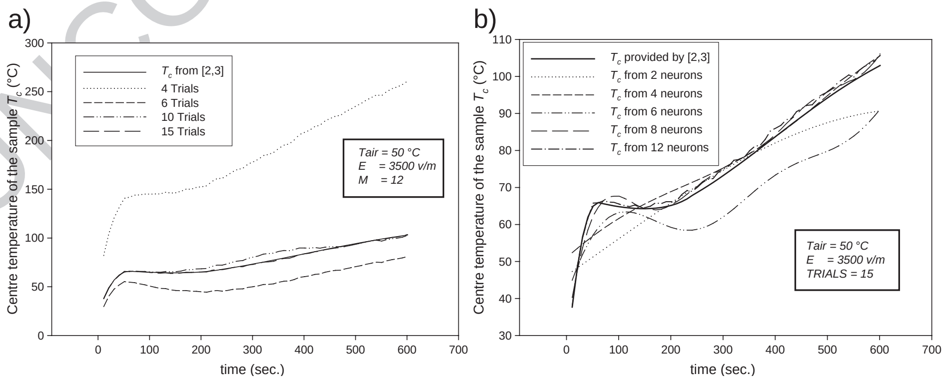

Fig. 4. Predictive behaviour of the proposed model for estimation of $T_c(t)$ from inputs E and T_{air}.

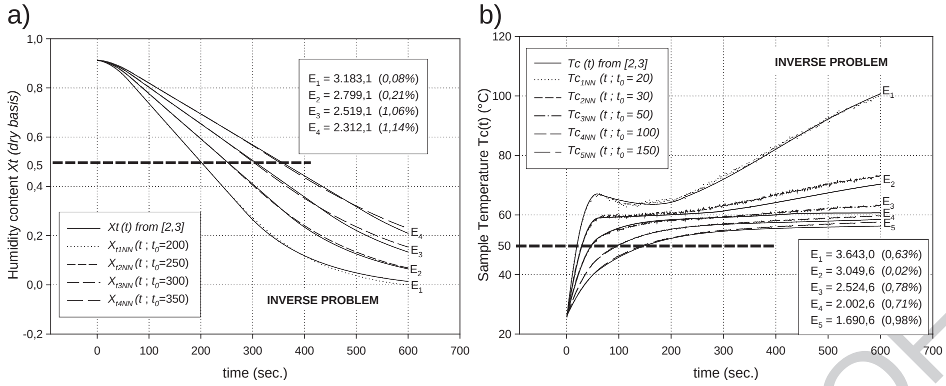

Fig. 5. Proposed model applications for estimation of (a) $X_t(t)$ and (b) $T_c(t)$, fitting $[T_{c0}, t_0]$ and $[X_{t0}, t_0]$.

The model performance for the prediction of $X_t(t)$ from the input parameters (E and T_{air}) has been evaluated as a function of the number of trials and RBF neurons. For this purpose, different neuron weight matrixes and training curves have been autonomously generated and learned. Fig. 3a shows the neural model prediction of $X_t(t)$ as a function of the number of trials while Fig. 3b illustrates the neural model estimation of $X_t(t)$ versus the neuron weights' number. From this figure, one can conclude that the higher the trials and neurons number, the more accurate performance for the neural model, which is normally expected. Yet, Fig. 3 also shows that 10 trials for the neural network training and 10 neuron weights are sufficient to precisely predict the temporal evolution of $X_t(t)$, and that the model is much more sensitive to trial amount than to neurons number for smooth curves such as $X_t(t)$.

The same assessment has been carried out for to the other involved curve: $T_c(t)$. Fig. 4a shows the behaviour of the temperature neural model prediction versus the number of training trials for the RBF neural network, while Fig. 4b represents the prediction convergence of this model versus the number of RBF neuron weights. From Fig. 4, it can be concluded, unlike that for $X_t(t)$, that 15 trials during the learning stage and 12 RBF neurons are sufficient to obtain a good prediction of the temporal evolution for $T_c(t)$ in the considered drying process. Again the neural model seems to be very sensitive to the training process as expected from Fig. 3. Nevertheless, it is clear from Fig. 4b that the temperature prediction is more sensitive to the number of neurons than in the case of moisture content. This may be due to the fact that temperature time evolution is less linear than moisture content evolution which implies the need of more slots for the considered time interval $t \in [0, T]$.

Finally, the capabilities of the proposed model to give accurate solutions for the inverse problem are evaluated and analysed. Eqs. (8)–(11) have been applied to solve the inverse problem in the drying process to find the optimal value of E that matches the desired target point $[T_{c0}; t_0]$ or $[X_{t0}; t_0]$. In order to test the resolution of inverse problems, the temperature (T_{c0}) and moisture content (X_{t0}) targets have been kept constant and different values for t_0 have been considered, and the results are shown in Fig. 5.

In Fig. 5a, the optimum value for E, the moisture content target objective (0.5 dry basis), the drying curves for both the drying model [2,3] and the neural network architecture are shown. Likewise, Fig. 5b illustrates the temperature target (50 °C), the optimum value for E and the temperature curves provided by the drying and the neural model. As Fig. 5 shows, the matching error at the targets $[T_{c0}; t_0]$ or $[X_{t0}; t_0]$ is negligible, while the predictive identification of the drying and temperature curves is very precise.

4. Conclusions

In this paper, a model based on artificial intelligence has been applied to the design of a novel architecture which interconnects the adaptive characteristics of the RBF (Radial Basis Function) neural networks with the algebraic tools of the polynomial matrix equations. This neural architecture has been applied to model and predict the sample temperature and moisture content evolution in microwave-assisted drying processes during a long time interval, where the air temperature and the electric field are considered as the input parameters of the process. Additionally, the use of RBF and polynomial networks has allowed the resolution of the inverse problem by finding the optimal value of the microwave electric field that forces $T_c(t)$ or $X_t(t)$ to reach a target value, T_{c0} or X_{t0}, at a desired instant t_0. The obtained results for both the temperature and moisture content prediction and the inverse calculation show that this modelling technique is very precise provided that a proper training process and neuron number estimation is carried out. The main advantage of the proposed learning-based model is to provide a closed solution for the described inverse problem, which is difficult to solve by the conventional drying models based on differential equations. Although this neural architecture has been tested on a very particular microwave-assisted drying model, the obtained conclusions can be readily extended to other drying models or techniques due to the adaptive capabilities of neural networks.

References

[1] G.V. Barbosa-Cánovas, H. Vega-Mercado, Dehydration of Foods, Chapman & Hall, 1996.

[2] J.L. Pedreño-Molina, J. Monzó-Cabrera, A. Toledo-Moreo, D. Sánchez-Hernández, International Communications in Heat and Mass Transfer 32 (2005) 323.

[3] J. Monzó-Cabrera, A. Díaz-Morcillo, J.L. Pedreño-Molina, D. Sánchez-Hernández, Microwave and Optical Technology Letters 32 (2002) 465.

[4] E.K. Murphy, V.V. Yakovlev, FDTD-backed RBF neural network technique for efficiency optimization of microwave structures, 9 AMPERE Conf. on Microwave and RF Heating, Loughborough, U.K., 2003.

[5] X.G. Wang, W. Liu, L. Gu, C.J. Sun, C.E. Gu, C.W. de Silva, Development of an intelligent control system for wood drying processes, in: Proc. of Int. Advanced Intelligent Mechatronics, vol. 1, IEEE/ASME, 2001, p. 371.

[6] Zhang Qinghua, S.X. Yang, Modeling and parameter optimization of rough rice drying using artificial neural networks, Proceedings of the International Joint Conference on Neural Networks, IJCNN '02, vol. 1, 2002, p. 794.

[7] R. Platon, M. Amazouz, Modeling of a drying process using subtractive clustering based system identification, in: IFSA World Congress and 20th NAFIPS International Conference, Joint 9th, vol. 5, 2001, p. 2994.

[8] J. Monzœ-Cabrera, et al., International Communications in Heat and Mass Transfer 27 (8) (2002) 1101.

[9] D. Broomhead, D. Lowe, Complex Systems 2 (1998) 322.

[10] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, Oxford, 1995, p. 254.