Agricultural Sciences, 5(2) 97-106 (2000)

Neural Network Assisted Experimental Designs for Food Research

S. Sreekanth[1], C. Chen[2], S.S. Sablani[4], H.S. Ramaswamy[2*], and S.O. Prasher[3]

[1]Pratt and Whitney Canada, Longueuil, PQ, [2]Department of Food Science and Agricultural Chemistry/ [3]Department of Biosystems Engineering, Macdonald Campus of McGill University, Ste. Anne-de-Bellevue, PQ H9X 3V9, [4]Department of Bioresource and Agricultural Engineering, College of Agriculture, Sultan Qaboos University, P.O. Box 34, Al Khod 123, Sultanate of Oman

استخدام الشبكة العصبية في تصاميم التجارب البحثية في مجالات الأغذية

سري سريكانث وكورين تشين و شيام سبلاني و هوساهالي راماسوانى و شيف براشار

خلاصــة: هــذه الورقة ستقوم بإعطاء نبذه عن مقدرة الشبكة العصبية الاصطناعية في التنبؤ بعامل المعطيات الكلي من عامل المعطيات الجزئي ترتبط ببعض معطيات التجارب المستخدمة عموما. مخطط العامل الجزئي و العامل الكلي مثل L_8 و L_9 و L_{18} ومخطــط بوكــس و بنكن في الأخذ بعين الاعتبار في الأصل مع وجود بعض التغيرات الحسابية L_{8+6} و L15 و L_{9+1}. العــامل الكلــي (٣ عوامل X ٥ مستويات) و العامل الجزئي خرجوا بتوظيف ست عشره معادلة رياضية مختلفة (أربع في كل فئة: الخطية ب أو بدون تقاطعات – غير خطية ب أو بدون تقاطعات) مختلف الشبكات العصبية الصناعية أجريت وتم اختيار المخطط الافضل لكل معادلة باعتبار المقدرة في التنبؤ بالعامل الجزئي. في حالات مختلفة معدل الخطأ النسبي مع مخطط L_{18} (الــذي يحــتوي علــى معطــيات داخله اكثر من أي مخطط أخر) كان اكثر من أي مخطط جزئي آخر اصغر عموما. الشبكة العصبية الاصطناعية اختارت بوكس و بنكن L_{18} و L_{15} للتنبؤ بعامل المعطيات الكلي بشكل معقول بنسبة خطأ تقل عن ٥% . مخطط L_{8+6} يعمل جيدا مع تجارب مختلفة في حالات روائية مختلفة.

ABSTRACT: The ability of artificial neural networks (ANN) in predicting full factorial data from the fractional data corresponding to some of the commonly used experimental designs is explored in this paper. Factorial and fractional factorial designs such as L_8, L_9, L_{18}, and Box and Behnken schemes were considered both in their original form and with some variations (L_{8+6}, L_{15} and L_{9+1}). Full factorial (3 factors x 5 levels) and fractional data were generated employing sixteen different mathematical equations (four in each category: linear, with and without interactions, and non-linear, with and without interactions). Different ANN models were trained and the best model was chosen for each equation based on their ability to predict the fractional data. The best experimental design was then chosen based on their ability to simulate the full-factorial data for each equation. In several cases, the mean relative errors with the L_{18} design (which had more input data than other models) were even higher than with other smaller fractional design. In general, the ANN assisted L_{8+6}, Box and Behnken, L_{15} and L_{18} designs were found to predict the full factorial data reasonably well with errors less than 5%. The L_{8+6} model performed well with several experimental datasets reported in the literature.

Keywords: neural network, factorial data, fractional data, food, prediction.

Most experimental approaches in food research involve evaluating how an output parameter (dependent variable) is influenced by several other factors (independent variables). Whether it is intended for the purpose of understanding the influence of test factors, for selecting test conditions for an optimized output or for developing predictive models, experimental approaches in data gathering generally involve a full or fractional factorial design. Three possible situations (Taguchi, 1987a,b) can arise in this exercise: (a) the functional relationship between the independent factors and the dependent factors are unknown, (b) the functional relationship is known in an approximate way, but some of the parameters need to be determined, such as in dimensional analysis and (c) the relationship is known completely. In cases (a) and (b),

*Corresponding author.

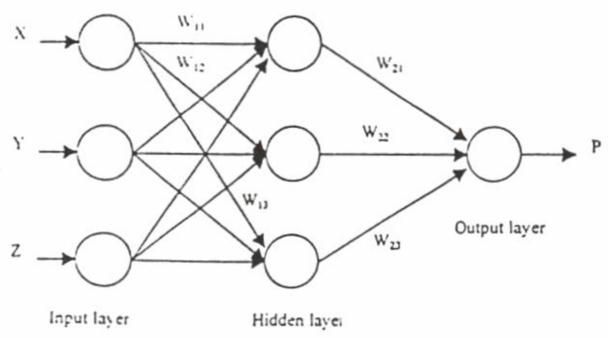


Figure 1. A multi-layered artificial neural network showing a typical configuration.

experiments have to be performed either in order to identify the form of the dependence function or to estimate the parameters.

Food and pharmaceutical research experiments can be expensive and time consuming. It is not always economical or practical to conduct test runs with a full factorial design and therefore, it is desirable to minimize the number of experiments while limiting loss in information. Several researchers have tried to construct fractional factorial experimental designs which provide a reasonably accurate picture of systems. A review of the various experimental designs was provided recently by Kim and Kalb (1986). A comprehensive treatment of various aspects of experimental design is also given by Taguchi (1997a,b). These experimental designs have been successfully applied in various fields (Ramaswamy *et al.*, 1988; Hachigan, 1989; Matulis *et al.*, 1995; Ashie *et al.*, 1996; Cano *et al.*, 1997) and in general, they are based on a limited number of carefully chosen combinations from a full factorial design, such that the statistical significance of the influence of independent variables on the output could be evaluated. Simplicity is often achieved in these designs, if it is known that there are no interactions between test variables.

In recent years, the concept of artificial neural networks (ANN) has gained popularity in many fields of engineering and science. ANN concepts have been used in applications such as pattern recognition (Ding and Evans, 1994) and for a variety of prediction problems (Sablani *et al.*, 1995, 1997; Balasubramanian *et al.*, 1996). In these problems, one or more dependent variables are predicted for a given set of independent factors. It has been reported that ANNs perform better than conventional regression analysis (Maureen and Charles, 1992; Bharath and Drosen, 1994). The feature of the ANN that makes it attractive for many of the above applications is its ability to learn the relationship between input and output variables.

The overall purpose of the present work was to explore the potential of ANN to generalize the input/output relationship in the design of experiments. The specific objective was to explore the ability of ANN models developed from various fractional designs in predicting the 3-factor x 5-level full factorial data to see if any particular fractional design provided better performance with several test cases. Since five level full-factorial experimental data were hard to find in the literature, these data were generated from several representative mathematical equations to provide broader testing conditions for the ANN models. The study involved a systematic evaluation of the performance of ANN with different fractional designs using simulated data which were later verified using some experimental datasets.

Methodology

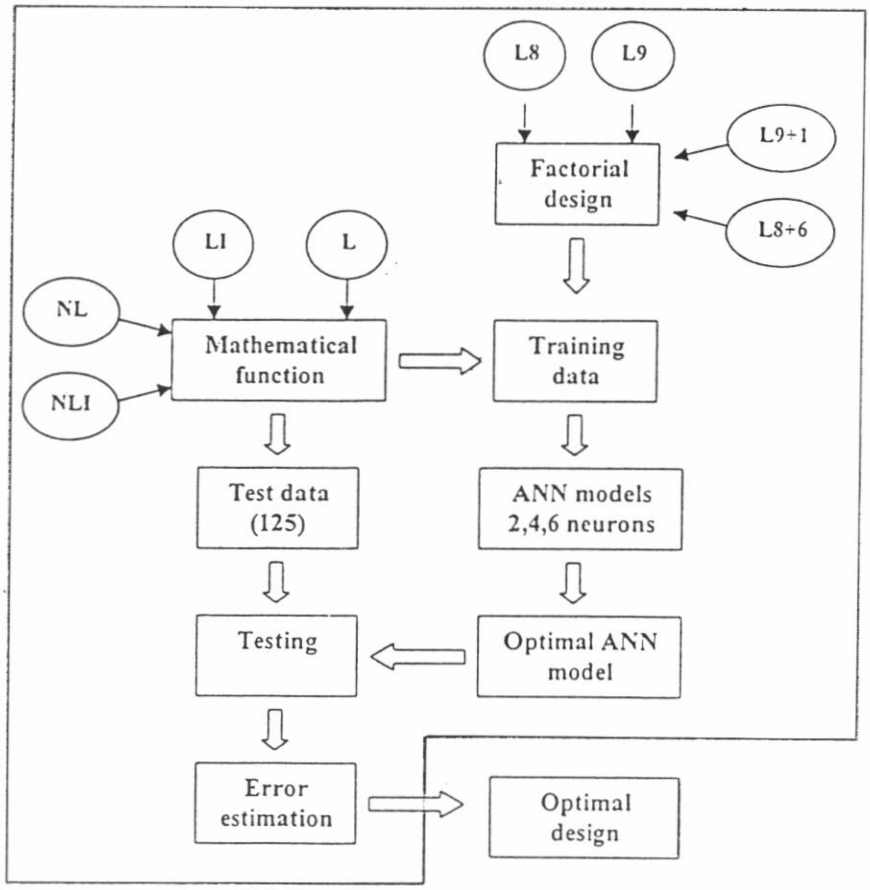
ANN MODEL. Only a brief description of the ANN structure is presented here, but additional details can be obtained in the literature (Maureen and Charles, 1992; Bharath and Drosen, 1994; Sablani *et al.*, 1995, 1997). Figure 1 shows a schematic of an ANN model. It consists of the processing elements (neurons) arranged in layers (input, hidden and output layers). The transfer of signals from one element to other elements is scaled by the connection weight. Mathematically, the function of a processing element can be represented as:

$$X_1 = \sigma\left(\sum_j X_j W_{ij} + b_i\right) \qquad (1)$$

where X is the output from a processing element, W the weight factor, i the processing element from the current layer, j the processing element from the preceding layer, and s is any non-linear function which has a continuous first derivative.

Network training (learning) is an important phase in the development of an ANN model during which the network weights are adjusted to map the input dataset to the output dataset. Error signals associated with the output elements are transmitted using a back-propagation algorithm. The weights are adjusted in each interconnection so as to minimize the error in the network response. This procedure is repeated over the entire learning dataset for a specified number of times (learning runs), chosen by trial and error. In principle, if sufficient number of these input/output combinations are used for learning/training, such a trained ANN should be able to predict the output for new inputs. These learning sets can be compiled from experimental data or obtained from computer simulation.

The following steps describe the general procedure employed in this study to obtain an optimized ANN-



assisted experimental design (Figure 2): (1) Data required for the different fractional designs (as well as for the full-factorial design for the purpose of verification) involving 3 factors at 5 levels were first generated using selected mathematical equations; (2) Only three-layer ANN models were considered with one input layer, one hidden layer and one output layer. The number of processing elements in the hidden layer was varied from 2 to 6 and ANN models were trained with the *fractional datasets* for each mathematical equation; (3) The optimal ANN model (for a given design and mathematical equation) was selected based on the performance of the ANN models to accurately predict the desired output of the fractional model; (4) The optimized ANN models (with number of elements in the hidden layer set at 2, 4 or 6) were then used to predict the full-factorial data for each design. These were then compared with their corresponding data from the full-factorial design previously computed using the mathematical equations (see step 1). The best design was selected as the one yielding the lowest prediction error; (5) The whole process was repeated for different mathematical equations. The best design for a mathematical equation was identified based on consistent performance with all these equations; and (6) Finally, the developed model was tested by applying experimental data.

FRACTIONAL DESIGNS. In the employed fractional designs, the levels were represented by numbers 1 to 5, with level 1 corresponding to the minimum value of the factor and level 5 for the maximum. The three independent variables were denoted x, y and z. The simplest fractional design with 3 factors is L_8, which corresponds to a full factorial design with 2 levels. The L_8 design is shown in the form of an orthogonal array in Table 1. The L_8 can be expected to perform poorly well with non-linear data. For the L_8 design only extreme values of the factors were used while for the L_9 design different combinations of the 3 level design were employed (Table 1). The L_{18} is a more rigorous design allowing various interactions. This consists of 18 experimental test runs. The Box and Behnken (1960) design consists of 13 experimental test runs (12 corresponding to the center of each edge of a cube and one corresponding to the center of the cube itself which is actually repeated 3 times for a 3-level scheme and seven times for a 5-level scheme to obtain an estimate of the experimental error). The other three designs used in the present work were L_{9+1}, L_{8+6}, and L_{15}. These three designs were variants of the L_9 and L_8 designs. In L_{9+1}, an additional combination (5,5,5) was added to accommodate the extreme condition (Table 1). The design of L_{8+6} (14 experimental data sets) was an extension of the L_8 design in which six experiments were

TABLE 1

Orthogal arrays of levels in different designs employed in the study.

L_8	L_9	L_{9+1}	L_{8+6}	Box & Behnken	L_{18}
1 1 1	1 1 1	1 1 1	1 1 1	1 1 3	1 1 1
1 1 5	1 3 3	1 3 3	1 1 5	1 5 3	1 1 5
1 5 1	1 5 5	1 5 5	1 1 1	1 3 1	1 3 1
1 5 5	3 1 5	3 1 5	1 1 5	1 3 5	1 3 3
5 1 1	3 3 1	3 3 1	5 5 1	3 1 1	1 5 3
5 1 5	3 5 3	3 5 3	5 5 5	3 1 5	1 5 5
5 5 1	5 1 3	5 1 3	5 5 1	3 5 1	3 1 3
5 5 5	5 3 5	5 3 5	5 5 5	3 5 5	3 1 5
	5 5 1	5 5 1	1 1 3	5 1 3	3 3 1
		5 5 5	3 3 3	5 5 3	3 3 5
			3 3 1	5 3 1	3 5 1
			5 5 3	5 3 5	3 5 3
			3 3 3	3 3 3	5 1 1
			3 3 5	3 3 3	5 1 3
				3 3 3	5 3 3
					5 3 5
					5 5 1
					5 5 5

added at mid levels of the factors (incidentally, these points represent the 8 vertices and 6 face centers of a cube). The L_{15} design was an extension of L_{8+6} with the center of the cube (3,3,3) included. These three designs do not have any mathematical basis. Although each appear to be similar to several of the fractional factorial models, the L_{9+1}, L_{15}, and L_{8+6} were selected more on an intuitive basis.

MATHEMATICAL EQUATIONS. In order to identify a robust fractional design, several (16 in total) linear and nonlinear mathematical functions were considered. Some of these equations accounted for two-way interaction effects between the variables. In the following equations, P is the dependent variable and x, y, and z are the independent variables with numerical values ranging from 0 to 1.

Linear without interaction (L):

$$P = x + y + z \quad (2)$$

$$P = 100x + 75y + z \quad (3)$$

$$P = 100x + 5y + 3z \quad (4)$$

$$P = 100x - 75y + z \quad (5)$$

Linear with interaction (LI):

$$P = x + y + z + xy + xz + yz \quad (6)$$

$$P = 100x + 75y + z + 175xy + 100xz + yz \quad (7)$$

$$P = 100x + 5y + 3z + 105xy + 103xz + yz \quad (8)$$

$$P = 100x - 75y + z + 100xy + xz + yz \quad (9)$$

Nonlinear without interaction (NL):

$$P = x^{1.5} + y^{0.5} + e^{z} \quad (10)$$

$$P = 15x^{3} + 10y^{0.1} + e^{3z} \quad (11)$$

$$P = 15x^{3} + \frac{10}{4y^{2}+1} + 20\,ln\,(z+1) \quad (12)$$

$$P = 15x^{3} + \frac{10}{4y^{2}+1} + \frac{15}{1+4e^{-4z}} \quad (13)$$

Nonlinear with interaction (NLI):

$$P = x^{1.5} + y^{0.5} + e^{z} + x^{1.5}y^{0.5} + y^{0.5}e^{z} + x^{1.5}e^{z} \quad (14)$$

$$P = 15x^{3} + 10y^{0.1} + e^{3z} + 15x^{3}y^{0.1} + x^{3}e^{3z} + y^{0.1}e^{3z} \quad (15)$$

$$P = 15x^{3} + \frac{10}{4y^{2}+1} + 20\,ln\,(z+1) + \frac{15x^{3}}{4y^{2}+1} + \frac{20\,ln\,(z+1)}{4y^{2}+1} + 20x^{3}ln\,(z+1) \quad (16)$$

$$P = 15x^{3} + \frac{10}{4y^{2}+1} + \frac{15}{1+4e^{-4z}} + \frac{15}{1+4e^{-4z}} \quad \frac{1}{4y^{2}+1} + \frac{15x^{3}}{4x^{2}+1} + \frac{15x^{3}}{1+4e^{-4z}} \quad (17)$$

In linear equations (without interactions), the number of prominent variables were varied from 1 to 3 in different equations by employing some multiplication factors. In the fourth equation, which apparently indicates two variables to be prominent, the effect of one reduces the effect of the other variable to a certain extent. In the case of linear equations with interactions, the number of significant interactions and the main variables were different in each equation. Various non-linear terms, such as logarithmic and exponential, were incorporated with varying degrees of exponents into the non-linear equations. The coefficients for the different variables were selected such that the magnitude of each term relative to other terms was in the desired range. While formulating these equations, an attempt was made to incorporate functions that are typically encountered in food processing situations (exponential, quadratic, reciprocal, logarithmic etc).

ANN PROGRAM. NeuralWorks Professional II/Plus, version 5.23 (NeuralWare Inc., Pittsburgh, PA) was employed for ANN modeling. A standard back-propagation algorithm with tangent hyperbolic transfer function and normalized cumulative delta learning rule was applied as the basic architecture of network.

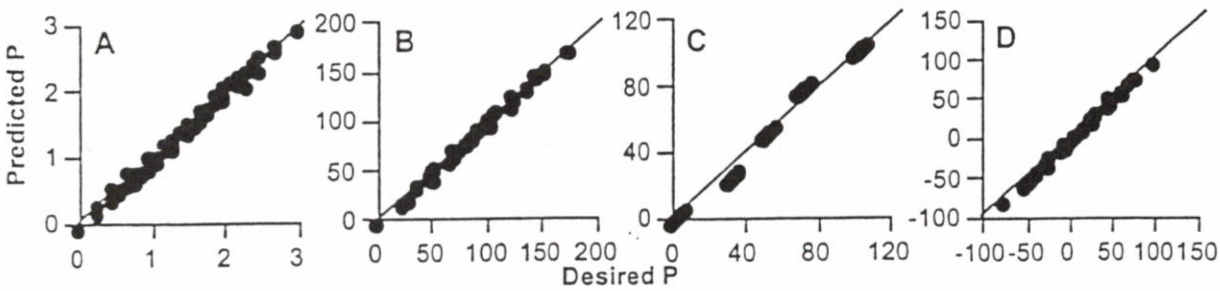


Figure 3. Performance of the ANN model with the L_{8+6} design for linear equations without interactions: A, equation ; B, equation 3; C, equation 4; and D, equation 5.

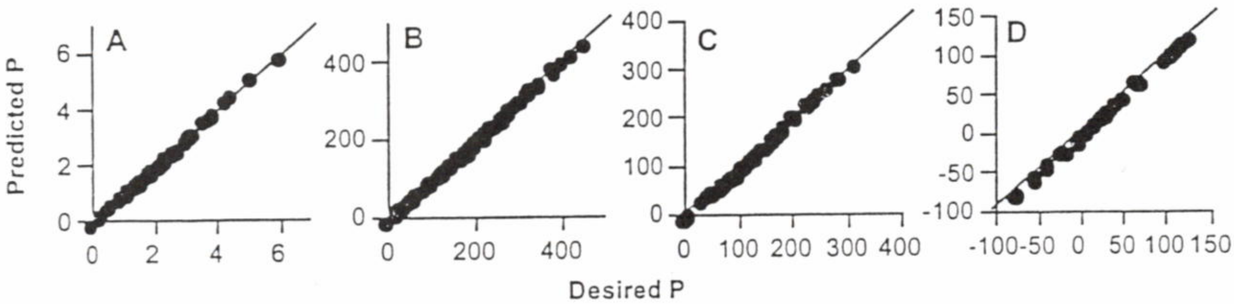


Figure 4. Performance of the ANN model with the L_{8+6} design for linear equations with interactions: A, equation 6; B, equation 7; C, equation 8; and D, equation 9.

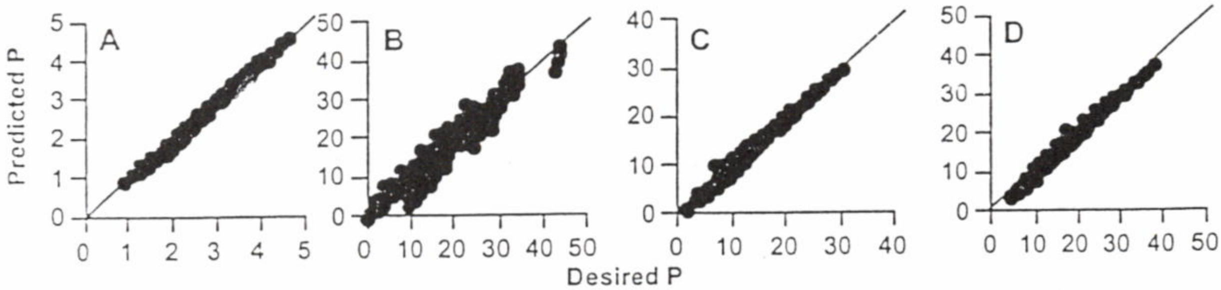


Figure 5. Performance of the ANN model with the L_{8+6} design for linear equations without interactions: A, equation 10; B, equation 11; C, equation 12; and D, equation 13.

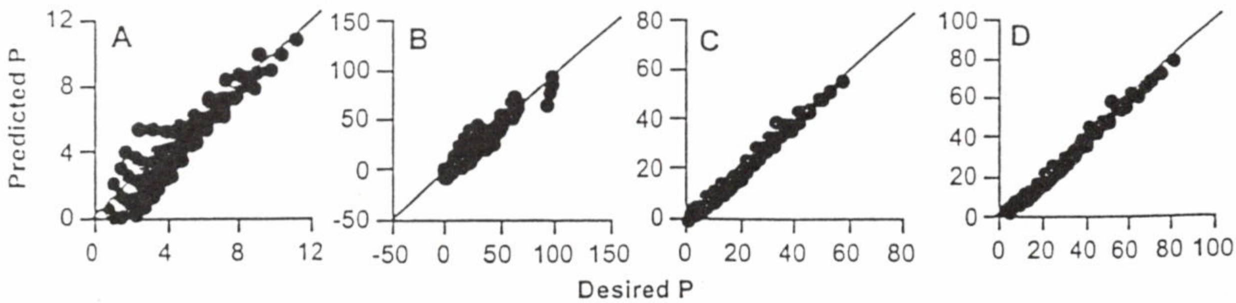


Figure 6. Performance of the ANN model with the L_{8+6} design for linear equations with interactions: A, equation 14; B, equation 15; and C, equation 16; and D, equation 17.

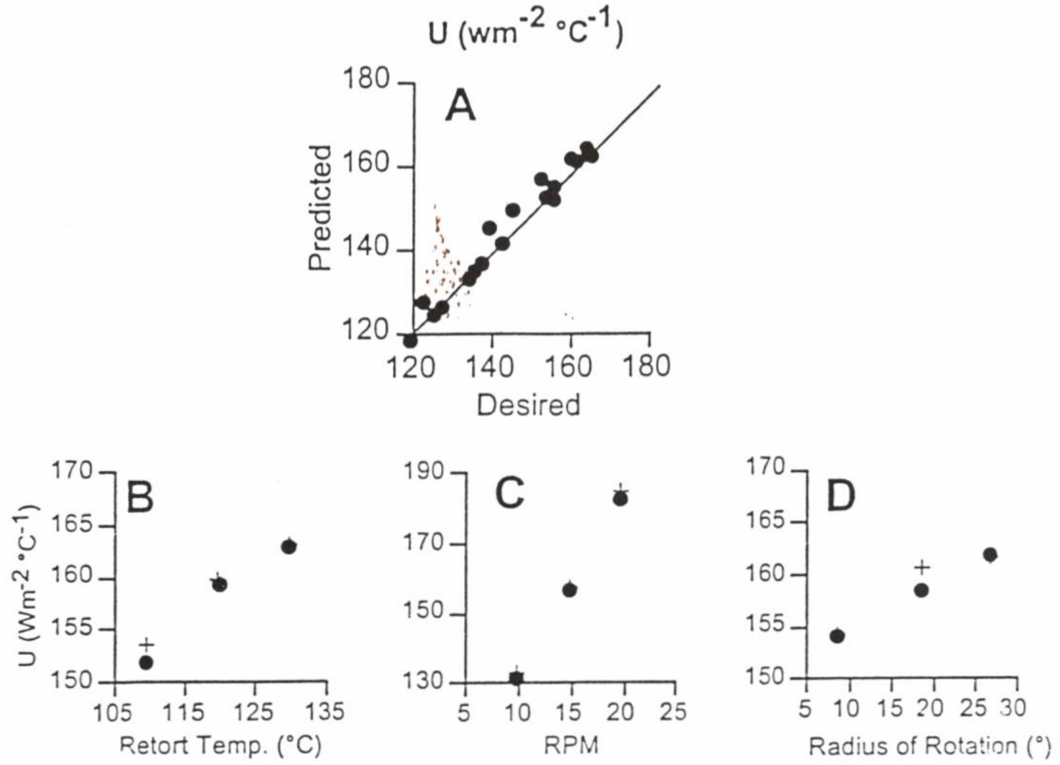


Figure 7. Prediction of the overall heat transfer coefficient (U; test dataset 1) using the ANN model with L_{8+6} design: A, desired vs predicted; B, mean effect of retort temperature on U; C, mean effect of RMP on U; and D, mean effect of radius of rotation on U, + symbols correspond to experimental values and closed circles predict values.

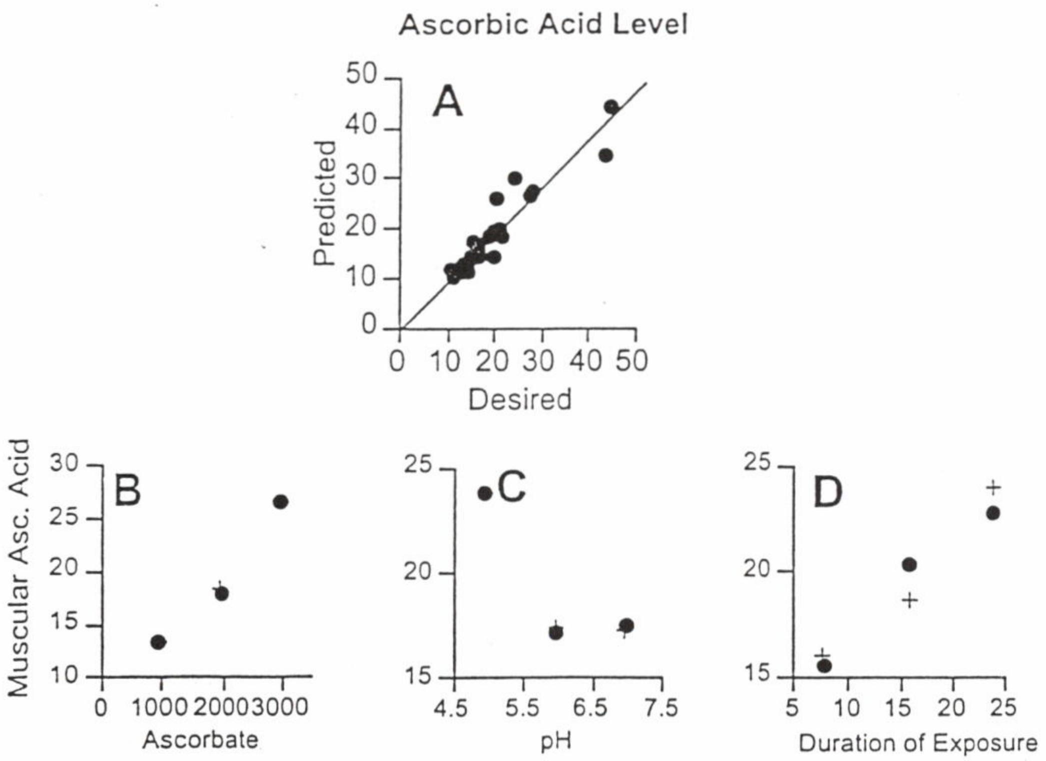


Figure 8. Prediction of muscle ascorbic acid pick-up (test data set 4) using the ANN model with L_{8+6} design: A, desired vs predicted ascorbic acid levels; B, mean effect of ascorbic acid concentration on ascorbic acid concentration; C, mean effect of pH on ascorbic level; and D, mean effect of duration of exposure on ascorbic acid levels, + symbols correspond to experimental values and closed circles predict values.

ERROR PARAMETERS. The prediction performances of different ANN models were evaluated using the mean relative error (E_R) which is defined as,

$$E_R = \frac{1}{N}\sum_{n=1}^{N}\frac{\left|P_d - P_p\right|_n}{(P_d)\max - (P_d)\min} x100 \qquad (18)$$

where P_d and P_p represented the desired and predicted values of the dependent variable. The standard deviation of the relative errors is denoted by S_R.

Results and Discussion

DEVELOPMENT OF THE MODEL. Independent variables x, y and z were selected at five levels (0, 0.3, 0.5, 0.7 and 1.0) in the range of 0 to 1. Several ANN models were trained using datasets consisting of three input columns (x, y, z) and an output column (P_d) for different fractional designs and mathematical functions. Three layer ANNs (1 input, 1 hidden and 1 output layer) with hyperbolic tangent functions were used in all the cases. ANN models were trained until performance reached its optimal level. Three different ANN configurations were considered for optimization by varying the number of processing elements (2, 4, and 6) in the hidden layer. ANNs with more than six neurons were not considered because, for all mathematical equations tested, the model performances were fairly good with 6 or lesser number of neurons. This was true for the number of hidden layers also (1 layer was adequate). It should, however, be noted that the ANN specifications are problem dependent and a 6-neuron single hidden layer network model is not necessarily the best configuration for all problems.

Although the single hidden-layer ANN configuration was optimal for all designs studied, the number of neurons in the hidden layer depended on the type of mathematical function and the design used. Once the optimal ANN configuration was identified for each of the fractional designs, full-factorial data were generated for each of the equations, using this optimal ANN model. The full-factorial data consisted of 125 data points. These ANN-predicted values were compared with corresponding exact values of P generated from the mathematical functions. Table 2 shows the error parameters calculated for each design and equation. For L_8, L_9, and L_{9+1}, the error levels were higher for non-linear equations compared to linear equations. However, the other four designs maintained the errors below a 5% level for most of the linear and non-linear equations. Because L_9 is a subset of L_{9+1}, it is reasonable to expect that errors with L_{9+1} will be less than or equal to that of L_9. The results also indicate that adding an extra data point as input for the model need not always be beneficial. Contrary to expectations, the L_{18} design did not perform better than Box and Behnken, L_{8+6} or L_{15} designs with more than 50% of the equations. Thus, it is clear that the performance of the four fractional designs are generally comparable and that it is not always necessary to use an L_{18} design to optimize performance. The L_{8+6}, while being comparable to other models, has the additional advantage of a lower number of experiments. Only this design was evaluated further for performance testing.

Figures 3-6 show the combined plots of predicted values P_p, using the ANN-assisted L_{8+6} design, plotted against the desired value of P_d for all the equations. The error parameters in Table 2 are good indicators of the performance of an ANN model. Nevertheless, they are average values and reflect only the overall performance. Figures 3-6 provide a graphic representation of the performance of the model at all different levels. These figures are plotted for L_{8+6} design. With few exceptions, the performance of the ANN model was good with both linear and non-linear equations. At present, the reason for the inconsistency with some mathematical relationships is not very clear.

For linear equations, the model does not really require the 14 data pairs to estimate the coefficients algebraically (without ANN) and once these coefficients are estimated, the resulting equations could compute the full factorial data with 100% accuracy. In this sense, ANN models for the linear equations may seem

TABLE 2

Error values for ANN models from different designs.

Eq. #	L_8	L_9	L_{9+1}	L_{8+6}	L_{15}	Box & Behnken	L_8
2	2.4*/1.7**	2.2/1.8	2.9/2.5	1.8/1.3	2.0/1.6	4.6/2.6	1.0/0.8
3	4.8/3.1	3.4/2.7	2.4/2.1	1.6/1.4	1.6/1.4	1.7/1.3	1.2/1.1
4	2.6/1.9	2.3/1.9	2.2/1.9	2.5/2.9	2.5/2.9	2.3/2.5	2.5/3.0
5	4.0/2.5	2.8/3.4	2.6/3.1	1.5/1.4	1.5/1.4	1.4/1.4	1.6/1.4
6	4.9/3.0	2.3/2.3	2.28/2.0	0.9/0.9	1.0/1.0	2.0/1.8	0.6/0.6
7	6.7/4.6	3.2/2.8	4.6/3.3	0.8/0.8	0.9/0.9	1.5/1.3	0.9/1.3
8	7.9/6.0	3.9/3.1	3.4/2.9	1.0/1.0	0.7/0.8	1.9/1.6	1.0/1.0
9	3.0/3.4	6.7/5.7	2.9/2.1	0.8/1.0	0.8/1.0	1.2/1.1	0.7/1.0
10	5.0/3.6	4.4/3.9	3.8/3.5	1.4/1.1	1.6/1.3	1.7/1.4	1.5/1.25
11	11.2/7.9	7.3/5.9	6.3/5.5	5.3/4.0	3.3/3.1	2.4/2.2	2.4/2.2
12	11.4/9.3	4.7/4.1	6.7/5.6	2.1/2.2	0.9/1.2	1.2/1.4	1.5/1.6
13	9.6/7.6	7.0/5.6	7.7/7.3	1.7/2.0	2.0/2.2	1.4/1.4	1.8/1.7
14	6.0/3.8	3.4/3.3	6.1/5.7	2.4/2.4	1.8/1.8	1.5/1.4	1.3/1.2
15	9.0/7.3	5.5/5.5	4.9/5.2	4.6/4.5	4.1/4.1	4.1/4.1	5.7/6.3
16	8.9/7.6	6.9/5.9	6.0/5.4	1.7/1.7	1.2/1.1	1.7/1.9	1.1/1.0
17	7.9/6.6	8.0/7.6	7.1/6.1	1.4/1.4	1.1/1.1	1.7/1.8	1.7/1.6

*Mean relative error (E_R).
**Standard deviation of errors (S_R).

unnecessary and may add a certain level of inaccuracy compared to algebraic computations. However, if one looks at it from the experimental design point of view, where the equation to be fit is not known *a priori*, the ANN approach is deserved of merit. The fact that the ANN-assisted models would perform equally well with non-linear equations, makes them more useful from the experimental design point of view.

It should be noted that conventional analyses such as ANOVA, with fractional designs permit analysis of main and interaction effects. Although the ANN-assisted design does not directly permit such an analysis, the generated data could be subjected to statistical analyses. Since the emphasis of the ANN-assisted design is to accurately predict the full factorial data, the model provides a better relationship between the independent and dependent variable. Furthermore, in the present analysis, some mathematical equations have been used which are encountered in food processing applications such as reaction kinetics, thermal processing, drying, freezing etc. However, they are by no means exhaustive. If for a given situation, other mathematical equations, more representative of the physical problem, could be constructed, they should be used instead for developing the ANN model to develop a better design. Experiments could then be performed according to this optimal design.

TESTING OF THE MODEL. Having identified the L_{8+6} design as the optimal one for the different equations considered, an attempt was made to test the performance of the design with published experimental data. Unlike mathematical data, experimental data will have errors associated with measurements. By using the individual values of several replicates, an estimate of experimental variations could be found. In the present study, only the mean values of the test data for each conditions were used as inputs for the purpose of training. This means that the neural network model will only predict the mean values associated with the different testing conditions. It should however, be noted that it is possible to include the experimental scatter while training the ANN model. Four different sets of experimental data were taken from the literature. All sets consisted of 3-factor 3-level data corresponding to 27 different experimental conditions. L_{8+6} design data were extracted from these data, for each set, and the optimal neural network model was developed based on the method described earlier.

TEST DATASETS 1 AND 2. First and second datasets correspond to experimental analysis of heat transfer to liquid particle mixtures in cans during end-over-end thermal processing (Sablani and Ramaswamy, 1996). End-over-end thermal processing refers to commercial sterilization of canned products in a rotary retort. Cans, stacked vertically in a cage, will undergo end-over-end agitation as the cage is rotated during the heat treatment. This induces forced convection in the cans resulting in product mixing which also improves the rate of heat transfer and hence product quality. When dealing with particulate foods in cans undergoing agitation, there are two levels of heat transfer. The first one is the transfer of heat from the heating medium in the retort (water or steam) to the liquid inside the can (governed by the overall heat transfer coefficient, U) and the second one is the transfer of heat from the liquid inside the can to the particles (governed by the fluid-to-particle heat transfer coefficient, h). In the experimental set-up, the overall and fluid-to-particle heat transfer coefficients were measured as a function of process temperature (110 to 130°C), rotational speed (10 to 20 rpm) and radius of rotation (9 to 27 cm). Two separate neural network models were developed, one for the overall heat transfer coefficient and the second one for the fluid-to-particle heat transfer coefficients. The optimal neural network model for both models consisted of 6 neurons and one hidden layer. The full factorial data were predicted from the neural network model and compared with measured data. Figure 7a shows the predicted heat transfer coefficient plotted against measured, indicating that the prediction was good with an R^2 value of 0.967 and a mean relative error of 3.4% with 4.4% standard deviation. In Figures 7b-d, the mean effect plots of each input on the overall heat transfer coefficient are shown. In these figures, the effect curves were derived from both predicted and measured data, and again demonstrate a good fit.

TEST DATASET 3. The third dataset (Ramaswamy and Tung, 1990) corresponds to the measurement of surface heat transfer coefficient (h) associated with water immersion heating media as influenced by the temperature (105 to 115°C), overpressure (70 to 140 kPa) and flow rate (10 to 20 standard cubic feet per minute). Overpressure thermal processing is used for processing of foods in flexible and semi-rigid containers to protect the integrity of packages during heating and cooling by reducing the tendency of occluded gasses inside the package to expand due to heat. Because of the thin profile nature of these containers, heat transfer will be rapid, process time will be short and hence product quality will be generally superior. In this study, the heat transfer ability of the heating medium in the retort was evaluated as a function of temperature, overpressure and flow rate. The optimal neural network model consisted of 10 neurons and 1 hidden layer. The results for this case provided an R^2 value of 0.83 and the mean relative error of the prediction was 6.24% with the standard deviation of 9.63%.

TEST DATASET 4. As a fourth set of data (Thed *et al.*, 1993), the measurement of ascorbate absorption by live channel catfish was used as influenced by ascorbate concentration (1000 to 3000 ppm), pH (5 to 7) and duration of exposure (8 to 24 hr). The optimal neural network model in this case consisted of 6 neurons and 1 hidden layer. The results shown in Figures 8a-8d, provide good comparison with the measured data with an R^2 of 0.89 and a mean relative error of 4.3% with a 6.75% standard deviation. Almost a perfect fit was observed between experimental and predicted values for ascorbate concentration and pH, while the values of exposure time were reasonably well predicted. However, it should be noted that the trend of expected and predicted values with respect to exposure time was quite different indicating occasional deviations in the predictive behavior of ANN-based experimental designs.

TEST DATASET 5. In the above experimental datasets, the models involved only three factors and three levels. Hence, it may be argued that there is no assurance that the developed approach would accurately predict multivariate functions with more than three factors and more than three levels with combinations of linear and non-linear interactions. In order to accommodate these, a new set of data from a recent publication (Chen and Ramaswamy, 1999) was used. In this study, the rheological properties of tapioca starch were evaluated as a function of concentration, pH, temperature and cooking time, each with five levels. Using a central composite design, experiments were carried out to compute the consistency coefficient (m) and flow behavior index (n) of the power law model:

$$\sigma = m\gamma^{n} \quad (19)$$

where s is the shear stress and g is shear rate.

The following predictive equations involving temperature (T), concentration (C), pH and cook time (t) were developed using a second order response methodology:

$$\begin{aligned} m = & -0.183 + 0.195C - 0.0079T + 0.051\,pH \\ & -0.0028t\ 0.047\,(C-4)^2 + 1.2*10^{-4}(T-50)^2 \\ & -8.4*10^{-4}(t-20)^2 - 0.0037(C-4)(T-50) \\ & (R^2 = 0.94, S_{y,x} = 0.0664) \end{aligned} \quad (20)$$

$$\begin{aligned} n = & 1.057 - 0.114C + 0.055T - 0.05pH \\ & -0.023(C-4)^2 - 0.017(pH-6)^2 \\ & +0.035(t-20)^2 - 0.002(T-50)(C-4) \\ & (R^2 = 0.91, S_{y,x} = 0.0532) \end{aligned} \quad (21)$$

These equations involved four factors at five levels with second order interactions and were considered more complex. ANN models were trained taking three factors at a time using L_{8+6} data subsets and full factorial data were predicted from the trained model. These are compared with individual values from the experimental study (using the statistical model Equations 20 and 21) and demonstrated excellent prediction for each factor at the four levels as a function of the other three variables. All these models had 1 hidden layer and 10 neurons and the predicted mean relative errors ranged from 2-3% at different pHs, 3-5% at different concentrations, 3-5% at different temperatures and 2-4% at different cook times. The R^2 values ranged from 0.92 - 0.98.

Conclusion

A neural network approach was evaluated for the design of experiments. Various 3-factor experimental designs and several linear and non-linear mathematical functions were used in the development of the method. Three different ANN configurations were considered. Results indicated that the L_8, L_9, and L_{9+1} designs failed to provide results of reasonable accuracy. The remaining four designs showed consistently better performance for all equations. The errors associated with these four designs were less than 5% for most of the equations. The L_{18} design, although predicting the full factorial data with reasonable accuracy, did not perform better than L_{8+6}, L_{15} and Box and Behnken designs in several equations. The L_{8+6} design was verified with experimental data related to food research and it performed fairly well in most of the cases.

References

Ashie, I.N.A., B.K. Simpson, and H.S. Ramaswamy. 1996. Control of endogenous enzyme activity in fish muscle by inhibitors and hydrostatic pressure using RSM. *Journal of Food Science* 61:350-356.

Balasubramanian, A., R.C. Panda, and V.S. Ramachandra Rao. 1996. Modeling of a fluidised bed drier using artificial neural network. *Drying Technology* 14:1881-1889.

Bharath, R. and J. Drosen. 1994. *Neural Network Computing*. McGraw-Hill, Inc. New York, U.S.A.

Box, G.E.P. and O.W. Behnken. 1960. Some new three level designs for the study of quantitative variables. *Technometrics* 2:455-476.

Cano, M.P., A. Hernandez, and B. de Ancos. 1997. High pressure and temperature effects on enzyme inactivation in strawberry and orange products. *Journal of Food Science* 62:85-88.

Chen, C.R. and H.S. Ramaswamy. 1999. Rheology of Tapioca Starch. *Food Research International* 32:319-325.

Ding, K. and M.D. Evans. 1994. Shape feature extraction and classification of food material using computer vision. *Transactions of the ASAE* 37:1537-1545.

Hachigian, J. 1989. An experimental design for determination of D-values describing inactivation kinetics of bacterial spores: Design parameters selected using computer simulation. *Journal of Food Science* 54:720-726.

Kim, J.S. and J.W. Kalb. 1996. Design of experiments: An overview and application example. *Medical Device and Diagnostic Industry* 18:7-12.

Matulis, R.J., F.K. Mckeith, J.W. Sutherland, and M.S. Brewer. 1995. Sensory characteristics of frankfurters as affected by fat, salt and pH. *Journal of Food Science* 60:42-47.

Maureen, C. and B. Charles. 1992. *Understanding Neural Networks: Computer Explorations*. The MIT Press, London, England.

Ramaswamy, H.S. and M.A. Tung. 1988. Heating rates of thin profile packages in two pilot scale steam/air retorts. *Canadian Institute of Food Science and Technology* 21:316-322.

Ramaswamy, H.S. and M.A. Tung. 1990. Surface Heat Transfer Coefficients for Water Immersion Heating Media in Vertical Retorts. In: *Engineering and Food. Preservation Process and Related Techniques*, E.E.L. Spiess, and H. Schubert (Editors) 2:85-94. Elsevier Applied Science Publishing. New York, U.S.A.

Sablani, S.S. and H.S. Ramaswamy. 1996. Particle heat transfer coefficients under various retort operating conditions. *Journal of Food Process Engineering* 19:403-424.

Sablani, S.S., H.S. Ramaswamy, and S.O. Prasher. 1995. A neural network approach for thermal processing applications. *Journal of Food Processing and Preservation* 19:283-301.

Sablani, S.S., H.S. Ramaswamy, S. Sreekanth, and S.O. Prasher. 1997. Neural network modeling of heat transfer to liquid particle mixtures in cans subjected to end-over-end processing. *Food Research International* 30:105-116.

Taguchi, G. 1987a. *System of Experimental Design*. Vol. 1. American Supplier Institute, Inc., NY.

Taguchi, G. 1987b. *System of Experimental Design*. Vol. 2. American Supplier Institute, Inc., NY.

Thed, S.T., M.C. Erickson, and R.L. Shewfelt. 1993. Ascorbate absorption by live channel catfish as a function of ascorbate concentration, pH, and duration of exposure. *Journal of Food Science* 58:75-78.

Received 13 March 2000.
Accepted 18 September 2000.