ELSEVIER

Food Chemistry 50 (1994) 157–165

0308-8146/94/$7.00

Prediction of quality and origin of black tea and pine resin samples from chromatographic and sensory information: evaluation of neural networks

Keith I. Tomlins* & Clifton Gay

Natural Resources Institute, Central Avenue, Chatham Maritime, Kent, ME4 4TB, UK

(Received 5 May 1993; revised version received and accepted 8 July 1993)

The accuracies of neural network and statistical methods were similar for classifying the origin of black teas from their phenolic composition. When the data are non-normal, as was the case for the pine resin samples, the neural network offered a significant improvement. Neural networks were less accurate than stepwise multiple regression as a model for predicting black tea score and price from their chemical composition and sensory attributes. The accuracy improved and the training time was reduced when training variables chosen by stepwise multiple regression were selected. An advantage of the neural network model is that a single model could predict several parameters simultaneously. The selection criterion of neural networks could be estimated by inspection of the most positive weights derived from two-layer trained networks.

INTRODUCTION

Neural networks have been the focus of much interest with a great deal of the attention having been centred on speech and image recognition. They are modelled on the neurons in the brain and the synapses that connect them; this enables a computer to simulate many of the brain's abilities. Rather than being programmed for a particular application, neural networks can 'learn' during the training process to arrive at a generalised solution. They are also thought to be more robust to noise and inconsistencies in the data than many other pattern recognition methods (Lippmann, 1987 ; Kamal *et al.*, 1989; Anon, 1992).

In chemical analysis, neural networks have been used to classify sugar alditols from their NMR spectra (Thomsen and Meyer 1989), in spectroscopic calibration (Long *et al.*, 1990; Gemperline *et al.*, 1991; Song *et al.*, 1992); to identify partially methylated alditol acetates by their mass spectra (Sellers *et al.*, 1990); in pattern recognition of chromatographic data (Long *et al.*, 1991); in an electronic odour detector (Chang *et al.*, 1991; Persaud, 1992); and an electronic wine taster (Anon 1992).

In sensory science, neural networks have been proposed as a method of advancing the understanding of complex unstructured tasks involving human insight and judgement (Galvin & Waldrop 1990) and have been used to model the sensory determination of colour quality of tomato and peaches (Thai & Shewfelt, 1991).

Artificial neural networks appear to offer advantages over current pattern recognition methods. The aim of this study is to compare neural network methods with statistical procedures (regression and multivariate) used at NRI (McDowell *et al.*, 1991; Taylor *et al.*, 1992) used for pattern recognition and quality prediction from chromatographic and sensory data.

MATERIALS AND METHODS

Neural network software

A commercially available feed-forward neural network (Brainmaker V2.3, California Scientific Software, CA, USA) was used. It offered the choice of several transfer functions (sigmoidal, linear, linear threshold, gaussian, step) and the ability to alter the number of hidden layers and the number of neurons in each hidden layer.

The network software was run using a 25 MHz 386SX IBM-compatible computer (Elonex PC-320X) fitted with 4 megabytes (MB) RAM, a 3·5 inch floppy and 40 MB hard drive.

Statistical analysis

Statistical analysis (regression, multivariate) was carried out using either Statgraphics (STSC Inc, MD, USA) or

* To whom correspondence should be addressed.

Genstat (Numerical Algorithms Group Ltd, Oxford, UK).

Sampling method for evaluating the neural network and statistical models

Sampling was by the hold-out method (Yoon *et al.*, 1993). Each data set was randomly divided into four quarters, which, in turn, were withheld. Using the remaining three quarters, neural network and statistical models were developed to predict the outcome for each withheld quarter.

Pattern recognition data sets

Four data sets (detailed below), obtained either by high performance liquid chromatography (HPLC) or gas chromatography (GC), were evaluated.

(a) A set of 77 black teas, comprising 12 phenolic constituents (theaflavins, flavonol glycosides and unknowns, possibly thearubigins) from seven origins (Assam, Northern India, Bangladesh, Sri Lanka, Kenya, Tanzania and Malawi); chemical analysis was carried out by HPLC (McDowell *et al.*, 1991).

(b) A set of 96 black teas, comprising 37 phenolic constituents (theaflavins, flavonol glycosides and unknowns, possibly thearubigins) from 23 tea estates (African Highlands, Bargang, Bulwa, Deundi, Diyyagami West, Kangaita, Kapchorua, Kaporet, Limbuli, Litein, Lupembe, Makwasa, Mukumbani, Nyankoba, Ogembo, Phulbari, Ragalla, Ratelshoek, Tenduet, Tombe, Torokahuna, Toromityana, Tshivase); chemical analysis was carried out by HPLC (McDowell *et al.*, 1991).

(c) The same data set as in (b), except the teas were classified into nine countries (Kenya, Tanzania, Uganda, India, Bangladesh, Sri Lanka, Malawi, Republic of South Africa, and Zimbabwe) instead of estates.

(d) A set of 158 xylem resin samples (*Pinus caribaea* var. *caribaea*, *P. caribaea* var. *bahamensis*, *P. caribaea* var. *hondurensis*) comprising six resin acid constituents from nine provenances in Zimbabwe; chemical analysis was carried out by GC (Coppen *et al.*, 1993).

Regression data sets

Three data sets obtained either by HPLC, GC or sensory evaluation were evaluated.

(e) A subset of 35 black teas from those used in (a). These varied by price as assessed by professional tea tasters.

(f) A subset of the remaining 42 black teas from those used in (a). These varied by score as assessed by professional tea tasters.

(g) A set of 30 black teas from six regions (Kenya, Assam, Uganda, Sri Lanka, Malawi, Bangladesh) which were assessed by a sensory panel at NRI using 12 attributes (flowery, earthy, smoky, brown, red, bright, clarity, sour, bitter, astringent, strength, lively). These were compared with the score (strength and colour, brisk and bright, quality and flavour, milk) as determined by professional tea tasters.

RESULTS AND DISCUSSION

Artificial neural networks are thought to have the ability to 'learn' during a training process where they are presented with a sequence of stimuli (inputs) and a set of expected responses (outputs). Learning is said to happen when the artificial neural network arrives at a generalised solution for a class of problems. Compared with biological intelligence, the size of the artificial neural network used in this study has been reported to be about halfway between a worm and a cockroach (Lawrence, 1991).

The basic processing element of an artificial neural network is a neuron or node (Fig. 1). These are analogous to theoretical models of neurons in biological systems. Each node has a series of weighted inputs, w_i, which may be either external signals or the outputs from other neurons. The inputs are equivalent to synapses and the weights represent the strengths of synaptic connections. These may be positive or negative in sign, corresponding to excitatory or inhibitory inputs in the system. The sum of the weighted inputs is transformed by a transfer function (linear or non-linear).

A network of such neurons is shown in Fig. 2. This is an example of a feed-forward network where the output from each neuron is not dependent on its previous values. Feed-forward networks consist of an input layer, a hidden layer and an output layer. As an example, the inputs are the amounts of each chemical component $a_1 \ldots a_n$. The input nodes transfer the weighted input signal to the nodes in the hidden layer. A connection between node i in the input layer and node j in the hidden layer is represented by the weighting factor w_{ji}. Hence, there is a vector of weights w_j for each of J

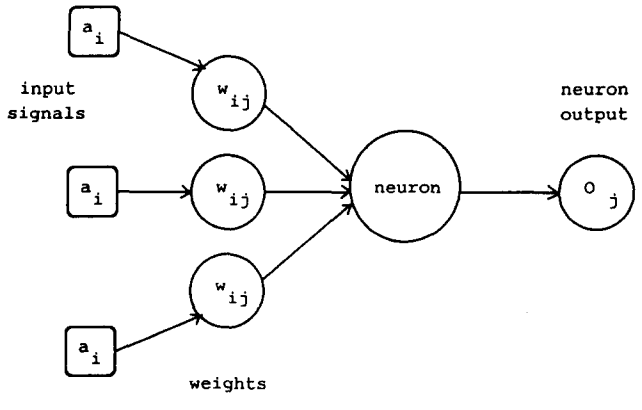
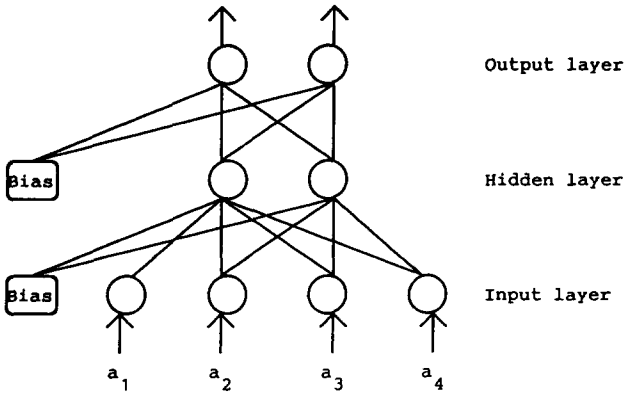


Fig. 1. Schematic diagram for an artificial neural network neuron or mode.



Where: $a_1 \ldots a_n$ are input variables
Bias input compensates for non-zero offsets in the network

Fig. 2. Architecture for a three-layer back-propagation neural network.

neurons in the hidden layer. These weights are adjusted in the learning process. Each layer also has a bias input to adjust for any non-zero offsets in the data. The output of each hidden neuron is a function of the sum of that neuron's weighted inputs.

During the learning procedure, a series of input patterns with their corresponding expected output values are presented to the network in an iterative fashion while the weights are adjusted. The error in the expected output is back-propagated through the network to determine adjustments to the weights. The training process is concluded when the desired level of precision between the expected output and the actual output is achieved.

Methods of training neural networks

The neural network software offered a variety of options concerning training the network. These were assessed to evaluate their influence on the efficiency of training.

Transfer function

The sum of the weighted inputs to each neuron is transformed by a transfer function (sigmoidal, linear, linear threshold, gaussian and step). Of these functions only the sigmoidal function with a gain of 1 (Fig. 3) was suitable for successfully training a network. A particular advantage of this function is that it caters for non-linear relationships. This agreed with other similar work (Long *et al.*, 1990, 1991).

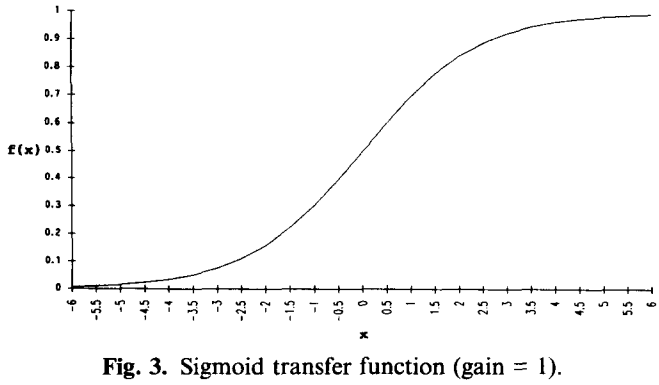


Fig. 3. Sigmoid transfer function (gain = 1).

Number of hidden layers and neurons

The number of hidden layers and neurons in each layer could be varied to a maximum nine hidden layers and 512 neurons in any model. If they are too few, the network fails to learn. Adding more neurons has been reported as being similar to increasing the number of principal components in principal components regression (Long *et al.*, 1990). However, too many neurons and the network is slow to train and may memorise the training set rather than encoding the generalised solution.

In this study, a series of trial and error runs revealed that optimisation was best achieved by starting with a simple two-layer network, followed by inserting hidden layers and neurons until no further improvement in the network performance was noted. Most data sets evaluated required the addition of one or two hidden layers.

Methods of determining the completion of training

Judging the completion of training is crucial. An inadequately trained network may make guesses; if overtrained, it may memorise instead of generalising. The neural network software used in this study offered two methods of assessing the completion of training. These were training until the network could predict the value of every training fact to within a specified degree of accuracy (usually 90% or greater), or by periodic evaluation of testing facts that are separate from the training set.

Of the two routines, the latter was preferred because this tested the ability of the network to generalise and predict rather than its ability to recall. Figure 4 shows an example comparing the two learning methods. In this instance, after 900 runs (single pass through the entire training set), the network could predict all the training facts with an accuracy of 90% or greater. The test set indicates that the network achieved a training accuracy of between 55 and 65% after 20 runs and that this did not improve with prolonged training. Therefore the latter method could be used to reduce the training time with no loss of performance.

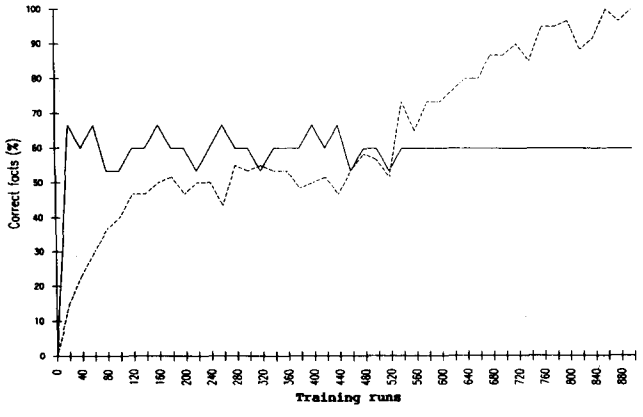


Fig. 4. Compatrison of neural network training techniques. (—) Test set, (—) training set.

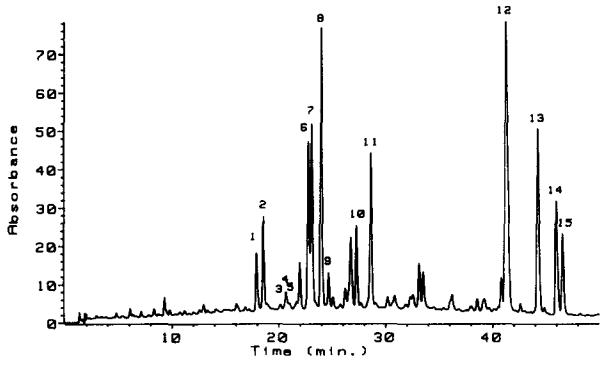


Fig. 5. HPLC chromatogram of phenolic compounds in black tea. Where 1 and 2, flavonol glycoside (X1-4); 3–5, unknown (X5-7); 6, rutin (FG1); 7, isoquercetin (FG2); 8, quercitin glycoside (FG3); 9, flavonol glycoside (X8); 10 and 11, kaempferol glycosides (FG4 and FG5); 12, flavonol aglycone and theaflavin (F, TF1); 13, theaflavin-3-gallate (TF2); 14, theaflavin-3'-gallate (TF3); 15, theaflavin-3,3'-digallate (TF4).

Optimum number of output neurons

Where there is more than one possible outcome, the neural network can be either trained to predict one output fact at a time (i.e. a single output neuron), or it can be trained to offer several outcomes (multiple output neurons). Of the training sets evaluated in this study little difference between the network designs was noted. Multiple output neural networks were preferred as these simplified the training.

Classification of tea and resin samples from their chromatographic data

Three data sets, derived from HPLC or GC, were used to compare the effectiveness of neural networks with statistical methods currently used at NRI.

Figures 5 and 6 show example HPLC or GC chromatograms for the black tea (Sets a, b, c) and resin acids (Set d).

Considering the statistical procedure, the discriminant functions (DF) derived using canonical variates analysis (CVA) were used to predict the origin of the unknown samples based on their chemical composition. This prediction can be gauged by applying the DF as an allocation rule to the samples included in each set.

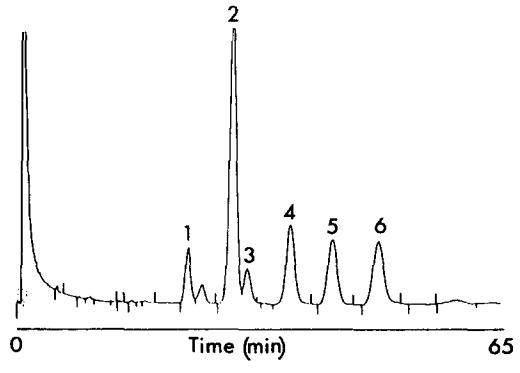


Fig. 6. GC chromatogram of methyl esters of resin acids in *Pinus caribaea*. Where 1, pimaric acid; 2, levopimaric + palustric acid; 3, isopimaric acid; 4, dehydroabietic acid; 5, abietic acid; 6, neoabietic acid.

Table 1. Classification of 77 black teas (Set a) by neutral network and statistical methods[a]

Result	Actual						
	A	B	K	M	NI	SL	T
Neural network model							
A	61	13		5	60	40	
B	28	87			20		
K			80	5	20		100
M				85			
NI	6						
SL	6			5		50	
T			20			10	
Statistical model							
A	45	17		6	40	20	
B	22	70		6	40		
K			40				100
M				88			
NI	22	13			20		
SL	11			20	60		
T			60				

[a] A, Assam; B, Bangladesh; K, Kenya; M, Malawi; NI, Northern India; SL, Sri Lanka; T, Tanzania.

For each data set the optimum neural network architectures were 11, 11, 11, 7 for Set (a), 37, 37, 23 for Set (b), 37, 37, 9 for Set (c) and 6, 100, 30, 9 for Set (d); where in each series the first and last numbers refer to the number of input and output neurons, respectively, and intermediate numbers refer to the number of neurons in each hidden layer. In each case, one or two hidden layers were required for the network to train effectively. The addition of hidden layers suggests that the relationship between origin and chemical composition for each data set is non-linear.

Tables 1–4 show the comparative performance of the neural network and statistical methods in classifying the samples in each set using the hold-out procedure. For each set, the neural network was more effective in correctly classifying the samples than the statistical procedure (Table 5). However, the improvement varied, being relatively small for Sets b (64–68%) and c (89–91%) and larger for Sets a (58–69%) and d (41–76%). The improved performance for the resin acid data (Set d) might be explained by its non-normality, where the chemical characteristics are grouped into several discrete groups as opposed to a continuous response (Birks & Kanowski, 1988). In the case of the chemical composition of the tea, processing may result in a more normal distribution of the data.

Inspection of Tables 1–4 also indicates that both techniques met similar difficulties in classifying certain samples. For example, in Set (a) neither model could correctly classify the black teas from Northern India or Tanzania. Also where mis-classifications occurred, they tended to be similar for both methods.

Other studies have reported equivalent verdicts. Neural networks have been reported to out-perform discriminant analysis methods in predicting bond-ratings and stock-price performance (Yoon *et al.*, 1993)

Table 2. Classification of 96 black teas (Set b) by estate[a]

Result	Actual																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	23
Neural network model																					
1	50			33									50								
2																					
3		100	100																		
4				34																	
5						100															
6								33													
7																					
8	50							34													
9																					
10										83		17	50	8				30			
11								33													
12												83									
13										17											
14														61	29						
15				33										31	64						
16																94					
17																	100		33		
18																		30		33	
19															7	6			67		
20																		30			
23																				67	

	Actual																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	23
Statistical model																						
1																						
2							100															
3			67					67														
4				67																		
5	50				100	50			100													
6																						
7														8	7							
8	50		33	33				33														
9						50								8	7							
10										83								33				
11																						
12												83	50							50		
13															7							
14														62	7							
15														14	58							
16																100						100
17		100									100				14		100					
18													50					67		50		
19										17									100			
20														8								
23												17									100	

a 1, African Highlands; 2, Bargang; 3, Bulwa; 4, Deundi; 5, Diyyagami West; 6, Kangaita; 7, Kapchorua; 8, Kaporet; 9, Limbuli; 10, Litein; 11, Lupembe; 12, Makwasa; 13, Mukumbani; 14, Nyankoba; 15, Ogembo; 16, Phulbari; 17, Ragalla; 18, Ratelshoek; 19, Tenduet; 20, Tombe; 21,Torokahuna; 22, Toromityana; 23, Tshivase.

and tsetse-fly distribution in Zimbabwe (Ripley, 1993). Other statistical classification procedures, such as K nearest neighbour (KNN), have been reported to be as effective as neural networks (Ripley, 1993) and derive the solution in a considerably shorter time. However, in the pattern recognition of jet fuel chromatographic data (Long *et al.*, 1991) and simulated active sonar waveforms (Chen, 1991), neural networks offered improved results over the KNN and soft independent modelling of class analogy (SIMCA) techniques.

Examination of the weight matrix from a trained neural network has been proposed as a method of determining which input neurons are more vital for the determination of the network output values (Long *et al.*, 1991; Thai & Shewfelt, 1991). Inhibitory weights with a negative value indicate the regions which are not contributing to the outcome of the network. In particular, two layer networks have been proposed as this simplifies the matrix (Thai & Shewfelt, 1991).

Table 6 shows a two-layer weight matrix for Set (a);

Table 3. Classification of 96 black teas (Set c) by country of origin[a]

Result	Actual								
	K	T	U	I	B	S	M	SA	Z
Neural network model									
K	85	14							
T	10	72	12						
U	5		88					50	
I				100					
B					100				
S		14				86	20		
M						14	80		
SA									
Z								50	
Statistical model									
K	90								
T	5	86	12	4					
U			88				20		
I				89					
B				4	100				
S	5	14		4		86	20		
M						14	60		
Z/SA								100	

[a] K, Kenya; T, Tanzania; U, Uganda; I, India; B, Bangladesh; S, Sri Lanka; M, Malawi; SA, Republic of South Africa; Z, Zimbabwe

the weights greater than 1·0 are highlighted in bold type. This suggests which input neurons (phenolic constituents) the network associates with each black tea origin. The network indicates that discrimination involves all the chemical components. The most positive weight for each tea was as follows: TF4 (Assam), FG3 (Bangladesh), TF2 (Kenya, Northern India), X1,4

Table 4. Comparison of neural network and statistical models for classifying 158 samples of resin acids (Set d)[a]

	Actual provenance								
	1	2	3	4	5	6	7	8	9
1	79					6		14	
2		66							
3		16	83	18					
4		6	6	76	26				
5	5	6	11	6	69		5		
6	11				5	88	11	14	6
7	5	6				6	79		12
8								65	6
9					5		5	7	76
Statistical model									
1	63		6				11	21	17
2	5	44		18	11		5		
3		17	56	29	21				6
4	5	6	22	35	26		11		
5		27	16	18	37		11		
6	12					35	15	21	24
7	5	6				35	21		6
8	11					18	11	37	12
9					5	11	15	21	35

[a] 1, *P. caribaea* var. *caribaea*; 2–5, *P. caribaea* var. *bahamensis*; 6–9, *P. caribaea* var *hondurensis*.

Table 5. Percent correct classification of black tea and resin acid samples by neural and statistical methods (Sets a–d)

Data set	Commodity	Correct %	
		Neural	Statistical
(a) Country	Tea	69	58
(b) Estate	Tea	68	64
(c) Country	Tea	91	89
(d) Provenance	Pine	76	41

(Malawi) and FG5 (Sri Lanka and Tanzania). In the case of the Assam, Northern India and Kenya teas, the theaflavin components TF2 and TF4 were important in their classification. This selection is supported by commercial tea tasters who assign high scores for brightness and briskness in these teas and that the theaflavin components correlate with these attributes (McDowell et al., 1991). It should be noted that this matrix relates only to linear relationships. The optimum network required the addition of two hidden layers, suggesting that non-linear relationships are necessary for optimal discrimination between the regions. Similar trends were noted for the other data sets.

A further factor to consider when assessing the weight matrix of a neural network is that the weights are assigned only within the context of the training set used. For example, the authors have trained a simple network to classify fruit and vegetables by shape and size (i.e. round, large, small, etc.) and colour (i.e. red, green, yellow, etc.). In this training set, tomatoes were described as round, small and red. Since tomatoes were the only red commodity in the set, inspection of the weight matrix revealed that the network used only this attribute to classify them, ignoring the shape and size characteristics. However, the size and shape of tomatoes does become important in their classification when strawberries are included in the training set, as these are also red, but have a different shape.

Table 6. Weight matrix for country prediction of 77 black teas by phenolic composition (Set a)[a]

	A	B	K	M	NI	SL	T
X1,4	**1·37**	−7·99	0·22	**6·23**	**1·09**	0·22	−3·16
X8	−1·44	−3·00	**1·22**	**2·65**	−1·18	−0·12	−0·16
X5,7	0·10	0·25	−3·19	−1·10	0·85	0·12	−1·88
FG1,2	**2·32**	−2·05	**2·84**	**5·96**	−0·70	−2·41	−1·37
FG3	−2·74	**2·53**	**1·62**	**1·01**	−1·50	**2·50**	**1·89**
FG4	−2·25	−4·53	−1·30	0·04	−2·30	**3·41**	**1·89**
FG5	−0·19	−2·81	**1·37**	−5·19	0·74	**4·32**	**5·00**
F,TF1	0·03	−4·25	**2·35**	−0·70	−2·09	−4·80	0·42
TF2	−2·96	−1·79	**3·22**	−3·30	**3·06**	−0·52	−4·27
TF3	−0·41	−0·69	**1·39**	−2·05	0·70	−2·90	−0·93
TF4	**3·36**	−3·01	−1·43	−6·06	−0·61	1·23	−3·63

[a] X1,4, flavonol glycoside; X5,7, unknown; X8, flavonol glycoside; FG1,2, rutin and isoquercetin; FG3, quercetin glycoside; FG4 and FG5, kaempferol glycosides; F,TF1, flavonol aglycone and theaflaven; TF2, theaflavin-3-gallate; TF3, theaflavin-3'-gallate; TF4, theaflavin-3,3'-digallate; A, Assam; B, Bangladesh; K, Kenya; M, Malawi; NI, Northern India; SL, Sri Lanka; T, Yanzania.

Regressions of black tea for price and score as determined by professional tea tasters and phenolic constituents by their chemical composition

Three data sets, derived from HPLC, GC and sensory evaluation were used to compare the effectiveness of neural networks against statistical regression methods (multivariate stepwise) currently used at NRI.

Considering the statistical analysis of Sets e, f and g, stepwise regression (McDowell *et al.*, 1991), using an *F*-ratio of 4 as an inclusion and exclusion criterion for all potential phenolic components or sensory attributes, was used. For Sets e and f, the phenolic compounds TF2, FG5, F,TF1, TF4 and FG5 were suggested respectively. A more complex stepwise regression analysis on the intra-origin modelling effect (inclusion and exclusion *F*-ratio of 4) on Sets e and f suggested the phenolic components F,TF1 and X8, respectively; this model allows for the fact that the professional tea taster may know the origin of each tea before tasting them. Considering the sensory evaluation of the teas by a trained consumer panel (Set g), stepwise regression indicated the sensory panel attribute red as a predictor of the professional tasters terms strength and colour, and bitter as predictors of the professional tasters attributes of brisk and bright, quality and flavour and with milk.

For each data set, the optimum network architectures were 12, 12, 1 for Set e and 12, 12, 1 for Set f. Seven additional input neurons were included in these models to cater for the intra-regional effects; the network architectures were 19, 19, 1 for Set (e; intra-origin), and 19, 19, 1 for Set (f; intra-origin). For Set g the architecture used was 12, 12, 4. In each case, one or two hidden layers were required for the network to train effectively.

For both the neural and statistical models, the closeness of fit of the predicted results to the true results was estimated by linear regression and comparing the correlation coefficient (R^2).

Table 7 shows the correlation for predictions based on the statistical and neural network models. For all the data sets compared, the statistical regression method was more accurate and consistent in predicting the price and score of the black teas from the chemical and sensory panel data. It is noteworthy that the inclusion of the intra-origin effect for score improved the fit and that this was most marked for the neural model.

The neural network models were retrained using input neurons that corresponded to those selected by the stepwise regression procedure. The correlations are shown in Table 7. The results show an overall improved predictive capability of the network. Additionally, limiting the size of the network also reduced the training time.

Neural networks have been compared with linear regression methods in sensory evaluation of tomato and peach (Thai & Shewfelt, 1991), in engineering (Nui *et al.*, 1991) and in spectroscopic analysis (Gemperline *et al.*, 1991). The statistical regression methods had a slight advantage in numerical accuracy, but the neural computing techniques involved fewer steps during the analysis phase; this was particularly so when more than one parameter was predicted. This study has found similar results and suggests that stepwise regression could be used to 'prune' or 'trim' redundant input neurons.

Table 7. Comparison of statistical and neural network regression models

Data set	Correlation (R^2)		
	Statistical model	Neural model (all inputs)	Neural model (stat. inputs)
(e) General	74·8	70·3	76·0
Intra-origin	72·7	71·9	75·8
(f) General	32·1	7·4	26·2
Intra-origin	67·8	62·2	56·5
(g) Strength and colour	49·9	10·2	38·9
Bright and brisk	32·1	4·8	19·8
Quality and flavour	26·8	17·9	14·2
With milk	40·3	27·2	22·7

Two-layer networks as a method of reducing the number of input neurons

Examination of the weight matrix from a trained neural network has been proposed as a method of determining which input neurons are not contributing to network output (Long *et al.*, 1991; Thai & Shewfelt, 1991); inhibitory weights with a negative value indicate the regions which are not contributing to the outcome

Table 8. Weight matrix of phenolic components in black tea for the prediction of price[a]

	General model	Intra-regional
X1,4	1·17	−0·08
X5,7	−2·22	−0·51
X8	−0·38	−0·86
FG1,2	0·11	0·23
FG3	−0·41	0·76
FG4	−1·09	−0·40
FG5	−0·89	−0·69
F,TF1	1·89	2·44
TF2	1·94	2·17
TF3	1·84	0·27
TF4	1·67	−0·99
A	—	0·93
B	—	−0·10
K	—	−0·90
M	—	0·02
NI	—	1·06
SL	—	−0·62
T	—	−1·42
Φ	−1·08	−1·23

[a] X1,4, flavonol glycoside; X5,7, unknown; X8, flavonol glycoside; FG1,2, rutin and isoquercetin; FG3, quercetin glycoside; FG4 and FG5, kaempferol glycosides; F,TF1, flavonol aglycone and theaflavin; TF2, theaflavin-3-gallate; TF3, theaflavin-3'-gallate; TF4, theaflavin-3,3'-digallate; A, Assam; B, Bangladesh; K, Kenya; M, Malawi; NI, Northern India; SL, Sri Lanka; T, Tanzania; Φ, bias.

Table 9. Weight matrix of phenolic components in black tea for the prediction of score[a]

	General model	Intra-regional
X1,4	0·64	0·38
X5,7	−0·08	−0·80
X8	−0·58	−1·38
FG1,2	−2·37	0·05
FG3	−0·36	−0·17
FG4	0·65	−0·42
FG5	1·82	0·66
F,TF1	0·96	0·33
TF2	0·37	0·41
TF3	0·36	−0·28
TF4	2·74	0·59
A	—	1·16
B	—	−0·62
M	—	−0·85
NI	—	0·27
Φ	−2·14	0·59

[a] X1,4, flavonol glycoside; X5,7, unknown; X8, flavonol glycoside; FG1, 2, rutin and isoquercetin; FG3, quercetin glycoside; FG4 and FG5, kaempferol glycosides; F,TF1, flavonol aglycone and theaflavin; TF2, theaflavin-3-gallate; TF3, theaflavin-3'-gallate; TF4, theaflavin-3,3'-digallate; A, Assam; B, Bangladesh; M, Malawi; NI, Northern India; Φ, bias.

of the network. In particular, two-layer networks have been proposed, as this simplifies the examination of the matrix (Thai & Shewfelt, 1991).

Examples of weight matrices from trained two-layer networks are given for each of the above data Sets (e, f, g) in Tables 8–10.

Considering Table 8 (Set e) for the prediction of tea price from the phenolic composition, in the general model, the network weights for F,TF1, TF2, TF3 and TF4 had the largest positive values compared to F,TF1, TF2 and FG5 as selected by stepwise regression. In the intra-regional model, F,TF1 and TF2 were the most important compared to F,TF1 as chosen by stepwise regression.

Considering Table 9 (Set f) for the prediction of tea score from the phenolic composition, in the general model, the network weights for FG5 and TF4 had the largest positive values and agreed with those chosen by stepwise regression. In the intra-regional model, the network chose the same phenolic constituents as in the general model whereas stepwise regression selected X8.

Regarding Table 10 (Set g) for the prediction of tea score from sensory panel attributes, the largest weight for strength and colour was red and for all the other score terms the dominant weight was for bitter except for milk, where strength had the same weight. This is in good agreement with those sensory attributes chosen by stepwise regression.

This suggests that the largest positive weights selected by inspection of the matrix from a two-layer neural network are similar to those chosen by stepwise regression methods. This may provide a method for simplifying the architecture of neural networks.

Table 10. Weight matrix of black tea sensory attributes for the prediction of score

Sensory panel attributes	Professional tea-taster attributes			
	Strength and colour	Bright and brisk	Quality and flavour	With milk
Bitterness	−0·62	2·74	4·70	1·95
Brightness	0·46	2·31	0·79	1·52
Brownness	−0·77	−1·74	−0·69	−0·67
Clarity	−1·31	−2·09	−2·46	−2·12
Earthy	0·31	0·12	0·42	0·87
Flowery	1·28	1·02	1·38	0·96
Liveliness	0·92	0·14	−0·32	−0·77
Redness	1·70	−1·47	−1·31	−0·60
Smoky	1·36	0·27	−0·71	−0·17
Sour	−0·73	1·20	0·92	0·23
Strength	0·41	2·06	0·21	1·96
Astringency	0·49	−0·86	−0·23	−0·45
Φ	−1·87	−0·71	−0·97	−0·67

Training time required by neural networks

The time necessary to train a neural network increases with the number of examples in the training set and the number of neurons in the model. Of the training sets used, the training time varied between 1·5 and 60·0 min but, once trained, prediction by the network is fast; only one pass through the network is needed. The computer used can also make a large difference; for example, a set that required 70 min training using an IBM XT needed only 4·5 min using an Elonex PC-320X.

Ripley (1993) has shown that comparable statistical procedures require considerably less time to form a model than that by neural networks. In this study, where the training set is small and the neural architecture simple, neural networks can be trained in a short time (1·5 min). Furthermore, in regression analysis where the outcome of several parameters is predicted, a simplified single neural network model may replace the need for a number of statistical models.

CONCLUSION

Neural networks performed as effectively as statistical methods for classifying black tea samples by their phenolic composition. In the case of the xylem resin samples, the neural network offered a vast improvement; it is thought this may be because of the non-normality of the resin acid composition.

Compared to regression methods for the prediction of black tea score and price from their chemical and sensory panel attributes, in general, the neural network was less accurate. The accuracy of the neural network improved and the training time reduced when it was trained using the variables chosen by stepwise regression. Where accuracy is not the most important criterion (for example, in quality control), the neural network could have the advantage in predicting several parameters simultaneously.

Neural networks have been described as a 'black

box' and this clearly is a disadvantage for understanding the basis on which networks make their selection (Ripley, 1993). The most positive weights derived from two-layer trained neural networks approximated to the optimal variables selected by stepwise regression and this could be used as a method for assessing their selection criteria.

It is suggested that, alongside statistical procedures, neural networks are a useful tool in pattern recognition and regression of chromatographic (GC and HPLC) and sensory data. Furthermore, neural networks are simpler to use by the newcomer, as a sound knowledge of statistical methods is not required for their operation.

REFERENCES

Anon. (1992). Designing an electronic wine-taster using neural networks. *Electronic Engng.*, **March**, 37–41.

Birks, J. & Kanowski, P. (1988). Interpretation of the composition of coniferous resin. *Silvae Genetica*, **37**, 29–39.

Chang, S. M., Iwasaki, Y., Suxuki, M., Tamiya, E., Karube, I. & Muramatsu, H. (1991). Detection of odorants using an array of piezoelectric crystals and neural network pattern recognition. *Analytica Chimica Acta*, **249**, 323–9.

Chen, C. H. (1991). On the relationships between statistical pattern recognition and artificial neural networks, *Int. J. Pattern Recognition and Artificial Intelligence*. **5**, 655–61.

Coppen, J. J. W., Gay, C., James, D., Robinson, J. M. & Mullin, L. J. (1993). Xylem resin composition and chemotaxonomy of three varieties of *Pinus caribaea*. *Phytochemistry*, in press.

Galvin, J, R., Waldrop Jr, H. L. (1990). The future of sensory evaluation in the food industry. *Food Technol.*, **44**, 95–6.

Gemperline, P. J., Long, J. R. & Gregoriou, V. G. (1991). Nonlinear multivariate calibration using principle components regression and artificial neural networks. *Anal. Chem.*, **63**, 2313–23.

Kamal, N., Karna, N. & Breen, D. M. (1989). An artificial neural networks tutorial: part 1—basics. *Neural Networks*, *1*, 4–22.

Lawrence, J. (1991). *Introduction to Neural Networks*, ed S. Luedeking. California Scientific Software, CA, USA, p. 128.

Lippmann, R. P. (1987). An introduction to computing with neural networks. *IEEE ASSP Magazine*, **April**, 4–21.

Long, J. R., Gregoriou, V. G. & Gemperline, P. J. (1990). Spectroscopic calibration and quantitation using artificial neural networks. *Anal. Chem.*, **62**, 1791–7.

Long, J. R., Mayfield, H. T. & Henley, M. V. (1991). Pattern recognition of jet fuel chromatographic data by artificial neural networks with back-propagation error. *Anal. Chem.*, **63**, 1256–61.

McDowell, I., Feakes, J. & Gay, C. (1991). Phenolic composition of black tea liquors as a means of predicting price and country of origin. *J. Sci. Food Agric.*, **55**, 627–41.

Nui, A., Li, D. L., Wan, L. & Wee, W. G. (1991). A comparative study of model fitting by using neural networks and regression. *Proc. SPIE-Int. Soc. Optical. Eng.*, **1469**, 495–505.

Persaud, K. C. (1992). Electronic gas and odour detectors that mimic chemoreception in animals. *Trends in Anal. Chem.*, **11**, 61–7.

Ripley, B. D. (1993). Statistical aspects of neural networks. In *Networks and Chaos-Statistical and problematic aspects*, ed O. Barndorff-Nielsen, J. L. Jensen & W} S. Kendall, Chapman and Hall, UK, London.

Sellers, J., York, W., Albersheim, P., Darvill, A. & Meyer, B. (1990). Identification of the mass spectra of partially methylated alditol acetates by artificial neural networks. *Carbohydrate Res.*, **207**, c1–c5.

Song, X. H., Chen, Z. & Yu, R. Q. (1992). Artificial neural networks applied to the classification of complex fluorides based on the transition emission of the europium (II) ion. *Chemometrics and Intelligent Lab. Systems*, **16**, 213–19.

Taylor, S., Baker, D., Owour, P., Othieno, C. & Gay, C. (1992), A model for predicting black tea quality from the carotenoid and chlorophyll composition of fresh green tea leaf, *J. Sci. Food Agric.*, **58**, 185–91.

Thai, C. N. & Shewfelt, R. L. (1991). Modeling sensory colour quality of tomato and peach: Neural networks and statistical regression. *American Soc. Agric. Eng.*, **34**, 950–5.

Thomsen, J. U. & Meyer, B. J. (1989). *J. Magn. Reson.*, **84**, 21–217.

Yoon, Y., Swales, G. & Margavio, T. M. (1993). *J. Opl. Res. Soc.*, **44**, 51–60.