INT. J. PROD. RES., 2004, VOL. 42, NO. 1, 101–118

Taylor & Francis
Taylor & Francis Group

Applying artificial neural networks and virtual experimental design to quality improvement of two industrial processes

X. SHI†*, P. SCHILLINGS† and D. BOYD†

Artificial neural networks (ANNs) are powerful tools to model the non-linear cause-and-effect relationships inherent in complex production processes, usually for process and quality control. This paper substantiates the concurrent application of ANNs and virtual design of experiments to quality improvement. For a chemical manufacturing process and a printed circuit board machining process, respectively, empirical ANN models were constructed and validated using historical data, which were further used to predict the outputs of well-designed process settings. The predicted results were then used to perform statistical tests and identify the significant factors and interactions that affect the quality-related output variables. For the production of a resin intermediate, it was revealed that the combination of low water concentration and an appropriate ratio of raw materials increases both the yield and product quality in a synergistic manner. For the machining of printed circuit board slot by a milling cutter, it was concluded that a high forwarding speed was preferred for the better quality of the milled surface. For both cases, the preliminary conclusions lead to the directions of further real-world experiments for quality improvement. The data mining approach integrating ANNs and virtual design of experiments showed great potential to achieve a better understanding of process behaviour and to improve the process quality efficiently.

1. Introduction

To improve the quality of any complex manufacturing process, it is important first to achieve quantitative understanding of process behaviour, such as possible cause-and-effect relationships inherent in the process. In addition to statistically designed experiments, great importance is attached to full utilization of historical data, which were gathered either during the course of production or from ill-designed experiments. Such data, existing in the data repository of companies, are referred to as happenstance data, so called because they were not collected for a well-designed experiment under controlled conditions. However, these happenstance data possess potential value to quality improvement efforts if they contain information on process behaviour. Although happenstance data would be better used to ask questions rather than to reach firm conclusions (Pyzdek 1999), if coupled with appropriate data-mining approaches, they can be used to model the functional relationships in the process and thus draw preliminary conclusions indicating the direction for improvement.

Revision received June 2003.

†Department of Industrial and Management Engineering, 230 Roberts Hall, Montana State University, Bozeman, MT 59717-3900, USA.

*To whom correspondence should be addressed at: 416 Cobleigh Hall, Western Transportation Institute, Bozeman, MT 59717-3900, USA. e-mail: Xianming_S@erc.montana.edu

International Journal of Production Research ISSN 0020–7543 print/ISSN 1366–588X online © 2004 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/00207540310001602937

Artificial neural networks (ANNs) provide non-parametric, data-driven, self-adaptive approaches to information processing. They are powerful in tackling complex, non-linear problems and have been successfully used to model, predict, control and optimize non-linear systems (Takahara *et al.* 1997, Hussain 1999, Nesic and Vrhovac 1999, Inamdar *et al.* 2000, Kim and Kim 2000, Lou and Nakai 2001). ANNs are gaining favour in applications as diverse as forecasting, signal processing, pattern recognition and classification, process control, and decision-support (Goodacre and Kell 1993, Wienke and Buydens 1996, Bode 1998, Schmitt *et al.* 1998, Zhang *et al.* 2001). This may be attributed to their distinguishing features and to the advantages that they hold over traditional, model-based methods (Zhang *et al.* 1998). First, ANNs are robust and can produce generalizations from experience even if the data are incomplete or noisy. Second, ANNs can learn from examples and capture subtle functional relationships among case data. Prior assumption about the underlying relationships in a particular problem, which in the real world are usually implicit or complicated, need not be made. Third, ANNs provide universal approximation functions flexible in modelling linear and non-linear relationships.

The ANN paradigm adopted in this study was the multiplayer feedforward neural network, of which a typical architecture is shown in figure 1. The nodes in the input and output layers consist of independent variables and response variable(s), respectively. One or two hidden layers are included to model the dependency based on the complexity of relationship(s). For a feedforward network, signals are propagated from the input layer through the hidden layer(s) to the output layer, and each node in a layer is connected in the forward direction to every node in the next layer. Every node simulates the function of an artificial neuron. The inputs are linearly summated utilizing connection weights and bias terms and then transformed via a non-linear transfer function.

For the training of the networks, we adopted the error back-propagation (BP) algorithm. All the connection weights and bias terms for nodes in different layers are initially randomized and then iteratively adjusted based on certain learning rules. For each given sample, the inputs are forwarded through the network until they reach the output layer producing output values, which are then compared with the target values. Errors are computed for the output nodes and propagated back to the connections stemming from the input layer. The weights are systematically modified to reduce the error at the nodes, first in the output layer and then in the hidden layer(s). The changes in weights involve a learning rate and a momentum factor and

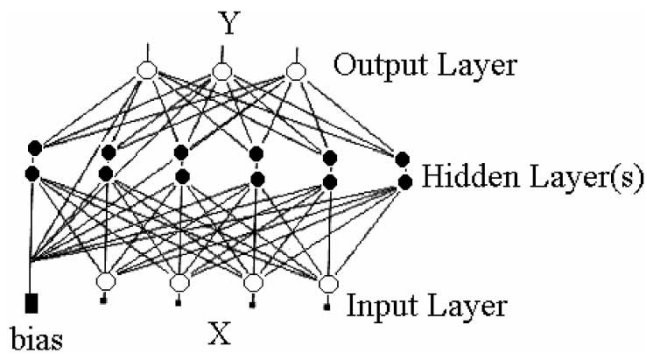


Figure 1. Typical multiplayer feedforward neural network architecture.

are usually in proportion to the negative derivative of the error term. It may take a few thousands of rounds, repeating the feedforward and error back-propagation, before the predicted output gets very close to the target value. The learning process is continued with multiple samples until the prediction error across all samples in the training data is minimized to a reasonable range or stabilized (convergence). Thereafter, knowledge of the network remains encoded in the refined connection weights and bias terms that can be used to recall any trained sample or, if well generalized, to predict unknown input–output pairs. For more details, see Rumelhart *et al.* (1986) and McAuley (1997).

Design of experiments (DOE) (Hicks and Turner 1999) is a well-established methodology that enables the analyst to draw inferences or test hypotheses about an entire population based on a few sampled observations. Usually the advantage of DOE is to minimize the number of experiments required to determine the effects of various factors on the response of a system. The data are obtained via a statistical scheme such that subsequent analysis of variance (ANOVA) can be performed to separate the contribution to the response by various factors from that by natural random errors.

In recent efforts of modelling and predicting functional relationships, response surface methodology (RSM) is the most widely used framework associated with well-designed experiments (Myers and Montgomery 1995, Alauddin *et al.* 1997). Both ANN and classical RSM can be used for response surface modelling. Response is predicted under various process settings within the range of the data used to construct the model. As opposed to ANN, however, classical RSM requires the specification of a polynomial function, usually quadratic, to fit the relationship and thus is incapable to approximate more complicated non-linear relationships. For RSM, sensitivity analysis of input variables is difficult to perform due to the presence of cross interactions (Lou and Nakai 2001). For the same industrial process, the ANN model has been found to fit the data better and to have higher predictive capability than RSM (Spedding *et al.* 1997).

Although designed experiments are usually costly and time consuming, they are collected under tightly controlled conditions and produce valuable information with less noise. Therefore, researchers have attempted to establish robust ANN models through the use of designed data alone (Coit and Smith 1995, Ko *et al.* 1999, Chen 2001) or a combination of designed data and production data (Coit *et al.* 1998). Regardless of the data source used to construct the ANN models, as long as the models can be validated, they can be applied to response surface modelling and be used to optimize process settings (Kustrin *et al.* 1998).

In this paper, we present two case studies to illustrate the concurrent application of ANNs and virtual design of experiments (VDOE) to quality improvement. In one case, we studied the non-linear effects of reaction conditions on yield and colour for an intermediate in a chemical process, in which statistical experimental design was performed to obtain the data. While in the other case, we studied the non-linear effects of the process variables and state variables on the surface quality of printed circuit boards (PCBs) machined by a milling cutter, in which experiments were performed under prespecified conditions but without reasonable randomization.

An ANN was used as a data mining approach to abstract the useful information from existing designed data or happenstance data, in other words, to deduce reliable data from noisy data. From an existing set of data composed of a quality-associated response variable and independent variables that may contribute to the response, we

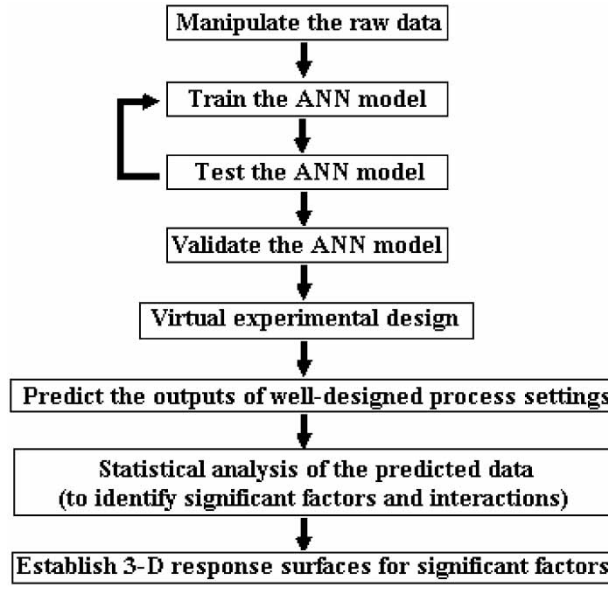


Figure 2. Data mining process diagram.

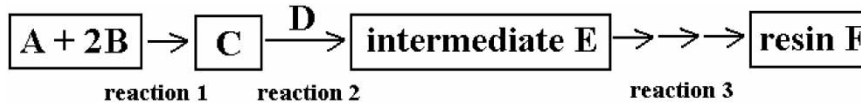


Figure 3. Reaction sequence for the preparation of resin F.

randomly chose two small sets of data as test data and validation data, respectively, and used the rest as training data. The test data were used to monitor the performance of the model during training. Then, the validation data were used to measure the performance of the trained model. Once the empirical ANN model was validated, it was used to predict the outputs of well-designed process settings. The predicted data were further used to perform statistical tests and to identify the significant factors and interactions that affect the quality-associated output. For the significant factors, the established mathematical models were also used to construct the three-dimensional response surfaces. The data mining process is illustrated in figure 2. With the better understanding of the process behaviour, relevant cause-and-effect relationships were quantified and the direction for quality improvement was discussed.

2. Chemical manufacturing process

2.1. *Modelling of the chemical process*

We used the data of Taylor *et al.* (1997) to identify the process parameters that affect the yield and colour of a chemical intermediate at a production plant and to model the corresponding relationships between the process parameters and the responses. As shown in figure 3, the chemical intermediate E for a polymer resin F was manufactured through the reaction of monomer D with monomer C, which was synthesized by the reaction of monomer A with two equivalents of monomer B. Six process parameters potentially affecting the yield and colour of E as suggested by

previous experimentation were chosen by Taylor *et al.* (1997) in an initial screening experiment, i.e. per cent solids, temperature, water concentration, stoichiometry (molecular ratios) of B to A in reaction 1, reaction time and catalyst concentration. The quality targets were high yield of E and low colour reading for E, which were reflected by analytic yield and yellowness index (YI) values, respectively.

Taylor *et al.* performed a 2^{6-1} fractional factorial design for screening studies, yielding 32 different combinations of the extreme values for the six factors. In addition, five replicate trials at the centrepoint of the design (all factors set to the mean of their ranges) were run to supplement the information. The order of conducting the 37 experiments was randomized and table 1 shows the data used to train, test and validate ANN models. The goal of our work was to investigate the scenario from a different point of view.

For the training of feedforward neural networks, we used a modified BP algorithm as described below, using a sigmoid function in equation (1) as the non-linear transfer function and the sum of the mean squared error in the output layer (SMSE) as the convergence criteria. The training of the networks was performed in batch mode. The values of learning rate and momentum factor α were initialized at 0.9 and 0.7, respectively, and they were automatically adjusted during the training process to avoid the trap of local minima while maintaining the features of the BP algorithm. All the data for input and output were normalized based on equation (2), where X_i and NX_i are the i-th value of factor X before and after the normalization, and $X_{\min}$ and $X_{\max}$ are the minimum and maximum of factor X, respectively. The program was written in C language:

$$f(x) = (1 + e^{-x})^{-1} \tag{1}$$

$$NX_{\mathrm{i}} = (X_{\mathrm{i}} - X_{\min} + 0.1)/(X_{\max} - X_{\min} + 0.1). \tag{2}$$

With the 34 training samples in table 1, we established two mathematical models, NN-1 and NN-2, to quantify the relationships between the six investigated factors and the response, analytic yield or YI, respectively. The models were also tested and validated, and table 2 lists the parameters and performance of the ANN models for the chemical process. The number of hidden layers and nodes in them are generally related to the complexity of the relationship. The more complex the relationship, the more layers and nodes are necessary. Usually one or two hidden layers are enough to approximate the reality. In our work, the selection of layers and nodes took into consideration driving the SMSE as small as possible and the training process as efficient as possible.

From the learning results, we can see that the established ANN models have good 'memory' and the trained matrices of interconnected weights and bias reflect the hidden functional relationships very well. Because the test and validation errors of the models were small, the models are reliable for prediction of analytic yield and YI of the intermediate E synthesized under any other combination of process parameters as long as they are within the range we investigated.

2.2. *Prediction and statistical analysis*

To identify the factors and interactions significantly affecting the responses, we varied the level of each parameter at its extreme and centrepoint values and performed a virtual 3^6 factorial design, yielding 729 different combinations for the six factors. Three levels of each factor were explored so that there were enough

	Investigated factors						Responses	
Run no.	Solids (%)	Temperature (°C)	Water added (ppm)	Stoichiometry[s] (%)	Reaction time (min)	Catalyst[c] (mol%)	Analytic yield (%)	YI
1	40	100	500	98	30	10	64	3.9
2	40	100	500	102	10	10	43	7.0
3	10	140	0	98	10	10	99	2.1
4	10	100	0	102	30	1	88	4.0
5	40	100	500	98	10	1	59	3.5
6	40	100	0	102	30	10	92	4.0
7	40	140	500	98	30	1	60	4.3
8	40	140	0	102	30	1	88	4.4
9	40	100	0	98	30	1	94	2.1
10	40	100	500	102	30	1	39	7.4
11	10	140	500	102	10	10	44	7.4
12	10	140	500	98	30	10	64	3.3
13	10	100	500	98	30	1	59	3.9
14	25	120	250	100	20	5.5	80	2.0
15	25	120	250	100	20	5.5	82	2.3
16	40	140	500	102	10	1	40	7.4
17[t]	10	140	0	98	30	1	94	2.5
18	40	100	0	98	10	10	98	1.7
19	40	140	500	102	30	10	44	7.8
20	10	100	500	98	10	10	64	3.5
21	40	100	0	102	10	1	88	3.6
22	10	140	500	102	30	1	40	7.8
23	40	140	500	98	10	10	64	3.9
24	10	100	0	102	10	10	92	3.6
25	25	120	250	100	20	5.5	79	2.2
26	10	100	0	98	10	1	94	1.7
27	40	140	0	98	30	10	99	2.5
28	40	140	0	98	10	1	94	2.1
29[v]	10	140	500	98	10	1	60	3.9
30	10	140	0	102	10	1	83	4.0
31	10	140	0	102	30	10	93	4.4
32	25	120	250	100	20	5.5	82	2.3
33	10	100	500	102	30	10	43	7.4
34	10	100	0	98	30	10	98	2.1
35[t]	40	140	0	102	10	10	93	4.0
36	25	120	250	100	20	5.5	82	2.3
37	10	100	500	102	10	1	39	7.0

[t]Samples 17 and 35 were used as test data; [v]sample 29 was used as validation data; [s]weight per cent excess or deficiency in B from the nominal ratio of B/A = 2.0; [c]mole per cent based on D concentration.

Table 1. Experimental data to train, test and validate ANN models for the chemical process.

	Response	Topological structure	Training SMSE	Testing SMSE	Validation SMSE
NN-1	analytic yield	6-6-1	0.0085	0.019	0.015
NN-2	YI	6-6-1	0.0074	0.032	0.013

Table 2. Parameters and performance of the ANN models for the chemical process.

degrees of freedom (d.f.) left for us to investigate the two- and three-way interactions. While other choices of fractional factorial design such as 3^{6-1}, 3^{6-2} and 3^{6-3} were available, we chose the 3^6 factorial design because the virtual experiments were not limited by real sources other than the computer time. Owing to the nature of virtual experiments, there are no uncontrolled factors (noise) contributing to the responses, and therefore it was not necessary to randomize the order of conducting the 729 experiments. For the same reason, every combination was unreplicated. Levels and controllability of reaction parameters used for prediction of the chemical process are shown in table 3. For the designed (virtual experiment) data, the responses were predicted using the established ANN models and the resulting virtual data were used to determine the effects of the investigated factors on analytic yield and YI.

We used SAS to perform the ANOVA procedure, which was applied to test the statistical hypotheses about the significance of each factor and their two- and three-way interactions to the dependent variable, analytic yield or YI. ANOVA tables for the dependent variables, analytic yield and YI, are summarized in table 4. Of key importance in the tables are the two columns, F ratio and p value, to test the reliability and efficiency of the models. The F ratio works as a surrogate of the signal-to-noise ratio, where the noise is the residual error indicating contribution to the response by higher order terms not included in the three-way ANOVA model. We may consider the p value as the smallest level α at which the data are significant, in other words, it indicates the risk of incorrectly not including the terms affecting the dependent variable. From table 4, the large F ratios and the small p values suggested that the models were highly reliable and included the terms significantly affecting analytic yield and YI.

	Solids (%)	Temperature (°C)	Water added (ppm)	Stoichiometry (%)	Reaction time (min)	Catalyst (mol%)
Prediction levels	10	100	0	98	10	1
	25	120	250	100	20	5.5
	40	140	500	102	30	10
Precision (±) [Taylor *et al.* 1997]	1.00	3.00	5.00	0.25	0.50	0.50

Table 3. Levels and controllability of reaction parameters used for prediction of the chemical process.

	Source	Df	Sum of squares	Mean square	F	p
Analytic yield	model	232	310 173.31	1336.95	901.9	< 0.0001
	error	496	735.24	1.48		
YI	model	232	2688.72	11.59	403.39	< 0.0001
	error	496	14.25	0.029		

Table 4. Summary of ANOVA tables for the dependent variable analytic yield and YI.

Source		d.f.	Sum of squares	Mean square	F^3	p
one way[1]	solids	2	16.20	8.10	5.46	0.0045
	temperature	2	302.84	151.42	102.15	< 0.0001
	water	2	232 465.41	116 232.70	78 412.10	< 0.0001
	stoichiometry	2	51 781.15	25 890.57	17 466.10	< 0.0001
	reaction time	2	19.34	9.67	6.52	0.0016
	catalyst	2	4457.58	2228.79	1503.57	< 0.0001
Two ways[2]	{all}	60	19 457.88	324.30	218.78	< 0.0001
Three ways[2]	{all}	160	1672.91	10.46	7.05	< 0.0001
Error		496	735.24	1.48		
Total		728	310 908.55			

[1]All main effects listed, whether significant or not.
[2]Only part of them were found significant and the details ignored due to the space limit.
[3]Calculated using the residual error.

Table 5. Detailed ANOVA results for the analytic yield model.

Source		d.f.	Sum of squares	Mean square	F^3	p
One way[1]	solids	2	0.17	0.08	2.92	0.0548
	temperature	2	47.32	23.66	823.53	< 0.0001
	water	2	1012.19	506.09	17 615.7	< 0.0001
	stoichiometry	2	1318.25	659.13	22 942.3	< 0.0001
	reaction time	2	51.57	25.79	897.58	< 0.0001
	catalyst	2	1.22	0.61	21.32	< 0.0001
Two ways[2]	{all}	60	206.97	3.45	120.07	< 0.0001
Three ways[2]	{all}	160	51.01	0.32	11.10	< 0.0001
Error		496	14.25	0.029		
Total		728	2702.96			

[1]All main effects listed, whether significant or not.
[2]Only part of them were found significant and the details ignored due to the space limit.
[3]Calculated using the residual error.

Table 6. Detailed ANOVA results for the YI model.

We also obtained the components of the ANOVA tables for analytic yield and YI, and the detailed results for them are shown in tables 5 and 6, respectively. Any factor or factorial interaction was considered significant if they had $p < 0.05$. From table 5, we can see that all the six factors, per cent solids, temperature, water concentration, stoichiometry of B to A, reaction time, and catalyst concentration were significant to the analytic yield. This conclusion differs from the results of Taylor *et al.* (1997), in which they concluded that only water, stoichiometry and catalyst were significant. The difference can be partly explained by the fact that we used the residual error to calculate the F ratios, instead of the 'pure error' calculated from the five replicated trials set to centrepoint settings. In addition, we used the data generated by the ANN model, instead of the raw data themselves, for the statistical tests, which enabled us to test all the 20 three-way interactions between the six factors, instead of only 10 used by Taylor *et al.* The virtual DOE gave us advantages to explore more solution space and to reveal a 'bigger picture' of the functional relationships. Nonetheless, if not applied with caution, the modelling process may also distort the information contained in the data used to construct the model.

	Solids	Temperature	Water added	Stoichiometry	Reaction time	Catalyst
Solids	–	–	–	+	–	–
Temperature		–	+	+	–	+
Water added			–	+	–	–
Stoichiometry				–	–	–
Reaction time					–	+
Catalyst						–

+, Interaction was significant for both responses at 1% level; –, otherwise.

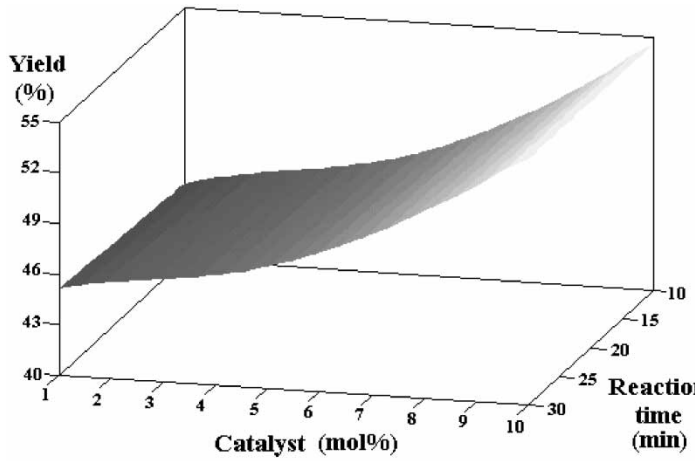
Table 7. Symmetric significance matrix of two-way interactions for both analytic yield and YI.

In that our goal was to produce the intermediate E with low yellowness index in a high yield, we also studied the effects of the six factors and their interactions on YI. From table 6, we concluded that all factors except the per cent solids were significant to YI. Since the per cent solids factor was involved in some of the significant interactions, its effect could not be ignored. To achieve the goals simultaneously, we focused on the interactions that were significant for both analytic yield and YI. Table 7 shows the symmetric significance matrix of two-way interactions for both responses at 1% level. In addition to the six two-way interactions that were significant for both responses, the three-way interactions between water concentration, stoichiometry and temperature or catalyst concentration were found significant for both responses at 1% level.

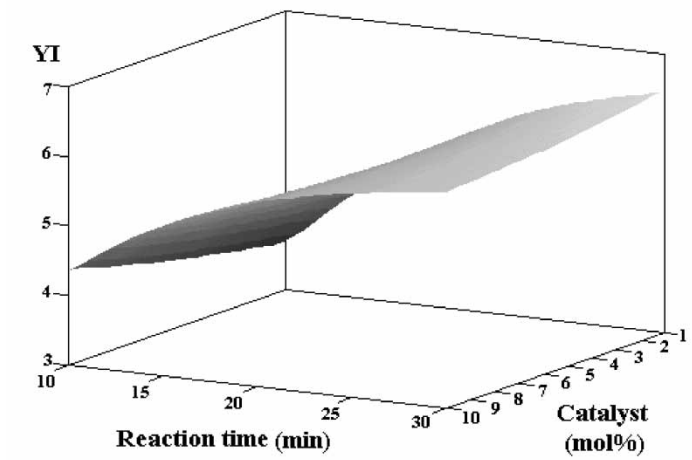
2.3. *Response surface modelling*

As suggested by the statistical tests, all six investigated factors were significant for both analytic yield and YI, and their interactions were complicated. Therefore, we applied the established ANN models to response surface modelling by varying the levels of two factors and setting the other factors at certain levels. For instance, figure 4 shows the ANN predicted response surface of (a) analytic yield and (b) YI over reaction time and catalyst concentration, with the per cent solids, temperature, water concentration and stoichiometry set to 25%, 120°C, 500 ppm and 100%, respectively. We concluded that under the specified settings the analytic yield of intermediate E increased as the catalyst concentration increased, which was consistent with the fact that the chemical reaction was incomplete with a yield less than 55% Owing to the strong interaction between the catalyst concentration and reaction time, the effect of the catalyst concentration on YI was not intuitive and the relationship changed with the reaction time. While the effect of reaction time on yield was not apparent, the shorter reaction time produced intermediate E with lower YI, i.e. higher quality.

Figure 5 shows the ANN predicted response surface of (a) analytic yield and (b) YI over water concentration and stoichiometry, with the per cent solids, temperature, reaction time, catalyst concentration set to 25%, 120°C, 30 min and 10 mol%, respectively. We concluded that the interaction of water concentration and stoichiometry strongly affected the yield and YI when the other four factors were set to the specified levels. The simultaneous adoption of low water concentration and low stoichiometry not only increases the yield of intermediate E dramatically, but also



(a)



(b)

Figure 4. ANN predicted response surface of (a) analytic yield and (b) YI over reaction time and catalyst concentration, with the per cent solids, temperature, water concentration and stoichiometry set to 25%, 120°C, 500 ppm and 100%, respectively.

improves its quality by reducing its YI in a synergistic manner. Using the established model, we predicted that a zero water concentration and a stoichiometry of 100 would give a 'next-to-optimal' yield of 97.27% and a yellowness index of 1.77. Better results were expected by introducing the variation of the other four factors, the per cent solids, temperature, reaction time, and catalyst concentration, even though these four factors were found not significantly affecting the responses when both the water concentration and stoichiometry were set to low levels. With the knowledge in process behaviour achieved, we should be able to design a few real world experiments within the neighbourhood of low water concentration and low stoichiometry to further search for the reaction conditions that optimize both analytic yield and YI.

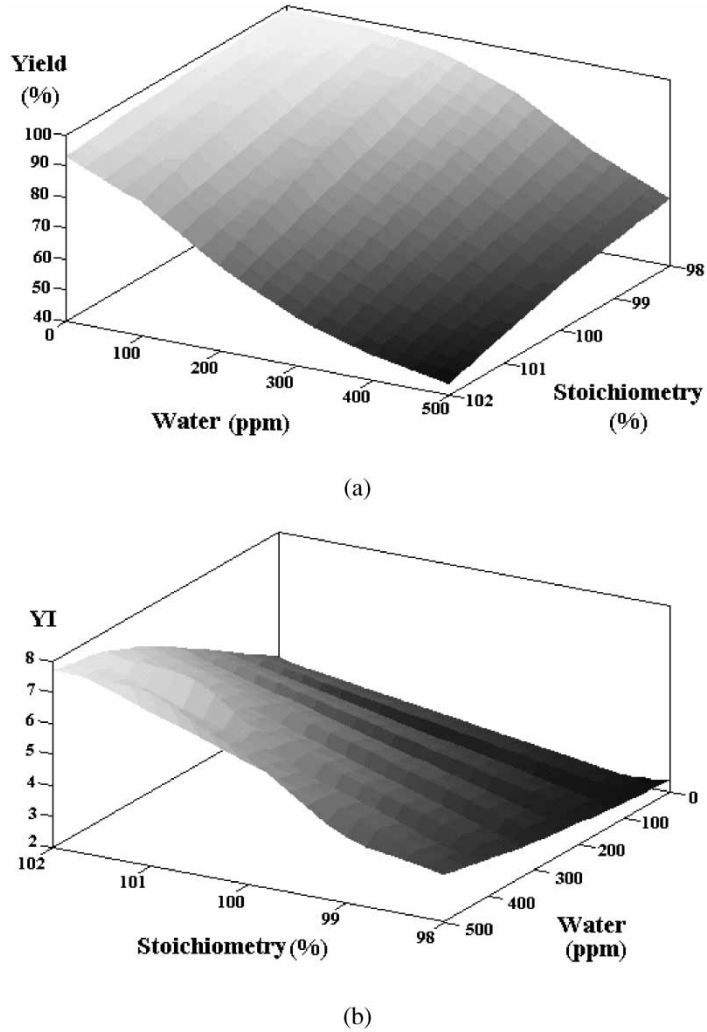


Figure 5. ANN predicted response surface of (a) analytic yield and (b) YI over water concentration and stoichiometry, with the per cent solids, temperature, reaction time, catalyst concentration set to 25%, 120°C, 30 min, and 10 mol%, respectively.

3. PCB machining process

3.1. *Modelling of the machining process*

The machining process of PCBs by a milling cutter involved several process and state variables that may affect the machining quality of PCBs and it was desirable to model the corresponding relationships between the significant independent variables and the quality-related responses. Figure 6 shows a typical milling cutter used in this study, which was a toothed spinning tool that could cut slots in the PCBs. The PCB was generally composed of several layers and the milling cutter was first employed to drill a slot into the PCB layers by spinning at a certain speed and forwarding at another speed. Then, the same cutter was employed to track along the slot again to make the slot surface smooth, during which the spinning and forwarding speeds were the same as the drilling process. After the PCB was machined, threadlike burrs and

Figure 6. Typical milling cutter used for machining PCBs.

'white spots' appeared on the slot surface, and the white spots were glistening spots reflected by the roughness of the surface. The quality target was high surface finish of the PCB slot, which was indicated by the size of burr and white spot at a specific depth. The smaller the burr and white spot, the better the machining quality of the slot.

We used the unpublished data provided by a company specialized in manufacturing hard alloy tools, which is in Guandong, P. R. China. The company performed 32 experiments to offer different combinations of two state variables, the number of PCB layers (NL) and depth of the checkpoint on the drilled slot (DP), and three process variables, spinning speed (V), forwarding speed (F) and milling cutter diameter (D). The state variables differed from the process variables in that they were not controllable for a specific PCB. As shown in table 8, the state variables varied at four different levels and the process variable varied at two different levels. The responses of interest were the height of the burr (HB), and length (LS) and width (WS) of the white spot at the checkpoint on the drilled slot. The order of conducting the 32 experiments was not randomized and the historical data thus were happenstance data.

For the training of feedforward neural networks, we used the same algorithm as described in the previous case study and all the data for input and output were normalized based on equation (2). With the 30 training samples in table 8, we established three mathematical models, NN-3, NN-4 and NN-5, to quantify the relationships between the five investigated factors and the response, HB, LS or WS, respectively. The models were also tested and validated. Table 9 lists the parameters and performance of the ANN models for the machining process.

From the learning results, we can see that the established ANN models have good 'memory' and the trained matrices of interconnected weights and bias reflect the hidden functional relationships very well. Because the test and validation errors of the models were small, the models are reliable for the prediction of burr and white spot sizes for a PCB slot machined under any other combination of the five state and process variables, as long as they are within the range investigated.

3.2. *Prediction and statistical analysis*

To identify the factors and interactions significantly affecting the responses, we varied the state variables at their original four levels and varied the level of each

	Investigated factors					Responses		
Run no.	NL	DP (mm)	V (krpm)	F (m/min)	D (mm)	HB (mm)	LS (mm)	WS (mm)
1	1	0*	32	0.48	1.7	0.05	0.06	0.45
2	1	2.8	32	0.48	1.7	0.40	0.04	0.40
3	2	5.6	32	0.48	1.7	0.05	0.12	0.45
4	2	7.4	32	0.48	1.7	0.95	0.05	0.35
5	3	0*	32	0.48	1.7	0.05	0.06	0.35
6	3	2.8	32	0.48	1.7	0.6	0.08	0.23
7	4	5.6	32	0.48	1.7	0	0.05	0.40
8	4	7.4	32	0.48	1.7	0	0.05	0.40
9	1	0*	32	0.48	1.3	0.10	0.05	0.20
10	1	2.8	32	0.48	1.3	0.05	0.05	0.18
11	2	5.6	32	0.48	1.3	0.50	0.08	0.40
12[t]	2	7.4	32	0.48	1.3	0.20	0.07	0.50
13	3	0*	32	0.48	1.3	0.10	0.06	0.25
14	3	2.8	32	0.48	1.3	0.10	0	0
15	4	5.6	32	0.48	1.3	0.10	0.09	0.42
16	4	7.4	32	0.48	1.3	0.20	0.10	0.42
17	1	0*	40	0.36	1.7	0.05	0.05	0.05
18	1	2.8	40	0.36	1.7	0.04	0.10	0.20
19	2	5.6	40	0.36	1.7	0.03	0.15	0.35
20	2	7.4	40	0.36	1.7	0.03	0.15	0.65
21	3	0*	40	0.36	1.7	0.06	0.08	0.40
22	3	2.8	40	0.36	1.7	0.05	0.10	0.35
23	4	5.6	40	0.36	1.7	0.40	0.10	0.40
24	4	7.4	40	0.36	1.7	0.50	0.10	0.50
25	1	0*	40	0.36	1.3	0.30	0.06	0.35
26	1	2.8	40	0.36	1.3	0.40	0.07	0.50
27[v]	2	5.6	40	0.36	1.3	0.10	0.13	0.50
28	2	7.4	40	0.36	1.3	0.10	0.12	0.45
29	3	0*	40	0.36	1.3	0.50	0.08	0.35
30	3	2.8	40	0.36	1.3	1.00	0.05	0.50
31	4	5.6	40	0.36	1.3	0.30	0.10	0.32
32	4	7.4	40	0.36	1.3	0.20	0.15	0.60

[t]Sample 12 was used as test data; [v]sample 27 was used as validation data;
*DP = 0 defined the starting point of the drilled slot.

Table 8. Experimental data to train, test and validate ANN models for the machining process.

	Response	Topological structure	Training SMSE	Testing SMSE	Validation SMSE
NN-3	HB	5-5-4-1	0.0055	0.015	0.087
NN-4	LS	5-9-1	0.0030	0.004	0.002
NN-5	WS	5-4-4-1	0.0057	0.026	0.012

Table 9. Parameters and performance of the ANN models for the machining process.

process variable at its extreme and centrepoint values, which was a virtual $4^2 \times 3^3$ factorial design yielding 432 different combinations for the five factors. Owing to the nature of virtual experiments, there are no uncontrolled factors contributing to the responses, and therefore the order of conducting the 432 experiments was not necessarily randomized. That is also the reason why every combination was unreplicated. For the designed data, the responses were predicted using the established ANN

models and the resulting virtual data were used to determine the effects of the investigated factors on height of the burr (HB), and length (LS) and width (WS) of the white spot.

We used SAS to perform the ANOVA procedure, which was applied to test the statistical hypotheses about the significance of each factor and their two- and three-way interactions to the dependent variable, MH, LS or WS. Compared with the previous case, the F ratios are smaller and the p values larger in the ANOVA tables, which suggest that the functional relationships between the independent variables and responses are much more complex in the machining process and higher order terms ignored by the three-way ANOVA models may contribute to the responses. The three-way ANOVA models can be used to identify the significant factors and interactions since the p values are under the 5% level. Analysis of the components of the ANOVA tables indicated that all the five factors, number of PCB layers (NL), depth of the checkpoint (DP), spinning speed (V), forwarding speed (F) and milling cutter diameter (D), were significant to the three responses. Since our goal was to produce the PCB slot with low values in not only HB, but also LS and WS, we focused on the interactions that were significant for the three responses simultaneously. At the 5% level, there were two two-way interactions significant for all the responses, namely the one between the number of PCB layers (NL) and depth of the checkpoint (DP), the other between the forwarding speed (F) and milling cutter diameter (D). In addition, only the three-way interaction between depth of the checkpoint (DP), forwarding speed (F) and milling cutter diameter (D) was significant for all the responses at 5% level.

3.3. *Response surface modelling*

As suggested by the statistical tests, all five investigated factors were significant for the size of both the burr and white spot and their interactions were complicated. Therefore, we applied the established ANN models to response surface modelling by varying the levels of two factors and setting the other factors at certain levels. For instance, figure 7 shows the ANN predicted response surfaces of burr height (HB), white spot length (LS) and white spot width (WS) over forwarding speed (F) and milling cutter diameter (D) for a slot of depth in 0 mm in a four-layer PCB, with spinning speed (V) set to 40 krpm. We concluded that under the specified settings, the burr height and white spot length of the machined slot significantly decreased as the forwarding speed increased, whereas the influence of forwarding speed on the white spot width was less apparent. When machining the slot at low forwarding speeds, the burr height tended to decrease as the cutter diameter increased, and this relationship vanished at high forwarding speeds. Because of the strong interactions between forwarding speed and milling cutter diameter, the effects of the cutter diameter on the sizes of white spot were not intuitive and the relationships changed with the forwarding speed. Figure 7 also indicated that under the specified settings operating a milling cutter of diameter in 1.35–1.45 mm at a forwarding speed of 0.48 m/min gave very low values in HB and LS simultaneously, whereas the white spot width was still more than 0.25 mm.

Figure 8 differs from figure 7 in that the depth of checkpoint (DP) was 7.4 instead of 0 mm, while every other factor remained the same. We concluded that under the specified settings, the burr height of the machined slot significantly decreased as the forwarding speed increased, whereas the influence of cutter diameter was not significant. Increasing the forwarding speed and the cutter diameter tended to decrease

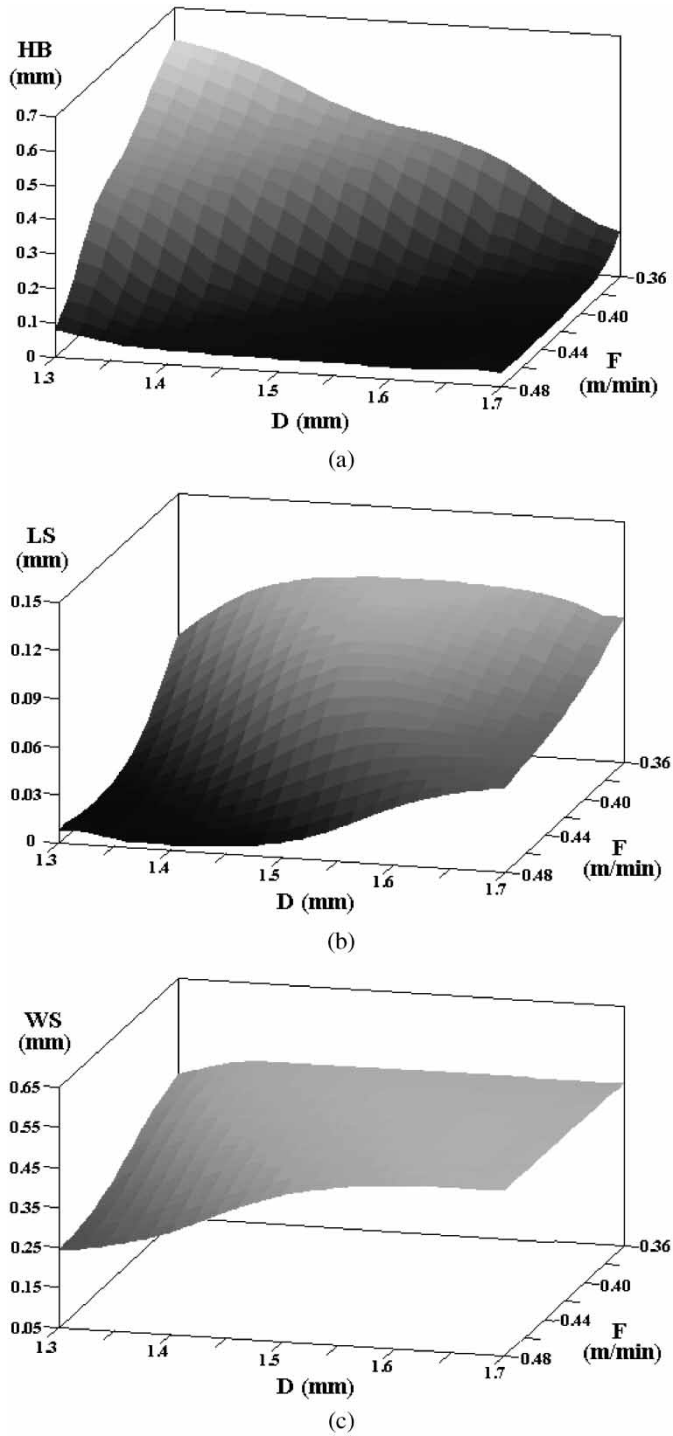


Figure 7. ANN predicted response surface of (a) burr height (HB), (b) white spot length (LS) and (c) white spot width (WS) over forwarding speed (F) and milling cutter diameter (D) for a slot of depth in 0 mm in a four-layer PCB, with spinning speed (V) set to 40 krpm.

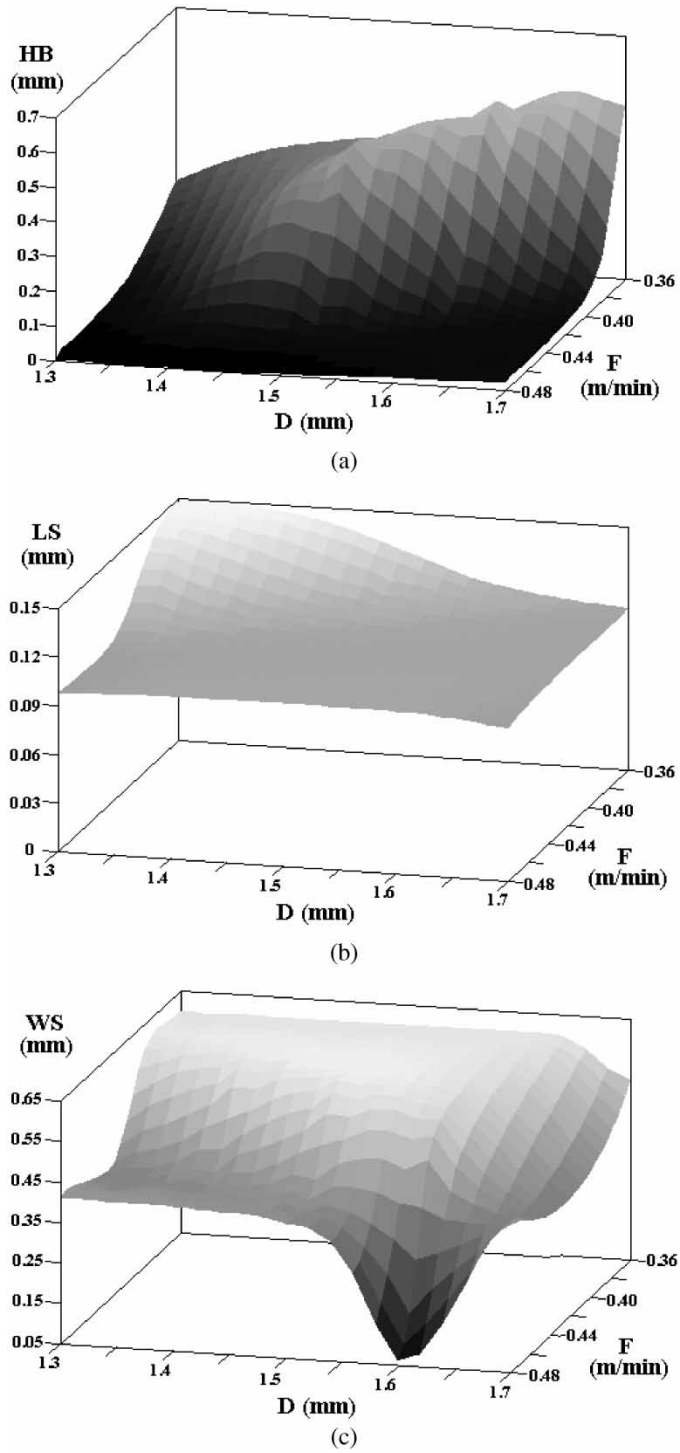


Figure 8. ANN predicted response surface of (a) burr height (HB), (b) white spot length (LS) and (c) white spot width (WS) over forwarding speed (F) and milling cutter diameter (D) for a slot of depth in 7.4 mm in a four-layer PCB, with spinning speed (V) set to 40 krpm.

the white spot length and the tendencies vanished when the cutter was in large diameter and/or was operated at high forwarding speed. While the effects of forwarding speed and cutter diameter on the white spot width were more complex, the smallest width could be achieved with a combination of a forwarding speed of 0.48 m/min and a diameter of 1.6 mm. Such a combination gave very low values in HB and WS simultaneously, whereas the white spot length was still more than 0.09 mm.

Therefore, we concluded that a high forwarding speed was generally preferred to decrease the sizes of burrs and white spots, namely, to achieve better surface quality for the machined slot. Although better results were expected by introducing the variation of the other factors, it was impossible to achieve low values in HB, LS and WS simultaneously within the investigated scopes. With the knowledge in process behaviour achieved, however, we can design a few real-world experiments with reasonably high forwarding speed to search further for the process parameters that minimize the size of both the burr and white spot for a PCB slot under certain state conditions.

4. Conclusions

For the production of intermediate E, the combination of low water concentration and low stoichiometry of B to A not only increases the yield of intermediate E dramatically, but also improves its quality by reducing its yellowness in a synergistic manner. These conclusions differ from the earlier reference, and give a more comprehensive picture of the functional relationships inherent in the chemical manufacturing process.

For the machining of PCB slot by a milling cutter, a high forwarding speed was generally preferred to decrease the sizes of burrs and white spots, namely, to achieve better surface quality of the slot. It was found impossible to achieve low burr heights, white spot lengths and white spot widths simultaneously within the investigated scopes and more real-world experiments are necessary to ensure better quality.

ANNs were successfully applied to model the complex functional relationships inherent in two industrial processes using existing data and the established models were used to predict the quality-associated responses, to identify the significant factors and interactions, and to achieve better understanding of the process behaviour. The preliminary conclusions derived from historical data lead to the direction of further real-world experiments for quality improvement.

References

Alauddin, M., Baradie, M. A. and El, Hashmi, M. S. J., 1997, Prediction of tool life in end milling by response surface methodology. *Journal of Materials Processing Technology*, **71**, 456–465.

Bode, J., 1998, Decision support with neural networks in the management of research and development: Concepts and application to cost estimation. *Information and Management*, **34**, 33–40.

Chen, M. C., 2001, Tolerance synthesis by neural learning and non-linear programming. *International Journal of Production Economics*, **70**, 55–5.

Coit, D. W., Jackson, B. T. and Smith, A. E, 1998, Static neural network process models: Considerations and case studies. *International Journal of Production Research*, **36**, 2953–1968.

Coit, D. W. and Smith, A. E., 1995, Using designed experiments to produce robust neural network models of manufacturing processes. Proceedings of the Fourth Industrial Engineering Research Conference (IERC), Nashville, TN, USA.

GOODACRE, R. and KELL, D. B., 1993, Rapid and quantitative analysis of bioprocesses using pyrolysis mass spectrometry and neural networks: application to indole production. *Analitica Chimica Acta*, **279**, 17–26.

HICKS, C. R. and TURNER, K. V., 1999, *Fundamental Concepts in the Design of Experiments*, 5th edn (Oxford: Oxford University Press).

HUSSAIN, M. A., 1999, Review of the applications of neural networks in chemical process control — simulation and online implementation. *Artificial Intelligence in Engineering*, **13**, 55–68.

INAMDAR, M., DATE, P. P., NARASIMHAN, K., MAITI, S. K. and SINGH, U. P., 2000, Development of an artificial neural network to predict springback in air vee bending. *International Journal of Advances in Manufacturing Technology*, **16**, 376–381.

KIM, D. J. and KIM, B. M., 2000, Application of neural network and FEM for metal forming processes. *International Journal of Machine Tools and Manufacture*, **40**, 911–925.

KO, D. C., KIM, D. H. and KIM, B. M., 1999, Application of artificial neural network and Taguchi method to perform design in metal forming considering workability. *International Journal of Machine Tools and Manufacture*, **39**, 771–785.

KUSTRIN, S. A., ZECEVIC, M., ZIVANOVIC, L. J. and TUCKER, I. G., 1998, Application of neural networks for response surface modeling in HPLC optimization. *Analitica Chimica Acta*, **364**, 265–273.

LOU, W. and NAKAI, S., 2001, Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH, and water activity. *Food Research International*, **34**, 573–579.

MCAULEY, D., 1997, The back-propagation network: learning by example [http://psy.uq.oz.au/~brainwav/Manual/BackProp.html].

MYERS, R. H. and MONTGOMERY, D. C., 1995, Response Surface Methodology: Process and Product Optimization Using Designed Experiments (New York: Wiley).

NESIC, S. and VRHOVAC, M., 1999, A neural network model for CO_2 corrosion of carbon steel. *Journal of Corrosion Science Engineering* [http://www.cp.umist.ac.uk/JCSE/)1,Paper 6].

PYZDEK, T., 1999, Virtual-DOE, data mining and artificial neural networks [http://www.QualityAmerica.com/knowledgecente/articles/PYZDEKneural.htm].

RUMELHART, D. E., HINTON, G. E. and WILLIAMS, R. J., 1986, Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland and the PDP Research Group (eds), *Parallel Distributed Processing* (Cambridge, MA: MIT Press), pp. 318–362.

SCHMITT, J., UDELHOVEN, T., FLEMMING, H. C. and NAUMANN, D., 1998, Stacked spectral data processing and artificial neural networks applied to FTIR and FT-Raman spectra in biomedical applications. *Proceedings of SPIE 3257*, pp. 45–58.

SPEDDING, T. A. and WANG, Z. Q., 1997, Study on modeling of wire EDM process. *Journal of Materials Processing Technology*, **69**, 18–28.

TAKAHARA, J., TAKAYAMA, K., ISOWA, K. and NAGAI, T. 1997, Multi-objective simultaneous optimization based on artificial neural network in a ketoprofen hydrogel formula containing *O*-ethylmenthol as a percutaneous absorption enhancer. *International Journal of Pharmaceutics*, **158**, 203–210.

TAYLOR, R. W., BARREN, J. P., NICK, R. J. and CAWSE, J. N., 1997, Non-linear effects on yield and colour for an intermediate in an industrial process. *Polymer Testing*, **16**, 75–89.

WIENKE, D. and BUYDENS, L., 1996, Adaptive resonance theory based neural network for supervised chemical pattern recognition (FuzzyARTMAP), Part 1: Theory and basic properties. *Chemometrics and Intelligent Laboratory Systems*, **32**, 151–164.

ZHANG, G., PATUWO, B. E. and HU, M. Y., 1998, Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, **14**, 35–62.

ZHANG, G. P., PATUWO, B. E. and HU, M. Y., 2001, A simulation study of artificial neural networks for non-linear time-series forecasting. *Computers and Operations Research*, **28**, 381–396.