AU J.T. 3(3): 171-180 (Apr. 2000)

A Six-Sensor Monitor for Identification of Oriental Sauces Using an Artificial Neural Network*

David T. Win
Faculty of Engineering, Assumption University
Bangkok, Thailand

Abstract

A handheld monitor consisting of six tin-oxide semiconductor sensors was used for detecting organic volatiles in oriental sauces. Identification of the sauces by data processing with an artificial neural net was performed on sensor peak height data recorded and was found to be better than 90% efficient. The use of duplicate data sets allowed the identification of 14 sauces. Net training was slower with a decreased number of sensors, larger net size (number of test samples) and increased sensor response pattern similarity.

Samples tested were Cornwell's, Kikkoman (sr) salt reduced, Yeo's, Pearl River Bridge (PRB), Zu Miao, Farmland soy sauces; Songai and Pearl River Bridge mushroom (MushPRB) soy sauces; Spiral and Pureharvest (PH) tamari sauces; Pureharvest (PH) shoyu sauce; Ayam and Tiparos fish sauces; and Ayam teriyaki sauce. Sauces with very high volatile contents, Kikkoman (sr), the two tamari sauces, shoyu sauce, and teriyaki sauce, were diluted 1:10 just before testing. The results indicate the possibility of development into an on-line monitor for quality control during the production of sauces.

Keywords: *Semiconductor gas sensors, battery-powered monitor, gas analyzer, artificial neural network.*

Introduction

A number of reviews on portable monitors for gases, liquids and solids have been published (Lopez-Avila and Hill 1997) due to the interest in on-site analytical techniques. The present paper is on an investigation concerning application of a six-sensor handheld gas monitor using tin-oxide semiconductor sensors for identifying oriental sauces by monitoring the volatile organic constituents in the headspace.

Possible applications of portable field analyzers for environmental monitoring, quality control, quality assurance and food technology have been described (Meuzelaar 1997). Development of tin-oxide MOS (metal oxide semiconductor) for specific gas detection was described by Di Benedetto *et al.* (1998). Effect of gas characteristics on sensing properties of thick film tin oxide based sensors has been studied by Shimizu *et al.* (1998). A number of Taguchi tin-oxide MOS based handheld analyzer applications have been examined at the University of Tasmania. These include beer identification using six sensors and ethanol determination using two sensors. A Langmuir model based linear calibration method for ethanol determinations using two Taguchi MOS sensors was also developed by Di Benedetto *et al.* (1998).

Applications attempted in other laboratories include discrimination of meat products, sausages and hams, and bacterial strains, aromatic or pathogenic, using six Taguchi tin-oxide MOS sensors (Vernat-Rossi *et al.* 1996); identification of gases using a

** Part 1 of the Report on Research conducted by the author at the University of Tasmania.*

temperature modulated single tin-oxide MOS sensor (Kato *et al.* 1997) and a commercial cooking oil tester (Mittal 1996).

Another interesting application is odor sensing electronic noses which have been developed. These normally consisted of multi-sensor arrays and artificial neural net (ANN) with back propagation or fuzzy logic (Garnder *et al.* 2000; Nanto *et al.* 1996; Vlachos *et al.* 1996). Fuzzy c-means algorithm, FCMA in conjunction with a radial basic function (RBF) neural network was claimed to improve ANN performance (Wang and Xie 1996). A novel variation is the electronic tongue such as that used for beverage tasting and multi-component analysis of polluted water (Legin *et al.* 1997).

Soy sauce production usually consisted of two main fermentation stages: (i) *solid state fermentation* where the soybeans were exposed to air, and (ii) *solution fermentation* where the above partially fermented beans were further fermented in salt solution. Various volatiles were produced, which presumably contributed to the aroma and taste of the sauces (Roling *et al.* 1996).

Total sensor response is the accumulation of the responses to each of the individual sauce volatiles; the total response for each sensor gave six data peaks corresponding to the six sensors. This pattern or "fingerprint" of the sauces was then identified via the neural network or via statistical MAP functions. Use of more sensors would give better differentiated "fingerprint" and increased sensors should therefore yield improved sauce identification efficiency.

In general, commercial multi-sensor monitors, such as AromaScan from Meeco Holdings Pty. Ltd. which uses a 32 conduction polymer array (Lardner 1997), or Alpha MOS Fox from Analytical Equipment Co., which uses 12 sensors in 55 different combinations (Fotheringham 1997), can be considered to be better than the present six-sensor monitor. However the increased identification efficiency with increased sensors could not be expected to be linear indefinitely. A saturation point would exist, above which increased use of sensors would be of no real practical significance. It could even be economically unsound. Hence the use of the minimum number of sensors just sufficient for efficient performance is desirable. It was decided to test the capability and efficiency of the in-house built (University of Tasmania) six-sensor monitor, complemented with ANN, for identifying oriental sauces.

The use of the gas analyzer depends on the sensors chosen and an Artificial Neural Net for data processing. MOS type semiconductors and conducting polymers (CP) are commonly used types of gas sensors. Higher sensitivity, better day-to-day reproducibility, longer service life, less replacement cost, greater humidity tolerance, and easier regeneration on saturation were the advantages of MOS sensors over CP sensors (Fotheringham 1997). Thus, MOS sensors were chosen.

In ANN, usual threshold functions are step-functions or hard-limiters. In real neurons, information was encoded in terms of firing frequency rather than pulse presence or absence. To represent this in ANN, the step function was 'softened' or 'squashed'; fuzzy logic and sigmoid function were examples. However, back propagation employed in the first ANN softwares is still widely used. Design, theory, and applications of neural networks are discussed thoroughly by Lawrence (1994). A summary of Brainmaker, California Scientific Software, used in this work is dsecribed elsewhere (Di Benedetto *et al.* 1998).

The present work used a six-sensor MOS array in the FIA mode similar to that described before (Di Benedetto *et al.* 1998). A back propagation ANN was used for pattern recognition and discrimination. Fourteen local and imported oriental sauces readily available at Australian supermarkets were tested.

Experimental

Reagents and Solutions

Kikkoman (sr) salt reduced and Yeo's soy sauces; Pureharvest and Spiral tamari sauces; Pureharvest shoyu sauce and Ayam teriyaki sauce were diluted 1:10 just before testing. Cornwell's, Pearl River Bridge (PRB), Zu Miao and Farmland soy sauces; Songai and Pearl River Bridge mushroom soy sauces; and Tiparos and Ayam fish sauces were used straight. 50 mL of sample sauce liquids were taken in 100 mL volumetric flasks.

The Monitor

The hand-held monitor, similar to that described before by the author (Tin 1999), consisted of six Taguchi tin-oxide MOS sensors TGS 880, 825, 824, 822, 813, 800 (referenced as sensors 1, 2, 3, 4, 5 and 6, respectively) from Figaro, Osaka, Japan, and a Tattletale data-logger which allowed array options of 2, 3, 4, 5 sensor combinations in addition to the full six-sensor array for on-site delayed or timed measurements. It was used to sample well-stabilized saturated headspaces at a constant flow rate of 1 .L/min; a rate of 0.33 volume changes per second or one volume change every three seconds. The analog output voltage was captured on the data logger (Tattletale data logger) and real-time voltage readings were simultaneously displayed on an Acer/Extensa 355 computer via an RS232C connector. Collected data files were then downloaded to the computer and saved as text files in Txtools, transferred to a spreadsheet, plotted out as graphs on the VDU and used for data analysis.

Procedures

A head-space sampling method was used. The sampling tube was inserted for 10 seconds, into the well-stabilized saturated headspace, without shaking, to a point about 5 cm above the liquid surface of a 50 mL sample, contained in a 100 mL volumetric flask.

The responses from the six sensors were recorded on the Tattletale data logger, while being monitored in real time on the computer screen. The recorded sensor responses for the samples were downloaded onto the computer and saved as text files in Txtools. Single, duplicate or triplicate measurements, using 2, 3, 4, 5 or 6 sensor arrays, in different sensor combinations, were done on each of the 14 sauce samples. Peak height measurements were made manually and entered in Excel files, which were then loaded into the ANN Brainmaker (from California Scientific Software Inc., CA, USA), where Brainmaker files were created and the network was trained. The process was repeated on chosen test sample sauces; peak height data were loaded onto the Brainmaker, where Running Fact files were created and fed to the above trained net for sauce identification.

Results and Discussion

Net Training Parameters

The data, saved as text files in Txtools, were trained in the ANN using 200 hidden neurons, a Sigmoid nonlinear saturation transfer function, a tolerance level of 10%, and a learning rate of 1.000.

It was found that 200-220 hidden neurons gave the shortest net training times, at 10% tolerance and 1.000 learning rate. Increased hidden neurons tend to push neural nets towards memorizing rather than predicting, and excessive use of hidden neurons was generally discouraged. Unlike stock market cases, where daily predictions based on previous data were required, sauce identification demanded the net to memorize the input patterns and identify the test samples by comparing test patterns to input patterns, embedded in the net memory. Memorizing rather than predicting was necessary. Therefore use of such a high number of hidden neurons, in order to promote net memory, was justified in this case.

There were two important sets of variables in the Neural nets; the set of sauce sensor responses, which depended on the number of sauces used and the individual variations of these responses, which depended on repeated measurements. The tolerance level, used in net training, referred to the latter. A variation of 10% was tolerated by the trained net, i.e. any response pattern or fingerprint within 10% of a sample pattern would be recognized as the sample.

ANN Net Training

The ease of net training should depend on number of sauce samples, number of sensors, response similarity of sensors in the array and number of repeats. The number of samples was fixed at 14, and only the latter three factors would affect net training.

Number of Sensors: More sensors in an array would give better pattern recognition factors, and should consequently make net training faster. This was observed in Table 1, where six sensor arrays had net training times of just over a minute, as compared to around two minutes for 3 or 4 arrays, and over 10 minutes for two sensor arrays.

Response Similarity of Sensors: Obviously similar response patterns would make it harder for the net to recognize and memorize the differences.. This should make net training more difficult. For example in Table 1, the three-sensor array (2, 4, 6) had less similar responses among its three sensors than the (1, 2, 3) array. For a net size of 14 x 2 = 28, the net training times were two minutes for the former and five minutes for the latter. The (2, 6) two-sensor array with similar responses between the two sensors could not be trained fully at all.

Repeated Measurements: Using single measurement for each sauce, a net size of 14x1=14, net training was achieved for sauce sets of five members or less. If six or more members were used, one member, usually the sixth, was ignored during training, and was consequently identified incorrectly. Thus, only 13 of the 14 facts, patterns created by 14 sauce responses, were recognized (Table 1). Tiparos fish sauce, the sixth sauce item, was always identified incorrectly (identifications marked B, bad identifications) when using the 14 x 1 = 14 nets of 2, 3 and 6 sensor arrays. Thus identification of sauce samples were limited to a maximum of five, if using single measurements.

Increased repeats should promote net training by improved statistical significance, brought about by increased data. This, should make the net more effective at identification. Thus all 14 sauces, including Tiparos, were identified correctly when using duplicate, triplicate or quintuplet data with 3, 4, 5 or 6 sensor combinations (Table 1). However, the larger data set should make training times longer. As shown in Table 1, in the six-sensor array, training times for the 14 x 5 = 70 fact net, quintuplet data, was over two min and the corresponding 14 x 1 = 14 fact net, single data, was 1.5 min.

As shown in Table 1, only 25 facts were recognized in the duplicate data, 14 x 2 = 28 fact nets of two-sensor arrays. Three facts of two sauces were ignored in turn. Similarly, only 38 facts were recognized in the 42 fact nets and only 63 facts were recognized in the 70 fact net.

Multiplicates obtained by copying single measurement data were sufficient to obtain efficiently trained nets which were capable of identifying the 14 sauces. Thus, practical multiple measurements were unnecessary for net training purposes.

Sauce Identification

Single Peak Height Measurements: With a six-sensor array, using single measurements of the head space volatiles for each sauce, no identification difficulties were encountered if less than six sauces were analyzed at any one instance, with corresponding small nets. For example, any five of the 14 sauces set, could be identified by using appropriate five fact nets. Fig. 1 shows successful identification of five low volatile sauces, and Fig. 2 shows efficient identification of 1:10 diluted high volatile sauces, using six sensor arrays, single data set and five fact nets. If more than five were attempted, one sauce, usually the sixth, was ignored during net training as mentioned above, with consequent error in identifying that particular sauce, Tiparos fish sauce in this case. It was interesting to note that even Tiparos was identified correctly in the five fact net (Fig. 1).

The five-sample limit can be attributed to statistical insufficiency of presented data. In

ANNs used for market predictions, previous variations on prices of items, such as gold, silver, copper, were used to obtain a trained net. In the case of sauces, each sauce would be an item, analogous to gold, silver or copper; and the repeated measurements would be analogous to price variations. Thus it was decided to try multiple data.

Duplicate Peak Height Measurements: As mentioned earlier, in 28 fact nets, only 25 of the 28 facts of duplicate data were recognized. The sixth and twelfth sauces of the set were ignored at alternate runs during net training, but identification was not inhibited. Identification of 14 sauces was achieved. Response peak heights of 14 sauces to six sensors used for net training and identification were shown in Table 2. Fig. 3 shows efficient 14 sauce identification, using six sensors and 28 fact nets. Similarly, efficient 14 sauce identification was observed with all 6, 5, 4, 3 sensor arrays, using 28 fact net (Table 1). The three-sensor arrays (2, 4, 6), (1, 3, 5), (1, 2, 3) and (4, 5, 6) were all efficient.

Since the number of sauces, 14, was an obviously very high figure, which none of the sauce companies would produce, it was important for the net to be capable of identifying less number of sauces. Identification of Cornwell's and Kikkoman (sr) soy sauces with low and high volatile contents, respectively, Tiparos fish sauce and Pearl River Bridge mushroom soy sauce, selected for their different nature, were tested. Fig. 4 shows correct identification of the selected four different sauces, using a three-sensor array, 28 fact nets. The five-sauce limits for single data set 14 fact nets mentioned earlier was useful in itself, but identification was restricted to the five sauces used in net training. The 14 sauce, 28 fact trained net would give a variety of options concerning sauce numbers and sauce combinations.

Triplicate Measurements: Triplicate and quintuplet data, 14 x 3 = 42 and 14 x 5 = 70 fact nets of six-sensor arrays were also efficient at 14-sauce identification. However, the identification fractions did not improve over the three- sensor array 28 fact nets. Thus three-sensor arrays with duplicate data were the minimum requirements for efficient identification.

Measurement Difficulties: Continued fermentation of sauces on exposure to the atmosphere and variable headspace build up, were the two main difficulties in obtaining reproducible sensor responses, in addition to the well known difficulties, such as humidity and temperature variations. Hence these factors had to be controlled by refrigeration, working in a temperature, humidity regulated room, and by giving enough time for build up of saturated equilibrated headspace, between repeated measurements.

Possible Applications

One possible application area was quality control or quality assurance. Proactive quality standards like the ISO 9000, required continuous assessment all along the production stream, as opposed to the reactive standards which simply required checking the finished product output. Thus small, efficient on-line testers would be in great demand in the future.

Any variation from the norm would have sensor responses different from the standard values. Thus, if these variations could be detected by the ANN trained net, then any sub-standard quality sample batch would be identified. Hypothetical variations of ±0.5%, ± 1%, ±2%, ±5%, ±7% and ±10%, in the Cornwell's soy sauce test response data were fed into the three-sensor (2,4,6) array duplicate data 28 fact ANN trained nets, for identification. Fig. 5 shows that distinct identification pattern differences were observed, starting from 1% response variation. The sharpest changes were in the Zu Miao component of the identifications. Thus this change could be used to set the required quality standard of Cornwell's soy sauce. Therefore the present six-sensor tester had a potential for being developed into an online tester for oriental sauce production. This possibility was being explored.

Another application of importance to Asia was process monitoring in soy sauce production. A tester, similar to the six-sensor, possibly with less sensors, could be used to follow fermentation during soy sauce production, and optimum production parameters could be determined. Work on similar lines had been done on Japanese *miso* (soy paste). Lipid membranes were used as taste transducers in a multi-channel taste sensor, an electronic tongue analogous to the electronic nose.

Advantages Over Multi-sensor Arrays

The main advantages of the six-sensor MOS array over multi-sensor MOS arrays were comparable identification capabilities at comparatively lower cost and the minimized error probability due to less peak measurements. Thus in sauce identification, the six-sensor tester, being much less costly, was preferable to the much more costly multi-sensor monitors. It also had a potential for development into an online quality assurance tester, for proactive, rather than reactive, conformity to standards, in accordance with ISO 9000 quality standards; and a soy sauce production monitor.

Conclusion

The six-sensor monitor, coupled with Brainmaker ANN software, has been shown to identify up to 14 oriental sauces efficiently. The monitor is of comparatively lower cost, and it could be developed into an online tester for quality assurance in sauce production, and also extended for use in other fermentation industries. It therefore offers an alternative method in sauce manufacturing and wine and beer production to the more sophisticated but costly multi-sensor monitors presently available commercially.

Acknowledgments

The author is indebted to ARC for a grant and to Assumption University, Bangkok, for granting a leave of absence.

References

Di Benedetto, L.T.; Alexander, P.W.; and Hibbert, B.D. 1998. A field-portable gas analyzer with an array of six semiconductor sensors. Part 1: Quantitative determination of ethanol. Part 2: Identification of beer samples using artificial neural networks. Field Anal. Chem. Technol. 2: 135-43.

Fotheringham, D. 1997. Analytical Equipment Co. Anal. Test.Technol.12: 44.

Garnder, J.W.; Liobet, E.; Hines, E.L.; Mottram, T.T.; and Bartlett, P.N. 2000. Fuzzy ARTMAP-based electronic nose data analysis. Sensors and Actuators, B6: 183-90.

Kato, Y.; Yoshikawa, K.; and Kitora, M. 1997. Temperature-dependent dynamic response enables the qualification and quantification of gases by a single sensor. Sensors and Actuators, B40: 33-7.

Lardner, R.. 1997. Meeco Holdings Pty. Ltd. Anal. Test. Technol.12: 22-3.

Lawrence, J. 1994. Introduction to Neural Networks: Design, Theory and Applications. Calif. Sci. Software Press, Nevada City, CA.

Legin, A.; Vlasov, Y.; Rudnitskaya, A.; Selezenev, B.; Di Natale, C.; Nacagnano, A.; Davide, F. and D'Amico, A. 1997. Multicomponent analysis on polluted waters by means of an electronic tongue. Sensors and Actuators, B44: 423-8.

Lopez-Avila, V.; and Hill, H. 1997. Gas sensors. Anal. Chem. 69: 289R - 305R

Meuzelaar, H. (ed.) 1994. Field-portable analytical instrumentation. Trends in Anal. Chem. 13: 252-94.

Mittal, G. 1996. Personal Communications. Univ. Guelph, Guelph, Ontario, Canada.

Nanto, H.; Kondo, K.; Habara, M.; Douguchi, Y.; Waite, R.I. and Nakazumi, H. 1996. Identification of aromas from alcohols using a Japanese-lacquer-film-coated quartz resonator in conjunction with pattern recognition analysis. Sensors and Actuators B35-36: 183-6.

Roling, W.F.M.; Apriyantono, A.; and van Verseveld, H.W. 1996. Kecap production. J. Ferm. and Bioeng. 81: 275-8.

Shimizu, Y.; Maekawa, T.; Nakamura; Egashira, M. 1998. Effects of gas diffusivity and reactivity on sensing properties of thick film SnO_2-based sensors. Sensors and Actuators, B46: 163-8.

Tin, D.W. 1999. Assumption University and Sensor Development. AU J.T. 3: 51-7.

Vernat-Rossi, V.; Talon, R; Denoyer, C.; Berdague, J. L; Garcia, G. 1996. Rapid discrimination of meat products and bacterial strains using semiconductor gas sensors. Sensors and Actuators, B37: 43-8.

Vlachos, D.; Fragoulis, D.K.; and Avaritsiotis, J. 1996. An adaptive neural network topology for degradation compensation of thin film tin oxide gas sensors. Sensors and Actuators, B45: 223-8.

Table 1. ANN net training and identification data

No. Sensors	Net Size	Net File	Training Time	Facts Used	Facts Trained	Sauces	Run File	Out File	Remark
6	14x1=14	s6n14x~1.net	1:32	14	13	14	r14s6t~1.in	r14s6n14.out	B
6	14x2=28	s6n14x~2.net	1:44	28	25	14	r14s6t~1.in	r14s6n28.out	G
6	14x3=42	s6n14x~3.net	1:29	42	38	14	r14s6t~1.in	r14s6n42.out	G
6	14x5=70	s6n14x~5.net	2:13	70	63	14	r14s6t~1.in	r14s6n70.out	G
6	14x2=28	s6n14x~2.net	1:44	28	25	4	r4s6t~1.in	r4s6n28.out	G
6	5x1=5	5LSau6.net	0:34	5	5	5	5LSau6.in	5LSau6.out	G
6	5x1=5	5HSau6.net	0:44	5	5	5	5HSau6.in	5HSau6.out	G
5	14x2=28	s5n14x~2.net	1:10	28	25	14	r14s5t~1.in	r14s5n28.out	G
4	14x2=28	s4n14x~2.net	2:19	28	25	14	r14s4t~1.in	r14s4n28.out	G
3(2,4,6)	14x1=14	s3n14x~1.net	2:04	14	13	14	r14s3t~2.in	r14s3n14.out	B
3(2,4,6)	14x2=28	s3n14x~2.net	2:06	28	25	14	r14s3t~2.in	r14s3n28.out	G
3(1,3,5)	14x2=28	s3an14x2.net	3:19	28	25	14	r14s3a~1.in	r14s3a28.out	G
3(1,2,3)	14x2=28	s3bn14~1.net	5:07	28	25	14	r14s3b~1.in	r14s3b28.out	G
3(4,5,6)	14x2=28	s3cn14~1.net	3:39	28	25	14	r14s3c~1.in	r14s3c28.out	G
2(1,3)	14x1=14	s2n14x~1.net	10:36	14	13	14	r14s2t~1.in	r14s2n14.out	B
2(1,3)	14x2=28	s2n14x~2.net	43:07	28	25	14	r14s2t~1.in	r14s2n28.out	B
2(2,6)	14x2=28	s2an14x2.net	70% Trained	28	25	14	r14s2a~1.in	r14s2a28.out	B

G = Good. Correct Identification B = Bad. Incorrect Identification.

Table 2. Sample net training and identification six sensor responses

No.	Sauce/Sensor	Training Peak Heights / mV 1	2	3	4	5	6	Identification Peak Heights / mV 1	2	3	4	5	6
1	Cornwell's	160	170	50	120	50	180	150	160	50	115	60	180
2	Spiral Tamari	750	760	230	520	160	560	730	740	220	540	150	50
3	Kikkoman(sr)	540	550	150	330	100	360	520	530	150	320	100	350
4	PH Tamari	50	60	20	40	20	50	50	60	20	40	20	50
5	Yeo's	230	240	50	150	50	180	220	230	50	150	60	200
6	Tiparos	120	130	50	120	50	120	130	140	50	120	50	130
7	Ayam Fish	170	180	60	160	30	130	160	170	60	150	30	130
8	Songai	70	80	30	60	30	80	80	90	30	60	30	80
9	PH Shoyu	350	360	100	260	60	250	330	340	100	260	60	250
10	PRB	100	110	20	70	20	80	100	110	20	70	20	80
11	MushPRB	70	80	20	50	20	50	80	90	20	60	20	50
12	Teriyaki	320	330	100	230	80	230	320	330	100	220	80	220
13	Zu Miao	170	180	50	110	60	150	160	170	50	100	60	150
14	Farmland	120	130	30	80	40	120	110	120	30	70	50	110

Legend: PH = Pureharvest PRB = Pearl River Bridge MushPRB = Pearl River Bridge Mushroom Soy

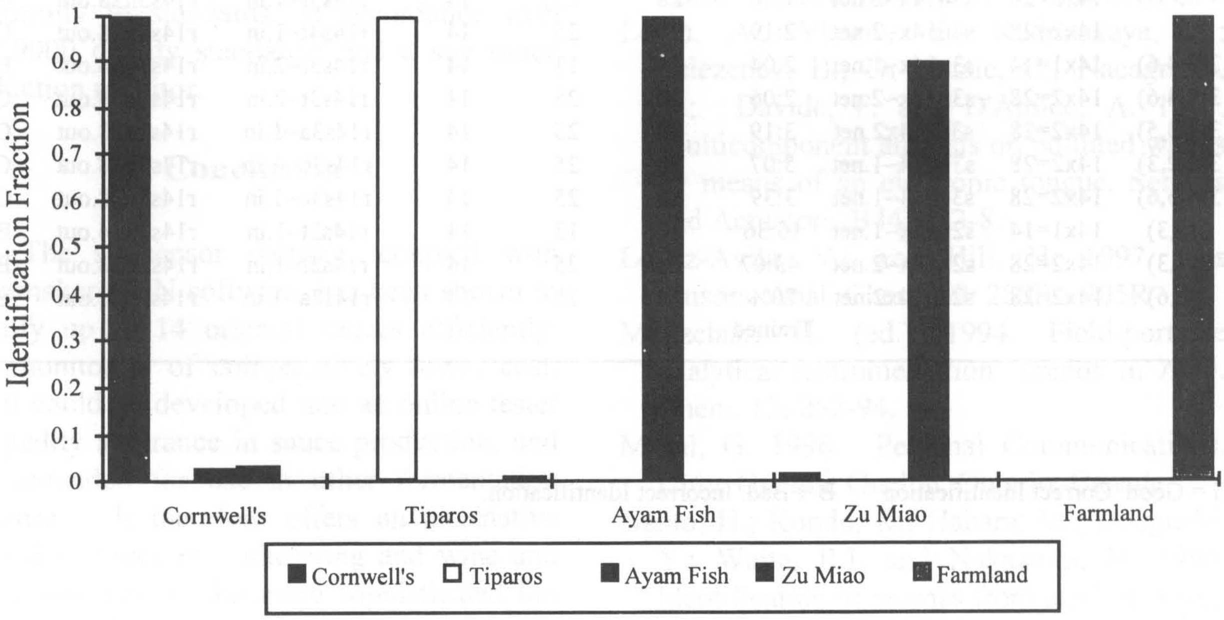

Fig. 1. Identification of five low volatile sauces 6 sensors.
Single Data Set; 5 Fact Net

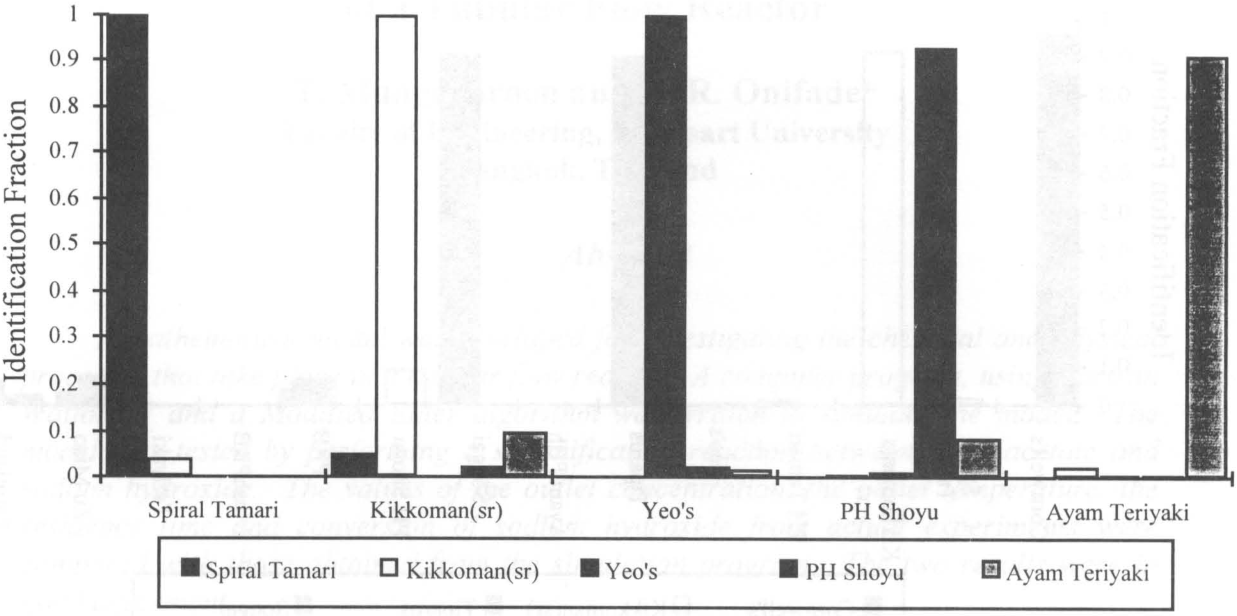

Fig. 2. Identification of five 1:10 diluted high volatile sauces.
6 Sensors; Single Data Set; 5 Fact Net

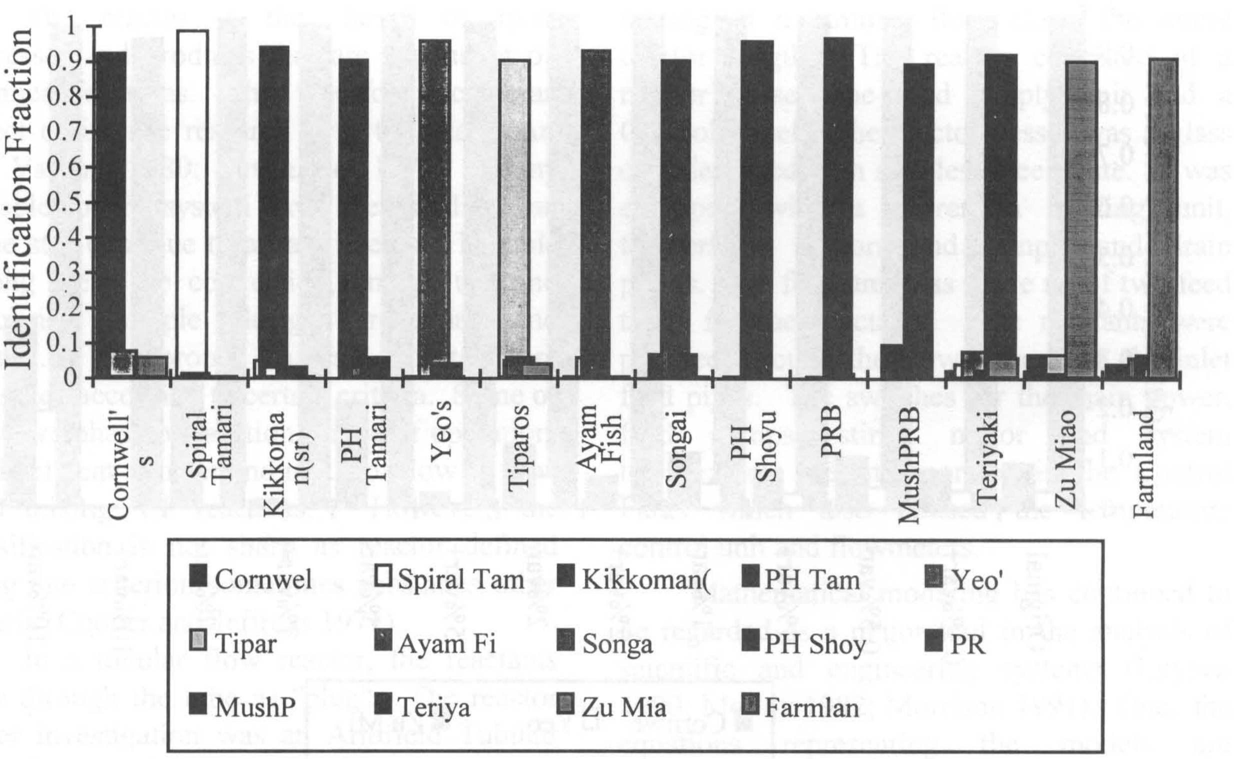

Fig. 3. Identification of fourteen sauces, six sensors.
Duplicate Data Set; 28 Fact Net

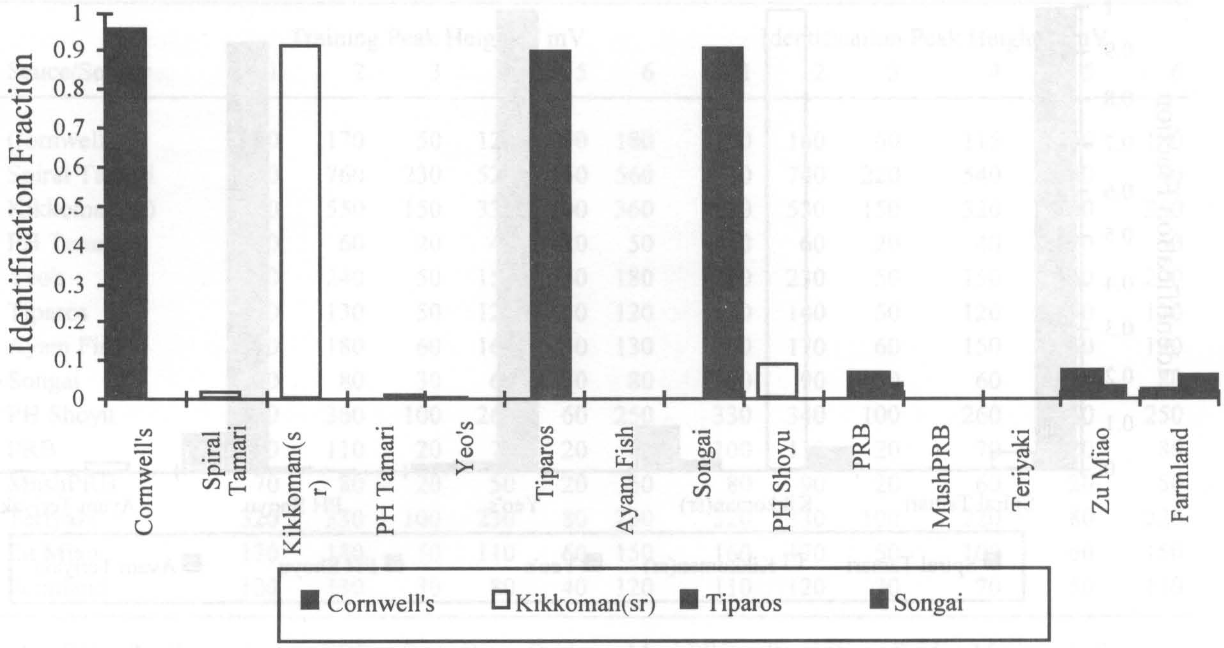

Fig. 4. Identification of four selected sauces three sensors (2, 4, 6).
Duplicate Data Set; 28 Fact Net.

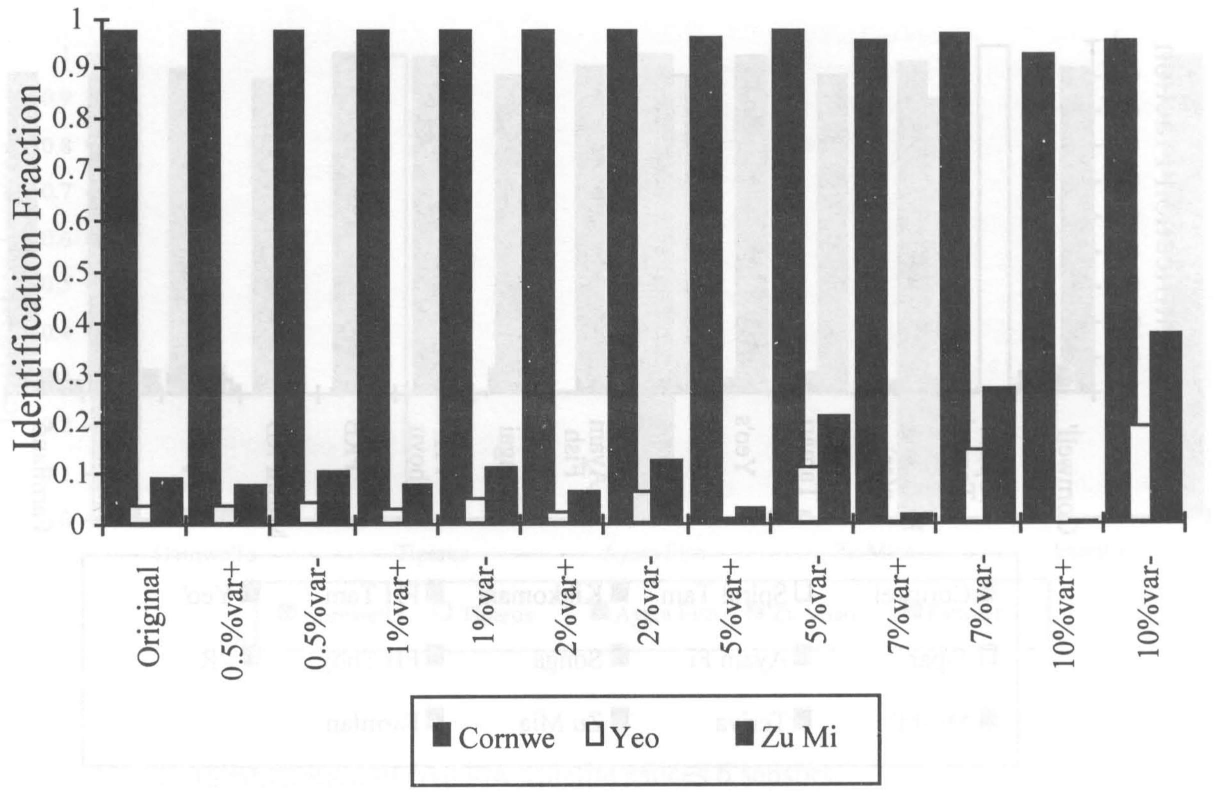

Fig. 5. Hypothetical variations of Cornwell's sauce.
Pro-active quality assurance possiblity