This is to certify that the

dissertation entitled

Electronic Sensory Array Incorporating Artificial Neural Network Algorithms for Rapid Identification and Quantification of Escherichia coli and Salmonella enterica Serovar Typhimurium and Their Volatile Metabolites

presented by

Ubonratana Siripatrawan

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Packaging

Bruce Harte
Major professor

Date May 9, 2002

| DATE DUE | DATE DUE | DATE DUE |
|---|---|---|
| JUN 2 8 2004 | | |
| SEP 2 4 2007 | | |
| JAN 2 7 2006 | | |
| NOV 2 | | |
| | | |
| | | |
| | | |
| | | |
| | | |

# ELECTRONIC SENSOR ARRAY INCORPORATING ARTIFICIAL NEURAL NETWORK ALGORITHMS FOR RAPID IDENTIFICATION AND QUANTIFICATION OF *ESCHERICHIA COLI* AND *SALMONELLA ENTERICA* SEROVAR TYPHIMURIUM AND THEIR VOLATILE METABOLITES

By

Ubonratana Siripatrawan

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirement
for the degree of

DOCTOR OF PHILOSOPHY

School of Packaging

2002

## ABSTRACT

### ELECTRONIC SENSOR ARRAY INCORPORATING ARTIFICIAL NEURAL NETWORK ALGORITHMS FOR RAPID IDENTIFICATION AND QUANTIFICATION OF *ESCHERICHIA COLI* AND *SALMONELLA ENTERICA* SEROVAR TYPHIMURIUM AND THEIR VOLATILE METABOLITES

**By**

**Ubonratana Siripatrawan**

A rapid method to identify and quantify *E. coli* and *Salmonella enterica* serovar Typhimurium and their specific volatile metabolites was developed using an array of 12 nonspecific metal oxide electronic sensors incorporating artificial neural network algorithms. The metabolic volatile compounds used as indicators of *E. coli* and *Salmonella* Typhimurium in the samples were identified using solid phase microextraction coupled with gas chromatograph/mass spectrometer (SPME/GC/MS). Principal Component Analysis (PCA) was used for data exploration and dimensional reduction. The Mahalanobis distance metric was determined based on Discriminant Factor Analysis (DFA) for sample classification to differentiate volatiles in control samples from that containing the target microorganisms. Artificial neural networks were trained to identify and quantify *E. coli* and *Salmonella* Typhimurium and their volatile metabolites.

The neural networks were shown to be capable of correlating voltametric responses with number of *E. coli* with low mean square errors. The electronic sensor array was found to be satisfactorily correlated with colony counting and

GC/MS methods. This technique provides a rapid, simple, and precise analysis of the biochemical composition of microbiological systems, for identification of potentially pathogenic microorganisms.

Dedicated to the Land, Kings, and People of Thailand

## ACKNOWLEDGMENTS

I would like to extend my intensity of respect and appreciation to King Chulalongkorn the Great and King Bhumibol Adulyadej the Great, for his impressive contributions to the progress of modern Thai education and for their keenly awareness of the importance of education and of the value of youth to the future of Thailand. I also would like to express my gratitude to the Royal Thai Government, Ministry of University Affairs, and Chulalongkorn University, for giving me an opportunity to further my studies in the United States to gain knowledge in order to help fulfill the betterment of Thailand.

I would like to express my highest respect and appreciation to my major professor Prof. Bruce R. Harte, Ph.D., Director, School of Packaging, who is recognized for being insightful, intelligent, humorous, and practical, for his time, kind thought, guidance, personal support, and for showing his appreciation throughout my research and endowed me with a global perspective. I am also highly grateful to my committee Prof. John E. Linz, Ph.D., Department of Food Science and Human Nutrition, for sharing his time, graciously providing me with all food microbiological facilities, and giving me the constructive suggestions on food microbiological research. I also wish to thank my committee Dr. Robert H. Clarke, and Dr. Hugh Lockhart, School of Packaging, for sharing their time serving as my committee. I am also indebted to the Center for Food and Pharmaceutical Packaging Research for the in-part support on my research.

I wish to express my sincere appreciation and gratitude to Mr. Matthew D. Rarick for his tireless helps, time, and constructive suggestions on all of the

microbiological experiments. Many thanks to my friend Antonis Kanavouras for his help and information on SPME. I also wish to thank Sakakibara sensei for being very kind and understanding.

My sincere thanks for all my best friends, especially Rujida & Patcharawat Uthaisombat, Korada (who went through late night working with me all these years), Napawan, Sunetra, P'Manee, Hiromi, Tatsuya, Derek, Eugene, John, and Stephen for all their helps, pep talk, and for helping make my stay here a memorable one.

I wish to thank my far away but always best friends, Araya, Duangjai, Wilai, Angsana, Suntareeya, Nucharin, Michio, and Kul for sending me love, gifts, and delicious Thai foods all these years. I also wish to thank all my friends from Agro-Industry 6, Kasetsart University, friends and colleagues in Food & Biotechnology, Chulalongkorn University. I also wish to thank Mary Ann Merrill, Colleen, and my friends in School of Packaging and Food Microbiology.

Finally, I would like to express a deepest appreciation and gratefulness to my wonderful parents, my caring sister and my protective brother for their constant love, inspiration, personal support, and patiently waiting for me to return home.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

*"It is the excitement, the honesty of a response to Nature, that guides our hand; and if this excitement is often so strong that one works without noticing that one is working, if brushstrokes sometimes come thick and fast like words in a conversation or letter, then one ought not to forget that it has not always been like that and that there will be many a depressing day barren of inspiration in the future"*

*-Vincent van Gogh*

Fresh fruits and vegetables are one of the most-rapidly growing food product categories in the United States. However, contamination of fresh fruits and vegetables, juices, and fresh prepared salads with foodborne pathogens is a major food safety concern. Raw fresh fruits and vegetables have been recognized as potential vehicles for transmission of pathogenic microorganisms known to cause human disease. The frequency of outbreaks epidemiologically associated with raw fruits and vegetables has increased in the United States as a result of change in dietary habits and increased import of food (Altekruse et al., 1997).

Foodborne pathogens cause an estimated 6.5-33 million cases of human illness and up to 9,000 deaths in the United States each year (ERS, 1999). Over 40 different foodborne pathogens, including bacteria, fungi, viruses, and parasites, are believed to cause human illnesses and $2.9-$6.7 billion in

treatment cost and cost revenue are annually attributed to foodborne illnesses (Economic Research Service, U.S. Department of Agriculture, 1999).

Diseases caused by foodborne pathogens range from mild to severe, including death. These diseases have significant direct and indirect economic impacts on society. Direct economic impacts include cost of treatment and lost productivity. Indirect economic impacts include cost of lost trade and lost consumer confidence, legal costs, and loss of market confidence (Acnab, 1997).

For fresh fruit and vegetables, maintenance of quality and safety are most important. The identification and implementation of effective control measures to attain these goals is a formidable challenge to regulatory agencies, agriculturists and food manufacturers. The short shelf life of many fresh fruits and vegetables predicates the need for refrigerated storage and rapid delivery to the retail market. Such prerequisites tend to preempt microbiological testing for contamination by foodborne pathogens. However, recent evidence on the inherent health risk associated with their consumption, underlines the need for renewed vigilance and stringent controls at all stages of food production, harvesting, processing, packaging, and distribution (D'Aoust, 1994).

While great effort has been made to prevent contamination of fruits and vegetables during post-harvest processing, packaging, and handling, much improvement is needed in effective microbial detection to ensure that products are not contaminated. Development of a rapid, specific, and economical method which can routinely be applied to packaged food plant products to allow fast and

accurate detection of microorganisms would help reduce human health problems, and diminish this threat to the fruit and vegetable industries.

Many approaches have been used in order to develop rapid, precise, and accurate techniques in order to potentially identify pathogenic organisms in food products. Novel, reliable, sensitive, and economical methods continue to be developed to allow rapid, inline, and accurate detection of hazardous organisms and their toxins. Development of new and improved methods continues to receive a great deal of research attention and a reasonable amount of private and public sector funding (CAST, 1999).

Many approaches have been used in order to develop rapid, precise, and accurate techniques in order to potentially identify pathogenic organisms in food products. Novel, reliable, sensitive, and economical methods continue to be developed to allow rapid, inline, and accurate detection of hazardous organisms and their toxins. Development of new and improved methods continues to receive a great deal of research attention and a reasonable amount of private and public sector funding (CAST, 1999).

The annual cost of routine industrial microbiological testing of foods amounts to many millions of dollars (Habermehl, 1984). With such expenditures, it is necessary to question the raison d'être for routine microbiological testing. Microbiological safety and quality testing in the food industry has been based upon traditional plate count methods. These conventional techniques are laborious and time consuming.

Measurement of product volatiles has been shown to have promise as a tool for distinguishing different types of food freshness (Jonsson et al., 1997; Bachingger et al., 1998; Varns and Glynn, 1979).

The selection of volatiles for possible use as incipient disease indicators has been discussed in terms of the composite rate of pathogenic destruction within food products (Borjesson, 1996; Jonsson, 1997).

The metabolic interaction of volatiles arising from pathogen growth is becoming more important in explaining the pathogenic contamination. Gas Chromatography (GC), Gas chromatography and Mass Spectrometry (GC/MS), High Performance Liquid Chromatography (HPLC), Radio-immunoassay, and Calorimetric tests, are among the techniques available for determining the identity and quantity of volatiles. However, none of these methods can completely meet the time, sensitivity, and accuracy requirements, particularly when used in routine quality control evaluation.

A technique is needed to provide the necessary specificity and rapid response time required throughout the food distribution channel. The electronic sensor array is an instrument capable of doing that. Multisensor array technology has been shown to be able to provide rapid, continuous monitoring of a wide array of different volatile chemicals. The instrument includes an array of different electronic chemical sensors with partial specificity and an appropriate pattern recognition system, capable of recognizing simple or complex volatiles.

Pattern recognition routines based on either statistical methods or on artificial neural networks are used to evaluate the sensor array responses.

With increasing pressure to achieve profit margins and with increasing labor costs, alternative methodologies which can be used more cost-effectively are needed by the food industry. Also, it is important to have simple, rapid feedback methods so that action may be taken quickly if evidence for product contamination exits. Tests capable of detecting relatively low numbers of specific microorganisms or their metabolites are needed. These tests must be both simple and exceptionally rapid, have low operating costs, and must provide data comparable with traditional data systems. The electronic sensor technique, with its array of gas sensors together with pattern recognition techniques is rapid, simple and inexpensive, compared to more conventional methods.

Conceptually, the electronic sensor array technique stands apart from all other conventional microbiological methods using rapid sequential analysis as a detectable signal. The electronic sensor array has considerable potential as a simple and cost-effective means of predicting the microbiological safety of foods, as a routine part of any HACCP program.

As a result of metabolic activity, microorganisms produce various metabolites and intermediates in their growth environment. The electronic sensor array can be qualitatively and quantitatively employed to determine the specific catabolic products of specific compounds such as fatty acids, glucose, amino acids, or other compounds produced by microorganisms.

The development of non-specific sensors, which are available commercially for odor detection, could offer a rapid and relatively simple technique for monitoring the metabolic activity of microorganisms. However, raw sensor responses from an array of non-specific metal oxide sensors are generally insufficient to discriminate between a series of samples. Therefore, pattern recognition methods are considered important to extract the information from the electronic voltametric responses.

Array sensors produce data that can be subjected to a variety of multivariate analyses. Data analysis depends on the information required, whether identification of substances (classification) or the concentrations of given compounds. The choice of data analyses is important and the subsequent pattern recognition process can ensure the success of electronic sensor array for a particular application. In theory, much information can be generated if the sensors are properly matched to the matrix of analytes encountered.

Multivariate data analysis can be used to recognize the characteristic variation patterns. One possibility which has potential for the characterization of these volatile compounds is based on chemometrics. In this study, neural network and chemometrics techniques were used as an experimental platform in addition to the instrumental methods.

Using electronic sensors to detect volatile complexes has potential to be a sensitive, fast, one-step method to monitor *E. coli* contamination in products. The analysis of specific volatile organic compounds produced by metabolic

processes of *E. coli* may provide a rapid method that will allow detection of *E. coli* in a packaged plant product. However, the sensor responses obtained from an array of nonspecific gas sensors are generally insufficient to discriminate between a series of samples.

Therefore, pattern recognition methods are considered important to extract the information from the electronic sensor voltametric responses. This research studied the possibility of using different pattern recognition/classification techniques for successfully extracting the information from an electronic sensor array as well as SPME/GC/MS.

A primary advantage of an electronic sensor array as a quality assurance tool for the food industry is speed of analysis, including data acquisition and interpretation. Rapid, significant data interpretation is possible using various multivariate data analyses. In this study, neural network algorithms were used to develop an integrated framework where feature extraction and predictive learning are iteratively performed with the goal being optimal approximation. Neural network is a massively parallel distributed processor constructed of processing units, which can store learned knowledge, making it available for use (Hikawa, 2001; Hykin, 1999).

The artificial neural network algorithms developed in this study was Multilayer perceptrons (MLP) based on a back propagation algorithm. A Multilayer perceptron network based on a back-propagation algorithm was used to predict the number of *E. coli* and concentration of specific volatile metabolites.

The predictive learning ability of a neural network model was assessed by comparison to a reference method using a mean square error of prediction. The electronic sensor array coupled with the neural network is a rapid, and simple technique which can be used to identify and quantify *E. coli* and the concentrations of volatile metabolites in food samples. Correlations of data from electronic sensor array technology and other commonly used methods were investigated.

# CHAPTER 2

# OBJECTIVES AND SIGNIFICANCE

## OBJECTIVES

The overall objectives of this research were to develop an electronic sensor array technique incorporating neural network algorithms to provide a simple and rapid method for detection of the volatile compounds produced by *E. coli* and *Salmonella enterica* serotype Typhimurium, and which could be correlated to the presence of these microorganisms in packaged plant food products. Specific objectives include:

1. To determine volatile compounds produced by *E. coli* and *Salmonella* Typhimurium in pure cultures using a gas chromatograph/mass spectrometer.
2. To optimize the electronic sensor array for analysis of volatile compounds produced by *E. coli* and *Salmonella* Typhimurium.
3. To develop an electronic sensor method incorporating pattern recognition techniques which can be used to detect *E. coli* and *Salmonella* Typhimurium in nutrient media and packaged food products.
4. To develop the optimal neural network algorithms necessary to facilitate electronic sensor array identification and quantification of specific microorganisms and their volatile metabolites.

This research is presented in 3 main sections. Section 1 deals with the experimental set-up including SPME/GC/MS and electronic sensor assessment techniques for determination of volatile compounds produced by *E. coli* in nutrient media. In Section 2, the volatile compounds produced by *Salmonella* Typhimurium in nutrient media is presented. Section 3 deals with determination of volatiles produced by *E. coli* in packaged alfalfa sprouts, respectively.

*E. coli* was selected as the test organism because of the scientific consensus that it can be used as an indicator of fecal contamination (Jay, 1996). In addition, *Salmonella* Typhimurium was selected because it has been implicated as the causative agent in many food poisoning illnesses, and outbreaks.

Alfalfa sprouts were used in this research because of the increased demand for sprouts due to their popularity as a health food, and because sprouts have been increasingly implicated in foodborne illness. Many have become aware of the potential for this food to be a vehicle for foodborne illness (U.S. Food and Drug Administration, 1999). This research also studied the distinctive possibility of using artificial neural network algorithms for prediction of number of bacteria and concentration of volatile metabolites from the electronic sensor array as well as SPME/GC/MS.

## SIGNIFICANCE

In this research, it was hypothesized that the electronic sensor array incorporating neural networks could be used as part of a detection system which relates the volatile profile in a packaged product to the number of organisms in the product, and to the barrier and compatibility characteristics of the package materials.

This research is expected to result in the development of an economical, specific, and real-time detection system for routine analysis of *E. coli* contamination in packaged plant food products. This system would diminish the pathogen induced threats to the fresh vegetable industries and help reduce human health problems. A similar system can be easily constructed for the automatic detection of *E. coli* in packaged plant food products for the food industries. The detection system can also be incorporated into a HACCP protocol in the food industries.

Use of the electronic sensor array method for rapid detection of *E. coli* and *Salmonella* Typhimurium contamination in packaged food products may help to reduce human health problems. Ultimately, with this technology we may be able to detect other pathogenic bacteria in packaged food products. The impact of this would be to decrease the number of foodborne illnesses, the cost of loss work time, and money spent on treatment. The electronic sensor array technique may be applicable to monitoring of volatiles from pathogenic microorganisms at critical control points in HACCP systems, and to enable remediation when results exceed critical limits. Use of this technique for indicator organisms at critical

control points, at locations where contamination is likely to be highest, and in final products can provide actionable information to verify process control and direct process improvement.

# CHAPTER 3

# LITERATURE REVIEW

## A. PATHOGENS ASSOCIATED RISKS IN FRESH VEGETABLES

### 1. Introduction

Microbial pathogens in food cause 6.5-33 million cases of human illnesses and up to 9000 deaths each year in the U.S. Foodborne illnesses cost the United States $5.6-$9.4 billion each year (ERS, 1999; The Council for Agriculture and Science Technology, 1994). Sources of pathogens including bacteria, fungi, parasite, and viruses, more than 90 percent of confirmed foodborne disease cases and deaths are attributed to bacteria (The Center for Disease Control and Prevention, 1990).

Fruits and vegetables can become contaminated with pathogenic microorganisms during growing, harvesting, post-harvest, handling, processing, and distribution (Beuchat, 1999; Lund, 1992). Bacteria normally reside in the intestinal tracts of animals, including humans, and may contaminate raw fruits and vegetables through contact with feces, sewage, untreated irrigation water or surface water (Cliver, 1997). Many microbial contaminants are part of the environment and vegetables may be inadvertently contaminated. Washing fruits and vegetables with contaminated water and handling of produce by infected workers, vendors and consumers help spread pathogenic microorganisms (Brackett, 1994; Nguyen, 1994). Increased global trade in raw fruits and

vegetables, as well as increased international travel, could also increase the risk of produce-associated diseases (Hedberg et al., 1994).

## 2. Alfalfa Sprouts

There has been an increased demand for sprouts due to their popularity as a health food. Alfalfa sprouts are food which has high nutritional value and are good source of vitamins and minerals. Outbreak investigations have identified several factors that affect the microbial safety of sprouted seeds. Contaminated seed has been the most likely source for most outbreaks. Seed contamination can occur at the farm, seed processor, or sprouting facility (NACMCF, 1999). Frequent failures to isolate pathogens from implicated seeds suggests that seed contamination may be intermittent, at very low levels, or unequally distributed within seed lots. However, even low levels of human pathogens are a concern because pathogens on seeds can grow quickly under the favorable conditions of the sprouting process (Itoh et al., 1998; Jaquette et al., 1996).

### 2.1. Microbial Ecology of Pathogens Associated with Sprout

Various surveys have revealed aerobic plate counts (APC) of 3-6 x $10^4$ CFU/g (Andrews et al., 1979), and 5-400 x $10^3$ CFU/g (Prokopowich and Blank, 1991) on alfalfa seeds. From surveys by Patterson and Woodburn (1980), alfalfa sprouts sold from retail outlets had APC of $10^8$-$10^9$ CFU/g. They suggested that it is important to note that high microbial level per se is not necessarily of public

health concern in sprouts. It is the presence and growth of pathogenic microbes that is the major concern. Most microbial analysis of seeds or sprouts have not found pathogenic bacterial species such as: *Salmonella* (Patterson and Woodburn, 1980; Prokopowich and Blank, 1991; Splittstoesser et al., 1983), *Bacillus cereus* (Sly and Ross, 1987; Splittstoesser et al., 1983; Patterson and Woodburn, 1980;), *Staphylococcus aureus* (Sly and Ross, 1987; Splittstoesser et al., 1983) and Listeria (Prokopowich and Blank, 1991).

Pathogenic bacteria can be introduced to sprouts by a number of pathways, including via seeds, water used during germination and sprouting, unsanitary production practices, or mishandling by the consumer (CDC, 1997a; Patterson and Woodburn, 1980). Seeds are suspected to be the most likely source of pathogenic contamination. Conditions during sprouting, e.g. time, temperature, water activity, pH, and nutrients, are ideal for growth of pathogenic bacteria such as Salmonella and *E. coli* (Itoh et al., 1998; Hara-Kudo et al., 1997; Jaquette et al., 1996; Prokopowick and Blank, 1991).

## 2.2. Sprout Associated Outbreaks

Enterohemorrhagic *E. coli* O157:H7, and various types of salmonella have been the causative agents of documented outbreaks of foodborne illness associated with sprouts (Jackson, 1998). Epidemiological and microbiological evidence suggest that seeds were the source of the pathogens. *E. coli* O157:H7 can reproduce rapidly during sprout production (Itoh et al., 1998; CDC, 1997). Reported outbreaks of salmonellosis associated with sprouted seeds suggest an

initial low level contamination of alfalfa seeds, followed by growth during sprouting (Mahon et al., 1997).

### 2.3. Seed Treatment

A number of researchers have investigated the ability of chlorine compounds to inactivate pathogenic bacteria on alfalfa seeds. Beuchat (1997) inoculated alfalfa seeds with Salmonella and subjected them to a variety of treatments (calcium and sodium hypochlorite, hydrogen peroxide, and ethanol) at different concentrations. Significant reductions in Salmonella populations were observed with most increased in concentration of the test chemical.

A study by Moline (1999) concluded that 2 % calcium hypochlorite showed the greatest reduction, but did not completely eliminate the natural microflora. Calcium hypochlorite solution up to 20,000 ppm did not affect percent germination of the alfalfa seeds. Beuchat (1999), however, found that at high concentrations, alfalfa seed rate of germination was slowed. FDA (1999) and NACMCF (1999) suggested that the sprout industry should pursue the use of calcium hypochlorite at 20,000ppm (a 2% solution) for soaking of the seeds prior to germination and growth. This seed treatment has the potential to substantially reduce microbial contamination of seeds which can be passed on though the growing sprouts.

## 3. Bacteria Associated with Fresh Vegetables

### 3.1. *Escherichia coli*

*Escherichia coli* is classified as an Enterobacteriaceae. It is a gram-negative, facultatively bacterium. *Escherichia coli* resides in the lower part of the intestinal tract of warm-blooded animals. Ordinarily this bacterium is harmless, but some strains are pathogens (Riemann and Bryan, 1979).

*E. coli* grow well on a large number of media and in many foods. These organisms have been reported to grow at temperature as low as -2 °C and 50 °C (Riemann and Bryan, 1979). In foods, growth is poor or very slow at about 5 °C although several authors have reported the growth of coliforms at 3-6°C, (pH 4.4-9.0). *E. coli* can be grown on media containing only an organic carbon source such as glucose and a source of nitrogen such as $(NH_4)_2SO_4$, along with other minerals. Consequently, these organisms grow well on nutrient agar and produce visible colonies within 24 hrs at 37°C (Jay, 1996).

#### 3.1.1. *E. coli* as a Fecal Indicator

It is, in general not feasible to examine every food or food product for the presence of hazardous organisms. The practice that has been in effect for many years and continues to be followed is to determine the sanitary quality of foods by their content of certain indicator organisms (Eley, 1992).

The use of *E. coli* as an indicator of water-borne pathogens was apparently first suggested in 1892 by F.Schardinger. Jay (1996) suggested that

since this organism is generally present in the intestinal tract, its presence outside the intestines can be regarded as contamination with fecal discharges of man or animals. Consequently, *E. coli* was employed as an indicator of fecal pollution in water and in foods (Choradash and Insalata, 1978).

The finding of large numbers of these organisms in foods and water is taken to indicate fecal pollution or contamination and the possibility that the etiologic agents of these diseases may be present. Whether or not intestinal pathogens are present, the presence of fecal matter in foods or water is undesirable (Jay, 1996).

### 3.1.2. Pathogenic *E. coli*

Currently, there are four recognized classes of enterovirulent *E. coli* (collectively referred to as the EEC group) that cause gastroenteritis in humans and are of public health importance. Among these are the Enterotoxigenic (ETEC) , Enteropathogenic (EPEC), Enteroinvasive (EIEC), and *Escherichia coli* O157:H7 (Enterohemorrhagic *E. coli*, EHEC) strains. These organisms have been isolated from numerous foods and water sources and have been incriminated in outbreaks of food and waterborne gastroenteritis in both the U.S. and other parts of the world (Shelton and Karns, 2001; Mead et al., 1999). Outbreaks of *Escherichia coli* O157:H7 have been linked to alfalfa sprouts (CDC, 1997), radish sprouts (Nathan, 1997), lettuce (Ackers et al., 1996), and apple cider (CDC, 1996).

### 3.2. Salmonella

Salmonella is a rod-shaped, facultative, nonsporeforming and gram-negative bacterium. Salmonellae are usually motile (nonmotile exceptions S. gallinarum and S. pullorum) by peritrichous flagella, usually catalase positive, oxidase negative, produce acid and sometimes gas from the fermentation of glucose, and reduce nitrates to nitrites. Most members of this family are found in the intestinal tract of man and animals as comensals or pathogens (Sharma and Carlson, 2000; Ekperigin and Nagaraja, 1998).

Environmental sources of the organism include water, soil, manufacturing, kitchen surfaces, animal feces, raw meats, raw poultry, and raw seafood. Contaminated foods are often of animal origin, such as beef, poultry, milk, or eggs, but any food may be a contaminate source. Vegetables may also become contaminated (Stanley, 1997). Agricultural practices, food processing, food distribution, and food service operations influence the transmission of salmonellae and the occurrence of outbreaks. Salmonellae can also survive for a long time in litter, soil, animal feces, trough water, and other substances in a farm environment (Ferretti et al., 2001; Pignato, 1996).

Salmonella have simple nutritional requirements and can grow on a large number of culture media and produce visible colonies well within 24 h at 37°C (Riemann and Bryan, 1979). They are generally unable to ferment lactose, sucrose, or salicin, although glucose and certain other monosaccharides are fermentable, with the production of gas. The optimum growth temperature for

salmonellae is 35-37 °C and, although there is some variation with serotype, salmonellae can grow at temperatures between 5 and 45 °C. Growth at temperatures below 10 °C, however, is very slow. Salmonellae grow between pH 4.5-9.0 with an optimum growth range of pH 6.5-7.5, pH above 9.0 and below 4.0 is bactericidal (Becker et al., 1986). Salmonellae do not compete well with food spoilage organisms, Escherichiae, or lactic acid bacteria (Jay, 1996)

The incidence of salmonellosis appears to be rising both in the U.S. and in other industrialized nations. *Salmonella enterica* serotype Typhimurium and *Salmonella* serotype Enteritidis are the most common in the United States (Centers for Disease Control and Prevention, July 16, 1999). Recently, the most common cause of food poisoning by a Salmonella species was due to S. Typhimurium. Salmonellosis is characterized by diarrhea, abdominal cramps, vomiting and nausea.

Salmonella have been isolated from many types of raw fruit and vegetables (Beuchat, 1996b; Wells and Butterfield, 1997). Outbreaks of salmonellosis have been caused by a diversity of fruits and vegetables, including alfalfa sprouts (Daussard, 1999; Jaquette et al., 1996), tomatoes (Zhuang et al., 1995), bean sprouts (Mahon et al, 1996), melons (Blostein, 1991), and unpasteurized orange juice (Cook et al., 1990).

Methods have been developed for many foods having a history of Salmonella contamination. Conventional culture methods require 5 days for

presumptive results, while rapid methods available today require only 2 days (Peplow et al., 1999; Steel et al., 1997).

## 4. Conventional Method for Foodborne Pathogen Detection

In microbiology the conventional growth-amplification techniques (e.g. colony counting) for the estimation of microorganisms are both tedious and time consuming, requiring 24-48 hours for a total viable count, while detection of a specific organism (e.g. *E. coli*) may require an additional 24-48 hours incubation. In conventional methods, to identify bacteria, the organisms must be isolated, including preenrichment and enrichment processes, from specimens and then identified using morphological, biochemical, and serological tests.

### 4.1. Colony Counting Method

For colony counting, the standard plate count is the most widely used method for determining the numbers of viable cells or colony forming unit (CFU) in food products (Jay, 1996). Plate counts are essentially direct counts of numbers of colonies in, or on agar plates, each colony being assumed to represent the progeny of one organism. In the colony counting method, food samples are homogenized, serially diluted in an appropriate dilutant, plated in a suitable agar medium, and then incubated at appropriate temperature and time. After inoculation and incubation, colonies are counted. The number of viable cells per unit amount of the original sample is then calculated (Penn, 1991; Kiss, 1984).

## 4.2. Identification of Foodborne Pathogens

The conventional methods of laboratory diagnosis usually consist of utilizing a range of culture media to encourage growth of the pathogen, even in the presence of many other contaminating bacteria (Eley, 1992).

For detection and identification of infective agents, the diagnosis of infectious diseases involves growth of microorganisms on plates or in broth (including preenrichment and enrichment media) to provide damaged organisms with an opportunity to repair physiologically and to repress populations of competitive flora whilst encouraging growth of pathogens. This usually requires several hours for incubation, followed by morphological colony, serological characterization, and biochemical testing for identification of the organisms.

Recently, a number of commercially available culture test kits have been introduced which facilitate more rapid detection of certain pathogens. One of these test kits utilizes a selective enrichment motility technique for detecting target microorganisms in foods. Following pre-enrichment, a tube containing a selective medium and an upper indicator medium, separated by a porous partition, is inoculated with the organism.

## B. SPME/GC METHODS FOR THE ANALYSIS OF VOLATILE COMPOUNDS

### 1. Introduction

Gas chromatographic (GC) analysis of compounds excreted by microorganisms in growth media is useful for the identification of microbial species. Gas chromatographic methods have been qualitatively and quantitatively employed to determine the specific catabolic products of various specific compounds such as fatty acids, glucose, amino acids, or other metabolic compounds produced by microorganisms. Gas chromatographic analysis of microbial cells or their products is useful not only for taxonomic differentiation of the organisms but also for diagnostic purposes.

Gas chromatography has been applied to pure cultures to identify bacterial species. However, analyses of volatile or partially volatile organic compounds with gas chromatography techniques usually begin with concentration of the analytes of interest though liquid-liquid extraction, purge-and-trap, headspace, or various other techniques. These procedures typically require long times, complicated equipment, and tremendous amounts of organic solvents.

The development of alternative analytical techniques that are rapid and simple has become increasingly important to reduce per sample time investment and to conduct real time analyses. Solid Phase Microextraction (SPME) is a rapid sampling technique which is well adapted to GC analysis of volatile compounds. SPME has been applied to the analysis of volatile and nonvolatile compounds in gaseous and liquid samples. SPME can eliminate the need for

solvents or complicated apparatus for concentrating volatile or nonvolatile compounds in liquid samples or in gas samples (Bao et al., 1999).

## 2. Solid Phase Microextraction

SPME is a simple, effective adsorption/desorption technique, which can eliminate the need for solvents or complicated apparatus for concentrating volatile or non volatile compounds in liquid samples or in gas samples. SPME can extract analytes from a variety of matrices by partitioning from a liquid or gaseous sample into an immobilized stationary phase. SPME is compatible with analyte separation/detection by gas chromatography or High Performance Liquid Chromatograph (HPLC), and provides linear results for a wide concentration of analytes (Song et al., 1998).

## 3. Theory of Solid Phase Microextraction

The principle of SPME is centered around the partitioning process of the analyte between the fiber coating and the sample. SPME eliminates preconcentration steps by directly extracting the analytes into a poly(dimethylsiloxane)-coated fiber. SPME is a multiphase (fiber coating, gas phase or headspace, and a homogeneous matrix such as pure water or air) equilibration process. The system consists of an aqueous phase with suspended solid particles having various adsorption interactions with analytes and a gaseous headspace.

The kinetics of the extraction process determines the speed of extraction. Mass transport theory is based on Fick's second law of diffusion describing mass balance in a dynamic system (Zhang and Pawliszyn, 1993).

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} \tag{3.1}$$

where *c(x,t)* is the concentration of the analyte at position *x* and time *t*, and *D* the diffusion coefficient of the analyte. For a one dimensional diffusion process, the mass of analyte absorbed by the polymeric coating at any given time, *M(t)*, can be calculated by;

$$M(t) = \int c(x,t)dx \tag{3.2}$$

All diffusion is assumed to behave according to Fick's law. Factors such as thermal expansion, swelling, and analyte/analyte interactions are assumed to be negligible (Gorecki and Pawliszyn, 1997).

Theoretically, SPME extraction is considered to be completed when the analyte concentration has reached distribution equilibrium between the sample matrix and the fiber coating (Gorecki et al., 1998).

## 4. Principle of Headspace Volatile Sampling by SPME

Headspace SPME is based on the equilibrium of analytes among three phases of the system. These three phases include the polymeric liquid coating, the headspace, and the aqueous solution (Zhang and Pawliszyn, 1993). SPME is an equilibrium method and once equilibrium has been reached, the

concentration of the analytes can be considered constant in all three phases. The limiting step in this process is considered to be the diffusion of the analytes through the system. For this reason the equilibrium time of the system must first be determined (Steffen and Pawliszyn, 1996; Goercki et al., 1998).

In general, volatile compounds require a thick coating, and a thin coating is most effective for adsorbing/desorbing semivolatile analytes. Full equilibration is not necessary for high accuracy and precision from SPME, but consistent sampling time and other sampling parameters are essential (Bao et al., 1999).

Desorption of an analyte from the SPME fiber depends on the boiling point of the analyte, the thickness of the coating on the fiber, and the temperature of the injection port. Some analytes can take up to 30 seconds to desorb, and cryogenic cooling may be required to focus these compounds at the inlet of the capillary column. Use of an inlet liner with a narrow internal diameter (e.g. 1mm) generally provides sharp peaks and can eliminate the need for cooling.

Heating a sample to elevated temperature provides energy for molecules to overcome energy barriers that tie them to the matrix. However, the absorption of analytes by the fiber coating is an exothermic process, which means that while the high temperature is good for the release of analytes from their matrix, it can adversely affect the absorption by the coating due to a decrease in the partition coefficients. As a result, there is usually an optimum temperature for headspace SPME sampling (Jelen et al., 1998; Forsyth and Dusseault, 1997).

## 5. Application of Solid Phase Microextraction

Solid Phase Microextraction (SPME) has been introduced as a modern alternative to current sample preparation technology, and has a wide range of applications. The use of a fiber for extraction can enhance the selectivity of the analysis because one may choose the stationary phase that best suits the analytes. By using headspace SPME, one can reduce matrix effects and interference present in the liquid sample (Gorecki and Pawliszyn, 1997).

In industry, practical uses of SPME can be found in environmental, food, pharmaceutical, clinical and forensic applications. SPME minimizes sample preparation and concentrates volatile analytes in a solvent-free manner. Headspace SPME is shown to be more sensitive than conventional headspace analysis of similar samples performed with an airtight syringe, and favorable to the commonly used purge and trap type analysis.

The SPME/GC technique has been utilized in such areas of food/flavor chemistry as pesticide determination in wine (Nogueira and Nascimento, 1999), examination of volatile compounds produced by tomato, and strawberry fruit (Song, et al., 1998), analysis of alcohols and esters in beer (Jelen et al., 1998), monitoring of volatiles of apple (Song et al., 1997), determination of methylcyclopentadienly manganese tricarbonyl, a gasoline antiknock additive in beverages (Forsyth and Dusseault, 1997), study of off-flavors in milk (Marsili, 1999), analysis of rancidity in dry-cured ham (Ruiz et al., 1998), detection of trace levels of taste and odor compounds in water (Bao et al., 1999) and

characterization of volatile metabolites emitted by microorganisms (Vergnais et al., 1998; Borjesson et al., 1993; Alugupalli et al., 1992).

Recently, the SPME technique has been applied to the analysis of bacterial metabolites. Vergnais et al. (1998) evaluated the solid phase microextraction for analysis of volatile metabolites produced by Staphylococci. It has been shown that SPME was able to extract esters and oxidation of free fatty acids from leucine catabolism of *Staphylococcus xylosus* and *Staphylococcus camosus*.

Nilsson et al. (1996) use headspace microextraction for the analysis of volatile metabolites emitted by *Penicillium* species. Headspace solid phase microextraction was used to collect volatile compounds emitted from six fungi of the genus *Penicillium* grown on yeast extract sucrose agar incubated at 25°C for 4 days. GC/MS was employed for the analysis of the profiles of volatile metabolites characteristic for each species. The results obtained by HS-SPME compared favorably with those obtained by Tenax adsorption. The results are in good accordance with earlier results obtained from the same fungi by collection on Tenax tubes and the formation of the tentatively identified metabolites is supported by biosynthetical considerations.

Guerzoni et al (1992) establish the HS-GC determination of the $CO_2$ produced by lactic acid bacteria. A rapid gas chromatographic and automatic methodology for the detection of viable microbial cells in food has been proposed (Song et al., 1997). This indirect method is based on the gas

chromatographic analysis of volatile metabolites in equilibrium in the headspace of sealed vials containing a Lactobacilli MRS-tomato juice broth medium inoculated with an aliquot of the sample under examination. The enzymatic activity of the strains was arbitrarily evaluated on the basis of the rate of $CO_2$ production in the headspace expressed as $\Delta\%CO_2$/hour when the initial load was $10^8$ cells/mL in the test medium (Guerzoni et al., 1992).

Alugupalli, Larsson, and Slosarek (1992) developed SPME/gas chromatography/mass spectrometry methods for detecting 2-decosanol, a alcohol characteristic of *Mycobacterium xenopi* in drinking water incubated at 42 °C for 12 weeks.

## C ELECTRONIC SENSOR ARRAY

### 1. Introduction

The conventional method used for foodborne pathogen detection relies on colony counting method. Major constraints of this method are that it is time consuming and laborious. Several research groups have been working to develop innovative and more rapid methods for detection of foodborne pathogens. They are based on different principles, such as immunoimmobilizaton, enzyme immunoassays, DNA probes, immuno-PCR, and hydrophobic grid membrane filters (Sharma and Carlson, 2000; Pignato et al., 1996).

The physicochemical properties of the products are also measured using conventional analytical equipment, such as gas chromatography and gas chromatography-mass spectrometry. Various substrates can be used to determine intermediary products produced as a result of catabolic activity of organisms. In turn, these can indicate various biochemical activities carried out by organisms in *vitro* and in *vivo* (Mitruka, 1975). However, complex mixtures may cause difficulty in manual interpretation of the GC patterns due to the large number of peaks.

The GC techniques are relatively complicated; therefore, trained personnel are required to use the methods most effectively. Extraction procedures, purification, derivative preparation, ion exchange, etc., are invaluable adjuncts to GC techniques (Kaipainen et al., 1997). Electronic sensors make an analysis of the volatile compounds emitted by a sample and

perform a classification process. The responses are then associated by statistical treatments and/or a neural network analysis and compared with an odor library resulting in a classification or prediction. The advantages of electronic sensor array over classic GC are that they are simpler to use and faster (Liden et al., 2000).

## 2. Electronic Sensor Array Technology

An electronic sensor system mimics the human olfactory system with sensitivity, reproducibility, and high levels of discrimination for the detection of simple and complex odors (Sberveglieri, 1998). Each chemical sensor represents a group of olfactory receptors and produces a time-dependent electrical signal in response to an odor. For data processing, pattern recognition analysis is the classification and memorizing which is the equivalent process in the final stage of the human olfactory process in the cerebral cortex of the brain (Bartlett et al., 1997).

The electronic sensor array is an analytical instrument combining an array of sensors, which can be specifically used for the analysis, identification, and recognition of complex odors and volatile organic compounds. The interaction of the volatiles on the sensing element causes changes in electrical resistance of the sensor. Since sensor kinetics are different, the data generated are converted into an odor fingerprint. This fingerprint is stored and can be used for comparison to standard samples. By storing the fingerprint of the odor, the electronic sensors can be used as a quality control, and a research and

development tool (Weber and Poling, 1996). Different pattern recognition techniques can be applied to the data including Principal Component Analysis, Hierachial cluster analysis, and Linear discriminant analysis (Sberveglieri, 1998; Barlett et al., 1997).

## 3. Sensor Technology

There are many types of sensors which can be used to detect specific gases and vapors. A basic requirement is that the sensors in an electronic sensor array must show partial sensitivity, i.e., that they can respond broadly to a range or class of gases rather than to a specific one. The optimum combination of sensors, sensor type and number will depend on the particular application (Lucas, Poling, and Bennincasa, 1998).

### 3.1. Metal Oxide Sensors

Metal oxides are semiconducting materials which are gas-sensitive. These sensors are comprised of a thin layer (50 μm) of an oxide film deposited on a ceramic tube or plate. The selectivity of the sensors is related to different catalytic amounts of a doping metal introduced as a trace impurity on the sensor surface. Reaction of an odorant with a sensor changes its conductivity (Lee et al., 2001; Mielle, 1996).

There are two main types of semiconductors: n-type semiconductors (mainly composed of zinc, tin or iron oxides) that respond primarily to reducing compounds, whereas p-type semiconductors (mainly nickel or cobalt oxide)

respond to oxidizing compounds. Other metal oxides, zinc oxide and tungsten oxide, are also available for this instrument.

These devices operate at elevated temperatures to avoid interference from water and to aid in rapid response and recovery times. The detection principle is based on measurement of the variation of the resistance of the metal oxide, thus, generating a measurable electronic signal.

Metal oxide sensors have sensitivities in the ppm, to ppb range (for special sensors) to a very broad range of chemical compounds. Due to their relatively low selectivity (all sensors respond in some measure to volatile compounds), the use of an array leads to a specific selectivity; i.e. a pattern or fingerprint. Metal oxide sensors are relatively resistant to humidity and to aging, and are made of particularly strong materials.

### 3.2. Conducting Polymer Sensors

Similar in principle to metal oxide sensors, a change in resistance of conducting polymer sensors is measured as a reaction to the sample vapor. Each individual device has a distinctive response characteristic to various volatiles. The polymers are very responsive to vapors from polar molecular species, however, they have very little or no response to alkane and non polar species. This makes them complementary to metal oxides which respond strongly to both types of species (Persaud and Pelosi, 1992).

The fabrication of these devices involves the deposition of a very thin film of an electrically conducting polymer material. They are based on heterocyclic

molecules, monomers, which are electropolymerized with various counter-ions in a solvent between two electrodes. These sensors work at ambient temperature with good discriminatory power. Typical sensitivities, however, are mostly in the ppm range, and the stability/ drift is higher than for other sensor technologies (Martin, Santos, and Agopito, 2001).

### 3.3. Quartz Crystal Micro-balance Sensors

These sensors consist of a piezoelectric quartz crystal oscillator coated with a sensing membrane such as acetyl cellulose or lecithin. Quartz crystal micro-balance sensors are based on a change in frequency due to the absorption of sample vapor on the sensor coating. These changes result in a distinct fingerprint for the sample. When the sensor is exposed to gas molecules, absorption and desorption occur from the coating (Ema et al., 1989). The gas, soluble in the coating, will increase the added mass on the crystal and decrease the frequency of the oscillation according to:

$$\Delta f = k \Delta m \qquad (3.3)$$

where $\Delta f$ is resonance frequency, $\Delta m$ is mass change caused by absorbed gas molecules, and K is a constant and refers to the basic resonance frequency and mass of quartz plate. The sensitivity and selectivity of the sensor depends on

the coating material selected and the quantity deposited (Stetter et al., 2000; Mielle, 1996).

## 4. Principle of Electronic Sensor Operation

The volatile chemical sensing system consists of three software modules; data acquisition and instrument control, data manipulation for extraction and preparation of signals, and data processing.

Data acquisition software samples the sensor array resistance at regular intervals storing the resultant data in the computer. As the resistance of the conducting polymers are inversely proportional to temperature, the temperature of the array is controlled and monitored. Sample temperature and sample humidity are also monitored. The signal is expressed as the percentage resistance change of each sensor compared to the initial sensor resistance (Persaud et al., 1999).

Each sensor element changes in resistance when exposed to volatile compounds. The degree of response to a given substance depends on the type of sensor used (Hatfield et al., 1994). The sensor response or sensitivity has been presented in a variety of ways, including the relative resistance value, $R_{gas}/R_{air}$ ; log relative resistance value; conductance difference, $G_{gas} - G_{air}$ or the fractional conductance change, $(G_{gas} - G_{air})/G_{air}$ , where $R_{gas}$, $R_{air}$, $G_{gas}$, $G_{air}$ are the resistance or conductance in gas or air respectively (Gardner et al., 1992; Gardner, 1991). The change in resistance when presented with a volatile

chemical relative to the base resistance in air $(R_{gas}- R_{air})/R_{air}$ is a convenient measure of response (Persaud et al., 1996).

### 4.3. Data Analysis

Data interpretation of multiple variables of several sensor responses requires the use of statistical interpretation for rapid verification of results. The premise of statistics is to reduce multi-dimensional data of several sensor responses to two (or three) dimensions.

The relevance of certain methods of data analysis with respect to sensor arrays was summarized by Ortega et al. (2000). They stated that methods such as discriminant analysis or partial least squares are parametric, i.e., they rely on a known probability distribution of the variables. Nonparametric methods such as principal components analysis have also proved useful (Kokot et al., 1998).

## 5. Application of Electronic Sensor Arrays

The electronic sensor array has been developed to fulfill a real need in the food industry for objective, automated, quality-monitoring sampling systems that can characterize the odor, and thus determine whether the production system is running to specification without requiring human sensory panels (Mille, 1996).

A considerable number of applications of the electronic sensors have been reported including, detection of contaminated soils (Getino et al., 1998), evaluation of the off-odor in salmon fillets stored at different temperatures (Luzuriaga and Balaban, 1998), evaluation of physiological maturity of tomatoes

(Maul et al., 1998), characterization of vinegar (Anklam et al., 1998), detection of boar-taint in meat (Bourrounes et al, 1995), estimation of fish freshness (Schweizer-Ber-berish et al, 1994), and quality estimation of ground meat (Winquist et al, 1993). In addition, the electronic sensor array technology meets a pronounced analytical need in biotechnology research and bioindustrial applications.

Electronic sensor array has the advantage of being non-invasive and allowing on-line monitoring in the off-gas effluents from a production system sensitive to microbial contamination. Other analytical techniques currently used in on-line bioanalysis, such as near-infrared spectroscopy and immunosensors, although providing real-time analyses, still lack the quality of non-invasiveness (Liden et al., 2000).

Bazemore and Rouseff (1998) reported that the electronic sensor could be used to discriminate between different seasonal cultivars (early-mid, and late), and different heat treatments (unpasteurized, 8s at 96 °C, 120 s at 96 °C, and 180 s at 96 °C). Bazzo et al. (1998) investigated the sensitivity of the electronic sensor for quality control of an edible oil. They reported that the electronic sensor could be used to evaluate the different qualities of an edible oil and provide good discrimination and stability for successful identification of oil quality. Repeated electronic sensor measurements also showed consistent identification and good recognition. There was good correlation between GC data, sensory panel results, and the electronic sensor responses.

Kaipainen et al. (1997) studied the possibility of partly substituting sensory panel work with the electronic sensors. The results showed that an electronic sensor array could rapidly classify volatiles from the samples. The electronic sensor result was better than the sensory panel description and GC/MS results. The electronic sensor was able to discriminate between acceptable and unacceptable sugar on the basis of volatile compounds.

Bachingger et al. (1998) used an electronic sensor array for on-line estimation of the glucose and ethanol concentrations in batch fermentation with *Saccharomyces cerevisiae.* They reported that the technique was a non-invasive, and real-time method which could significantly improve bioprocess monitoring and control.

## D. CHEMOMETRICS

### 1. Introduction

Chemometrics is an approach to analytical and measurement science based on the idea of indirect observation (Domingo et al., 2000; Mawatari et al., 1999; Luco, 1999; Wentzell, Andrews and Kowalski, 1997).

### 2. Principal Component Analysis (PCA)

PCA is an unsupervised technique commonly used in signal processing and pattern recognition. PCA is used to reduce the dimensionality of multivariate data whilst preserving most of the variance, and so is an excellent technique for observing the natural relationships between samples. The transformation is

designed in such a way that the data set may be represented by a reduced number of 'effective' features, while retaining most of the intrinsic data content. This means that the data set undergoes a dimensionality reduction (Haykin, 1999; Von der Malsburg, 1990; Oja, 1982).

The algorithm that is used to perform the principal components analysis on a data vector is as follows;

$$R = \mathbf{x}^T\mathbf{x} \tag{3.4}$$

where R represents a correlation matrix of data vector x and superindex T represents the transposed matrix. The correlation matrix R can be expressed in terms of its eigenvalues and eigenvectors as:

$$R = \sum_{j=1}^{m} \lambda_j q_j q_j^T , \qquad j = 1,2,\ldots,m \tag{3.5}$$

where $\lambda$ corresponds to the eigenvalues and the associated eigenvectors m-by-m matrix, $Q = [q_1 , q_2 ,\ldots, q_m ]$. The eigenvalue problem, commonly encountered in linear algebra, is recognized as;

$$Rq = \lambda q \tag{3.6}$$

For basic data representation,

$$a_j = q_j{}^T x \tag{3.7}$$

where $a_j$ are the projections of x onto the principal directions.

$$a = [a_1, a_2, \ldots a_j]^T \tag{3.8}$$

$$x = \sum_{j=1}^{m} a_j q_j \tag{3.9}$$

The set of principal components retained can be calculated as follows.

$$\hat{x} = \sum_{j=1}^{l} a_j q_j \tag{3.10}$$

As described by Malinowski (1991), PCA The latent variables can be written as a weighted sum of the original variables.

$$X = TP^T + E \tag{3.11}$$

where $T$ is a scores matrix, $P$ is a loadings matrix (the superindex $T$ indicates the transposed matrix), and $E$ is the residual error matrix. The results of PCA can be displayed in a score plot. In this plot, the scores of the different objects are plotted as a function of the latent variables. A loading plot shows which of the original variables make an important contribution to the latent variables and which of the original variables are well correlated (Lamberto and Saitta, 1995).

### 3. Discriminant Factor Analysis (DFA)

Discriminant Factor Analysis (DFA) is related to both multivariate analysis of variance and multiple regressions. Discriminant Factor Analysis can be used not only to test multivariate differences among groups, but also to explore which variables are most useful for discriminating among groups, if one subset of variables performs equally well as another, and which groups are most alike and most different (Hill and Engelman, 1992).

Classification is performed by assigning a pattern vector to the class with the closest Mahalanobis distance metric (Fung, 1995; Aishima, 1979), as given by

$$D_i^2 = \sum_{j=1}^{c}\sum_{i=1}^{n}(\bar{x}_{ij} - y_i)s^{-1}(\bar{x}_{ij} - y_i)^{-1} \qquad (3.12)$$

where $n$ is the dimensionality of the pattern, $c$ is the number of classes, $\bar{x}_j$ is the vector of means for class $j$, $y$ is the pattern vector being classified, and $s$ is the pooled variance-covariance matrix (Fisher and van Belle, 1993). A discriminant function is expressed as an equation as follows.

$$Z = a_1X_1 + a_2X_2 + a_iX_i + .... + a_mX_m \qquad (3.13)$$

In DFA, the samples are projected from their places in the complete measurement space into a suitable sub-space (Fisher and van Belle, 1993;

Aishima, 1979). For Discriminant functions, the variables are chosen according to their characteristics that differ between the groups. These variables are then linearly combined and weighted so that the groups are forced to be as statistically distinct as possible by choosing the linear combination of variates that maximizes the one-way analysis of variance F-test, which tests the equality of the means for the linear combinations.

## E. ARTIFICIAL NEURAL NETWORK

A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in 2 aspects: knowledge is acquired by the network from its environment thought a learning process, and interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge (Hykin, 1999). Neural networks have the advantage that they can handle nonlinear data and are more tolerant to noise of the system, and tend to produce lower prediction error rates than chemometric techniques (Packianather et al., 2000; Schryer and Mikkelsen, 2000).

### 1. Multilayer Perceptron (MLP)

The primary element of an ANN is the neuron. These neurons are arranged in input and output layers of one or more hidden processing layers (Nakamura and Yoshikawa, 2001). The most common neural network approach to regression-type problems is multilayer perceptrons (MLP). The popular

algorithm is known as a back propagation algorithm (Martin, Santos, and Agaptito, 2001). The majority of published work uses a feed forward net with back propagation for training. The inputs to a neuron include its bias and the sum of its weighted input. The output of a neuron depends on the neuron's inputs and on its transfer function (Devabhaktuni et al., 2001; Kimura and Nakano, 2000; Watanabe, 2000). In mathematical terms, a neuron *k* can be described by

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{3.14}$$

where $w_{k1},...,w_{km}$ are the weights of neuron k;

$u_k$ is the linear combiner output due to the input signals; $b_k$ is the bias; $\varphi(.)$ is the activation function; and $y_k$ is the output of the neuron.

$$y_j^l = \varphi_j(v_j(n)) \tag{3.15}$$

For MLP, the activation functions used in the network can be linear and/or a sigmoidal nonlinear function. A linear function can be calculated as

$$\varphi(v_j(n)) = \beta_o + \sum_{j=1}^{m} \beta_j x_j \tag{3.16}$$

where β is the regression coefficient and x is the input signal. The nonlinear sigmoidal activation functions frequently used include

a) logistic function:

$$\varphi_j(v_j(n)) = \frac{1}{1+\exp(-v(n))} \tag{3.17}$$

b) hyperbolic tangent function:

$$\varphi_j(v_j(n)) = \tanh(v_j(n)) \tag{3.18}$$

Figure 2.1 shows the architectural graph of multilayer perceptrons with one hidden layer and an output layer. The function signal is presumed to perform a useful function at the output of the network and at each neuron of the network through which a function signal passes, the signal is calculated as a function of the inputs and associated weights applied to that neuron (Hikawa, 2001; Tanaka and Hasegawa, 2001).

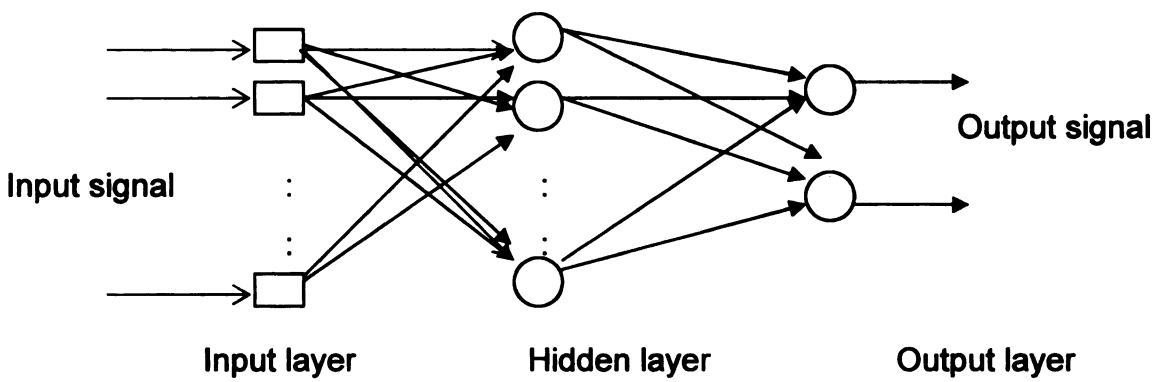


**Figure 2.1. Architecture of a multilayer perceptron with one hidden layer**

Neural networks have been applied to a wide variety of applications including, dynamic system control (Watanabe et al., 2001), electromagnetic optimization of microwave circuits (Bila et al., 1999), robotic manipulators (Terra and Tinós, 2001), and hydrologic events forecasting (Coulibaly et al., 2001).

# CHAPTER 4

# ELECTRONIC SENSOR ARRAY INCORPOATING ARTIFICIAL NEURAL NETWORK FOR ANALYSIS OF VOLATILE COMPOUNDS PRODUCED BY *ESCHERICHIA COLI* IN NUTRIENT MEDIA


## ABSTRACT

An electronic sensor array with 12 non-specific metal oxide sensors was evaluated for its ability to monitor volatile compounds in super broth alone and that inoculated with *E. coli* (ATCC25922) at 37 °C for 2-12 hours. Principal Component Analysis (PCA) was used for data exploration and dimensional reduction with 96% of the data accounted for. Using Discriminant Function Analysis (DFA), it was possible to differentiate super broth alone from that containing *E. coli* when cell number was $\geq 10^5$ CFU. The sensor array could also quantitatively discriminate between numbers of *E. coli* in samples. There was good agreement between the volatile profiles from the electronic sensor array and SPME/GC/MS. The potential to predict the number of *E. coli* and the concentration of specific metabolic compounds was investigated using Artificial Neural Network (ANN). Electronic sensor array incorporating neural network is a rapid and simple technique which can identify and quantify the number of *E. coli* in nutrient media.

## 4.1. INTRODUCTION

Many methods have been used to determine the presence of microorganisms in media or in food product including fiber optics, calorimetry, piezoelectric crystals, and flow cytometry. Gas chromatography/mass spectrometric methods have also been used for the detection and identification of microorganisms. These techniques have limitations (Erti & Mikkelsen, 2001).

The use of an array of sensors of varying affinities allows for the relative responses between the sensors to be used to produce a unique fingerprint of the volatile compounds. The volatile fingerprint generated by an electronic sensor array can be displayed graphically for quality control, or compared with a control, or other samples using neural network algorithms.

Electronic sensor technology has been used in the food and drink industries to detect taints and off-taste in food-products, more recently work has been reported on its application in environmental monitoring. Some researchers have reported on its use in assessing agricultural and sewage odors, by showing the relationships between sensor responses and odor concentration, whereas others have demonstrated the ability of a non-specific sensor array to differentiate between tainted and untainted waters for detecting off-flavor, monitoring bioprocess ingredients for detecting microbial contamination, and predicting different bacterial types and growth phases (Erti & Mikkelsen, 2001).

One popular technique characteristic of statistical methods and capable of deriving low dimensional representations is PCA. Discriminant Function Analysis can be used to build class classifiers. The ultimate goal of pattern recognition is

to learn classifier models whose expected performance on unseen data falls within acceptable bounds. This comes from the need to predict the degree of generalization and robustness of the classifier.

The objectives of this study were to identify volatile compounds generated by *E. coli* in super broth using SPME/GC/MS, to investigate the ability of the electronic sensor array incorporating chemometrics and neural networks to identify and quantify *E. coli.*

This research was chronologically conducted in three steps. The first step was data acquisition and collection. The second step was feature extraction using Principal Component Analysis (PCA) to explore the data under investigation, with data classification using Discriminant Factor Analysis (DFA). Predictions of the number of *E. coli* was done in the final step using three mathematical morphologies.

## 4.2. METHODOLOGY

### 4.2.1. Stock Culture Preparation

*E. coli* ATCC 25922 was obtained from the American Type Culture Collection (ATCC, Rockville, MD). *E. coli* was inoculated into Luria Bertani (LB) broth consisting of 10 g Bacto-tryptone, 5 g Bacto Yeast, and 5 g NaCl, and incubated at 37 °C in a gyrotory shaker (G-25 New Brunswick Scientific Corporation, New Brunswick, NJ) at 100 rpm. The *E. coli* suspension was then dispensed into sterile 125 mL GSA bottles, and centrifuged (Rotor-GSA model

RC 5 C Superspeed Centrifuge, Sorvall Instruments, Dupont Co., Haffman Estate, ILL) at 1600 x g for 10 min. The supernatant was decanted, and the resulting cell pellets were resuspended in a sterile 15 % glycerol solution. 1 mL aliquots of bacterial suspension were transferred into 1.5 mL microcentrifugal tubes and frozen using liquid nitrogen. The frozen cultures were stored at –80 °C.

### 4.2.2. Preparation of Test Solution

Super broth consisting of 32 g Tryptone, 20 g Yeast Extract, 5 g NaCl, 5 mL of 1 N NaOH, was used as a basal medium. Before use, *E. coli* was transferred from a stock culture to super broth and incubated at 37 °C for 10 hrs. Previous trials indicated that cell population reached $10^8$ CFU/mL at this time point. The culture was then diluted to the desired level of $10^2$ CFU/mL to allow for growth during the study. *E. coli* was inoculated at $10^2$ CFU/mL into super broth, and 5 mL of super broth were transferred into standard 20 mL headspace vials and sealed with PTFE-lined Teflon caps (Alpha M.O.S., Hillsborough, NJ) using a cap crimper.

All experiments were designed to use the headspace sampler and sealed sample vial system originally designed for use with gas chromatography and GC/MS. At time zero in the experiment, the cultures were allowed to grow in vials at 37 °C in a gyrotory shaker. The incubation times were varied, and the growth of the microorganism was measured. Samples were periodically

analyzed after incubation at 2, 4, 6, 8, 10, and 12 hours at 37 °C using a colony counting method, SPME/GC/MS, and electronic sensor techniques.

### 4.2.3. Colony Counting Method

The samples were serially diluted in sterile Butterfield's phosphate buffer. A series of dilutions was also prepared from the stock suspension. Serially diluted samples were plated in duplicate using 3M Petrifilm Aerobic Count Plates for determining total aerobic bacteria. The number of *E. coli* was determined using 3M Petrifilm *E. coli*/Coliform Count Plates containing Violet Red Bile nutrient agar and an indicator of glucuronidase activity for *E. coli*/Coliform. All plates were incubated at 37 °C for 48 ± 2 hrs. After incubation, colonies were counted and the viable cell count per unit amount of the original material calculated. Plate counts are essentially direct counts of numbers of colonies in or on plates. Each colony is assumed to represent the progeny of one organism. Plate counts are recorded as colony forming units (CFU).

### 4.2.4. Electronic sensor for monitoring of volatile compounds

An electronic sensor (Fox 3000, Alpha M.O.S., Hillsborough, NJ), with 12 metal oxide sensors (SYLG, SYG, SYAA, SYGH, SYGCTI, SYGCT, T301, P101, P102, P401, T702, and PA2) was used for monitoring changes in volatiles produced by *E. coli* in super broth medium. The system for volatile analysis, shown in Figure 4.1, combines a measurement chamber for generating the volatile compounds and a detection system. Electronic sensor analysis

conditions were determined by optimizing the following parameters, headspace generation temperature, and headspace generation time. The optimum conditions for detecting volatile compounds (Table 4.1) were then used for additional experiments. This instrument was linked to an autosampler unit capable of analyzing a total of 64 samples. The samples were placed in glass vials and sealed with crimped PTFE/metal septa.

The headspace volatile compounds from super broth and super broth inoculated with *E. coli* and incubated for 2, 4, 6, 8, 10, and 12 hours were monitored using the optimized condition. The samples were placed in the HS100 auto-sampler in arbitrary order. Prior to analysis, the vial was removed from the sample tray and placed in a temperature-controlled chamber. The automatic injection unit heated the samples to 35 °C with an incubation time of 300 second. The temperature of the injection syringe was 40 °C. The injector needle then removed 5000 µL of headspace and injected this into the sensor chamber. The delay time between two injections was 300 second.

Each injection was repeated, with separate samples (three times for all variations per day) for seven days. The electronic signals from the sensors were digitized, then transferred to the control computer. Resistance changes (difference in sensor resistance between air blank and odorous atmosphere) were recorded.

**Table 4.1. Optimum Conditions for Operating Electronic Sensors**

| | |
|---|---|
| **HS 100** | |
| Headspace generation time (s ) | 300 |
| Headspace generation temperature (°C) | 35 |
| Syringe temperature (°C) | 40 |
| Syringe type (mL) | 5 |
| Vial type (mL) | 20 |
| **SENSORS** | |
| Acquisition time (s ) | 180 |
| Acquisition period (s) | 1 |
| Delay (s) | 120 |
| Flow (mL/min) | 300 |
| Injection volume (μL) | 5000 |

### 4.2.5. Determination of Volatile Compounds Using SPME/GC/MS

Preliminary tests were done to determine the optimum SPME fiber (Supelco, Inc., Bellefonte, PA) for collecting the volatile compounds by comparing results from Polydimethylsiloxane fibers, Polydimethylsiloxane/ Divinylbenzene fibers, and Carboxen/Polydimethylsiloxane fibers. The results showed that the Carboxen/Polydimethylsiloxane fiber was the most suitable SPME fiber because it could collect more compounds produced by *E. coli* than other fibers. The Carboxen/Polydimethylsiloxane fiber was, therefore, used for all further experiments.

Before use, the SPME fiber was conditioned at 250 °C for 1 hr. The SPME process is shown in Figure 4.2. A 1 cm length of fused silica fiber, coated with a polymer is bonded to a stainless steel plunger and installed in a holder that looks like a modified microliter syringe. Volatile compounds were collected using a SPME device. The fiber is housed in a hollow, stainless steel needle. The plunger moves the fused silica fiber into and out of a hollow needle. To use the unit, the analyst draws the fiber into the needle, passes the needle though the septum that seals the sample vial, and depresses the plunger, exposing the fiber to the sample or the headspace above the sample. The fiber was pushed out of the needle and exposed to the headspace at 35 °C until equilibrium was reached.

For headspace sampling, 5 mL of liquid sample was placed into a 20 mL vial and the fiber was exposed to the head space of the media solution. Sampling temperature was 35°C, and sampling duration was 10 min., which was sufficient to permit the establishment of a near equilibrium for the compounds tested. Once sampling was finished, the fiber was withdrawn into the needle and transferred to the injection port of the gas chromatograph (HP-6890, Hewlett-Packard Co., Wilmington, DEL). Absorbed volatiles were desorbed from the fiber coating by inserting the SPME fiber through a predrilled septum (Thermogreen LB-2, Supelco Co., Bellefonte, PA) and into a glass lined, split-less injector port of a gas chromatograph. The GC column inlet was immersed in a liquid nitrogen bath during the desorption period, so that the volatiles were collected before entering the column for analysis. After 3 min of desorption, the fiber was

retracted and removed as was the liquid nitrogen. The oven was then closed and the GC was manually started.

A HP-6890 GC was used in analysis of the compounds. Volatiles were separated using a capillary column (SPB5, 30m x 0.1 mm id., 0.25 μm coating thickness). The carrier gas used was ultrapurified Helium (99.99% purity) at a flow rate of 0.5 mL $min^{-1}$. The temperature program was isothermal for 2 min at 40 °C and raised to 240 °C at a rate of 50 °C $min^{-1}$. Electron impact ionization (FCD-650, LECO Corp., St Joseph, MI), was used by the time-of-flight (TOF) mass spectrometer for volatile detection. Mass spectra were collected at a rate of 40 spectra/s over a range of 30-400 m/z. The ionization energy was 70 eV. Identification of volatile components was determined by comparison of collected mass spectra with those of authenticated standards and spectra in the National Institute for Standards and Technology (NIST) mass spectral library.

### 4.2.6. Multivariate data analyses

Multivariate data analysis techniques were used as pattern recognition tools and for multivariate calibration. The electronic responses and GC/MS data sets were obtained from a total of 7 days of consecutive analysis. Three replications of each variable were analyzed on each day to avoid growth of *E. coli* during the analytical operation. A Kruskal-Wallis was used to test if the samples tested each day were different. Descriptive statistics and a residuals plot were conducted to determine the normality of the data using SPSS Software for Windows Version 10.0 (SPSS Inc., Chicago, IL). The outliers were

determined and discarded from the data set. All matrix calculations were performed using the routines MATLAB Version 5.3 (Mathworks, Inc., Natick, MA) written by the author.

## 4.3. RESULTS AND DISCUSSION

### 4.3.1. Colony Counting Method

The population of *E. coli* grown in closed vials containing super broth is shown in Table 4.2. The data are from 15 replications. Prior to 6 hr incubation, the growth of *E. coli* was slow, probably because they were in the lag phase. After 12 hours, numbers increased from ~$10^2$ to 1.01 x $10^{10}$ CFU/mL.

**Table 4.2. Number of *E. coli* in super broth**

| Time (Hours) | Number (CFU/mL) |
|---|---|
| 2 | $3.85 \times 10^2$ |
| 4 | $3.56 \times 10^3$ |
| 6 | $6.29 \times 10^5$ |
| 8 | $5.34 \times 10^8$ |
| 10 | $6.89 \times 10^9$ |
| 12 | $1.01 \times 10^{10}$ |

### 4.3.2. Electronic Sensor Array Technology

#### 4.3.2.1. Data Collection

Figure 4.3 shows the comparison of sensor signal intensities during the acquisition period between super broth and super broth with *E. coli* and incubated for 6 hours. Each line represents a response from a sensor. In this study, twelve metal oxide sensors were used. The media had low responses from the sensor array. Different sensor signal intensities between samples are important to discriminate between samples. By choosing a growth medium with a minimum of volatile materials, the background contributed by the medium does not overwhelm the analytical signal. Growing the bacteria in sealed vials retains and collects volatiles in the headspace gases so they are not swept away by conventional aeration.

Figure 4.4 shows the average responses of the seven subgroups to the 12 sensors. The total operational testing time for each sample was 8 min. Therefore, only 3 samples from each variable (incubation time) were tested daily to avoid biochemical changes due to the growth of *E. coli* during the analytical period. The data from days 1 to 7 were statistically analyzed to determine if they were different by measuring the responses of each sensor to the samples from each day using nonparametric statistics (Kruskal-Wallis test). The results are shown in Table 4.3.

Kruskal-Wallis is equivalent to one-way ANOVA testing whether several independent samples are from the same population. The results (Table 4.3) show that samples tested each day were from the same population because all

results had a low Kruskal-Wallis statistics value with high probability ($p \geq 0.05$). There was no significant difference between samples on different days when the whole process, i.e., inoculation, growing, and analysis was repeated. The samples to be analyzed using SPME/GC/MS were tested in the same manner and the same conclusion was drawn.

**Table 4.3. Kruskal-Wallis test of samples with grouping variable (Day 1-7)**

| Sensors | Kruskal-WallisTest Statistics | Probability |
|---|---|---|
| SYLG | 0.421 | 0.981 |
| SYG | 0.329 | 0.988 |
| SYAA | 0.331 | 0.988 |
| SYGH | 0.084 | 0.999 |
| SYGCTI | 0.140 | 0.998 |
| SYGCT | 0.258 | 0.992 |
| T301 | 1.059 | 0.901 |
| P101 | 0.728 | 0.948 |
| P102 | 0.063 | 1.000 |
| P401 | 0.611 | 0.962 |
| T702 | 0.263 | 0.992 |
| PA2 | 0.770 | 0.942 |

#### 4.3.2.2. Pattern recognition of electronic sensor responses

**a) PCA**

Prior to PCA, the data matrix was standardized to zero mean and unit variance. This procedure ensures that the variables have the same weighting in the Principal Component (PC) model. The whole of the standardized data matrix was then submitted to PCA.

The most common criterion for choosing the number of principal components is the scree plot (Figure 4.5). In this case, 2 PCs were calculated and included in the model. The number of principal components is the number of substantively meaningful independent (uncorrelated) patterns, among the variables. The first PC accounts for a major fraction (83.21%) of the total variance of the data. The second principal axis is accounts for 12.60% of the variation not accounted for by the first factor.

Table 4.4 shows the eigenvalues and the percentage of variance explained by each PC. More than 83% of the data information was obtained by PC1and more than 12% was obtained by PC2. These two PCs together represent 96% of the information in the overall data set.

**Table 4.4. PCA of electronic sensor responses**

| PC | Eigenvalue | Explained Variance | Cumulative Variance |
|---|---|---|---|
| 1 | 11.043 | 9.985 | 83.205 |
| 2 | 0.638 | 1.487 | 95.599 |

Figure 4.6 shows a loadings plot of all samples. Sensors such as SYLG, SYGCTI, SYGCT, SYAA, and SYGH, had the highest negative loadings on PC1, but had positive loading on PC2. On the other hand, P101, T301, T702, PA2, and P401 had a high positive loading on PC1 and a positive loading on PC2. SYLG had a high positive loading on PC1 but low negative loading on PC2. On PC1, all sensors are important because of their high negative/positive loading. The score plot in Figure 4.7 shows 5 distinct clusters with positive and negative scores for both PCs. The variance is uniform and generates low within-group variations.

The control media and nutrient media inoculated with *E. coli* and incubated for 2 and 4 hours overlapped. The medium is indistinguishable from the early inoculants (number of *E. coli* were $3.85 \times 10^2$ and $3.56 \times 10^3$ CFU/mL, respectively), all of these points are grouped under the same area in the PCA score plot. This indicates that the electronic sensors were not able to detect a difference in the volatile metabolites at the early growth stage. Only after 6 hours ($6.29 \times 10^5$ CFU/mL) do the sensor responses become statistically distinguishable from the control medium. Groups of samples incubated for 6, 8,

10, and 12 hours occupied different areas, indicating that the electronic sensors differentiated between these samples. As a result, in the method reported here, *E. coli* could be detected when the number was above $10^5$ CFU/mL, or about $10^4$ times less than the densities found in mature cultures.

**b) DFA**

In order to perform the classification, the Mahalanobis distance was employed on relevant groups of samples. Canonical DFA was used in order to visualize the class separation.

The results are given in Table 4.5. Eigenvalues are indicative of the relative importance of the discriminant function in determining group separation. The separation of class-labeled samples is shown in Figure 4.8. The samples in the learning set were attributed to the groups whose average Mahalanobis distance was similar to the average value of the data points of a certain group.

From the DFA pattern (Figure 4.8) the data were classified into 5 groups. The control and nutrient media inoculated with *E. coli* and incubated for 2 and 4 hours overlapped. The medium is indistinguishable from the early inoculants, as all points were grouped under the same area in the DFA canonical plot. Samples incubated for 6, 8, 10, and 12 hours occupied different areas, indicating that the electronic sensors can differentiate between samples with different number of *E. coli*. As a result, *E. coli* could be detected when its number was above $10^5$ CFU/mL.

The technique developed from the learning set was applied to unknown samples (validation set). The classification result from unknown samples using discriminant analysis algorithm is presented in Table 4.6. Considering the percent correct classification, groups A, D, E, F, and G were always classified correctly (100 % correct). The misclassification groups were B classified as group A, and group C classified as group B. Percent corrects of sample in groups B and C did not reach 100 % because as shown in Figure 4.8 the electronic sensor could not discriminate samples from groups A, B, and C due to the low concentrations of volatile compounds produced by *E. coli*.

**Table 4.5. Canonical DFA analysis of electronic sensor responses**

| Discriminant Functions$^{th}$ | Eigenvalues | Cumulative Dispersion (%) |
|---|---|---|
| 1 | 228.444 | 85.90 |
| 2 | 34.714 | 99.00 |
| 3 | 1.981 | 99.70 |
| 4 | 0.663 | 100.00 |
| 5 | 0.031 | 100.00 |
| 6 | 0.001 | 100.00 |

**Table 4.6. Classification result* of the DFA algorithm for validation electronic sensor array data**

| | A | B | C | D | E | F | G | % Correct |
|---|---|---|---|---|---|---|---|---|
| **A** | 4 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **B** | 3 | 3 | 0 | 0 | 0 | 0 | 0 | **50** |
| **C** | 0 | 1 | 2 | 0 | 0 | 0 | 0 | **88** |
| **D** | 0 | 0 | 0 | 5 | 0 | 0 | 0 | **100** |
| **E** | 0 | 0 | 0 | 0 | 4 | 0 | 0 | **100** |
| **F** | 0 | 0 | 0 | 0 | 0 | 5 | 0 | **100** |
| **G** | 0 | 0 | 0 | 0 | 0 | 0 | 4 | **100** |
| Total | 7 | 4 | 2 | 5 | 4 | 5 | 4 | **91** |

***cases in row categories classified into column: control (A), super broth with E. coli after incubated for 2 (B), 4 (C), 6 (D), 8 (E), 10 (F), and 12 hrs (G).***

### 4.3.3. SPME/GC/MS

#### 4.3.3.1. Data Collection

The volatile compounds from the headspace of vials containing *E. coli* in super broth media incubated at 37 °C for 12 hours using SPME coupled to a gas chromatograph and a mass spectrometer was used to identify *E. coli* volatile metabolites. Figure 4.9 shows the chromatogram of volatile compounds from the headspace of the super broth media. Figure 4.10 shows the chromatogram of volatile compounds from the headspace of super broth inoculated with *E. coli* and incubated 37 °C for 6 hours.

Library searches of the mass spectrum were used to identify individual compounds from the chromatograph, based on their mass spectra. The mass spectrum of Indole from the headspace of samples inoculated with *E. coli* was determined from the Standard Mass Spectrum provided by NIST as shown in Figure 4.11. From library matches, the specific compounds produced by *E. coli* grown in super broth include, Dimethyl disulfide; Ethanol, O-acetimidoeyl; 2-Heptanone; Cyclopropane, pentyl; E-11,13-Tetradecadien-1-ol; Indole; and 2-Nonanone.

The basal media used in this research was super broth which has a high concentration of tryptophan. Trytophan is broken down into Indole due to the activity of the enzyme tryptophanase from *E. coli* (Figure 4.12). Tryptophanase degrades tryptophan to Indole, pyruvate, and ammonia (Moat & Foster, 1988). The alcohols and carbon dioxide detected in the super broth media are possibly from amino acid decarboxylation and deamination. Dimethyl disulfide may be present due to degradation of sulfur containing amino acids such as methionine and cycteine.

#### 4.3.3.2. Pattern recognition of GC/MS data

**a) PCA**

In this study, the concentration of the specific compounds could be determined using selective normalization. The normalized concentration of each compound was then standardized to zero mean and unit variances.

Figure 4.13 shows the average concentration of the seven compounds in seven sub groups. Each sub group contains 15 samples. The corresponding principal components score plot is presented in Figure 4.14. Samples with similar mass spectral pattern will have data points close to each other in the score plot. Accordingly, samples having in divergent mass spectra will be located further apart. Sample groupings are, therefore, easily identified in the score plot.

**b) DFA**

Data from SPME/GC/MS analysis was evaluated as explained in Section 4.3.2.2 b. The result obtained from DFA is shown in Figure 4.15. DFA was used to visualize the class separation between samples. Using DFA, six discriminant functions were found and 100% cumulative data dispersion was accounted for (Table 4.7). The first two functions have the highest eigenvalues and accounted for as much as 99.50% of the data. The data, for brevity however, were displayed using only two discriminant functions. From Figure 4.15, 5 sample groups accounted for the different numbers of *E. coli* in super broth due to the different incubation times.

The classification result from unknown samples using the discriminant analysis algorithm is presented in Table 4.8. Considering the percent correct classification, groups A, D, E, F, and G were always classified correctly (100 % correct). Group B was misclassified as group A, and group C. Group C was misclassified as group B. Percent corrects of sample in groups B and C did not

reach 100 % because (as shown in Figure 4.15) the electronic sensor could not discriminate between samples from groups A, B, and C due to low concentrations of volatile compounds produced by *E. coli*.

**Table 4.7. DFA analysis of SPME/GC/MS data**

| Discriminant Functions$^{th}$ | Eigenvalues | Cumulative Dispersion (%) |
|---|---|---|
| 1 | 203.928 | 98.10 |
| 2 | 3.016 | 99.50 |
| 3 | 0.770 | 99.90 |
| 4 | 0.112 | 100.00 |
| 5 | 0.089 | 100.00 |
| 6 | 0.005 | 100.00 |

**Table 4.8. Classification result* of the DFA algorithm for validation GC/MS data**

| | A | B | C | D | E | F | G | % Correct |
|---|---|---|---|---|---|---|---|---|
| **A** | 4 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **B** | 1 | 3 | 2 | 0 | 0 | 0 | 0 | **50** |
| **C** | 1 | 2 | 3 | 0 | 0 | 0 | 0 | **50** |
| **D** | 0 | 0 | 0 | 5 | 0 | 0 | 0 | **100** |
| **E** | 0 | 0 | 0 | 0 | 4 | 0 | 0 | **100** |
| **F** | 0 | 0 | 0 | 0 | 0 | 5 | 0 | **100** |
| **G** | 0 | 0 | 0 | 0 | 0 | 0 | 4 | **100** |
| Total | 6 | 5 | 5 | 5 | 4 | 5 | 4 | **86** |

***cases in row categories classified into column: control (A), super broth with E. coli after incubated for 2 (B), 4 (C), 6 (D), 8 (E), 10 (F), and 12 hrs (G).***

### 4.3.5. Artificial Neural Networks for Prediction of Number of *E. coli*

The root mean square error (RMSE) indicated that the standard deviation of the error between the predicted measured values can be used to evaluate the performance of calibration algorithms.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{I}(c_{predicted} - c_{true})^2}{n}} \qquad (4.2)$$

The neural network algorithm used was Multilayer perceptrons (MLP) based on back propagation. Figure 4.16 shows the effect of the number of neurons in the hidden layer on the mean square error (MSE) of the neural network to predict the number of *E. coli* (Figure 4.16)

The network architecture created for the *E. coli* data matrix comprises input layer, one hidden layer of neurons and one output layer. The transfer function in the hidden layer was a hyperbolic tangent sigmoidal nonlinear and a linear function was used in the output layer.

$$\varphi_j(v_j(n)) = a\tanh(v_j(n)) \qquad (4.5)$$

where $\varphi(.)$ is the activation function associated with the neuron and $v_j(n)$ is the induced local field of neuron *j*.

The training function used was a Levenberg-Marquardt (LM) algorithm.

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - [\boldsymbol{J}^T\boldsymbol{J} + \boldsymbol{I}]^{-1}\,\boldsymbol{J}^T\boldsymbol{e} \qquad (4.6)$$

$$J = \begin{bmatrix} \frac{\partial e_1}{\partial x_1} \frac{\partial e_1}{\partial x_2} \cdots \frac{\partial e_1}{\partial x_n} \\ \frac{\partial e_2}{\partial x_1} \frac{\partial e_2}{\partial x_2} \cdots \frac{\partial e_2}{\partial x_n} \\ \vdots \\ \frac{\partial e_n}{\partial x_1} \frac{\partial e_n}{\partial x_2} \cdots \frac{\partial e_n}{\partial x_n} \end{bmatrix} \quad (4.7)$$

where *e* is a vector of network errors (Choralambus, 1992). The LM algorithm is useful when accurate training is required.

The performance function performed during training feedforward neural networks was the mean sum of squares of the network errors (MSE).

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \quad (4.8)$$

The network outputs were plotted versus the target prediction of number of *E. coli* and concentration of the selected compounds. Figures 4.17 shows the predictions versus true values of numbers of *E. coli* in super broth by electronic sensor array data using ANN algorithms. The correlation coefficient, $R^2$, between the outputs and targets is a measure of how well the variation in the output is explained by the targets and outputs. $R^2$ is close to 1, which indicates a good fit. All calibration techniques predicted the number of *E. coli.*

## 4.4. CONCLUDING REMARKS

The electronic sensor array, which is based on adsorption/desorption reactions of volatiles on the surface of different electronic sensors, has been shown to discriminate between media without *E. coli* and media inoculated with *E. coli*. The influence of number of *E. coli* as a result of incubation time proved to have an effect on the electronic sensor responses. *E. coli* can be identified and quantified by their volatile metabolites. The electronic sensors therefore have potential to detect *E. coli* from their volatile metabolites. The results were shown to be reproducible when the entire analysis (i.e., inoculating, growing and analysis) was repeated.

While GC/MS is extremely useful in identifying the microorganisms from their volatile metabolites, it is not a panacea. Electronic sensor array is an equally good candidate to monitor changes in the composition of the gas phase of chemical products. The advantages of electronic sensor array over classic GC are that it is simpler and has higher speed. However, with the electronic sensor, the compounds are not analytically identified.

Multilayer perceptron neural network with back propagation algorithm showed the potential of the electronic sensors to predict number of *E. coli* and the concentration of selected compounds in unknown samples. The electronic sensor array was found to be satisfactorily correlated with colony counting and GC/MS methods as determined by using canonical correlations. The electronic sensor array coupled with neural networks can be used to identify and quantify the number of *E. coli* and volatile metabolites.

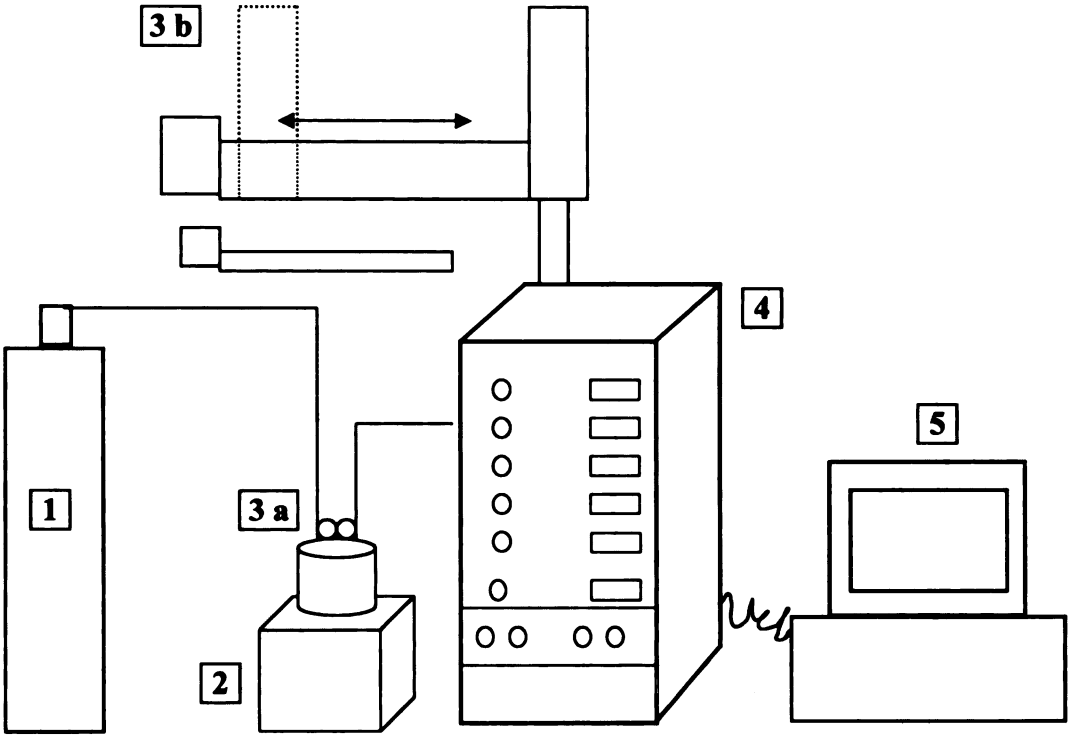


**Figure 4.1. Electronic sensor array system.**

(1) Synthetic air and a humidity regulator
(2) Measurement chamber to control the temperature of the heads pace
(3) Glass container with the (a) manual odorant sampler or (b) an autosampler (Block of 2 x 32 Vials)
(4) Electronic nose with 12 sensor channels
(5) Microcomputer for storing and processing sensor data

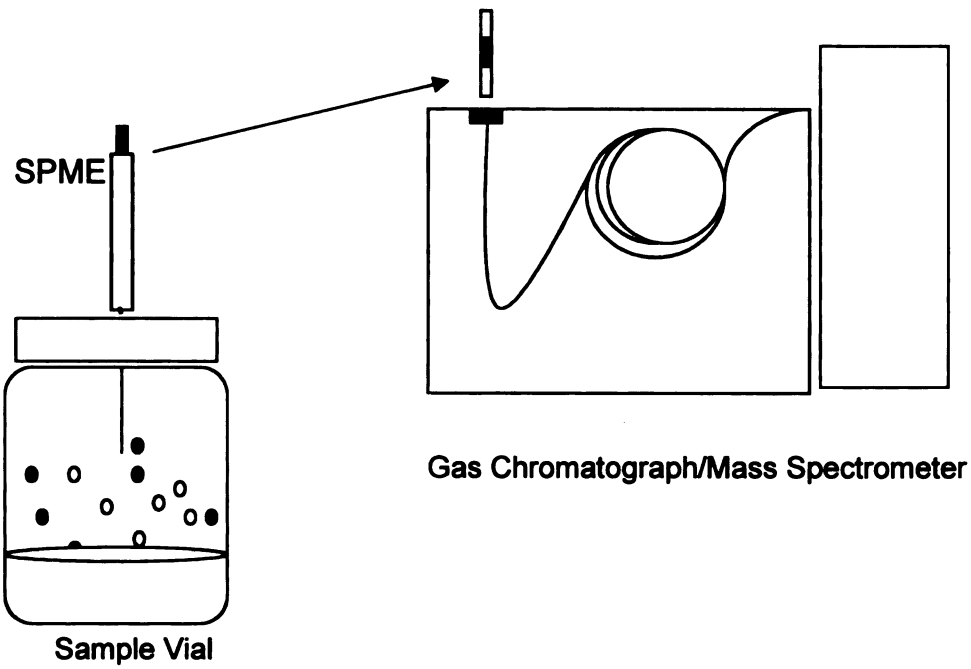


**Figure 4.2. SPME/GC/MS method for collecting volatile compounds from the headspace of sample.**

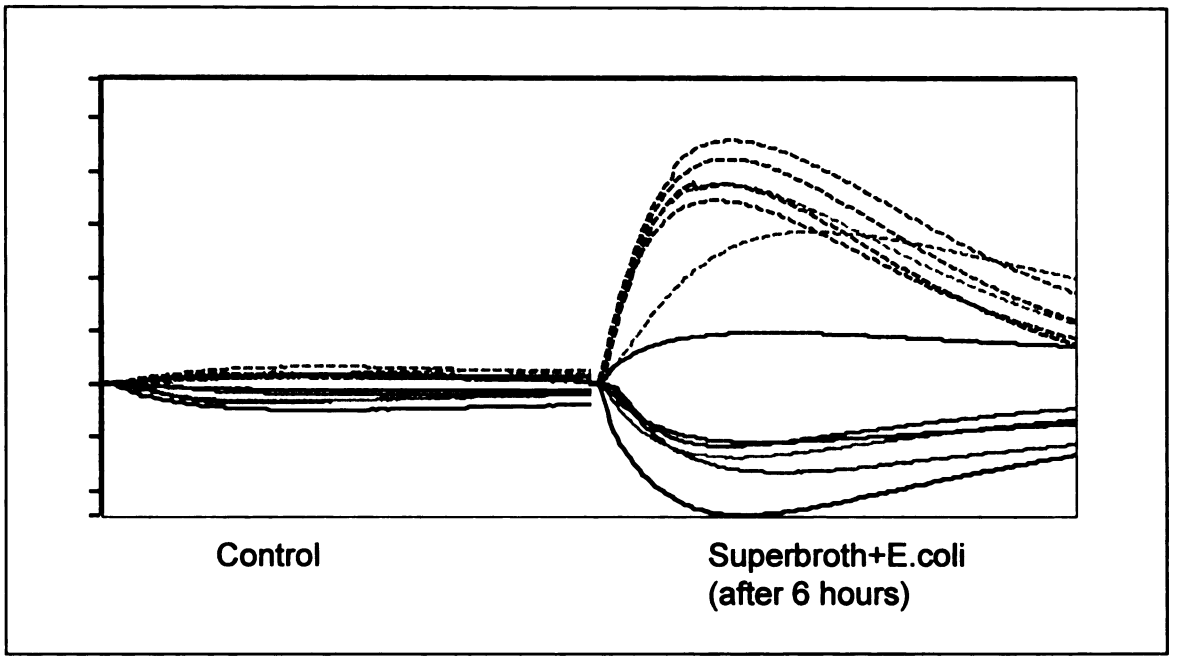


**Figure 4.3. Sensor responses from the headspace of super broth (control) and super broth inoculated with *E. coli* and incubated for 6 hours.**

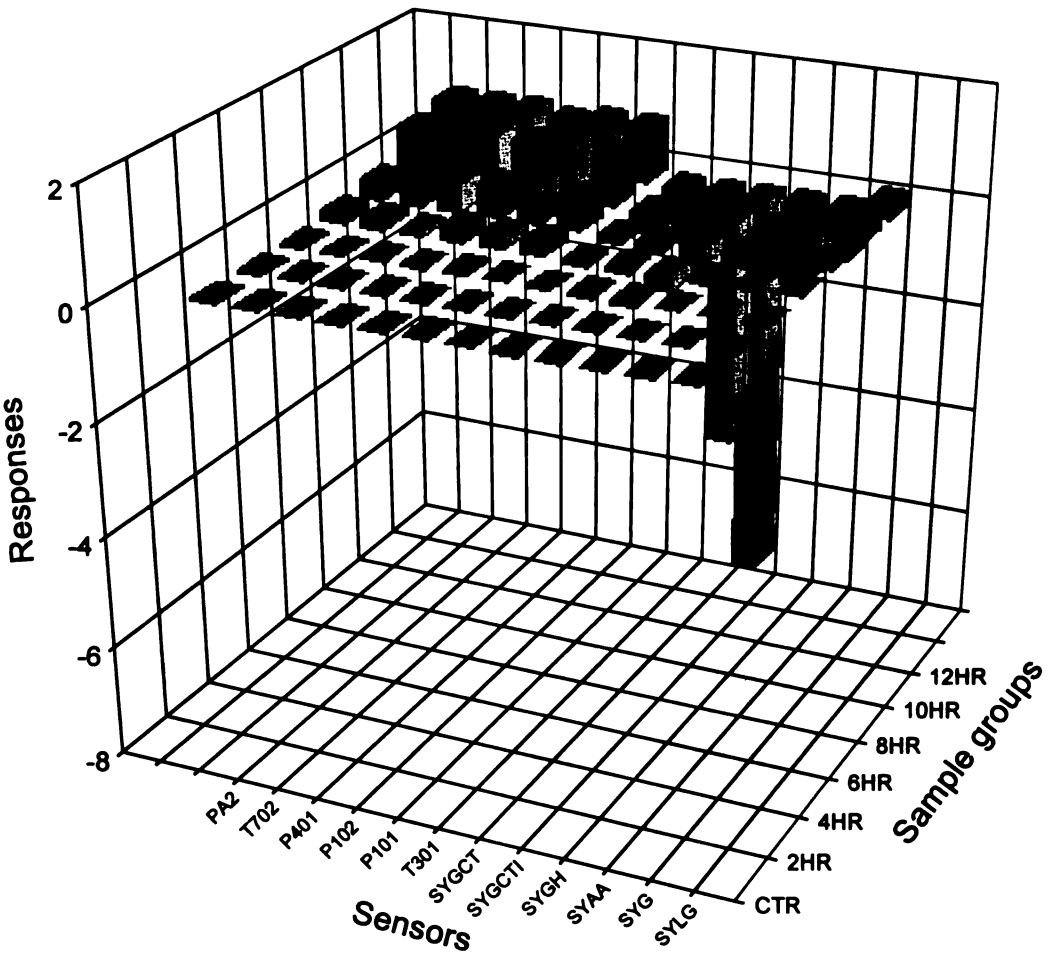


**Figure 4.4. Sensor responses of super broth and super broth inoculated with *E. coli* and incubated at 37 °C for 2, 4, 6, 8, 10, and 12 hours analyzed using 12 sensors.**

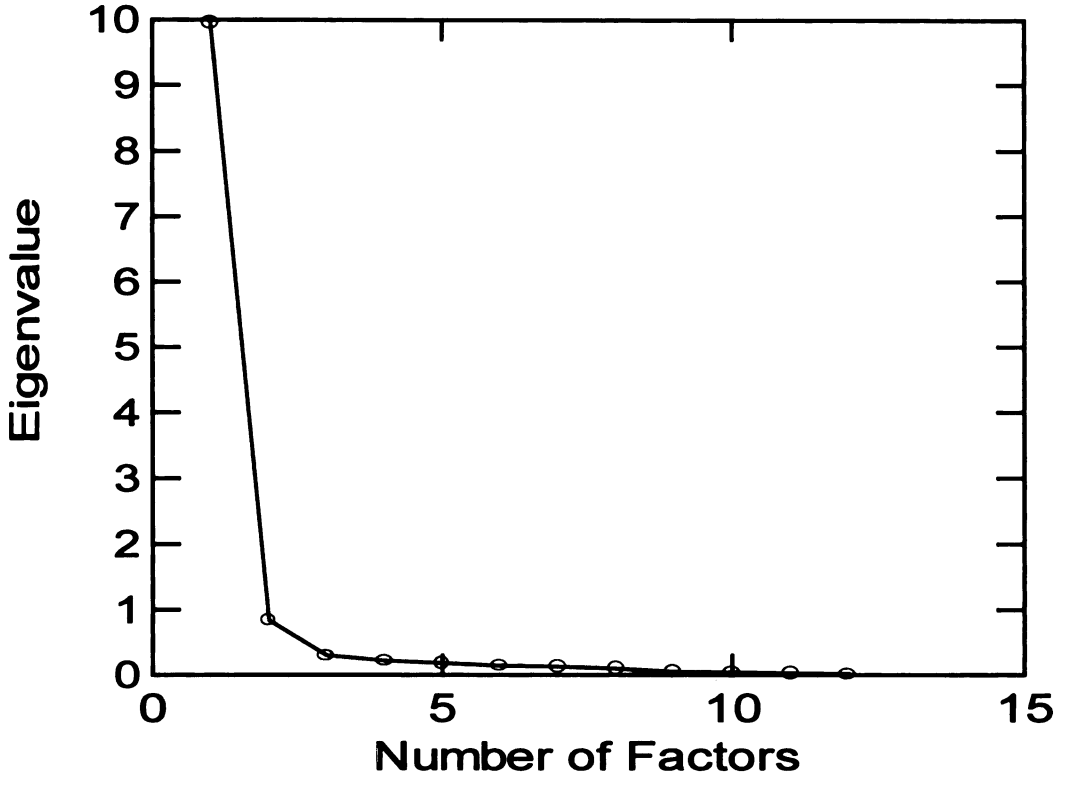


**Figure 4.5. Determination of PCs using a scree plot.**

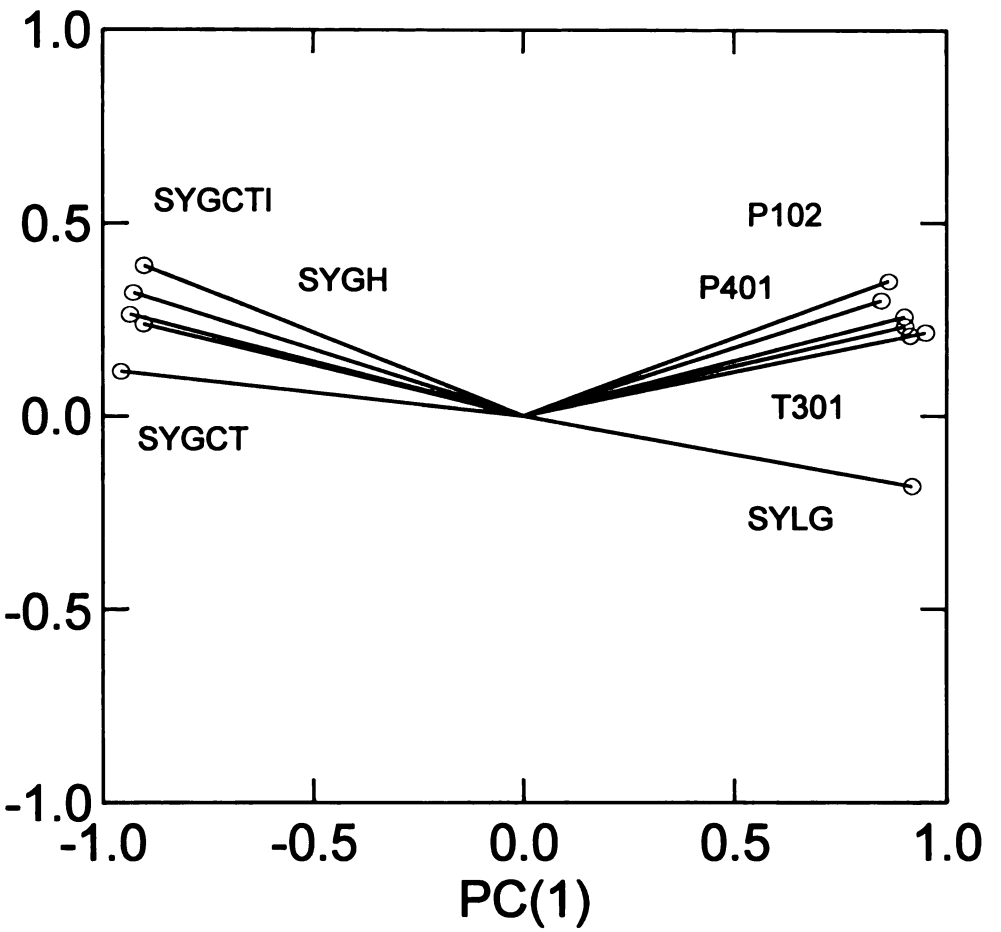


**Figure 4.6. PCA loadings plot of the data from electronic sensors.**

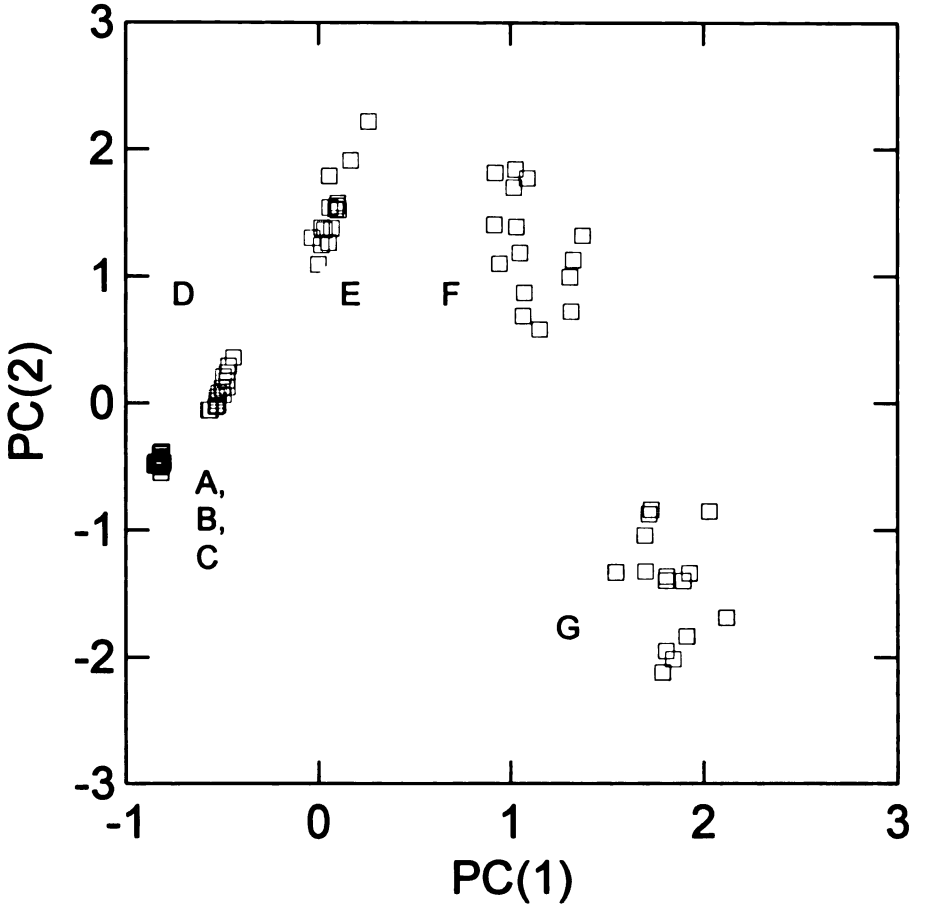


**Figure 4.7. PCA scores Plot of volatile compounds produced by *E. coli* in super broth using 12 electronic sensor. The samples are labeled after the following scheme: control (A), after incubated for 2 hrs (B), 4 hrs (C), 6 hrs (D), 8 hrs (E), 10 hrs (F), and 12 hrs (G).**

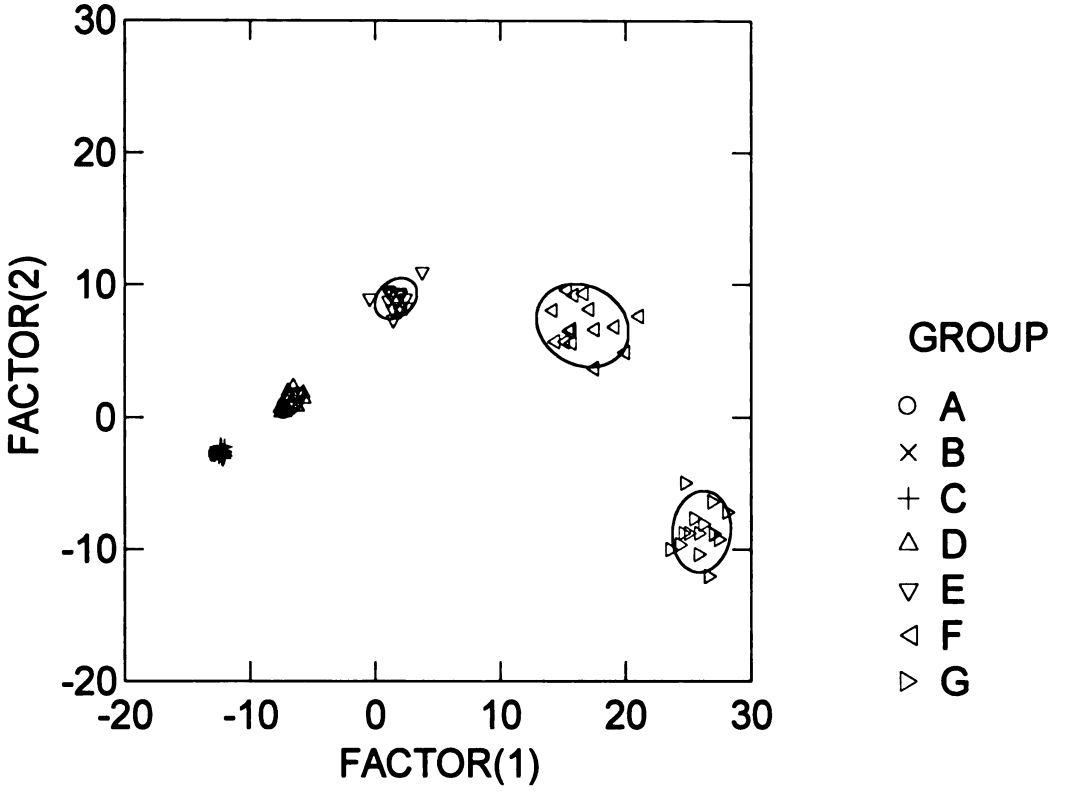


**Figure 4.8. DFA of volatile compounds of *E. coli* in nutrient media using electronic sensors. The samples are labeled using the following scheme: control (A), after incubated for 2 hrs (B), 4 hrs (C), 6 hrs (D), 8 hrs (E), 10 hrs (F), and 12 hrs (G).**

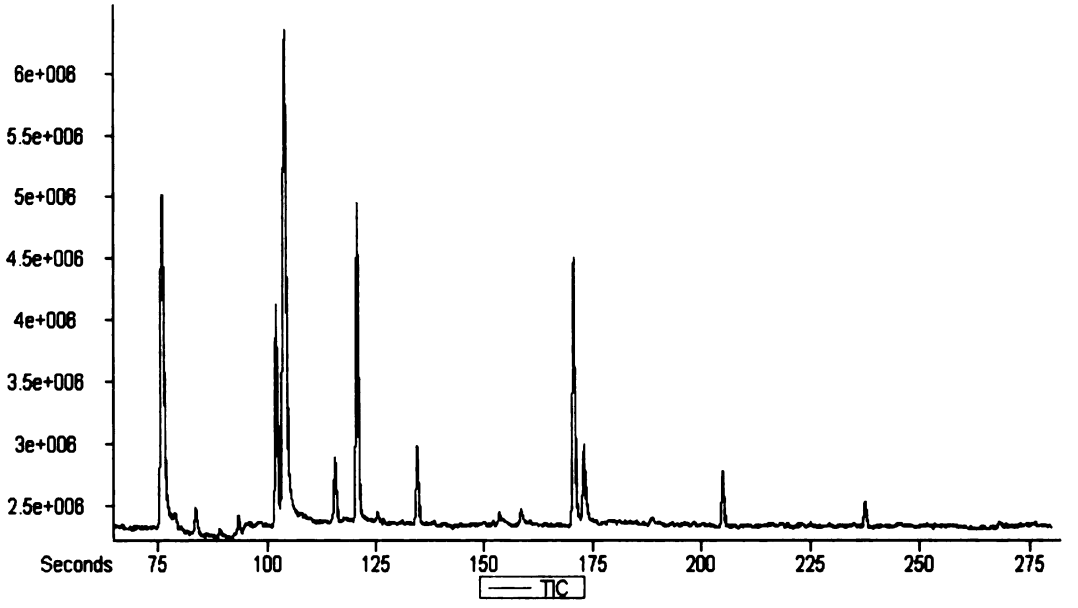


**Figure 4.9. Chromatogram of volatile compounds from headspace of super broth.**

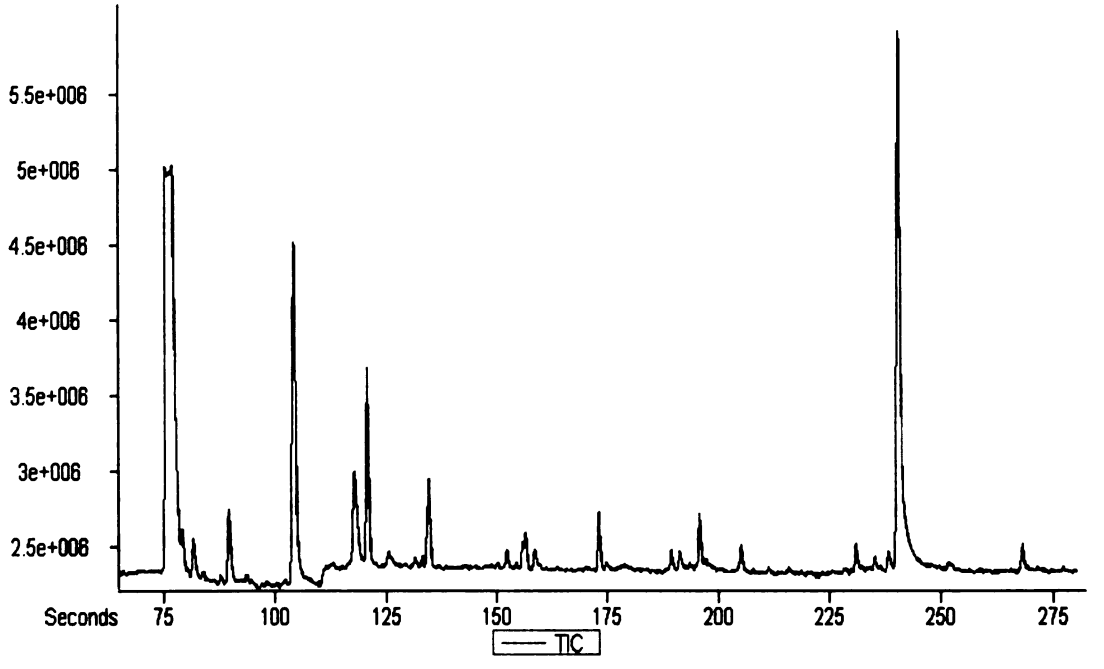


**Figure 4.10. Chromatogram of volatile compounds from headspace of super broth inoculated with *E. coli* incubated for 12 hours at 37°C.**

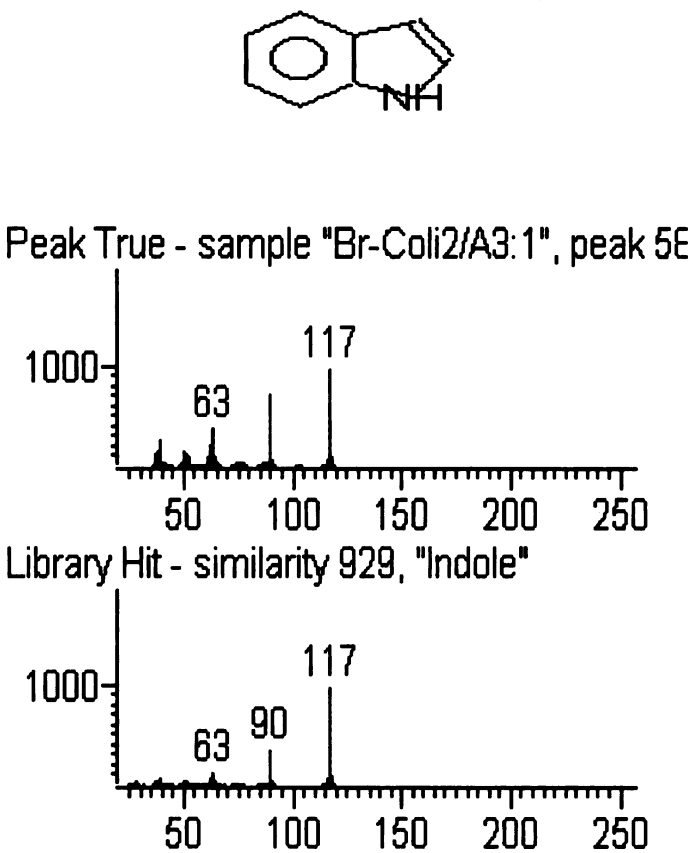


**Figure 4.11. Mass Spectrum of indole from super broth inoculated with *E. coli* compared with mass spectrum of indole from NIST.**

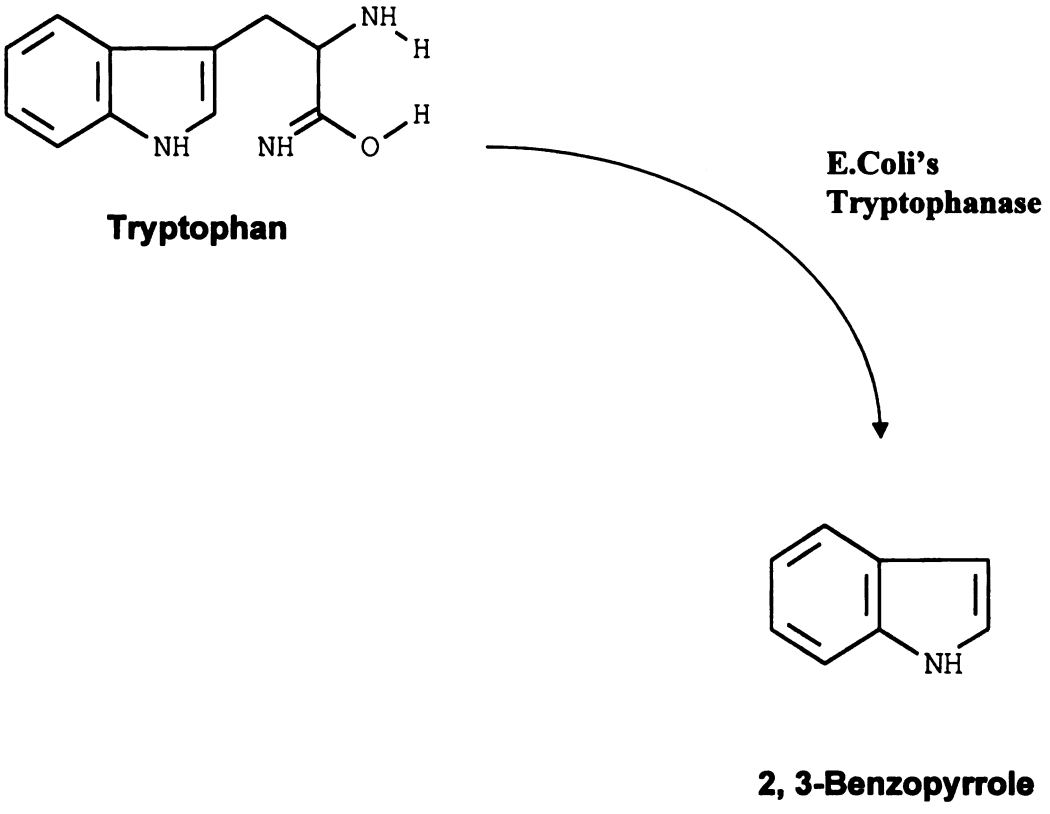


**Figure 4.12. *E. coli*'s catabolism of typtophan to indole.**

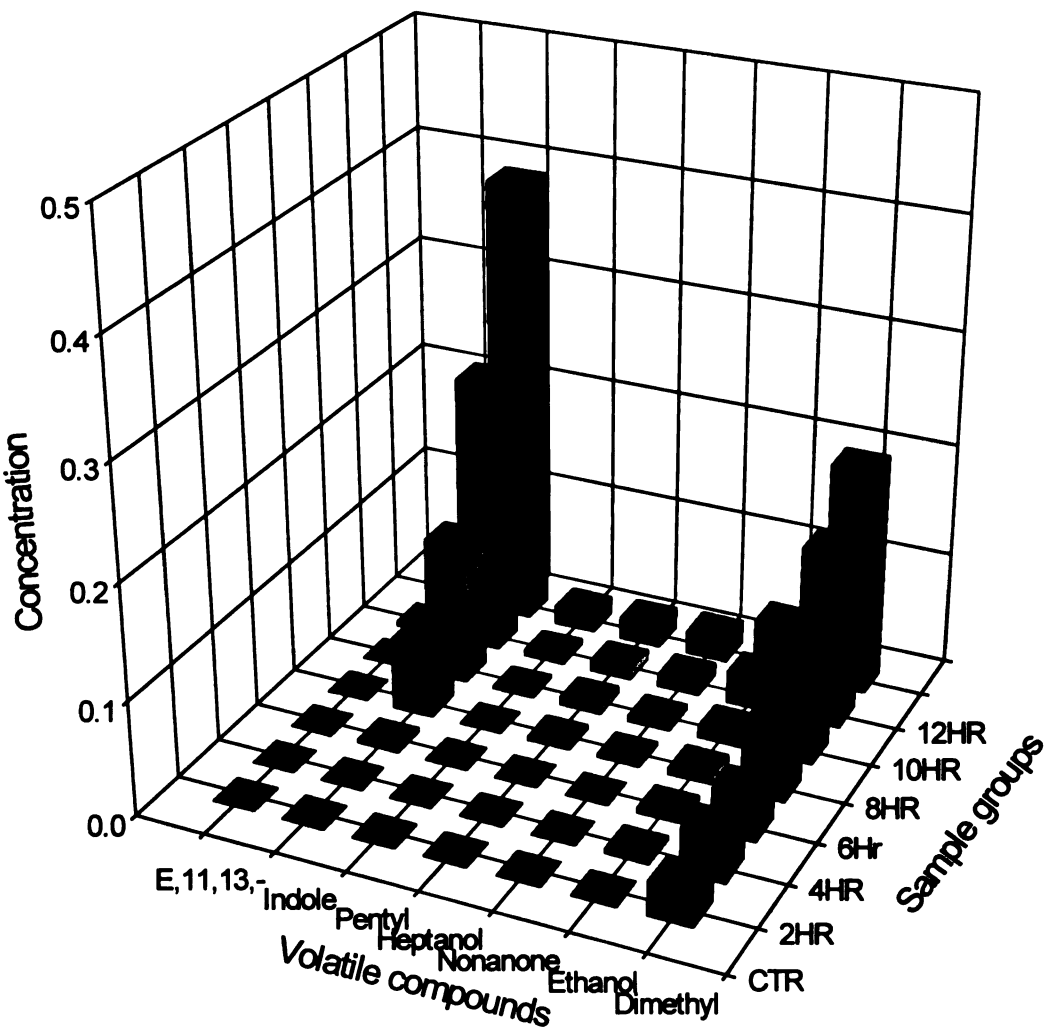


**Figure 4.13. Average concentrations of selected volatile compounds from headspace of super broth and super broth inoculated with *E. coli* and incubated at 37 °C for 2, 4, 6, 8, 10, and 12 hours analyzed using SPME/GC/MS.**

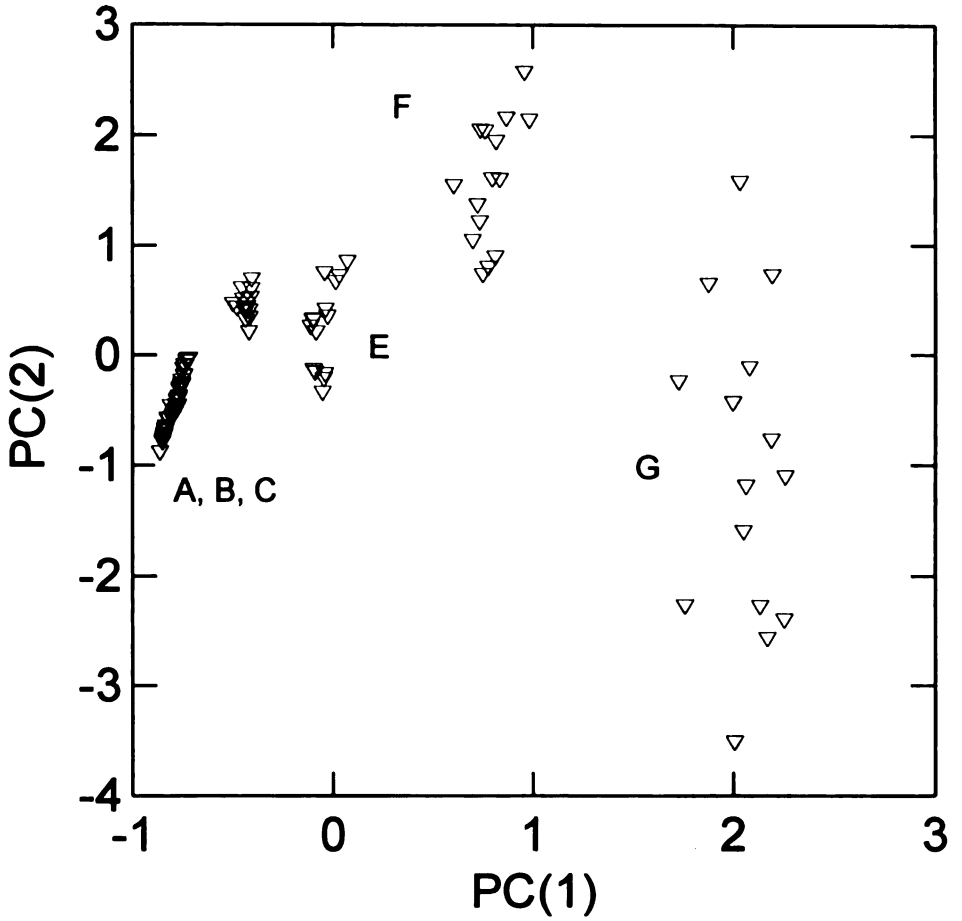


**Figure 4.14. Scores Plot of volatile compounds produced by *E. coli* in super broth using SPME/GC/MS. The samples are labeled using the following scheme: control (A), after incubated for 2 hrs (B), 4 hrs (C), 6 hrs (D), 8 hrs (E), 10 hrs (F), and 12 hrs (G).**

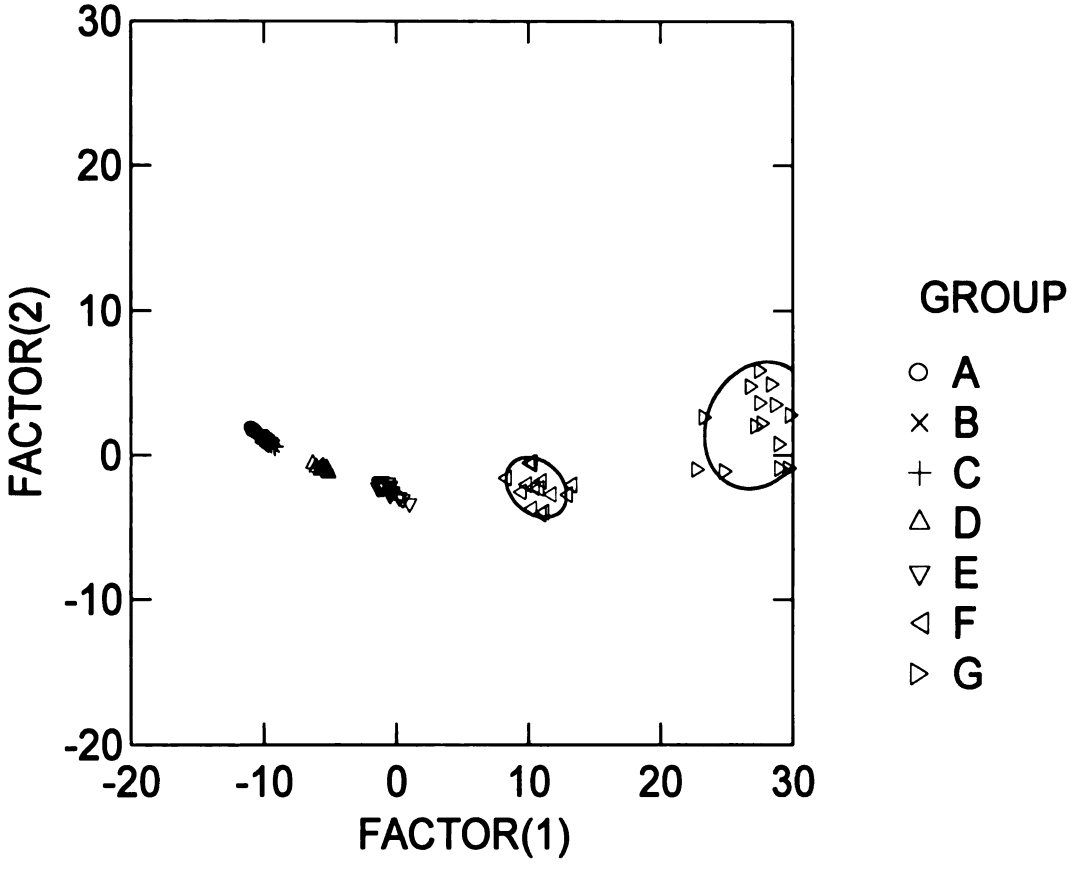


**Figure 4.15. DFA of volatile compounds produced by *E. coli* in super broth using SPME/GC/MS. The samples are labeled using the following scheme: control (A), after incubated for 2 hrs (B), 4 hrs (C), 6 hrs (D), 8 hrs (E), 10 hrs (F), and 12 hrs (G).**

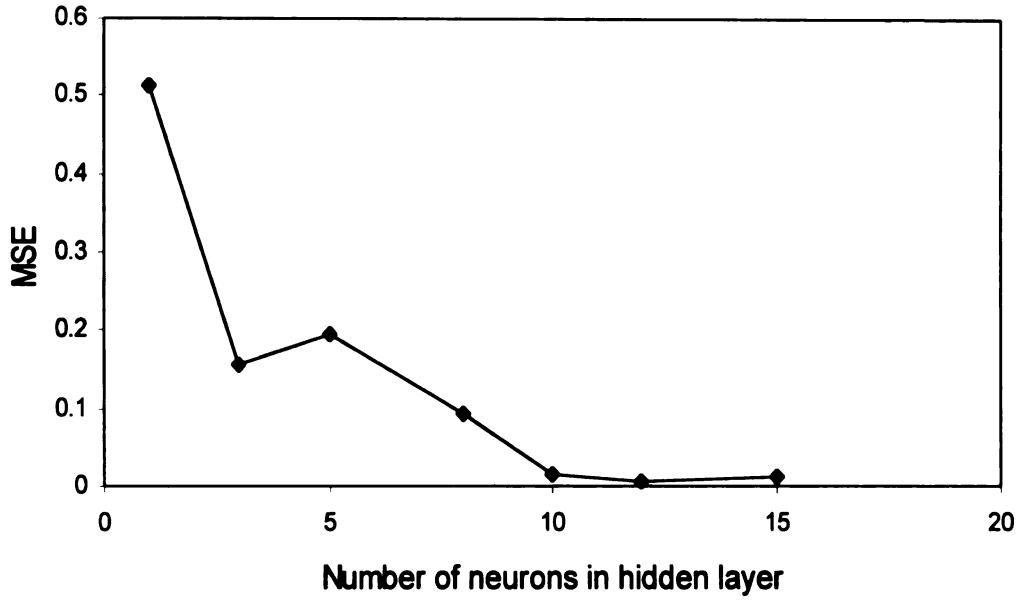


**Figure 4.16. Effect of number of hidden layers and number of neurons in hidden layer on the performance of the networks (MSE) of sensor responses *v.s.* number of *E. coli*.**

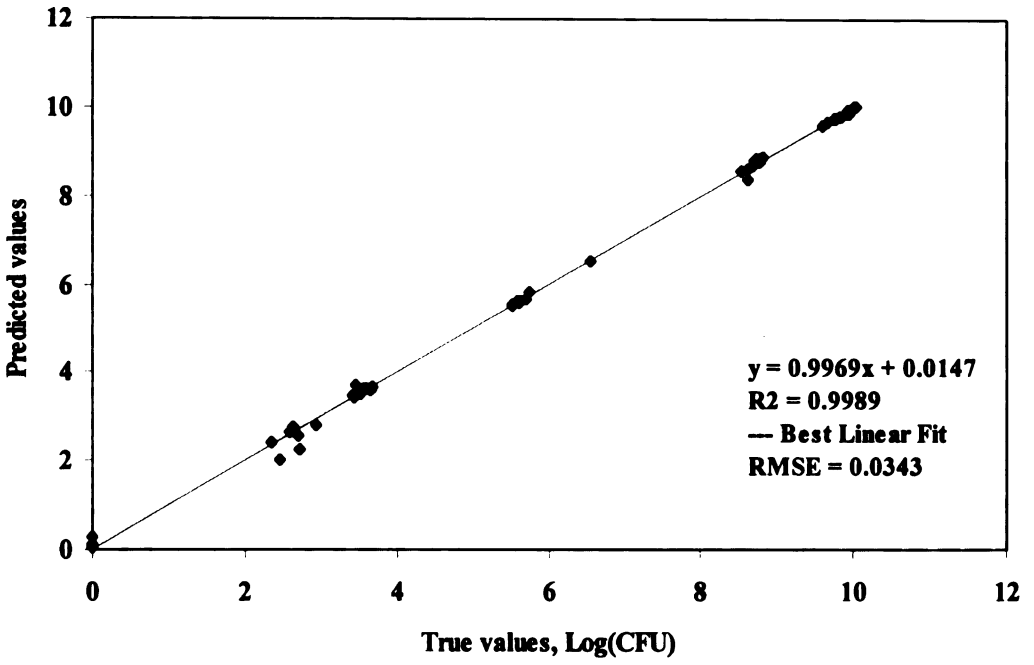


**Figure 4.17. The predicted *vs.* true numbers of *E. coli* using ANN, back propagation algorithm.**

# CHAPTER 5

# ELECTRONIC SENSOR ARRAY INCORPORATING ARTIFICIAL NEURAL NETWORK TO IDENTIFY AND QUANTIFY NUMBER OF *SALMONELLA ENTERICA* SEROTYPE TYPHIMURIUM AND VOLATILE METABOLITES IN NUTRIENT MEDIA


## ABSTRACT

A method was developed to predict the number of *Salmonella enterica* serovar Typhimurium in super broth and concentrations of specific volatile metabolites using a metal oxide electronic sensor array and neural network. To evaluate the complex data obtained from electronic sensors and SPME/GC/MS, Principal Component Analysis (PCA) was used for data exploration and dimensional reduction. Using Discriminant Factor Analysis (DFA) it was possible to differentiate super broth from that containing *Salmonella* Typhimurium. The multilayer perceptron neural network can be trained to identify and quantify *Salmonella* Typhimurium and volatile metabolites in nutrient media. Once trained, the networks were shown to be capable of correlating voltammetric responses with number of *Salmonella* Typhimurium. The electronic sensor array data was found to satisfactorily correlate with colony counting ($R^2$ = 0.866).

## 5.1. INTRODUCTION

There is a need for more rapid, precise, and accurate analyses of the biochemical composition of microorganisms for the identification of potentially pathogenic organisms. Electronic sensor array is an instrument-based technique which satisfies these requirements. Electronic sensor array technology has provided a new way for rapid simple and inexpensive analysis of volatiles.

Artificial Neural Network (ANN) has recently been introduced for modeling, simulation, and optimization of complex data. Fast, accurate, and reliable neural network models can be developed from measured/simulated data (Tanaka & Hasegawa, 2001; Bila et al., 1999). Artificial Neural Networks are among the possible nonlinear techniques which can be applied to the data collected in this study.

In this research, a method was developed to determine the presence of *Salmonella* Typhimurium in super broth using both an electronic sensor array and neural network system. This study demonstrates on a laboratory scale that the concentrations of five key volatile metabolites of *Salmonella* Typhimurium can be predicted from the electronic sensor responses using neural network algorithms. The predictive learning ability of an ANN model was assessed by comparison to a reference method using a mean square error of prediction.

## 5.2. METHODOLOGY

### 5.2.1. Stock Culture Preparation

The culture used was *Salmonella enterica* serotype Typhimurium obtained from Food Microbiological laboratory, Department of Food Science and Human Nutrition, Michigan State University, MI. *Salmonella* Typhimurium was inoculated into Luria Bertani (LB) broth consisting of 10 g Bacto-tryptone, 5 g Bacto Yeast, and 5 g NaCl, and incubated at 37 °C in a gyrotory shaker (G-25 New Brunswick Scientific Corporation, New Brunswick, NJ) at 100 rpm. The *Salmonella* Typhimurium suspension was then dispensed into sterile 125 mL GSA bottles, and centrifuged (Rotor-GSA model RC 5 C Superspeed Centrifuge, Sorvall Instruments, Dupont Co., Haffman Estate, ILL) at 1600 x g for 10 min. The supernatant was decanted, and the resulting cell pellets were resuspended in a sterile 15 % glycerol solution. 1 mL aliquots of bacterial suspension were transferred into 1.5 mL microcentrifugal tubes and frozen using liquid nitrogen. The frozen cultures were stored at –80 °C.

### 5.2.2. Preparation of Test Solution

Super broth consisting of 32 g Tryptone, 20 g Yeast Extract, 5 g NaCl, 5 mL of 1 N NaOH, was used as a basal medium. Before use, *Salmonella* Typhimurium was transferred from a stock culture to super broth and incubated at 37 °C for 10 hrs. *Salmonella* Typhimurium was inoculated at $10^2$ CFU/mL into super broth, and 5 mL of super broth were transferred into standard 20 mL headspace vials and sealed with PTFE-lined Teflon caps (Alpha M.O.S.,

Hillsborough, NJ) using a cap crimper. All experiments were designed to use the headspace sampler and sealed sample vial system originally designed for use with gas chromatography and GC/MS. At time zero in the experiment, the cultures were allowed to grow in vials at 37 °C in a gyrotory shaker. The incubation times were varied, and the growth of the microorganism was measured. Samples were periodically analyzed after incubation at 2, 4, 6, 8, 10, 12 and 14 hours at 37 °C using a colony counting method, SPME/GC/MS, and electronic sensor techniques.

### 5.2.3. Colony Counting Method

The samples were serially diluted in sterile Butterfield's phosphate buffer. A series of dilutions was also prepared from the stock suspension. Serially diluted samples were plated in duplicate using 3M Petrifilm Aerobic Count Plates. All plates were incubated at 37 °C for $48 \pm 2$ hrs. After incubation, colonies were counted and the viable cell count per unit amount of the original material calculated. Plate counts are essentially direct counts of numbers of colonies in or on plates. Each colony is assumed to represent the progeny of one organism. Plate counts are recorded as colony forming units (CFU).

### 5.2.4. Electronic sensor for monitoring of volatile compounds

An electronic sensor (Fox 3000, Alpha M.O.S., Hillsborough, NJ), with 12 metal oxide sensors (SYLG, SYG, SYAA, SYGH, SYGCTI, SYGCT, T301, P101, P102, P401, T702, and PA2) was used for monitoring changes in volatiles

produced by *Salmonella* Typhimurium in super broth medium. The optimum conditions for electronic sensor analysis were, 35 °C for headspace generation temperature, and 300 seconds for headspace generation time. This instrument was linked to an autosampler unit capable of analyzing a total of 64 samples. The samples were placed in glass vials and sealed with crimped PTFE/metal septa.

The headspace volatile compounds from super broth and super broth inoculated with *Salmonella* Typhimurium and incubated for 2, 4, 6, 8, 10, and 12 hours were monitored using the optimized condition. The samples were placed in the HS100 auto-sampler in arbitrary order. Prior to analysis, the vial was removed from the sample tray and placed in a temperature-controlled chamber. The vial temperature was held at 35°C whilst being spun in order to produce an equilibrated headspace. The time the vial remains in this chamber is the headspace generation time. The automatic injection unit heated the samples to 35 °C with an incubation time of 300 second. The temperature of the injection syringe was 40 °C. The injector needle then removed 5000 µL of headspace and injected this into the sensor chamber. The delay time between two injections was 300 second.

Each injection was repeated, with separate samples (three times for all variations per day) for seven days. The electronic signals from the sensors were digitized, then transferred to the control computer. Resistance changes (difference in sensor resistance between air blank and odorous atmosphere) were recorded.

### 5.2.5. Determination of Volatile Compounds using SPME/GC/MS

Preliminary tests were done to determine the optimum SPME fiber (Supelco, Inc., Bellefonte, PA) for collecting the volatile compounds by comparing results from Polydimethylsiloxane fibers, Polydimethylsiloxane/ Divinylbenzene fibers, and Carboxen/Polydimethylsiloxane fibers. The results showed that the Carboxen/Polydimethylsiloxane fiber was the most suitable SPME fiber because it could collect more compounds produced by *Salmonella* Typhimurium than other fibers. The Carboxen/Polydimethylsiloxane fiber was, therefore, used for all further experiments.

Before use, the SPME fiber was conditioned at 250 °C for 1 hr. A 1 cm length of fused silica fiber, coated with a polymer is bonded to a stainless steel plunger and installed in a holder that looks like a modified microliter syringe. Volatile compounds were collected using a SPME device. The fiber is housed in a hollow, stainless steel needle. The plunger moves the fused silica fiber into and out of a hollow needle. To use the unit, the analyst draws the fiber into the needle, passes the needle though the septum that seals the sample vial, and depresses the plunger, exposing the fiber to the sample or the headspace above the sample. The fiber was pushed out of the needle and exposed to the headspace at 35 °C until equilibrium was reached.

For headspace sampling, 5 mL of liquid sample was placed into a 20 mL vial and the fiber was exposed to the head space of the media solution. Sampling temperature was 35°C, and sampling duration was 10 min., which was sufficient to permit the establishment of a near equilibrium for the compounds

tested. Once sampling was finished, the fiber was withdrawn into the needle and transferred to the injection port of the gas chromatograph (HP-6890, Hewlett-Packard Co., Wilmington, DEL). Absorbed volatiles were desorbed from the fiber coating by inserting the SPME fiber through a predrilled septum (Thermogreen LB-2, Supelco Co., Bellefonte, PA) and into a glass lined, split-less injector port of a gas chromatograph. The GC column inlet was immersed in a liquid nitrogen bath during the desorption period, so that the volatiles were collected before entering the column for analysis. After 3 min of desorption, the fiber was retracted and removed as was the liquid nitrogen. The oven was then closed and the GC was manually started.

A HP-6890 GC was used in analysis of the compounds. Volatiles were separated using a capillary column (SPB5, 30m x 0.1 mm id., 0.25 μm coating thickness). The carrier gas used was ultrapurified Helium (99.99% purity) at a flow rate of 0.5 mL $min^{-1}$. The temperature program was isothermal for 2 min at 40 °C and raised to 240 °C at a rate of 50 °C $min^{-1}$. Electron impact ionization (FCD-650, LECO Corp., St Joseph, MI), was used by the time-of-flight (TOF) mass spectrometer for volatile detection. Mass spectra were collected at a rate of 40 spectra/s over a range of 30-400 m/z. The ionization energy was 70 eV. Identification of volatile components was determined by comparison of collected mass spectra with those of authenticated standards and spectra in the National Institute for Standards and Technology (NIST) mass spectral library.

### 5.2.6. Multivariate data analyses

Descriptive statistics and a residuals plot were conducted to determine the normality of the data using SPSS Software for Windows Version 10.0 (SPSS Inc., Chicago, IL). The data were explored and classified using Principal Component Analysis (PCA) and Discriminant Function Analysis (DFA). All matrix calculations were performed using the routines MATLAB Version 5.3 (Mathworks, Inc., Natick, MA) written by the author.

## 5.3. RESULTS AND DISCUSSION

### 5.3.1. Colony counting Method

The population of *Salmonella* Typhimurium grown in closed vials containing super broth is shown in Table 5.1. After 14 hours incubation, the numbers increased from $10^2$ to $6.47 \times 10^9$ CFU/mL.

**Table 5.1. Colony counts of *Salmonella* Typhimurium in Super broth incubated at 37 °C**

| Time (Hours) | Number (cfu/mL) |
|---|---|
| 2 | $6.15 \times 10^2$ |
| 4 | $4.76 \times 10^3$ |
| 6 | $2.77 \times 10^6$ |
| 8 | $2.01 \times 10^7$ |
| 10 | $3.25 \times 10^9$ |
| 12 | $5.36 \times 10^9$ |
| 14 | $6.47 \times 10^9$ |

### 5.3.2. Electronic Sensor Array Technology

#### 5.3.2.1. Data Collection

A total of 80 samples were analyzed in 8 subgroups including the control, 2, 4, 6, 8, 10, 12, and 14 hours of incubation. The sensor responses were given in units of the maximum change of the sensor's electrical resistance divided by the initial resistance ($\Delta R_{max}/R_o$).

Figure 5.1 shows the average sensor response intensity of all subgroups. The intensity patterns of all samples were different due to the difference in concentration of volatile metabolites generated by different numbers of *Salmonella* Typhimurium produced during the growth period from 2-14 hours.

Three samples from each variable (incubation time) were tested daily to avoid biochemical changes due to the growth of *Salmonella* Typhimurium during the operating period. The experiments were repeated for 5 days. The were statistically analyzed using the Nonparametric statistics, Kruskal-Wallis test as shown in Table 5.2. The results show that there was no significant difference between samples from the different days. Samples also analyzed using SPME/GC/MS and the same conclusion was found.

**Table 5.2. Kruskal-Wallis Test of samples with grouping variable (Day 1-5)**

| Sensors | Kruskal-WallisTest Statistics | Probability |
|---|---|---|
| SYLG | 0.166 | 0.997 |
| SYG | 0.146 | 0.997 |
| SYAA | 0.062 | 1.000 |
| SYGH | 0.251 | 0.993 |
| SYGCTI | 0.183 | 0.996 |
| SYGCT | 0.108 | 0.999 |
| T301 | 0.591 | 0.964 |
| P101 | 0.314 | 0.944 |
| P102 | 0.150 | 0.997 |
| P401 | 0.231 | 0.989 |
| T702 | 0.079 | 0.999 |
| PA2 | 0.194 | 0.996 |

#### 5.3.2.2. Pattern recognition of electronic sensor responses

**a) PCA**

Standardization was applied to this data matrix to transform the data into a more expedient form with size effects removed and variance scaled. PCA was then carried out on this pretreated data.

The data information was retained by PC1 (96.58%) and PC2 (1.44%). As a result, the first two PCs allowed 98.03 % of the relevant information from the data to be explained.

Figure 5.2 shows a loading plot of all samples. All sensors behaved in a similar way in that they had high loading on PC1 but zero or low loading on PC2. Sensors such as SYG, SYGCTI, SYGCT, SYAA, and SYGH had high negative loadings on PC1. SYLG, P102, P101, T301, T702, PA2, and P401, had high positive loading on PC1. The loadings displays facilitate an exploration of the contributions of the variables to each PC.

Sample patterns can be separated from each other using simple scatter plots of principal component scores. The score plot in Figure 5.3 shows 5 distinct clusters with positive and negative scores on both PCs. The medium is indistinguishable from the early stage of incubation (number of *Salmonella* Typhimurium were $6.15 \times 10^2$, $4.76 \times 10^3$, and $2.77 \times 10^6$ CFU/mL, respectively). All of these points are grouped together under the same area in the PCA score plot. Samples incubated for 8, 10, 12, and 14 hours were separated. The grouping of samples with high concentration of volatiles in the lower right quadrant was evident. Samples with a similar number of *Salmonella* Typhimurium (the same

incubation time) will have data points close to each other. Accordingly, samples having divergent numbers will be located further apart. Sample groupings are, therefore, easily identified in the score plot. Still, for many samples, it is difficult to predict whether a sample contains *Salmonella* Typhimurium.

PCA reduces a large data set of correlated variables to smaller numbers of uncorrelated components. Following its application, orthogonal projection directly leads to dimensionality reduction and possibly to feature selection.

**b) DFA**

In this study, DFA was used to determine if it is possible to separate two or more individual groups, given measurements for these individuals on several variables. Canonical DFA was used in order to visualize the class separation.

Since there were eight subgroups of treatment variables (control and 2-14 hours incubation times), seven discriminant functions were derived. Table 5.3 shows the eigenvalues and cumulative dispersions of the discriminant functions.

The separation of class-labeled samples is presented in Figure 5.4. The samples in the learning set were attributed into the groups whose average Mahalanobis distance was similar to the average value of the data points of a certain group, using the knowledge of the real qualitative groups as originally designated by the operator.

After this discriminant analysis has been performed, the same discriminant learning template can be applied to 'unknown' samples. The classification resulting from analysis of unknown samples using a discriminant analysis

algorithm is presented in Table 5.4. All sample groups were classified correctly (100 % correct).

**Table 5.3. Canonical DFA analysis of electronic sensor responses**

| Discriminant Functions[th] | Eigenvalues | Cumulative Dispersion (%) |
|---|---|---|
| 1 | 382.900 | 95.30 |
| 2 | 14.432 | 98.90 |
| 3 | 3.551 | 99.80 |
| 4 | 0.774 | 100.00 |
| 5 | 0.107 | 100.00 |
| 6 | 0.003 | 100.00 |
| 7 | 0.000 | 100.00 |

**Table 5.4. Classification result* of the DFA algorithm for validation electronic sensor array data**

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | % Correct |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **1** | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **2** | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | **100** |
| **3** | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | **100** |
| **4** | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | **100** |
| **5** | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | **100** |
| **6** | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | **100** |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | **100** |
| Total | 8 | 7 | 7 | 5 | 8 | 5 | 7 | 7 | **100** |

***cases in row categories classified into columns: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).***

### 5.3.3. SPME/GC/MS

#### 5.3.3.1. Data Collection

Headspace gases in the vials containing super broth and super broth inoculated with *Salmonella* Typhimurium and incubated at 37 °C for 2-14 hours were collected and analyzed. SPME coupled to a gas chromatograph and mass spectrometer enabled identification of *Salmonella* Typhimurium volatile metabolites.

Figure 5.5 shows the chromatogram of volatile compounds from the headspace of super broth inoculated with *Salmonella* Typhimurium, and

incubated at 37 °C. Library searches of the mass spectra were used to identify individual compounds from the chromatograph, based on their mass spectra. The specific compounds produced by *Salmonella* Typhimurium grown in super broth include Hydrogen sulfide, Ethanol, Carbon disulfide, Dimethyl cyclopropane, and 1-Propanol. The average concentrations of each compound of different sample groups are displayed in Figure 5.6. The higher the incubation time the higher the concentration of volatile metabolites.

#### 5.3.3.2. Pattern Recognition of GC/MS Data

**a) PCA**

The selected volatile compounds including Hydrogen sulfide, Ethanol, Carbon disulfide, Dimethyl cyclopropane, and 1-Propanol were used as measurement variables and were arranged into the data matrix with 80 rows and 5 columns.

In this study, no internal standard was added, and thus the concentration of the selected compounds could be determined using selective normalization. The normalized concentration of each compound was then standardized to zero mean and unit variances before submitted to PCA.

With PCA, PCs, total variance factor loading, and factor scores were calculated. The corresponding principal components scores plot is presented in Figure 5.8. Samples with similar mass spectral pattern will have data points close to each other in the score plot. The score plot shows the location of the 80 samples spanned by the two most significant PCs, which explain 97.68 % (PC1

96.64%, PC2 1.05%) of the total variation in the chromatogram. The score plot displays the relationships among the samples. Samples located near each other are considered similar, while dissimilar samples tend to be well separated. From the score plot, the 2 PCs are needed for sample separation.

Within the GC/MS data and PCA, it is difficult to distinguish the different number of *Salmonella* Typhimurium. Although overlapping, there was a trend toward increased total volatile concentration levels along PC1. The repeatability (10 repetitions per group) of the measurement of the individual samples is lower than that found with the electronic sensors.

**b) DFA**

DFA was used to visualize the class separation between samples. Using DFA, seven discriminant functions were found. Table 5.5 shows the eigenvalues and cumulative dispersion of the first five discriminant functions. From DFA (Figure 5.9), all samples were divided into 8 groups in the plot of samples on the first and second discriminant factors as shown in Figure 5.9. Each of 8 groups consists of an approximate similar number of bacteria and concentration of volatile metabolites. With SPME/GC/MS and DFA, all samples could be discriminated from control samples.

The classification result of unknown samples using discriminant analysis algorithm is presented in Table 5.6. All groups, except groups 6 and 7, were classified correctly (100 % correct). The percent corrects in groups 6 and 7 may not have reached 100 % because as shown in Table 5.1, the number of

*Salmonella* Typhimurium in super broth of groups 6 and 7 were slightly different and the same concentrations of volatile metabolites were expected.

**Table 5.5. DFA analysis of SPME/GC/MS data**

| Discriminant Functions$^{th}$ | Eigenvalues | Cumulative Dispersion (%) |
|---|---|---|
| 1 | 185.905 | 99.00 |
| 2 | 1.252 | 99.70 |
| 3 | 0.554 | 100.00 |
| 4 | 0.030 | 100.00 |
| 5 | 0.020 | 100.00 |

**Table 5.6. Classification result* of the DFA algorithm for validation of GC/MS data**

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | % Correct |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **1** | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | **100** |
| **2** | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | **100** |
| **3** | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | **100** |
| **4** | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | **100** |
| **5** | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | **100** |
| **6** | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | **90** |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | **90** |
| Total | 5 | 4 | 4 | 5 | 5 | 4 | 5 | 4 | **98** |

***cases in row categories classified into columns: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).***

### 5.3.5. Electronic Sensors and ANN for Prediction of number of *Salmonella* Typhimurium

The objective of this work was to investigate if the electronic sensor array technique could be used to predict number of *Salmonella* Typhimurium and concentration of *Salmonella* Typhimurium volatile metabolites in nutrient media using neural network.

Figure 5.10 shows the effect of the number of neurons in the hidden layer on the mean square error (MSE) of the neural network to predict the number of *E. coli* (Figure 5.10). The network errors did not decrease dramatically when used with more than 10 neurons.

The transfer function in the hidden layer was hyperbolic tangent sigmoidal nonlinearity and linear transfer function was used in the output layer. From a preliminary study, the Levenberg-Marquardt (LM) algorithm was able to obtain lower mean square errors than any other algorithms tested. Therefore, the training function used in this study was Levenberg-Marquardt algorithms.

Neural networks were trained using selected parameters in data sets from several cultivations and were subsequently validated on independent data sets for estimating the concentration variables. The best estimation accuracy was observed on the validation sets.

The performance of a trained network can be measured by the errors in the training, validation and test sets. The network outputs are plotted versus the targets (prediction of number of *Salmonella* Typhimurium). Regression analysis was performed between the network output and the corresponding targets.

Figure 5.11 shows the predictions versus true values of numbers of *Salmonella* Typhimurium in super broth using electronic sensor array data and neural network algorithms. The correlation coefficient, $R^2$, between the outputs and targets is a measure of how well the variation in the output is explained by the targets and outputs. From the result, $R^2$ (0.9999) is close to 1, which

indicates a good fit and, therefore, the network created can correctly predict the number of *Salmonella* Typhimurium (RMSE = 0.0290).

This study has shown that electronic sensor arrays and neural network can be considered to provide accurate information on the number of *Salmonella* Typhimurium. Electronic sensor array and ANN has potential in real-time detection of *Salmonella* Typhimurium.

## 5.4. CONCLUDING REMARKS

A method was developed to predict the number of *Salmonella* Typhimurium in super broth and concentrations of specific volatile metabolites using electronic sensor array and ANN. The presence of *Salmonella* Typhimurium in super broth is detected by the electronic sensors as changes in the biochemical composition of the headspace vapors. It is noticed that *Salmonella* Typhimurium can be identified and quantified by their volatile metabolites. The electronic sensors have potential to detect *Salmonella* Typhimurium from their volatile metabolite, by responding to an entirely different part of the complex chemical mixture.

PCA provides considerable advantage by reducing the data dimensionality without loss of information. PCA is a useful aid in the visualization of multivariate data. For this application, DFA is performed. Unlike PCA, DFA allowed for a better separation of the clusters than simply plotting the raw data, because the variance of the data set is preserved in a smaller number of factors.

Artificial neural network created in this study was trained to predict the number of *Salmonella* Typhimurium and volatile metabolites from analytical electrical responses. The algorithm used in this work was back-propagation with LM learning. The networks used consisted of a hidden layer containing biased neurons using sigmoidal transfer functions, and an output layer containing linear transfer function. The network structures used in this work were those that contained the minimum number of hidden neurons and hidden layers while still satisfactorily modeling the systems.

Considering the overall performance of the electronic sensor, good agreement was found between the electronic sensors array and colony counting methods and between electronic sensors and GC/MS. The electronic sensor technology is as good as the conventional GC/MS technique in predicting the number of *Salmonella* Typhimurium in the samples.

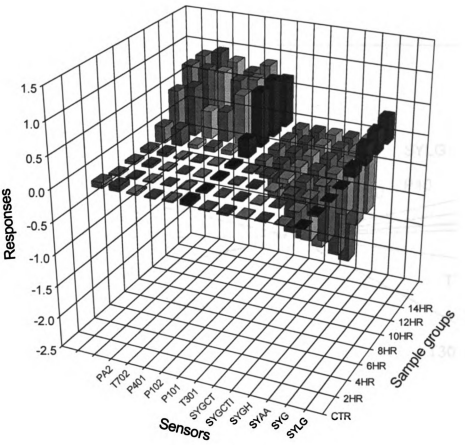


**Figure 5.1. Sensor responses of different samples groups (control and media inoculated with *Salmonella* Typhimurium and incubated for 2, 4 , 6, 8, 10, 12, and 14 hours) using 12 metal oxide sensors.**

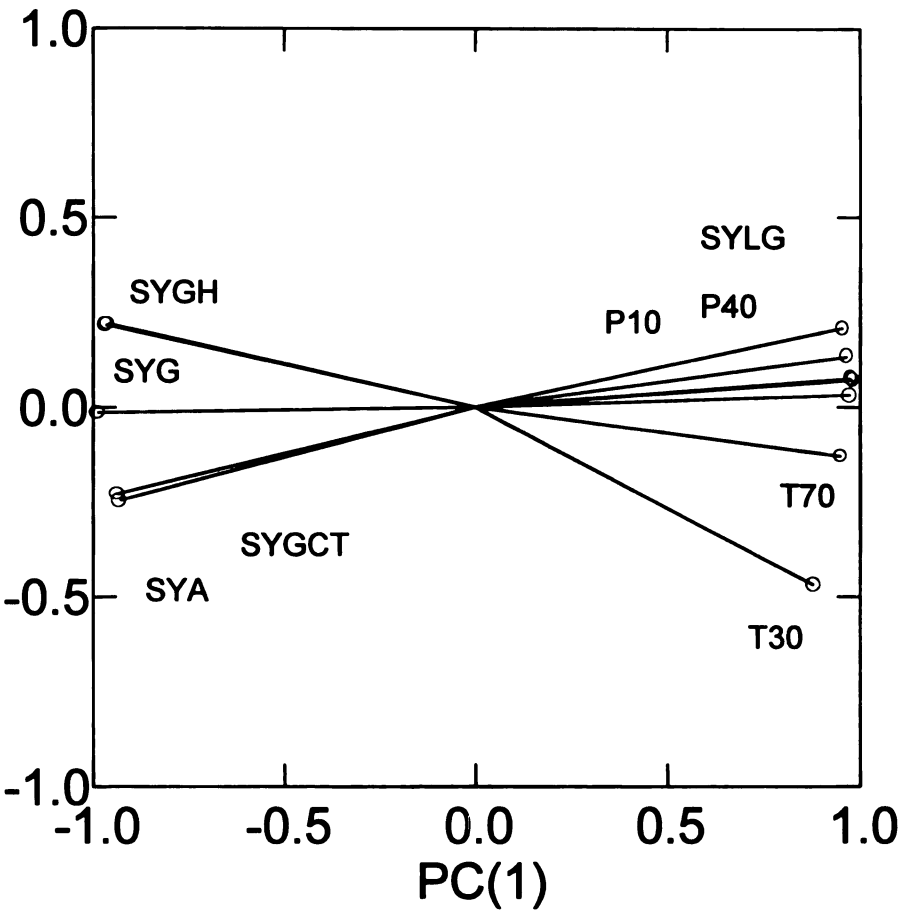


**Figure 5.2. PCA loadings plot of the data from electronic sensors with percent of total variance explained by PC1 96.584% and PC2 1.413%.**

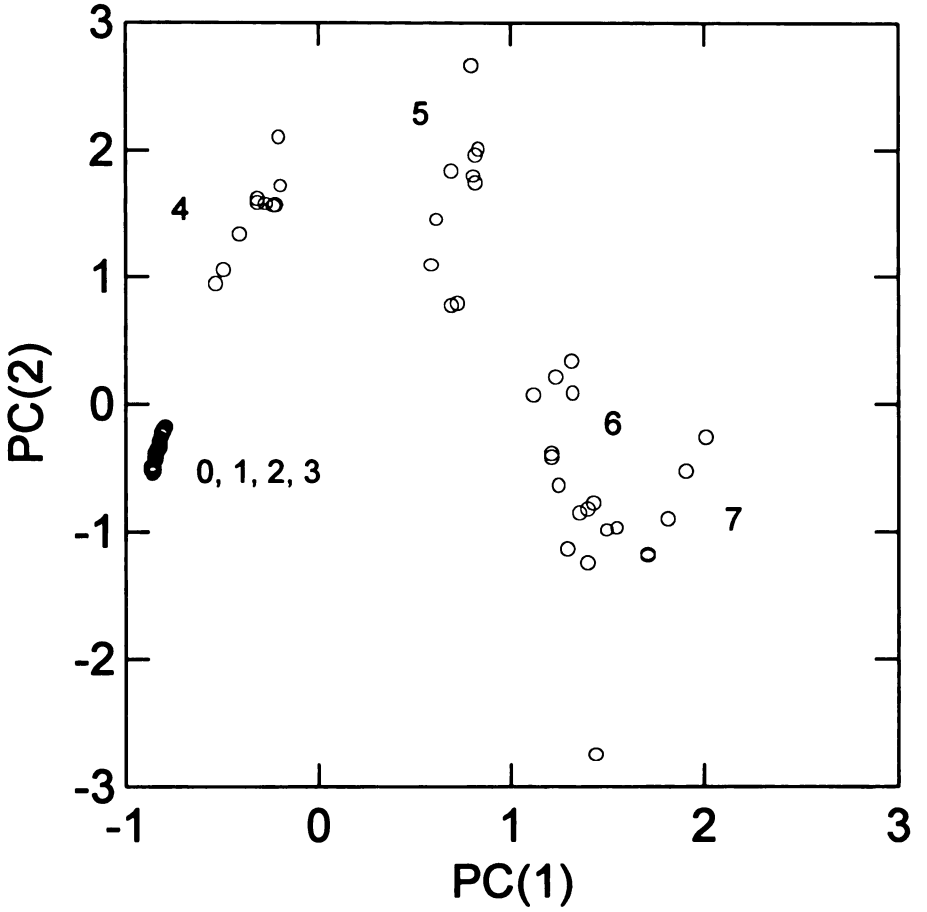


**Figure 5.3. PCA score plot of volatile compounds produced by *Salmonella* Typhimurium in super broth using 12 electronic sensors. The samples are labeled after the following scheme: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).**

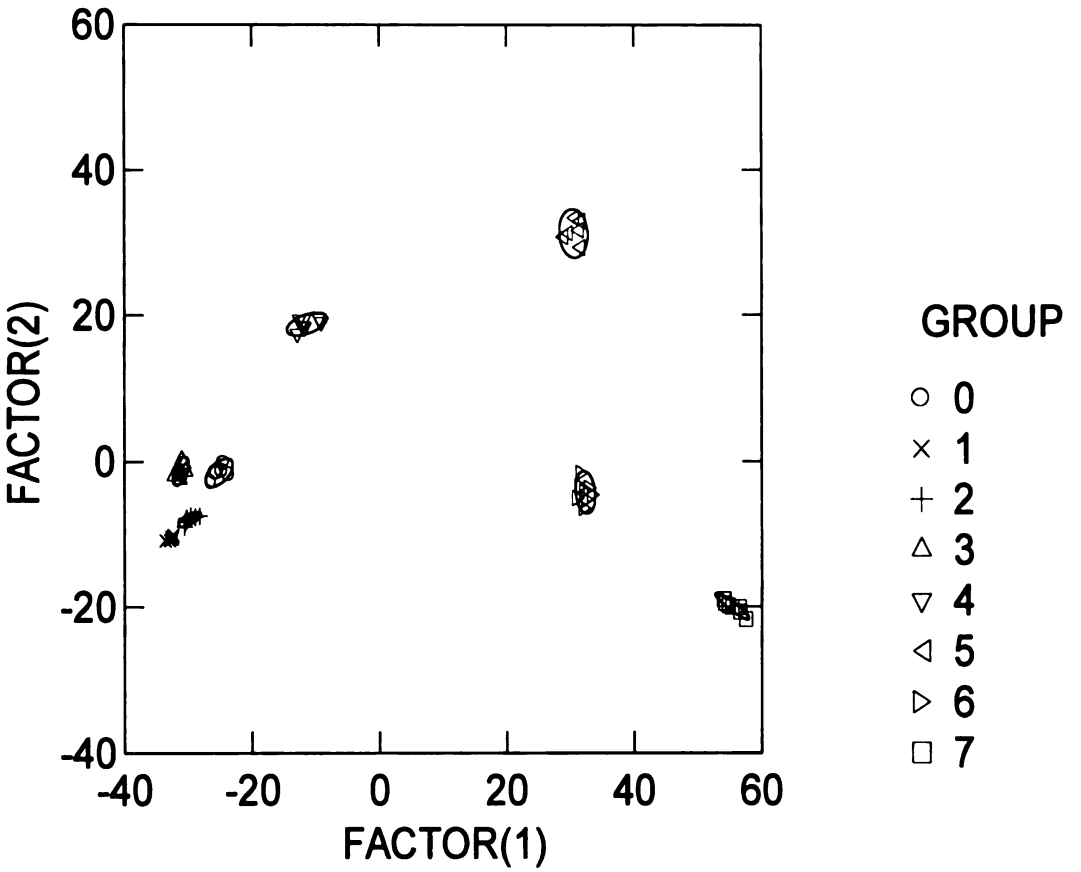


**Figure 5.4. DFA of volatile compounds of *Salmonella* Typhimurium in nutrient media using electronic sensors. The samples are labeled using the following scheme: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).**

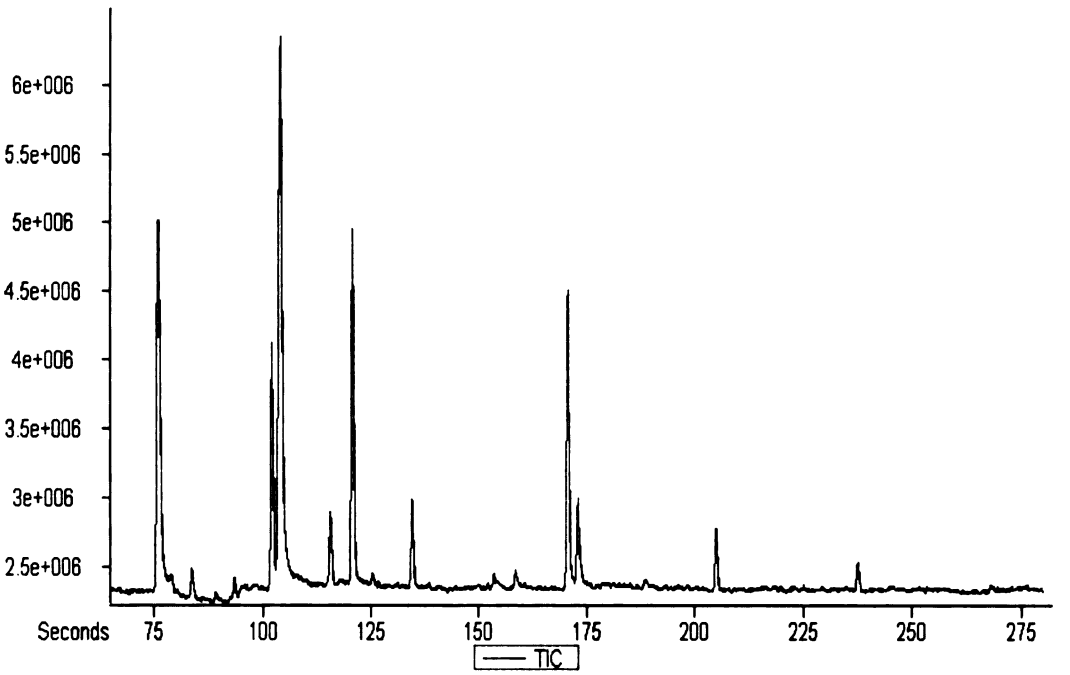


**Figure 5.5. Chromatograph of the volatile compounds from the headspace of super broth.**

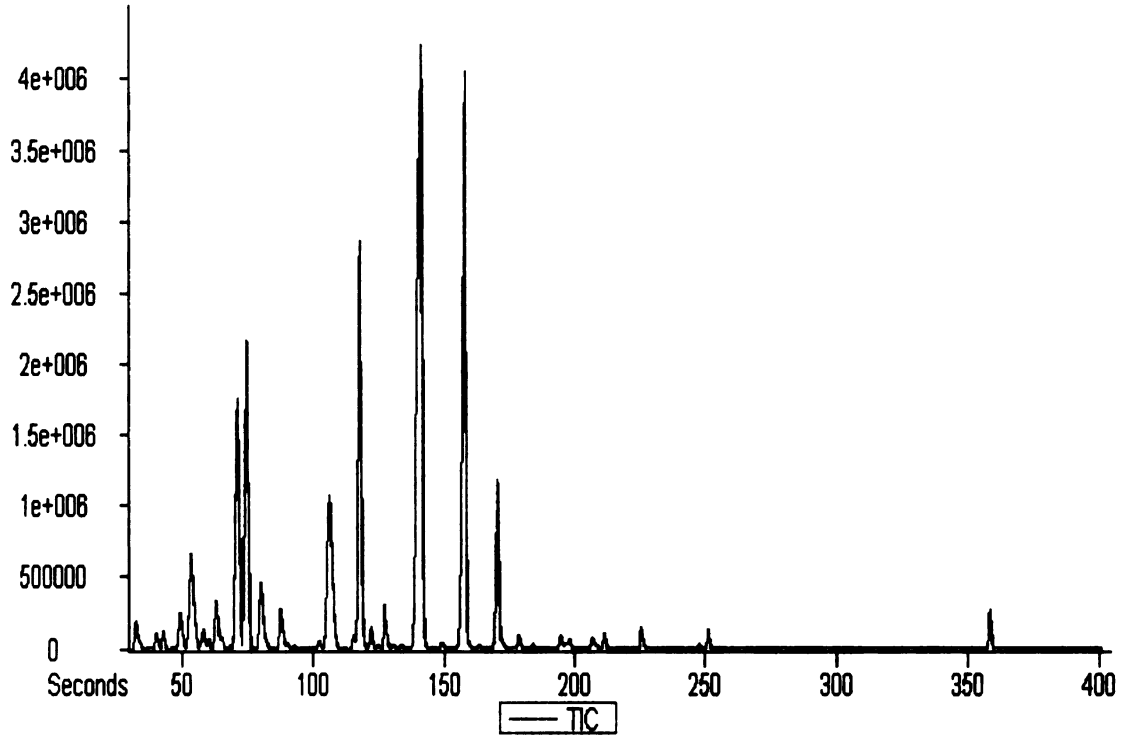


**Figure 5.6. Chromatograph of the volatile compounds from the headspace of super broth inoculated with *Salmonella* Typhimurium and incubated at 37 °C for 12 hours.**

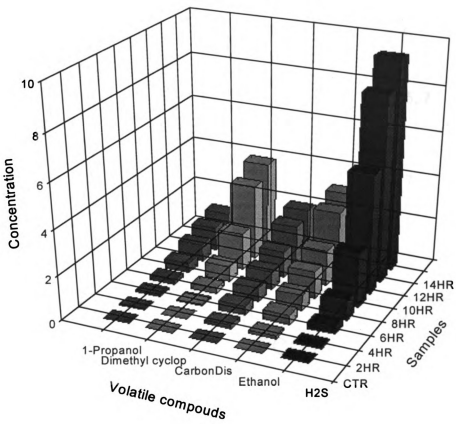


**Figure 5.7. Concentrations of volatile metabolites of different samples groups (control and media inoculated with *Salmonella* Typhimurium and incubated for 2, 4 , 6, 8, 10, 12, and 14 hours) using SPME/GC/MS.**

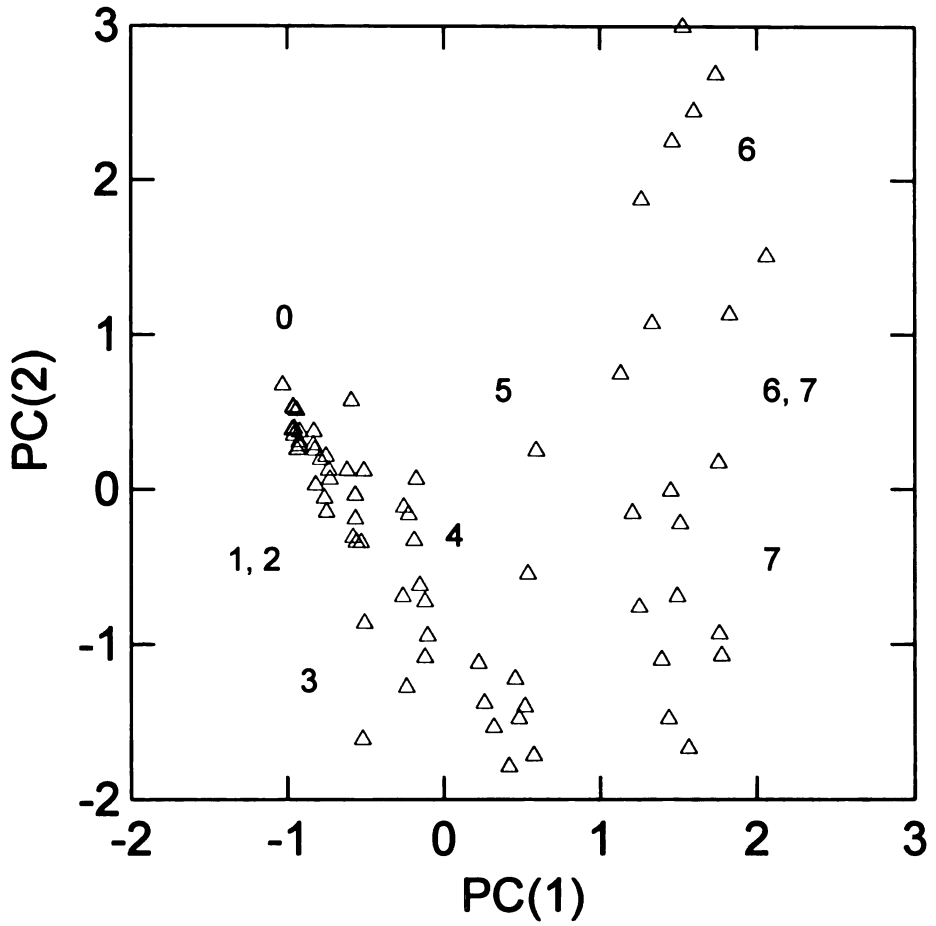


**Figure 5.8. Scores Plot of volatile compounds produced by *Salmonella* Typhimurium in super broth using SPME/GC/MS. The samples are labeled using the following scheme: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).**

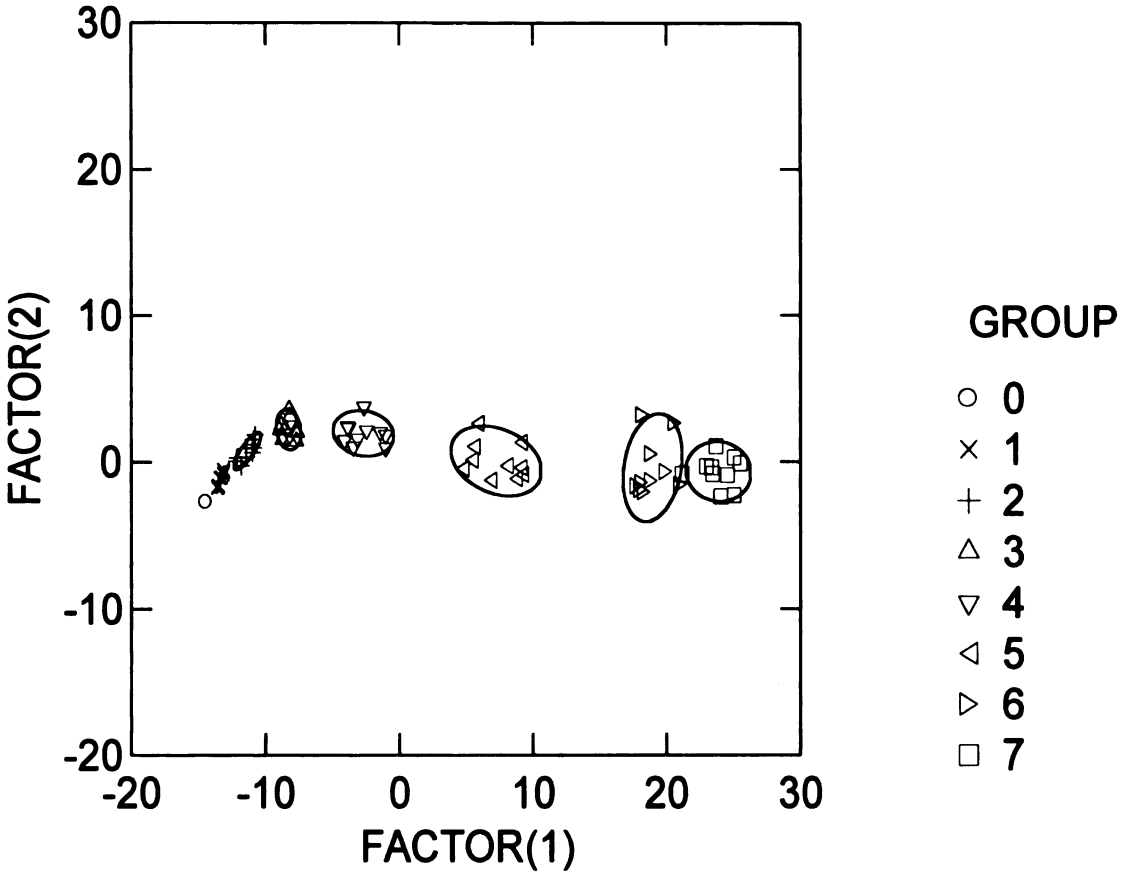


**Figure 5.9. DFA of volatile compounds produced by *Salmonella* Typhimurium in super broth using SPME/GC/MS. The samples are labeled using the following scheme: control (0), after incubated for 2 hrs (1), 4 hrs (2), 6 hrs (3), 8 hrs (4), 10 hrs (5), 12 hrs (6), and 14 hrs (7).**

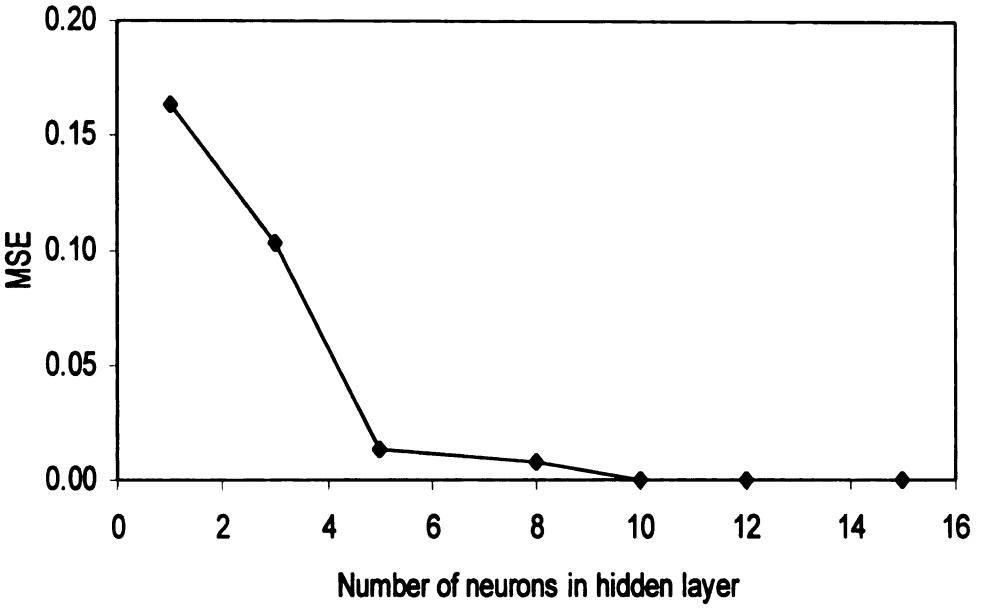


**Figure 5.10. Effect of number of hidden layers and number of neurons in hidden layer on the performance of the networks (MSE) of sensor responses *v.s.* number of *E.coli* .**

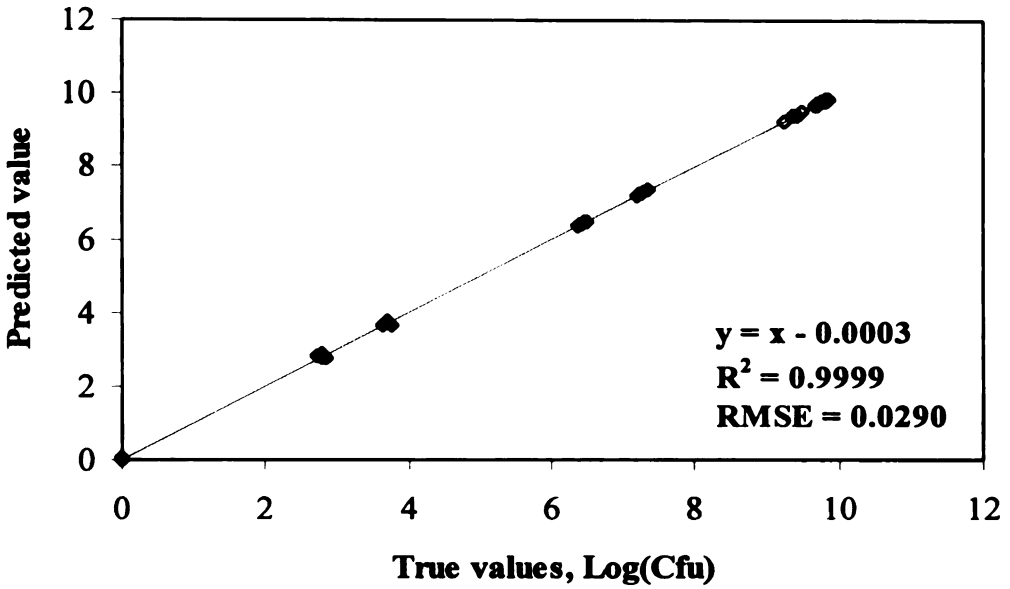


**Figure 5.11. The predicted *vs.* true numbers of *Salmonella* Typhimurium using ANN, back propagation algorithm.**

# CHAPTER 6

# ELECTRONIC SENSOR ARRAY INCORPORATING ARTIFICIAL NEURAL NETWORKS TO IDENTIFY AND QUANTIFY *ESCHERICHIA COLI* AND VOLATILE METABOLITES IN PACKAGED ALFALFA SPROUTS


## ABSTRACT

An array of 12 metal oxide electronic sensors incorporating neural network algorithms was used to identify and quantify *E. coli* and its volatile metabolites in packaged alfalfa sprouts. The metabolic volatile compounds used as indicators of *E. coli* were identified using solid phase microextraction coupled with gas chromatograph/ mass spectrometry (SPME/GC/MS). The volatiles from the headspace of packaged alfalfa sprouts were collected and analyzed using an electronic sensor array. The electronic sensor responses alone were not sufficient to distinguish samples with and without *E. coli*. Neural networks based on Multilayer perceptron (MLP) was trained to identify and quantify specific volatile compounds in mixtures under conditions where there were significant complications due to food components and volatile metabolites from other microflora in alfalfa sprouts. The networks were shown to be capable of correlating voltametric responses with number of *E. coli* with low mean square errors.

## 6.1. INTRODUCTION

Many approaches have been used to develop rapid, precise, and accurate techniques to potentially identify pathogenic organisms in food products. Novel, rapid, reliable, sensitive, and economical methods continue to be developed to allow rapid, in line, and accurate detection of hazardous organisms and their toxins. Development of new and improved methods continues to receive a great deal of research attention and a reasonable amount of private and public sector funding (CAST, 1999).

In this research, a method was developed to identify and quantify *E. coli* in packaged alfalfa sprouts using an electronic sensor array and neural networks. Alfalfa sprouts were studied because the National Advisory Committee on Microbial Criteria for Foods, NACMCF (1999) identified sprouts as a special problem because of the potential for pathogen growth during production. Also there is an increased demand for sprouts due to their popularity as a health food. Raw sprouts have been associated with at least eleven foodborne illness outbreaks since 1995. Exposure can occur due to seed contamination at the farm, seed processor, or sprouting facility. The processes used for the production of sprouted seed offer ample opportunity for cross contamination from a few seeds or sprouts to the entire production lot (FDA, 1999). *E. coli* was used as a target microorganism in packaged alfalfa sprouts because its presence in foods indicates fecal contamination, and possible pathogenic microorganisms (Eley, 1992).

Specific volatile metabolites were identified using Solid phase microextraction/Gas chromatograph/Mass spectrometry (SPME/GC/MS). Volatile *E. coli* metabolites can be used to indicate if *E. coli* is present in a product. The electronic sensors were used to analyze the volatile in the headspace above the samples.

Using electronic sensors to detect volatile complexes has the potential to be a sensitive, fast, one step method to monitor *E. coli* contamination in food products. Analysis of specific volatile organic compounds produced by the metabolic processes of *E. coli* may be able to be used to develop a rapid method that will allow detection of *E. coli* in a packaged plant product.

The primary advantage of the electronic sensor array in a quality assurance method is in its speed of analysis, including data acquisition and interpretation. Rapid, significant data interpretation is possible using various multivariate data analyses. In this study, the neural network algorithms were used to develop an integrated framework where feature extraction and predictive learning are iteratively performed with the goal being optimal approximation. Neural network is a massively parallel distributed processor constructed of processing units, which has propensity to store learned knowledge, thus making it available for use (Haykin, 1999). Neural networks have been used to solve many engineering problems ranging from image analysis (Tsuruta et al, 2000), dynamic systems (Watanabe et al, 2001), and robotic manipulation (Terra & Tinos, 2001).

The neural network algorithms developed in this study was Multilayer perceptrons based on a back propagation algorithm. The Multilayer perceptron network based on a back propagation algorithm was used to predict the number of *E. coli* and concentration of specific volatile metabolites.

## 6.2. METHODOLOGY

### 6.2.1. Preparation of Inoculated Vegetable

*E. coli* ATCC 25922 was obtained from the American Type Culture Collection (ATCC, Rockville, MD). *E. coli* was inoculated into Luria Bertani (LB) broth consisting of 10 g Bacto-tryptone, 5 g Bacto Yeast, and 5 g NaCl. The broth/culture was incubated at 37 °C in a gyrotory shaker (G-25 New Brunswich Scientific Corporation, NJ) at 100 rpm. The *E. coli* suspension was then dispensed into sterile 125 mL GSA bottles, and centrifuged (Rotor-GSA model RC 5 C Superspeed Centrifuge, Sorvall Instruments, Dupont Co., Haffman Estate, ILL) at 1600xg for 10 min. The supernatant was decanted, and the resulting pallets were resuspended in a 15 % glycerol solution. 1 mL aliquots of the bacterial suspensions were transferred into 1.5 mL microcentrifugal tubes and frozen using liquid nitrogen. The frozen culture was stored at –80 °C. *E. coli* ATCC 25922 was cultured in tryptic soy broth and incubated at 37 °C for 8 hrs in a gyrotory shaker. Before use, the culture was centrifuged at 6000 x g for 15 min. Broth was poured from the culture and the sedimented pellet was resuspended in sterile Butterfield's phosphate buffer. Preliminary experiments

were done to determine the population of *E. coli* necessary in the dipping suspension to result in an initial population of ~$10^5$ CFU/g on sprouts.

The alfalfa seeds (Natural Sprout Company, Springfield, MO) were soaked in 20,000 ppm of calcium hypochlorite prior to germination as advised by FDA (1999) and NACMCF (1999). Alfalfa sprouts were grown in a laboratory, under controlled conditions at the School of Packaging, Michigan State University. The sprouts were washed and drained several times. Alfalfa sprouts were placed in screened baskets, submerged in the suspension containing *E. coli* for 3 min and drained. The uninoculated control was similarly treated except sterile phosphate buffer was used in place of the inoculum. 50 g of sprouts were then dispensed into a commercial 1.5 mil, 6x8 $in^2$ LDPE bag (OTR = 62 cc/100 sq.in/day). The samples were incubated at 20 °C for 1, 2, and 3 days. The samples were analyzed using a colony counting method, gas chromatograph-mass spectrometer, and the electronic sensors.

### 6.2.2. Microbiological Analyses

The microbial cell count was determined on the date of inoculation and periodically throughout storage at days 1, 2, and 3. Sprout samples (25 g) were removed from the package and placed into sterile stomacher bags. Sterile phosphate buffer (225 mL) was added to the sample and the contents were homogenized by pummeling for 60 seconds in a Stomacher. A series of dilutions was prepared from the stock suspension, and Petri plates were inoculated with those dilutions expected to give countable colonies. Inocula consisting of each of

a dilution series were deposited on prepared plates in duplicate using 3M Petrifilm Aerobic Count Plates for determining aerobic bacteria and 3M Petrifilm *E. coli*/Coliform Count Plates containing Violet Red Bile nutrient as an indicator of glucuronidase activity for *E. coli*/Coliform. All plates were incubated at 37 °C for 48± 2 hrs. After inoculation, colonies were counted and the viable cell count per unit amount of the original sample was calculated.

### 6.2.3. Determination of Volatile Compounds Using SPME/GC/MS

Volatile compounds were collected using a Solid Phase Microextraction (SPME) (Supelco, Inc., Bellefonte, PA) device. The preliminary tests were done to determine the optimum SPME fiber for collecting the volatile compounds by comparing the results from the Polydimethylsiloxane fiber, Polydimethylsiloxane/Divinylbenzene fiber, and Carboxen/Polydimethylsiloxane fiber. The results showed that the SPME device most suitable was a Carboxen/Polydimethylsiloxane fiber. This fiber therefore, was used in further experiments. For headspace sampling, the fiber was exposed to the head space of the samples in the LDPE bag.

A gas chromatograph (HP-6890, Hewlett-Packard Co., Wilmington, DEL) was used to analyze the volatile compounds in the headspace of samples. Volatiles were separated using a capillary column (SPB5 30mx0.1 m id., 0.25 μm coating thickness). Helium gas was used as the carrier gas at a flow rate of 0.5 mL $min^{-1}$. The temperature program was isothermal for 2 min at 40 °C, and raised to 240 °C at a rate of 50 °C $min^{-1}$. Electron impact ionization (FCD-650,

LECO Corp., St Joseph, MI), was used by the time-of-flight (TOF) mass spectrometer to detect the volatiles. Mass spectra were collected at a rate of 40 spectras over a range of m/z 40-300. The ionization energy was 70 eV. Identification of volatile components was accomplished by comparison of collected mass spectra with those of authenticated standards and spectra in the National Institute for Standards and Technology (NIST) mass spectral library.

### 6.2.4. Electronic Sensors Procedure

The 12 metal oxide sensors (Fox 3000, Alpha M.O.S. Hillborough, NJ) were used to monitor the changes in volatiles produced by *E. coli* on the sprouts. The electronic sensors analysis conditions used in this study are shown in Table 6.1. Each injection was repeated, with separate samples. Injection was made with an HS100 autosampler.

The headspace volatile compounds from sprouts, and sprouts inoculated with *E. coli* and incubated for 0, 1, 2, and 3 days were monitored using the optimized condition. Injection of the headspace gas containing volatiles resulted in a measurable electronic signal. The electronic signals from the sensors were digitized, and then transferred to the control computer. Resistance changes (difference in sensor resistance between air blank and odorous atmosphere) were recorded.

**Table 6.1 Optimum condition for operating electronic sensors**

| | |
|---|---|
| Headspace generation time (s ) | 300 |
| Headspace generation temperature (°C) | 35 |
| Syringe temperature (°C) | 40 |
| Syringe type (mL) | 5 |
| Vial type (mL) | 10 |

### 6.2.5. Data Analysis

All matrix calculations were carried out using MATLAB 5.2 (Mathworks, Inc., Natick, MA) routines written by the author. Descriptive statistics were done using SPSS Version 10.0 (SPSS Inc., Chicago, IL). Graphical demonstrations were carried out with MATLAB.

## 6.3. RESULT AND DISCUSSION

### 6.3.1. Colony Counting Method

The cell counts of aerobic bacteria and *E. coli* on sprouts are shown in Table 6.2. All samples had a high number of total aerobic bacteria. However, *E. coli* was not found in the control samples. After 3 days incubation, the numbers of *E. coli* increased from $3.5 \times 10^5$ to $2.1 \times 10^8$ CFU/g.

**Table 6.2 Population of aerobic bacteria and *E. coli* on sprouts (cfu/g)**

| Storage time (day) | Sprouts | | Sprout inoculated with *E. coli* | |
|---|---|---|---|---|
| | Aerobic count | *E. coli* count | Aerobic count | *E. coli* count |
| 0 | $1.4 \times 10^7$ | - | $1.6 \times 10^7$ | $3.5 \times 10^5$ |
| 1 | $1.8 \times 10^8$ | - | $1.8 \times 10^8$ | $3.8 \times 10^5$ |
| 2 | $1.2 \times 10^9$ | - | $1.8 \times 10^9$ | $1.6 \times 10^7$ |
| 3 | $4.2 \times 10^9$ | - | $4.3 \times 10^9$ | $2.1 \times 10^8$ |

The alfalfa seeds were soaked in 20,000 ppm of calcium hypochlorite prior to germination as advised by FDA (1999) and NACMCF (1999). This treatment has the potential to substantially reduce microbial contamination which can be passed on to the growing sprouts. However, the total number of aerobic bacteria was as high as Log 7 (CFU)/g (Table 6.2). A study by Moline (1999) indicated

that 2 % calcium hypochlorite showed the greatest reduction but did not completely eliminate the natural microflora.

The number of aerobic bacteria, *E. coli* and salmonella on the alfalfa seeds was determined. The number of aerobic bacteria was ~1-2 Log (CFU)/g, while no *E. coli* or salmonella were found. The number of aerobic bacteria increased from ~1-2 Log (CFU)/g to ~8 Log (CFU)/g when the alfalfa sprouts were fully grown.

The conditions during sprouting, e.g. time, temperature, water activity, pH, and nutrient level, may have promoted the growth of microflora (CDC, 1997). Microorganisms on seeds can grow quickly under favorable conditions of sprouting (e.g. water activity, temperature, pH, time, and nutrients). Thus, the risk of foodborne disease associated with sprouts increases during sprouting (NACMCF, 1998).

Food composition may affect the type and extent of contamination by the nature of the raw product or ingredients that compose the food, thus providing nutrients and other intrinsic properties (e.g., pH, oxygen availability, presence of natural antimicrobials). Thus, the food source may allow, encourage, or discourage microbial growth. Food composition will also impact the presence and growth of other microorganisms (Ghazala, 1998; Nguyen & Carlin, 1994).

### 6.3.2. SPME/GC/MS

Headspace gases from the control samples (sprouts without *E. coli*) and sprouts inoculated with *E. coli* incubated at 20 °C for 1, 2, and 3 days were collected and analyzed using SPME coupled to a gas chromatograph and mass spectrometer. Library searches of the mass spectra were done to identify individual compounds, based on their mass spectra.

The volatile compounds present in the headspace of inoculated sprouts and absent in the headspace of uninoculated sprouts can be possible indicators of *E. coli* contamination. The specific compounds produced by *E. coli* on sprouts included Isopropyl alcohol; Ethanol, O-acetimidoyl-; Furan, 2-pentyl-; 2-Heptanone; Propyl cyclopropane; and Indole.

Trytophan is an amino acid found in alfalfa sprouts (Nguyen & Carlin, 1994) and is broken down into indole due to the activity of the enzyme tryptophanase from *E. coli*. Trytophanase degrades tryptophan to indole, pyruvate, and ammonia (Moat & Foster, 1988). In Figure 6.1 is shown the average concentrations of volatile compounds from the headspace of sprouts, and sprouts inoculated with *E. coli* at days 1, 2, and 3, respectively.

### 6.3.3. Electronic Sensor Array

#### 6.3.3.1. Data Collection

A total of 80 samples were analyzed (8 subgroups) which includes the control, and alfalfa sprouts inoculated with *E. coli* at time zero, and incubated for 1, 2, and 3 days. The sensor responses are given in units of maximum change of sensor's electrical resistance divided by the initial resistance ($\Delta R_{max}/R_o$).

Figure 6.2 shows the average sensor response intensity of the 8 subgroups (control sample and sprouts inoculated with *E. coli* and incubated at 20 °C for 0, 1,2, and 3 days). The intensity patterns of all samples were different due to the different concentrations of volatile metabolites generated by the different numbers of *E. coli* produced during the incubation period.

#### 6.3.3.2. Pattern Recognition of Electronic Sensor Array

**a) PCA**

Standardization was applied to the data matrix to transform the data into a more expedient form with size effects removed and variance scaled. PCA with Varimax rotation was then carried out on this pretreated data.

In this study, the correlation matrix of the data was used because the sensor responses were measured in the same unit. The results are displayed in a loading plot (Figure 6.3) and a score plot (Figure 6.4). The data information was retained by PC1 (43.08 %), PC2 (18.45 %), and PC3 (28.01 %). As a result, the first three PCs accounted for 89.45 % of the relevant information from the data.

Figure 6.3 shows a loading plot of all samples. The loadings display facilitates an exploration of the contributions of the variables to each PC. Sensors SYG, SYAA, and SYGH had high negative loadings on PC1, whereas sensors SYLG, T301, T702, and PC2 had high positive loadings on PC1. Sensors SYGCTI and SYGCT were the important sensors contributing to PC2 and sensors P101, P102, and P401 were the important variables contributing to PC3.

PCA reduces a large data set of correlated variables to smaller numbers of uncorrelated components. PCA was also employed to detect outliers. If any outlying samples were detected, they were removed from the data set. The result are displayed using a contour plot (Figure 6.4).

The score plot in Figure 6.4 shows several distinct clusters with positive and negative scores for PC1, PC2 and PC3. Samples in the same subgroups are grouped together under the same area in the PCA score plot. Samples having divergent numbers are located further apart. Many sample groups overlapped and thus, for many samples, it was difficult to predict whether a sample contained *E. coli*.

In this study, PCA did not provide sufficient discrimination. If no priori knowledge of class labels exists, it is difficult to differentiate between subgroups. However, once a PCA model has been developed (including scaling vectors, eigenvalues, loadings, and scores), the data can be evaluated using other data processing methods including multivariate calibration (Kresta et al., 1991) or multivariate statistical process control (Nomikos & MacGreger, 1995).

**b) DFA**

DFA was used to determine whether it is possible to separate two or more individual groups, given measurements for these individuals from several variables. The separation of class-labeled samples is presented in Figure 6.5.

The samples in the learning set were attributed to the groups whose average Mahalanobis distance was similar to the average value of the data points of a certain group, using the real qualitative groups as originally designated by the operator. Applying DFA, good classification was achieved. Samples from the first day of inoculation, and the control incubated for one day overlapped.

In this study, DFA (based on the Mahalanobis distance) was found to be a useful technique to describe and classify the training and validation of samples. DFA derives a projection that separates the different classes as far as possible, and compresses the individual classes as tightly as possible (Etemad & Chellappa, 1997).

### 6.3.4. Electronic Sensors and Artificial Neural network

The objective of using neural network in this study was to create a model whose expected performance on unseen data falls within acceptable bounds. This arises from the need to predict the degree of generalization and robustness of the model.

In this study, BR had better generalization than other algorithms. The Bayesian regularization activation function was, therefore, selected in order to

improve generalization. The training was stopped when the square error was relatively constant over several iterations.

The neuron model architecture created for *E. coli* comprises an input layer, one hidden layer of neurons and one output layer. The transfer function in the hidden layer was a hyperbolic tangent sigmoidal nonlinear function. A linear transfer function was used in the output layer. Linear and hyperbolic tangent sigmoidal functions were used as transfer functions in the network.

Training of the network was carried out by comparing the calculated target values and the desired output by calculation of sum square errors. Neural networks were trained using selected parameters in data sets from several cultivations and were subsequently validated on independent data sets for estimating the concentration variables. The best estimation accuracy was observed on the validation sets.

The performance of a trained network can be measured by the errors in the training, validation and test sets. The network outputs are plotted versus the targets (prediction of number of *E. coli*). Regression analysis was performed between the network output and the corresponding targets.

Figure 6.7 shows the predictions versus true values of numbers of *E. coli* in super broth using electronic sensor array data and neural network algorithms. The correlation coefficient, $R^2$, between the outputs and targets is a measure of how well the variation in the output is explained by the targets and outputs. A perfect fit, $R^2 = 1$, is indicated and, therefore, the network created can correctly predict the number of *E. coli* (RMSE = 0.0115).

This study has shown that electronic sensor arrays and neural network can provide accurate information on the number of *E. coli*. Electronic sensor array and neural network has potential in the real-time detection of *E. coli*.

## 6.4. CONCLUDING REMARKS

The volatile metabolites from the headspace of packaged alfalfa sprouts inoculated with *E. coli* were analyzed using SPME/GC/MS. The volatile compounds present in the headspace of inoculated sprouts and absent in the headspace of uninoculated sprouts can be used as possible indicators of *E. coli* contamination. The specific compounds produced by *E. coli* on sprouts include Isopropyl alcohol; Ethanol, O-acetimidoeyl-; Furan, 2-pentyl-; 2-Heptanone; Propyl cyclopropane; and Indole.

The electronic sensor was used to monitor changes in the composition of the gas phase of biochemical products from *E. coli* volatile metabolites directly from packaged alfalfa sprouts (without culturing in standard media). In this research, the electronic sensor array has shown potential to detect specific *E. coli* volatile metabolites, even though the sprouts contained high aerobic counts.

However, the capability of the electronic sensor array to detect the volatiles produced by *E. coli* occurred when the number of *E. coli* was higher than$10^5$ CFU/g. A low number of target organisms in a heterogeneous population of organisms may be present in food, and may be extremely significant in food poisoning. The proper selective growth enrichment stage may

be required prior to physical testing using the electronic sensor array, or SPME/GC/MS in order to reduce the level of false-negative results.

Principal component analysis was used to explore the data matrices obtained from the electronic sensors. PCA provides considerable advantage by reducing the data dimensionality without loss of information. Unlike PCA, DFA allows for better separation of the clusters than simply plotting the raw data, because variance in the data set is preserved in a smaller number of factors.

Multilayer perceptrons, based on the back propagation neural network algorithms created in this study, were trained to predict the number of *E. coli* and volatile metabolites from analytical electrical responses. The algorithm used in this work was back propagation with LM learning.

The networks used consisted of a hidden layer containing neurons using sigmoidal transfer functions and an output layer containing a linear transfer function. The network structures used in this work were those that contained the minimum number of hidden layers while still satisfactorily modeling the systems.

Electronic sensor array, with neural networks has the potential to be a rapid and real time method, and to detect target microorganisms in food products (such as alfalfa sprouts). Neural network algorithms can help identify and quantify *E. coli*. These results demonstrate that handling of data using pattern recognition/classification techniques has potential to extract more structural information from electronic sensor array and GC/MS data.

This system is not limited to the aforementioned applications. The system can be applied to other packaged food products or incorporated into HACCP protocols or quality control systems in the food industries.

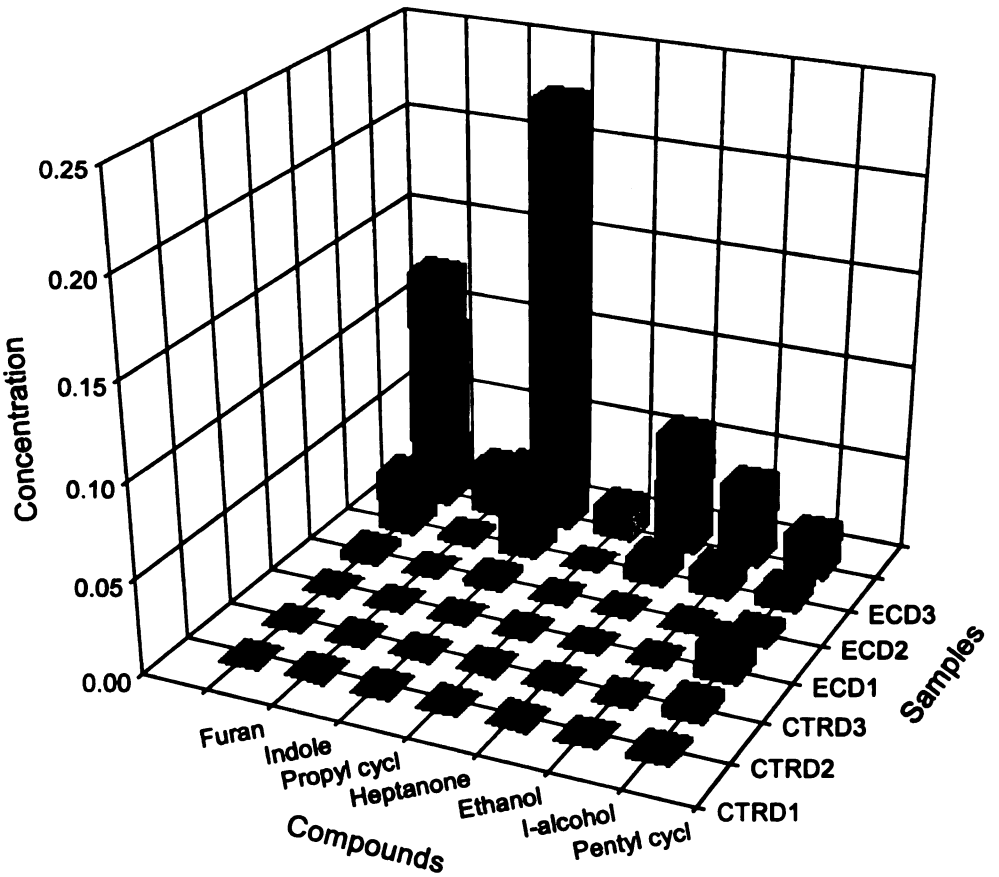


**Figure 6.1. Average concentration of specific volatile compounds associated with alfalfa sprouts (Control, CTR) and sprouts inoculated with *E. coli* (EC) in LDPE bags and incubated at 20 °C for 1, 2, and 3 days using SPME/GC/MS.**

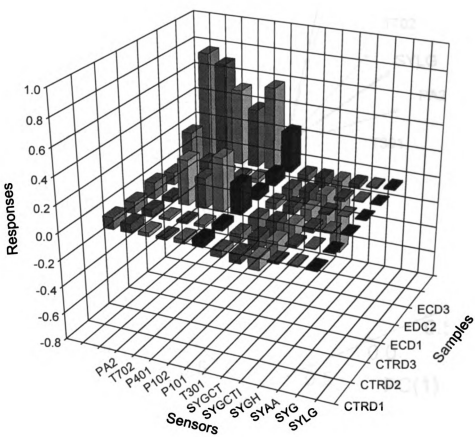


**Figure 6.2. Average sensor responses from the headspace of alfalfa sprouts (Control, CTR) and sprouts inoculated with *E. coli* (EC) in LDPE bags and incubated at 20 °C for 1, 2, and 3 days using 12 metal oxide sensors.**

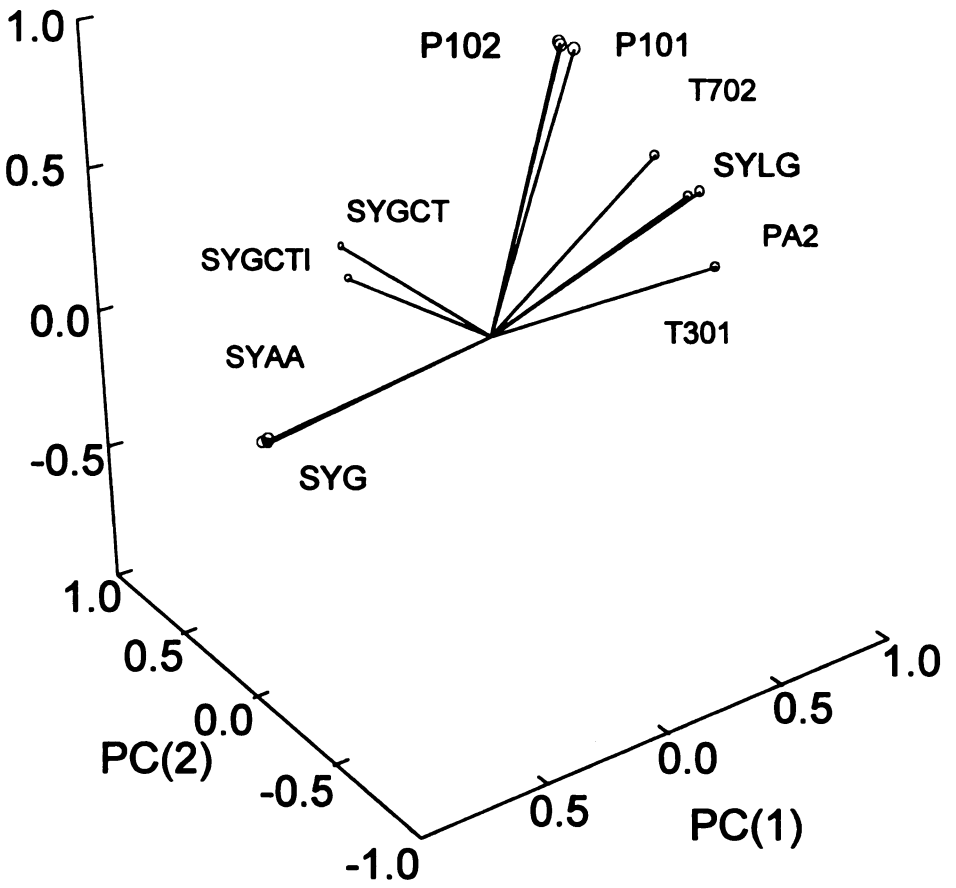


**Figure 6.3 PCA loading plot of data of alfalfa sprouts and sprouts inoculated with *E. coli* packed in LDPE bags and analyzed using 12 metal oxide electronic sensors**

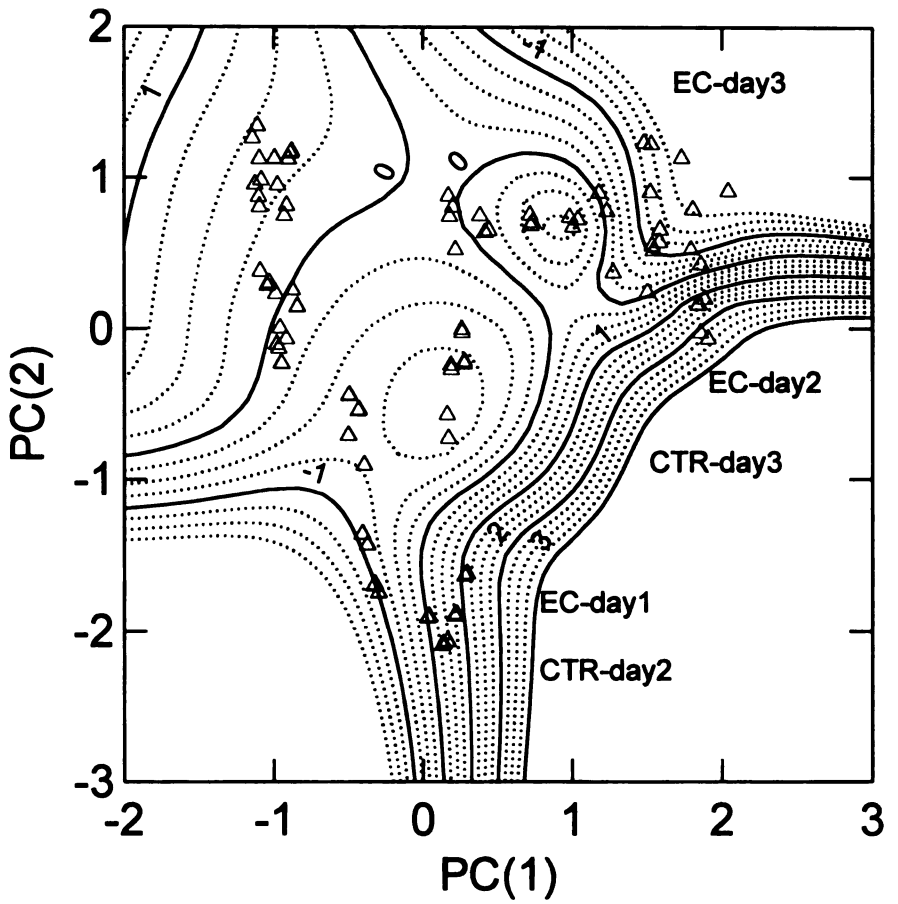


**Figure 6.4. Contour plot of PC scores on PC1, PC2, and PC3 from the data of alfalfa sprouts and sprouts inoculated with *E. coli* in LDPE bags on the first day of inoculation and after incubated at 20 °C for 1, 2, and 3 days.**

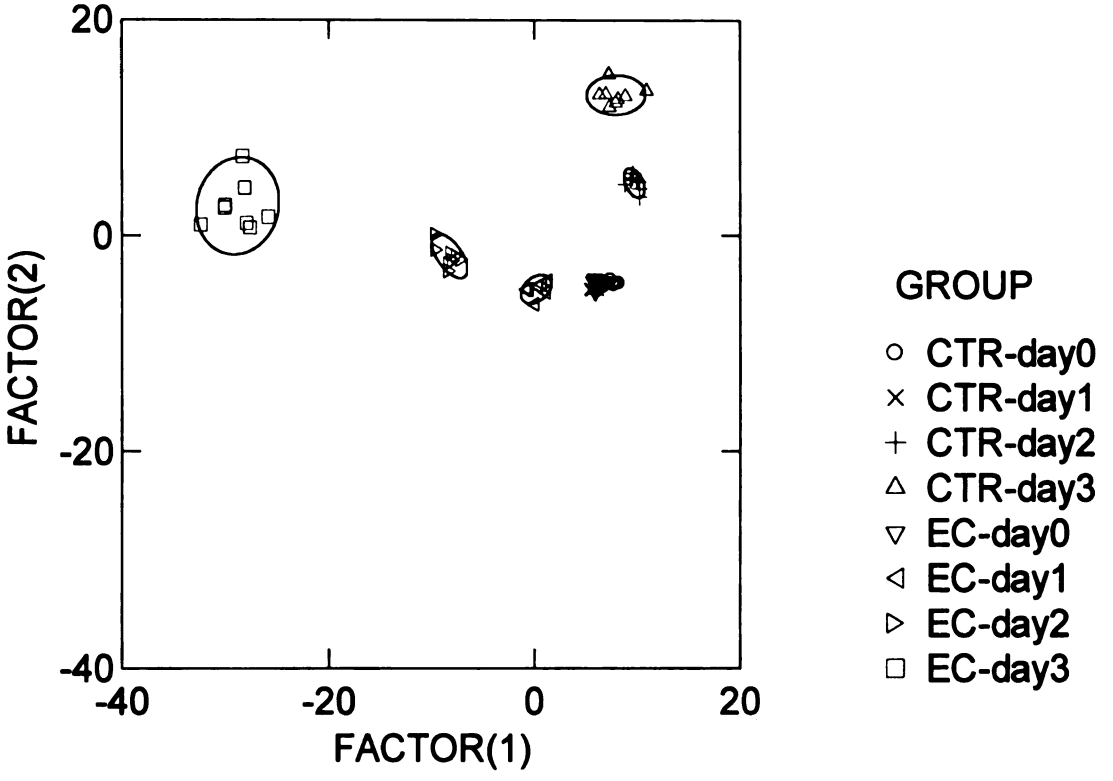


**Figure 6.5. DFA of volatile compounds using electronic sensor array from headspace of alfalfa sprouts and sprouts inoculated with *E. coli* (in LDPE bag) on the first day of inoculation and incubated for 1, 2, and 3 days.**

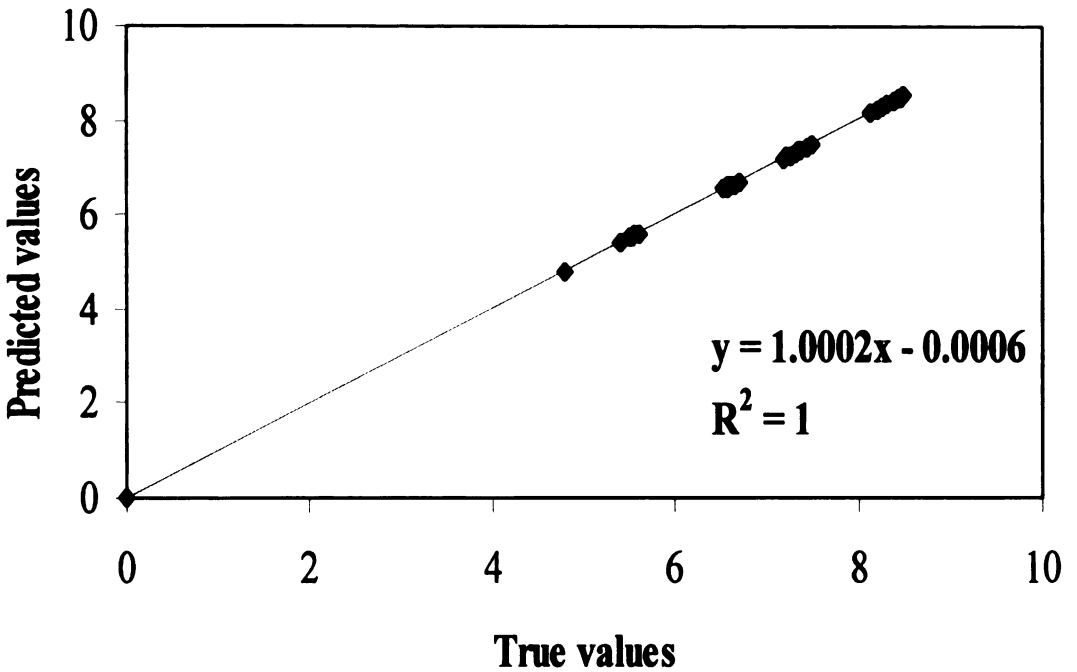


**Figure 6.6. The predicted *vs.* true numbers of *E. coli* using neural network-LM algorithm.**

# CHAPTER 7

# CONCLUSION

The preceding chapters demonstrated the potential utility of the electronic sensor array incorporating multidimensional data analyses including chemometrics and neural networks to identify and quantify the target microorganisms *E. coli* and *Salmonella enterica* serovar Typhimurium in control media and packaged alfalfa sprouts. The organism can be detected by identification of certain unique metabolites associated with the metabolic activity of the particular organism.

In chapter 4, the study was chronologically conducted in three steps. The first step was data acquisition and collection using SPME/GC/MS and the electronic sensor array. The second step was feature extraction using Principal Component Analysis (PCA) to explore the data, with data classification using Discriminant Factor Analysis (DFA).

Artificial Neural Network (ANN) technique showed the potential of the electronic sensors to predict the number of *E. coli* in unknown samples. The electronic sensor array coupled with neural networks can be used to identify and quantify the number of *E. coli*.

In chapter 5, a method was developed to predict the number of *Salmonella* Typhimurium in super broth, and concentrations of their specific volatile metabolites using electronic sensor array and neural networks. The

presence of *Salmonella* Typhimurium in super broth was detected by the electronic sensors as changes in the biochemical composition of the headspace vapors. The neural network algorithm created in this study was trained to predict the number of *Salmonella* Typhimurium and volatile metabolites from analytical electrical responses.

Considering the overall performance of the electronic sensor, good agreement was found between the electronic sensor array and colony counting methods and between electronic sensors and GC/MS. This study proves that electronic sensor array coupled with neural network can predict the number of *Salmonella* Typhimurium, and concentration of their volatile metabolites.

In Chapter 6, the volatile metabolites from packaged alfalfa sprouts inoculated with *E. coli* were analyzed using SPME/GC/MS. The volatile compounds present in the headspace of inoculated sprouts and absent in the headspace of uninoculated sprouts can be possible indicators of *E. coli* contamination. The electronic sensor was used to monitor changes in the composition of the gas phase of volatile metabolites directly from packaged alfalfa sprouts without culturing in standard media. The electronic sensor array has potential to detect the specific *E. coli* volatile metabolites, even though the sprouts contained high aerobic counts.

The Multilayer perceptrons based on back propagation neural network algorithms and created in this study were trained to predict the number of *E. coli* and volatile metabolites from analytical electronic responses.

Artificial neural networks have been found to be powerful and versatile techniques particularly suited to information processing. Their flexibility is a decisive asset compared with parametric techniques that require the assumptions of a specific hard model form. The results demonstrate that handling of data using pattern recognition/classification techniques has potential to open up new possibilities to extract more structural information from electronic sensor array.

While GC/MS is extremely useful in identifying the microorganisms from their volatile metabolites, it is not a panacea. GC/MS requires an experienced operator and has high maintenance. Because of these concerns, appropriate consideration must be given to properly quantify and normalize the data.

Electronic sensor array is an equally good candidate to monitor changes in the composition of the gas phase of chemical products. The advantages of electronic sensor array over classic GC are that it is simpler and is higher speed. However, with the electronic sensor, the compounds are not analytically identified.

This research has shown the distinctive possibility of using different pattern recognition/classification techniques for successfully extracting information from electronic sensor array. The electronic sensor array incorporating neural networks is a rapid and real time method to detect the target microorganisms in food products.

The electronic sensor technique can be economical, specific and real-time. The electronic sensor array has great potential in quality control, and process monitoring in the food industry both in principle and in practice.

This research has shown the potential of the electronic sensor array incorporating neural networks for pattern recognition/classification techniques to identify and quantify *E. coli* and their volatile metabolites in packaged alfalfa sprouts. This system is not limited only to the aforementioned applications. The system can be applied to other packaged food products, incorporated into a HACCP protocol or quality control system in the food industries.

# CHAPTER 8

# FUTURE PROSPECTS

As this research was pioneering in nature, much research remains. The future prospective works can be placed into 3 categories.

## 8.1. Electronic sensor array

a. Improve stability and sensitivity of the sensors. The electronic sensor array could not detect low levels of microorganisms and volatile metabolites in the samples, and thus, it is necessary to further improve the limit of detection and determination.

b. Since there is no universal electronic sensor technique at present that can solve all volatile sensing problems, the development of application-specific electronic sensor technology appropriate to the application is required.

c. Investigate and characterize sensor drift and create techniques to correct the specific problem.

d. Study the effects of selectivity and affinity of specific volatiles to the sensing elements of the electronic sensors to improve degree of sample discrimination.

e. Perform the previous research using samples with different, practical, and challenging scenarios.

## 8.2. Pattern recognition/Neural network

a. Study different pattern recognition techniques which can be automatically incorporated in the electronic sensing system. Several linear and nonlinear multivariate data analysis techniques should be further studied including Nonlinear PCA, Soft independent modeling of class analogy (SIMCA), and Cluster analysis (CA), for instance.

b. Incorporate online-monitoring. It is essential to maintain high quality in assembly recordings so that subsequent analyses can be successful. This requires some degree of on-line monitoring of signals, and particularly of separation.

c. Determine the optimum method for analyzing the neural networks. In addition to a procedure for amplitude covariation, trial-by-trial latency covariation should be studied which is more problematical. A number of methods can be used to determine response latency of an individual trial.

d. Simplify the depiction of the simulating model used for extracting and predictive learning of multivariate data from electronic sensors. The conclusion from the simulation is model-dependent. However, the simulation may well provide a different answer if another set of assumptions is chosen.

## 8.3. Distinctive applications

a. Determine the possible interactions between specific volatile compounds and packaging materials.

c. Apply this technique to other packaged food products.

d. Incorporate the techniques into a HACCP protocol or quality control system.

# BIBLIOGRAPHY

## BIBLIOGRAPHY


Acnab, W. B. 1997. A literature review linking microbial risk assessment, predictive microbiology, and dose-response modeling. Dairy Food and Environmental Sanitation. 17(7): 405-416.

Alugupalli, S., Larsson, L, and Slosarek, M. 1992. Application of gas chromatography-mass spectrometry for rapid detection of mycobacterium xenopi in drinking water. Applied and Environmental Microbiology. 11: 3538-3541.

Anklam, E., Lipp, M., Radovic, B., Chiavaro, E., and Palla, G. 1998. Characterization of Italian vinegar by pyrolysis-mass spectrometry and a sensor device (electronic nose). Food Chemistry. 61(1/2): 243-248.

Bachinger, T., Liden, H., Martensson, P., and Mandenius, C. 1998. On-line estimation of state variables in Baker's yeast fermentation using an electronic nose. Seminar in Food Analysis. 3, 85-91.

Bao, M., Griffini, O., Burrini, D., Santianni, D., Barbieri, K., and Maschini, M. 1999. Headspace solid-phase microextraction for the determinaiton of trace levels of taste and odor compounds in water samples. Analyst. 124(1): 459-466.

Barlett, P.N., Elliott, J.M., and Gardner, J.W. 1997. Electronic noses and their application in the food industry. Food Technology. 51(12):44-48.

Bazemore, R. and Rouseff, R. 1998. Discrimination of thermally treated orange juices by an electronic nose equipped with organic polymer sensors. Seminar in Food Analysis. 3: 59-66.

Bazzo, S, Loubet, F, and Tan, T. 1998. Quality control of edible oil using an electronic nose. Seminar in Food Analysis 3, 15-25.

Beuchat, L.R. 1995. Pathogenic microorganisms associated with fresh produces. J. of Food Protection. 59: 204-216.

Bila, S., Harkouss, Y., Ibrahim, M., Rousset, J., N'Goya, E., Baillargeat, D., Verdeyme, S., Aubourg, M., and Guillon, P. 1999. An accurate wavelet neural-network-based model for electromagnetic optimization of microwave circuits. J.CA Engineering. 9 (3): 297-306.

Borjesson, T. 1996. Electronic nose for odor classification of grains. Cereal Chemistry. 73(4): 457-461.

Borjesson, T.S., Stollman, U.M., and Schnurer, J.L. 1993. Off-odorous compounds produced by molds on oatmeal agar identification and relation to other growth characteristics. J.Agric.Food Chem. 41: 2104-2111.

Bourrounel, B., Talou, T., and Gaset, A. 1995. Application of a multi-gas-sensor device in the meat industry for boar-taint detection. Sensors and Actuators. 26-27: 250-254.

Busby, J.C., and Roberts, T. 1999. Bacterial foodborne disease: medical costs and productivity losses. USDA, Economic Research Service, Agricultural Economic Report No. 741.

Choradash, R.A. and N.F. Insalata. 1978. Incidence and pathological significance of escherichia coli and other sanitary indicator organisms in food and water, Food Technology. 62: 518-524.

Choralambus, C. 1992. Conjugate gradient algorithm for efficient training of artificial neural networks. IEEE Proceedings 139 (3): 301-310.

Cohen, J. 1988. Set correlation and contingency tables. Applied Psychological Measurement. 12: 425-434.

Cooley, W.W. & Lohnes, P.R. 1971. Multivariate data analysis. John Wiley & Son, Inc. New York.

Coulibaly, P., Bobée, B., & Anctil, F. 2001. Improving extreme hydrologic events forecasting using a new criterion for artificial neural network selection. Hydrological Processes. 15 (8): 1533-1536.

Council for Agricultural Science and Technology. 1994. Foodborne pathogens: risks and consequences. Task Force Report No.122.

Council for Agricultural Science and Technology. 1999. Foodborne Pathogens: Review of Recommendations. October.

D'Aoust, J. Y. 1994. Salmonella and the international food trade, International Journal of Food Microbiology 24:11-31.

Davis, H., Taylor, J.P., Perdue, J.N., Stelma, G.N., Humphrey, J.M., Rowntree, R., and Greene, K.D. 1988. A shigellosis outbreak traced to commercially distributed shredded lettuce, Am.J.Epidemiol., 128: 1312.

Devabhaktuni, V., Yagoub, M.C.E., Fang, Y., Xu, J., & Zhang, Q. 2001. Neural networks for microwave modeling: Model development issues and nonlinear modeling techniques. J. Microwaave CAE. 11: 4-21.

Dijksterhuis, G. 1995. Multivariate data analysis in sensory and consumer science: An overview of developments. Trends in Food Science & Technology. 6:206-210.

Domingo, C., Arcis, R.W., Osorio, E., Toledano, M., & Saurina, J. 2000. Principal component analysis and cluster analysis for the characterization of dental composites. Am.J.Epidemiol. 125: 2044-2048.

Eklov, T. 1998. Monitoring sausage fermentation using an electronic nose. Journal of the Science of Food and Agriculture. 76: 525-532.

Ekperigin, H.E., and Nagaraja, K.V.1998. Microbial food borne pathogens salmonella. Vet.Clin.N. Am. Food Anim. Pract. 14:17-29.

Eley, A.R. 1992. Chapter 2. Infective Bacterial Food Poisoning. Eley, A.R. in Microbial Food Poisoning. Chapman & Hall New York.

Erti, P. and Mkkelsen, S.R. 2001. Electrochemical biosensor for the identification of microorganisms based on lectin-lipopolysaccharide recognition. J. Chemistry. 7(2): 91-96.

Evans, A. S., and Brachman, P. 1991. Bacterial infections of humans: epidemiology and control. 2nd ed. Plenum Publishing, New York.

Extension Service, US Department of Agriculture. 1994. Complete guide to home canning. USDA/Extension Service. (Agriculture information bulletin no. 539).

Faber, N.M. 2000. Improved computation of the standard error in the regression coefficient estimates of a multivariate calibration model. J. Chem. 72: 4675-4676.

Fanazaki, N et al. 1995. Application of semiconductor gas sensor to quality control of meat freshness in food industry. Sensors and Actuators. 24-25: 797-800.

Favata, F., & Walker, R. 1991. A study of the application of neural networks to the traveling salesman problem. Cybernetics 64:463-468.

Ferretti, R., Mannazzu, I., Cocolin, L., Comi, G., and Clementi, F. 2001. Twelve-Hour PCR-Based Method for Detection of Salmonella spp. In Food. Applied and Environmental Microbiology. 67(2): 977-978.

Food and Drug Administration. 1998. Interim advisory on alfalfa sprouts. U.S. Department of Health and Human Services. T98-47. August 31.

Food and Drug Administration. 1999. Guidance for industry guide to minimize microbial food safety hazards for sprouted seeds. Center for Food Safety and Applied Nutrition, October 27.

Forsyth, D.S. and Dusseault, L. 1997. Determination of methylcyclopentadienly manganese tricarbonyl in beverages by solid phase microextraction. Food Additives and Contaminants 14(3): 301-307.

Fricker, C.R. and Tompsett, S. 1989. Aeromonas spp. In foods: a significant cause of food poisoning, Int.J. Food Microbiol. 9(17) 19-21.

Fung, W.K. 1995. Diagnostics in linear discriminant analysis. Journal of the American Statistical Association. 90(431): 952-956.

Gallagher, N.B., Wise, B.M., Stewart, C.W. 1996. Application of multi-way Principal components analysis to nuclear waste storage tank monitoring. J. Chem En. 20:739-744.

Gardner, J.W., and Bartlett, P.N. 1994. A brief history of electronic nose. Sensors and Actuators B. 18-19: 211-220.

Getino, J., Ares, L., Robla, J.I., Carmen, M.D., Sayago, H.I., Fernandez, M.J., Rodrigo, J., and Gutierrez, F.J. 1998. Gas sensor arrays for contaminated soils. (page 329-338) In Hurst, W.J. Electronic Noses & Sensor Array Based Systems: Design & Applications. Proceedings of the 5th International Symposium on Olfaction and the Electronic Nose. Technomic Publishing Company, Inc. PA.

Ghazala, S. 1998. Sous vide and Cook-chill processing for the Food Industry edited by, An Aspen Publication, Aspen Publishers, Inc. Maryland.

Gorecki, T., and Pawliszyn, J. 1997. Effect of sample volume on quantitative analysis by solid-phase microextraction: Part 1. Theoretical considerations. Analyst. 122(10): 1079-1086.

Gorecki, T., Kahled, A., and Pawliszyn. 1998. The effect of sample volume on quantitative analysis by solid-phase microextraction: Part 2. Experimental verification. Analyst. 123(10): 2819-2824.

Guerzoni, M.E., Gardini, F., Sinigaglia, M. and Guglielmi,G. 1992. HS-GC determination of lactic acid bacteria viable cells based on malolactic activity. Cerevisia and Biotechnology. Vol.17(1):19-23.

Hikawa, H. 2001. Multilayer neural network with on-chip learning based on frequency-modulated pulse signals and voting neurons. Electronics and Communications in Japan. 84 (1): 32-42.

Ho, J.L., Shands,K.N., Friedland,G., Eckind, P., and Fraser, D.W. 1986. An outbreak of type 4b Listeria monocytogenes infection involving patients from eight Boston Hospitals. Arch.Intern.Med.146-520.

Hodgins, D. 1995. The development of an electronic nose for industrial and environmental applications. Sensors and Actuators. 26-27: 255-258.

Hodgins, D., and Simmonds, D. 1996. The Electronic nose and its application to the manufacture of food products. Food Science and Technology Today. 11(3): 179-185.

Jay, J.M. 1996. Modern Food Microbiology-5$^{th}$ ed. Chapman & Hall, New York.

Jelen, H., Wlazly, K., Wasowicz, E., and Kaminski, E. 1998. Solid phase microextraction for the analysis of some alcohols and esters in beer: comparison with static headspace method. J.Agric.Food Chem. 46(3): 1469-1473.

Jonsson, A et al. 1997. Electronic nose for microbial quality classification of grains. International Journal of Food Microbiology. 35: 187-193.

Kaipainen, A, Ylisuutari, Lucus, Q, and Moy, L. 1997. A new approach to odor detection. International Sugar Journal. 99(1184): 403-408.

Kimura, M. & Nakano, R. 2000. Dynamical systems produced by recurrent neural networks. Systems and Computers in Japan. 31(4): 818-828.

Kresta, J.V., MacGregor, J.F., and Marlin, T.E. 1991. Multivariate statistical monitoring of process operating performance. Journal of Chemical Engineering. 69:35-47.

Lee, D., Jung, J., Lim, J., Hus, J., and Lee, D. 2001. Recognition of volatile organic compounds using sno2 sensor array and pattern recognition analysis. Sensors and Actuators, 77: 228-236.

Lucas, Q., Poling, J., and Bennincasa, V. 1998. Analysis with hybrid system PCM/CP/MOS. Seminar in Food Analysis. 3: 53-58.

Luco, J.M. 1999. Prediction of the brain-blood distribution of a large set of drugs from structurally derived descriptors using partial least square (PLS) modeling. J.Chem.Sci. 39(2): 364-404.

Luzuriaga, D.A. and Balaban, M.O. 1998. Electronic nose odor evaluation of salmon fillets stored at different temperatures. (page 162-169). In Hurst, W.J. Electronic Noses & Sensor Array Based Systems: Design & Applications. Proceedings of the 5th International Symposium on Olfaction and the Electronic Nose. Technomic Publishing Company, Inc. PA.

Madden, J. M. 1992. Microbial pathogens in fresh produce-the regulatory perspective. J. of Food Protection. 55: 821-823.

Malinowski, E.R. 1991. Factor Analysis in Chemistry 2nd ed. John Wiley & Sons, Inc. New York. Pp 350.

Martin, M.A., Santos, J.P., & Agapito, J.A. 2001. Application of artificial neural networks to calculate the partial gas concentration in a mixture. Sensor and Actuators. 77: 468-471.

Maul, F., Sargent, S.A., Balaban, M., Baldwin, E.A., Huber, D.J., and Sims, C.A. 1998. Aroma volatile profiles from ripe tomatoes influenced by physiological maturity at harvest: an application for electronic nose technology. J.Amer.Soc.Hort.Sci. 123(6): 1094-1101.

Mawatari, M., Nagashima, Y., Aoki, T., Hirota, T., & Yamada., M. 1991. Sensory characterization of beer flavor using cluster analysis. American Society of Brewing Chemists. 49(2): 59-64.

Mead, P.S., Sultsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P.M., and Tauxe, R.V. 1999. Food related illness and death in the United States. Emerg. Infect. Dis. 5:607-625.

Mielle, P. 1996. Electronic nose: towards the objective instrumental characterization of food aroma. Trends in Food Science & Technology. 7(12): 432-440.

Mitruka, B.M. 1975. Gas chromatographic methods for the analysis of microorganisms. In Gas chromatographic applications in microbiology and medicine, B.M. Mitruka, A Wiley Biomedical-Health Publication, John Wiley & Sons, New York

Morse, D.L., Pickard, L.K., Guzewich, J.J., Devine, B.D., and Shayegani M. 1990.Garlic-in-oil associated botulism: episode leads to product modification. Am J Public Health. 80:1372-3.

Nakai, S., Wang, A., Dou, J., Nakamura, S., Ogawa, M., Nakai, E., and Vanderstoep, J. 1999. Gas chromatography system for detection of e.coli and s.aureus contaminating salmon and hamburger. J Agric. Food Chem, 47(1) 576-583.

Nakamura, K., Yoshikawa, T. 2001. Evaluation of inside parameters for a rotation spreading associative network. Communications in Japan. 84 (10): 1-14.

National Advisory Committee on Microbial Criteria for Foods. 1999. Microbiological safety evaluations and recommendations on sprouted seeds. USDA/Center for Disease Control and Prevention. CPG 7120.28. October 26.

Nguyen-the, C. and Carlin, F. 1994. The microbiology of minimally processed fresh fruits and vegetables. Critical Reviews in Food Science and Nutrition. 34(4): 371-401.

Nilsson, T., Larsen, T., Montanarella, L., and Madsen, J. 1996. Application of headspace microextraction for the analysis of volatile metabolites emitted by Penicillium species. Journal of Microbiological Methods. 25: 245-255.

Nogueira, J.M.F. and Nascimento, A.M.D. 1999. Analytical characterization of madeira wine. J. Agric. Food Chem. 47: 566-575.

Nomikos, P., and MacGregor, J.F. 1995. Multivariate SPC charts for monitoring batch processes. J. Magnetic Resonance. 37(1).

Nouwen, J., Lindgren, F., Hansen, B., Karcher, W., Verharr, H., & Hermens, J. 1997. Classification of Environmentally occurring chemicals using structural fragments and Discriminant analysis. J. Magnetic Resonance. 31: 2313-2318.

Oja, E. 1982. A simplified neuron model analyzer. Journal of Mathematics. 15:267-273.

Okajima, K. 2001. An Infomax-based learning rule that generates cells similar to visual cortical neurons, Journal of Mathematical Biology. 14 (9): 1173-1180.

Peplow, M.O., Correa-Prisant, M., Stebbins, M.E., Jones, F., and Davis, P. 1999. Sensitivity, specificity, and predictive vales of three salmonella rapid detection kits using fresh and frozen poultry environmental samples versus those of standard plating. Applied and Environmental Microbiology. 65(3): 1055-1060.

Persaud, K.C., and Pelosi, P. 1992. Sensor arrays using conducting polymers for an artificial nose. Pg.238-256. In Sensors and Sensory Systems for an Electronic Nose. Gardner, J.W., and Bartlett, P.N. 1992. Kluwer Academic Publisher, Netherlands.

Pignato, S., Marino, A.M., Emanuele, M.C., Iannotta, V., Caracappa, S., and Giammanco, G. 1996. Evaluation of new culture media for rapid detection and isolatio of salmonellae in foods. Applied and Environmental Microbiology. 61(5): 1996-1999.

Prokopowich, D. and Blank, G. 1991. Microbiological evaluation of vegetable sprouts and seeds. J. of Food Protection. 54: 560-562.

Riemann, H. and Bryan, F.L. 1979. Food-borne Infections and Intoxications (Second editions) Academic press, Inc., New York.

Robacker, D. and Bartelt, R. 1997. Chemicals attractive to mexican fruit fly from Klebsiella pneumoniae and Citrobacter freundii cultures sampled by solid phase microextraction. Journal of Chemical Ecology, 23(12) : 2897-2915.

Ruiz, B.G., Vargas, R.G., and Villanova, R. 1987. Contamination on fresh vegetables during cultivation and marketing. International J. of Food Microbiology. 4: 285-291.

Ruiz, J., Cava, R., Ventanas, J., and Jensen, M. 1998. Headspace solid phase microextraction for the analysis of volatiles in a meat product: dry-cured ham. J.Agric.Food Chem 46(10) 4688-4694.

Sberveglieri, G. et al. 1998. A novel electronic nose based on semiconductor thin films gas sensor to distinguish different heat treatments of milk. Seminar in Food Analysis. 3: 67-76.

Schweizer-Berberich, P, Vaihinger, S, and Gopel, W. 1994. Characterization of food freshness with sensor arrays. Sensor Actuator, 18-19: 282-290.

Sharma, V.J. and Carlson, S.A. 2000. Simultaneous Detection of salmonella strains and Escherichia coli O157:H7 with fluorogenic PCR and single-enrichment-broth culture. Applied and Environmental Microbiology. 66(12): 5472-5476.

Song, J., Fan, L., and Beaudry, R. 1998 Application of solid phase microextraction and gas chromatography/time-of-flight mass spectrometry for rapid analysis of flavor volatiles in tomato and strawberry fruits. Journal of Agriculture and Food Chemistry. 46(8): 3721-26.

Song, J., Gardner, B.D., Holland, J.F, and Beaudry, R.M. 1997. Rapid analysis of volatile flavor compounds in apple fruit using spme and gc/time-of-flight mass spectrometry. J.Agric.Food Chem. 45: 1801-1807.

Spiegelman, C.H., McShane, M., Goetz, M.J., Motamedi, M., Yue, Q., and Cote, G.L. 1998. Theoretical justification of wavelength selection in PLS calibration. Anal Chem. 70:35-44.

Steele, M., McNab, W.B., Read, S., Poppe, C., Harris, L., Lammerding, A.M., and Odumeru, J.A. 1997. Analysis of whole-cell fatty acid profiles of verotoxigenic escherichia coli and salmonella enteritidis with the microbial identification system. Applied and Environmental Microbiology. 63(2): 757-760.

Steffen, A. and Pawliszyn, J. 1996. Analysis of flavor volatiles using headspace solid-phase microextraction. J. Agric. Food Chem. 44(5): 2187-2193.

Stetter, J.P., Strathmann, S., McEntegart, C., Decastro, M., and Penrose, W.R. 2000. New sensor arrays and sampling systems for a modular electronic nose. Sensors and Actuators. 69: 410-419.

Tanaka, T., and Saito, M. 1992. Quantitative properties of adaptive vector quantizers. Trans IEICE. 75: 1085-1092.

Tartakow, I. J. and Vorperian, J.H. 1981. Foodborne and Waterborne Diseases: Their Epidemiologic Characteristics. AVI Publishing Company, Inc., Westport, Connecticut.

Terra, M.H. and Tinós, R. 2001. Fault detection and isolation in robotic manipulators: A comparison among three architectures for residual analysis. Journal of Robotic Systems. 18(7): 357-374.

Tomasko, L., Helms, R.W., Snapinn, S.M. 1999. A discriminant analysis extension to mixed models. Statistics in Medicine. 18: 1249-1260.

Tsuruta, N., Taniguchi, R., and Amamiya, M. 2000. Hypercolumn model: A combination model for image recognition. Computers in Japan. 31(2): 49-61.

Turhan-Sayan, G., Leblebicioglu, K., & Ince, T. 1999. Electromagnetic target classification using time-frequency analysis and networks. Optical Tech. 21(2):63-69.

U.S. Department of Health and Human Services. 1999. Consumers Advised of Risks Associated with Raw Sprouts. U. S. Food and Drug Administration Morbidity and Mortality Weekly Report.

U.S. Food and Drug Administration. 1999. Salmonellosis outbreak prompts hydro harvest alfalfa sprout recall. U. S. Food and Drug Administration Morbidity and Mortality Weekly Report.

Van den Berg, W., and Lewis, C. 1988. Some properties of two measures of multivariate association. Psychometrica. 53: 109-122.

Vergnais, L., Masson, F., Montel, M., Berdague, L., and Talon, R. 1998. Evaluation of solid phase microextraction for analysis of volatile metabolites produced by staphylococci. J.Agricl.Food Chem.46(1):228-234.

Warburton, J. 1996. The electronic nose-the technology and application. Food Science and Technology Today. 10(2): 91-93.

Watanabe, K., Haba, T., Kudo, N., and Oohori, T. 2001. Fluctuation-driven learning rule for continuous-time recurrent networks and its application to dynamical system control. Computers in Japan. 32(3): 14-23.

Weaver, J., Saykin, A., Burr, B., Rlordan, I., and Macclender , A. 1994. Principal components analysis of functional MRI of memory. In: Abstracts form the 2nd annual meeting of the Society of Magnetic Resonance, San Francisco. Society of Magnetic Resonance. p808.

Weber, K. & Poling, J. 1996. Study of pure hopes with an electronic nose. The american society of brewing chemists. Chicago, Illinois.

Wentzell, P.D., Andrews, D.T., and Kowalski, B.R. 1997. Maximum likelihood multivariate calibration. J.Chem. 69: 2299-2311.

Winquist, F. 1993. Performance of an electronic nose for quality estimation of ground meat. Measurement Science and Technology. 9:1493-1500.

Zhang, A. and Pawliszyn, J. 1993. Headspace solid-phase microextraction. Anal. Chem. 65: 1843-1852.