

FREQUENCY RESPONSE BASED PERMITTIVITY SENSORS FOR MEASURING
AIR CONTAMINANTS

by

BRENTON R. WARE

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biological and Agricultural Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2012

Approved by:

Major Professor
Dr. Naiqian Zhang

Abstract

Permittivity, displayed when a dielectric material is exposed to an electric field, is a useful property for measuring impurities in a dielectric medium. These impurities often have a dipole moment different from the pure material, and the dipoles align through polarization and impede electric current. By measuring the resulting impedance in a known geometry, the permittivity can be determined.

Four permittivity sensors were utilized to measure contaminants that are associated with biofuels, specifically glycerol, ethanol, and ammonia. These sensors were based around either stainless steel or aluminum plates to ensure durability and reliability. By connecting each of these sensors to a signal generating control box, the gain and phase can be measured at 609 frequencies, from 10 kHz up to 120 MHz.

Data from each of the three contaminants were run through a method for detection. Measurements for ambient air and air with the contaminants were compared with a statistical analysis. Glycerol, ethanol, and ammonia each had significantly different measurements in the gain and phase data at a unique set of frequencies. Using a neural network analysis for detection resulted in a 95.8%, 93.9%, and 97.1% success rate for detecting glycerol, ethanol, and ammonia, respectively.

For ethanol and ammonia, where multiple concentrations were measured, regression methods were used to relate the frequency response data to the contaminant concentration. Stepwise regression, wavelet transformation followed by stepwise regression, partial least squares regression, and neural network regression were the four methods used to establish these relationships. Several regressions over-fit the data, showing coefficient of determination (R^2) values of 1.000 for training data, yet very low R^2 values for validation data. However, the best R^2 values of all the regressions were 1.000 and 0.996 for the training and validation data, respectively, from measuring ammonia.

Table of Contents

List of Figures	vii
List of Tables	ix
Acknowledgements.....	x
Dedication.....	xii
CHAPTER 1 - Introduction	1
CHAPTER 2 - Research Objectives	3
CHAPTER 3 - Review of Literature.....	4
3.1 Need for Research.....	4
3.1.1 Glycerol.....	4
3.1.2 Ethanol	4
3.1.3 Ammonia.....	5
3.2 Air Pollution Sensors	5
3.2.1 Sensors for Particulate Contaminants	5
3.2.2 Sensors for Gaseous Contaminants.....	6
3.3 Theory of Permittivity	8
3.4 Existing Permittivity Sensors.....	11
3.4.1 Water Permittivity Sensors	11
3.4.2 Biofuel Permittivity Sensors	11
3.4.3 Grain Permittivity Sensors.....	12
3.4.4 Soil Permittivity Sensors.....	12
3.5 Data Analysis and Statistical Methods	12
3.5.1 Stepwise Regression	14
3.5.2 Wavelet Transform	14
3.5.3 Partial Least Squares Regression	16
3.5.4 Neural Networks	17
CHAPTER 4 - Materials and Methods	19
4.1 Sensors	19
4.1.1 Probe Sensors.....	19
4.1.2 Box Sensors	21

4.2 Testing Chamber.....	24
4.3 Control Box.....	25
4.3.1 System Hardware.....	25
4.3.2 System Software.....	28
4.4 Sensitivity Tests.....	28
4.5 Experimental Procedure.....	29
4.5.1 Sample Preparation.....	29
4.5.2 Trial Procedure.....	32
4.5.3 Data Acquisition.....	34
4.6 Data Analysis.....	34
4.6.1 Qualitative Detection of Contaminant Presence in Air.....	34
4.6.2 Quantitative Detection of Contaminant Presence in Air.....	35
4.6.2.1 Stepwise Regression.....	37
4.6.2.2 Wavelet Transform.....	38
4.6.2.3 Partial Least Squares Regression.....	39
4.6.2.4 Neural Networks.....	40
CHAPTER 5 - Experimental Results and Discussion.....	43
5.1 Sensitivity Tests.....	43
5.1.1 Temperature.....	43
5.1.2 Testing Surface.....	44
5.1.3 Sensor Height.....	45
5.1.4 Sensor Radial Distance.....	46
5.1.5 Presence of a Thermometer.....	48
5.1.6 Cable Position.....	49
5.2 Glycerol Tests.....	50
5.3 Ethanol Tests.....	52
5.3.1 Qualitative Detection of Ethanol Presence in Air.....	52
5.3.2 Quantitative Detection of Ethanol Presence in Air.....	55
5.3.2.1 Small Probe Sensor.....	56
5.3.2.2 Large Probe Sensor.....	58
5.3.2.3 Small Box Sensor.....	60

5.3.2.4 Large Box Sensor.....	62
5.4 Ammonia Tests	65
5.4.1 Qualitative Detection of Ammonia Presence in Air	65
5.4.2 Quantitative Detection of Ammonia Presence in Air	67
5.4.2.1 Small Probe Sensor.....	68
5.4.2.2 Large Probe Sensor.....	69
5.4.2.3 Small Box Sensor.....	73
5.4.2.4 Large Box Sensor.....	75
5.5 Combined Tests	79
5.5.1 Recognizing Presence of a Contaminant	79
5.5.2 Identifying the Contaminant Present.....	80
CHAPTER 6 - Conclusions	82
6.1 Review of Research Objectives	82
6.1.1 Construction of Sensors	82
6.1.2 Investigate Sensor Sensitivity	82
6.1.3 Code for Analyzing Differences	82
6.1.4 Identifying Frequencies with Significantly Different Measurements.....	82
6.1.5 Code for Regression Analysis.....	82
6.1.6 Calculating R^2 and RMSE Values	83
6.2 Summary of Results.....	83
6.2.1 Glycerol Tests	83
6.2.2 Ethanol Tests.....	83
6.2.3 Ammonia Tests	84
6.2.4 Combined Tests	84
CHAPTER 7 - Recommendations for Future Research	85
References.....	86
Appendix A - Design of Sensors	90
Appendix B - Code for Sensor Sensitivity Tests	91
Appendix C - Code for Detection Tests.....	94
Appendix D - Neural Network Detection Program	96
Appendix E - Stepwise Regression Program	98

Appendix F - Wavelet Regression Program	101
Appendix G - Partial Least Squares Regression Program	104
Appendix H - Neural Network Regression Program	107
Appendix I - Supplementary Code	110
Appendix J - Significant Frequencies Found with Detection Method.....	115
Appendix K - Frequencies Selected in Stepwise Regression for Ethanol (Gain).....	131
Appendix L - Frequencies Selected in Stepwise Regression for Ethanol (Phase).....	132
Appendix M - Frequencies Selected in Stepwise Regression for Ammonia (Gain).....	133
Appendix N - Frequencies Selected in Stepwise Regression for Ammonia (Phase).....	134
Appendix O - Results with Neural Network Combined Test with Small Probe Sensor.....	135
Appendix P - Results with Neural Network Combined Test with Large Probe Sensor	137
Appendix Q - Results with Neural Network Combined Test with Small Box Sensor	139
Appendix R - Results with Neural Network Combined Test with Large Box Sensor	141

List of Figures

Figure 3.1: ZN-PD air particle sensor (Omron Corporation, 2008)	6
Figure 3.2: Ammonia sensor (Figaro USA, 2005).....	6
Figure 3.3: Volatile organic compound sensor (Figaro USA, 2005).....	7
Figure 3.4: Carbon dioxide sensor (Futurlec, 2004).....	8
Figure 3.5: GasFinder laser (Boreal Laser, 2009).....	8
Figure 3.6: Molecules with no electric field	9
Figure 3.7: Molecules with an external electric field.....	9
Figure 3.8: (a) The Haar wavelet, (b) a Daubechies wavelet, (c) a Biorthogonal wavelet, (d) a Coiflet wavelet, (e) a Symlet wavelet, adapted from Misiti et al. (2006).....	15
Figure 3.9: Wavelet decomposition, adapted from Misiti et al. (2006).....	16
Figure 3.10: Example of a neural network layer, adapted from Demuth and Beale (2004).....	18
Figure 4.1: Small probe sensor	20
Figure 4.2: Frequency response data for the small probe in ambient air	20
Figure 4.3: Large probe sensor	21
Figure 4.4: Frequency response data for the large probe in ambient air.....	21
Figure 4.5: Small box sensor (Shultz, 2009).....	22
Figure 4.6: Frequency response data for the small box in ambient air	23
Figure 4.7: Large box sensor (Shultz, 2009).....	23
Figure 4.8: Frequency response data for the large box in ambient air.....	24
Figure 4.9: Testing chamber	25
Figure 4.10: Block diagram of the control box, adapted from Tang (2009).....	26
Figure 4.11: Transfer function for gain and phase measurements.....	26
Figure 4.12: Ideal transfer gain and phase characteristics of AD8302 (Analog Devices, Inc., 2002)	27
Figure 4.13: Control box with power supply	28
Figure 4.14: Setup for sensitivity tests.....	29
Figure 4.15: Gain data for determining equilibrium time.....	33
Figure 4.16: Phase data for determining equilibrium time	33
Figure 4.17: Flowchart of detection program	35

Figure 4.18: Flowchart of stepwise regression program.....	38
Figure 4.19: Flowchart of wavelet regression program.....	39
Figure 4.20: Flowchart of PLS regression program.....	40
Figure 4.21: Flowchart of neural network regression program	42
Figure 5.1: Gain data for various temperatures in distilled water.....	43
Figure 5.2: Phase data for various temperatures in distilled water	44
Figure 5.3: Gain data on various table surfaces in distilled water	44
Figure 5.4: Phase data on various table surfaces in distilled water	45
Figure 5.5: Gain data for various sensor heights in distilled water	46
Figure 5.6: Phase data for various sensor heights in distilled water.....	46
Figure 5.7: Gain data for various sensor radial positions in distilled water	47
Figure 5.8: Phase data for various sensor radial positions in distilled water.....	47
Figure 5.9: Gain data for various thermometer positions in distilled water	48
Figure 5.10: Phase data for various thermometer positions in distilled water.....	48
Figure 5.11: Gain data for various cable positions in distilled water	49
Figure 5.12: Phase data for various cable positions in distilled water.....	49
Figure 5.13: Distribution of significant frequencies with glycerol, gain data	50
Figure 5.14: Distribution of significant frequencies with glycerol, phase data	51
Figure 5.15: Distribution of significant frequencies with ethanol, gain data.....	53
Figure 5.16: Distribution of significant frequencies with ethanol, phase data	54
Figure 5.17: Distribution of significant frequencies with ammonia, gain data	65
Figure 5.18: Distribution of significant frequencies with ammonia, phase data	66
Figure 5.19: Comparison of actual and predicted ammonia concentrations using wavelet transformation followed by stepwise regression on data from large probe sensor.....	72
Figure 5.20: Comparison of actual and predicted ammonia concentrations with PLS regression on data from large probe sensor.....	72
Figure A.1: Plate for the small probe.....	90
Figure A.2: Plate for the large probe	90

List of Tables

Table 4.1: Concentrations of ethanol used.....	31
Table 4.2: Concentrations of ammonia used.....	32
Table 4.3: Frequency ranges used in regression analysis	36
Table 5.1: Success rates of neural network for detecting glycerol	52
Table 5.2: Success rates of neural network for detecting ethanol.....	55
Table 5.3: R^2 values for ethanol with small probe sensor.....	57
Table 5.4: RMSE values for ethanol with small probe sensor.....	58
Table 5.5: R^2 values for ethanol with large probe sensor	59
Table 5.6: RMSE values for ethanol with large probe sensor	60
Table 5.7: R^2 values for ethanol with small box sensor.....	61
Table 5.8: RMSE values for ethanol with small box sensor.....	62
Table 5.9: R^2 values for ethanol with large box sensor	63
Table 5.10: RMSE values for ethanol with large box sensor	64
Table 5.11: Success rates of neural network for detecting ammonia	67
Table 5.12: R^2 values for ammonia with small probe sensor	68
Table 5.13: RMSE values for ammonia with small probe sensor.....	69
Table 5.14: R^2 values for ammonia with large probe sensor	70
Table 5.15: RMSE values for ammonia with large probe sensor	71
Table 5.16: R^2 values for ammonia with small box sensor	74
Table 5.17: RMSE values for ammonia with small box sensor.....	75
Table 5.18: R^2 values for ammonia with large box sensor	77
Table 5.19: RMSE values for ammonia with large box sensor	78
Table 5.20: Success rates of neural network for detecting contaminants with pooled data	79
Table 5.21: Success rates of neural network for identifying contaminants with pooled data	81

Acknowledgements

First, I would like to acknowledge and thank my major professor, Dr. Naiqian Zhang, for everything that he has done for me throughout my time as a Master's student. The opportunities that he has provided me have given me the best possible preparation for my future work. His high expectations have encouraged me to continuously improve and pursue excellence. I also appreciate his leadership in advising the six-time defending national champion KSU Robotics Team and placing so much emphasis on his philosophy, "We're not going there to compete. We're going there to win."

I also value the help and time of my committee members, Dr. Ronaldo Maghirang and Dr. Abigail Jager. Dr. Maghirang and Ms. Edna Razote have given me so much guidance in the background of air pollution and planning the tests. When it came to the statistical analysis, Dr. Jager was always extremely helpful in providing direction and debugging all the aspects of the statistical software.

Similarly, I have greatly appreciated the advice and support of the several advisors I have sought help from during my entire time at Kansas State University. Many thanks go to Dr. Donghai Wang (my undergraduate advisor), Dr. Gary Clark, Mr. Ed Brokesh, Dr. Christine Aikens, Dr. Chris Culberston, Mrs. Earline Dikeman, Dr. Keith Hohn, and Dr. Larry Erickson for sharing all of their wisdom and guidance with me.

While at the Department of Biological and Agricultural Engineering, I have appreciated the help that the staff, including Mr. Darrell Oard, Mrs. Barb Moore, and Mrs. Lou Ann Claassen, has provided me. As an undergraduate, I was also fortunate to work with Ms. Leidy Peña-Duque on her research of acid-functionalized nanoparticles.

I have also appreciated the friendship and help of the visiting scholars and fellow students in the Instrumentation and Control Laboratory, past and present, including Dr. Mandy Zhang, Dr. Yong Wei, Dr. Peng Li, Ms. Ning Tang, Dr. Tony Han, Dr. Joe Dvorak, Mr. Kevin Wang, Mr. Dan Bigham, Mr. Spencer Kepley, Mr. Jed Barker, and Mr. Marvin Petingco.

Finally, I would like to show my gratitude to my entire family and my good friends. My parents, Brad and Nancy Ware, have been so important to everything I have done throughout my entire life, and I cannot thank them enough for what they have done and continue to do for me. My brother, grandparents, and uncle have also given me so much encouragement and support to complete my studies. Lastly, my friends, especially DeeAnn Turpin, Erik Hellmer, Holly Kier, and James Meyer, are always there for me whenever I have needed help or just some cheering up.

Dedication

This work is dedicated to my grandfather, Colonel Richard M. Ware (December 12, 1919 – September 2, 2011). He was and continues to be a great inspiration in everything I do. Every day, I not only miss my grandfather, but a true friend and mentor. May the world never forget his energetic spirit, optimistic outlook, everlasting wisdom, and entertaining humor.

CHAPTER 1 - Introduction

With biofuels becoming much more common in transportation, the question arises about the air quality associated with both manufacturing and combusting these renewable fuels. Several contaminants, including glycerol, ethanol, and ammonia, are associated with these processes and should be quantified for compliance with air quality regulations and as a means of process quality control. Glycerol is a direct by-product of biodiesel synthesis; ethanol is the most common biofuel; ammonia is a by-product of anaerobic digestion, one of many methods of biofuel synthesis with any of several feedstocks. These three chemicals are volatile or semi-volatile liquids, so a certain volume of the chemical in the liquid state will produce a known concentration in a known volume of air upon evaporation.

Permittivity, the ability for a dielectric material to sustain an electric field, can be a useful property for measuring impurities within a dielectric medium. Often, these impurities have a dipole moment different from that of the pure medium; these dipole moments, when exposed to an external electric field, align through polarization and impede electric current. By measuring the resulting impedance in a known geometry, the permittivity can be determined. Historically, permittivity sensors have measured the salinity and moisture content of soils, quality of grains, and impurities in water and biofuels. Other major applications of these sensors are in chemical analyses and product quality control.

Factors such as frequency of the applied electric field and temperature at which the measurement is conducted have a great impact on permittivity measurements. As such, measurements need to be taken while considering several levels of these factors, especially frequency. Likewise, many properties of the sensors themselves can affect the permittivity measurements. With the sensor construction, the size and materials of the plates, the gap between plates, and the uniformity of the plates clearly have effects on the frequency response data (Tang, 2009; Shultz, 2009). External factors, such as the cable positioning or surrounding materials, may also have effects on the measurements.

This project utilized four different permittivity sensors to investigate their ability to measure the concentration of different air contaminants that are related to the manufacturing and combustion of biofuels. Two sensors were probe-style sensors constructed from stainless steel plates while the other two were box-style sensors made of aluminum plates. These sensors were connected, via coaxial cable, to a control box that recorded gain and phase data at 609 frequencies and transmitted the data through a serial cable.

After acquiring impedance spectra in the frequency domain, the contaminant concentrations were compared with multiple statistical methods. A set of programs was written for distinguishing between ambient air and air with contaminants on a qualitative basis. This program identified which frequencies had associated measurements significantly different between ambient air and air with contaminants present. Also, another program was written to use neural network analysis for detecting contaminants on a qualitative basis. Furthermore, programs were also written to perform regression analyses to determine varying concentrations. These methods include stepwise regression, wavelet transformation followed by stepwise regression, partial least squares regression, and neural network regression. Coefficient of determination (R^2) and root mean square error (RMSE) parameters were calculated to compare the utility of the various regression analysis tools.

CHAPTER 2 - Research Objectives

The overall objective of this work was to investigate the feasibility of using permittivity sensors for detecting and measuring air contaminants. These contaminants—glycerol, ethanol, and ammonia—were chosen because they are associated with the manufacturing and combustion of biofuels.

The specific objectives were as follows:

1. Design and construct probe-style permittivity sensors that are relatively inexpensive, rugged, and accurate.
2. Investigate the specific factors that affect sensor readings.
3. Write and implement SAS[®] and MATLAB[®] programs that analyze the differences in the frequency response data.
4. Identify the set of frequencies that had significantly different measurements between ambient air and air with contaminants.
5. Write and implement MATLAB[®] programs that perform regression analysis for ethanol and ammonia data, contaminants that were tested at multiple concentrations.
6. Calculate and compare R^2 and RMSE values for each regression analysis based on the different contaminants, sensors, datasets, and regression methods used.

CHAPTER 3 - Review of Literature

3.1 Need for Research

As the global petroleum reserves continue to approach depletion, an increased effort to manufacture and utilize biofuels has become a major topic of interest. According to the National Biodiesel Board (2007a), biodiesel production in the United States has continuously increased from 75 million gallons in 2005 to 250 million gallons and 450 million gallons in 2006 and 2007, respectively. Additionally, the 2010 biodiesel production in the European Union was greater than 9.5 million metric tons, or 2.9 billion gallons (European Biodiesel Board, 2012).

With traditional isooctane, the major component of gasoline, the major combustion product is carbon dioxide. However, the manufacture and combustion of biofuels may yield a variety of byproducts, including glycerol and ammonia.

3.1.1 Glycerol

In the biodiesel manufacturing process, triglycerides from vegetable oil or animal fat are reacted with alcohol through a process called esterification. Glycerol is the primary by-product of this process (Peterson et al., 2002). Peterson et al. (1997) analyzed the effect of glycerol on the performance of various types of biodiesel. Among the nine different types of biodiesel tested, the average composition of the fuels contained 1.1% glycerol by mass. Each had a significantly lower heat of combustion than traditional diesel, which has no glycerol. However, there was no correlation between the glycerol content and heat of combustion. Nonetheless, elevated levels of glycerol can plug storage tanks and fuel systems while deteriorating engine performance (National Biodiesel Board, 2007b).

3.1.2 Ethanol

Ethanol, the most well-known type of biofuel, is commonly produced through a fermentation process, whether aerobic or anaerobic. This allows adaptability in the manufacturing process so that a variety of feedstock and microorganisms can be used (McElroy, 2009). As a fuel source, ethanol is typically blended with traditional gasoline in a stated mixture; for example, E85 is 85% ethanol and 15% isooctane.

Vapors from ethanol, especially as a by-product of E85 fuel, is known to cause irritations to the eyes, nose, and throat (Clean Air Trust, 2008). Furthermore, exposure to these vapors can lead to lessened respiration, unconsciousness, coma, and death. The International Labour Organization sets a threshold concentration of 5,000 ppm (Stellman, 1998).

3.1.3 Ammonia

Ammonia is well-known to have detrimental effects to both human health and the environment. The U.S. Occupational Safety and Health Administration (2003) has set a permissible exposure limit of 50 ppm because of the potential respiratory effects and eye irritation. Furthermore, the immediately dangerous to life or health (IDLH) concentration is 300 ppm (U.S. OSHA, 2003). In the environment, ammonia is toxic to many species, especially aquatic animals; for instance, fish can become lethargic and die from excessive ammonia concentrations (Hargreaves and Tucker, 2004).

Harper et al. (2010) studied the effect of biofuel production on the emissions of ammonia. The biofuel farm under study was meant to capture biogas from swine manure and convert the biogas into biodiesel. The results of the study showed that the ammonia emissions from the biofuel manufacturing process were increased over a control farm by 38% and 48% during the summer and winter, respectively.

3.2 Air Pollution Sensors

Several types of sensors are used today for air quality measurements, and most of these sensors are meant to either measure particulate or non-particulate contaminants.

3.2.1 Sensors for Particulate Contaminants

With an emphasis on PM_{2.5} and PM₁₀, sensors constructed for measuring particulate contaminants require a sample be taken. For example, the Anderson RAAS sampler has a shielded inlet that leads to a cyclone for separating particles with an equivalent aerodynamic diameter of 2.5 microns or greater (BGI, 2003). The remaining particles are then collected on a filter, which is later weighed and converted into a concentration. However, this process of gravimetric analysis is complicated and requires a steady, precise flow rate (BGI, 2003).

Other types of particulate sensors use light scattering for measuring particle concentrations. As illustrated in Figure 3.1, one example of this is the Omron ZN-PD sensor, which uses a laser beam to scatter the particles (Omron Corporation, 2008). A light sensor near the laser beam measures the intensity of scattered light, which is correlated with the particulate concentration.



Figure 3.1: ZN-PD air particle sensor (Omron Corporation, 2008)

3.2.2 Sensors for Gaseous Contaminants

Sensors often measure the changes in conductive properties when gases are adsorbed into the metal oxide core of a sensor. Figaro USA (2005) manufactures the sensor highlighted in Figure 3.2, which can indicate the ammonia concentration by measuring the variance in the resistance of the alumina core. Similarly, Figure 3.3 shows a sensor for measuring the concentration of volatile organic compounds, such as ethanol, benzene, and *n*-hexane, which works by measuring the change in resistance of tin dioxide. The major drawback to these sensors is their sensitivity at low concentrations.

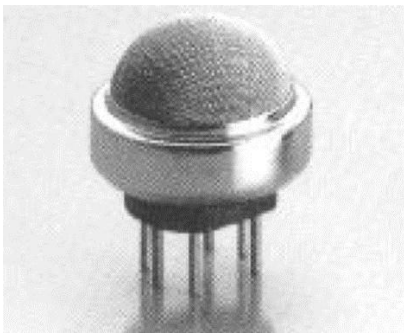


Figure 3.2: Ammonia sensor (Figaro USA, 2005)



Figure 3.3: Volatile organic compound sensor (Figaro USA, 2005)

Based on the technical datasheet (Figaro USA, 2005), the resistance of these sensors can be expressed by:

$$R_s = A \cdot [C]^\alpha \quad (\text{Equation 3.1})$$

where R_s = sensor resistance (Ω),

C = contaminant concentration (mol L^{-1}), and

A and α = constants that depend on the sensor, contaminant, temperature, and specific range of concentration (Ω and dimensionless, respectively).

Similar types of compact and versatile sensors, through a redox reaction specific to a target gas, produce a small voltage that is dependent on contaminant concentration. As displayed in Figure 3.4, a specific example of this is the MG811 carbon dioxide sensor from Futurlec (2004). The sensor utilizes the overall redox reaction $\text{Na}_2\text{O} + 2 \text{Li}^+ + \text{CO}_2 \rightleftharpoons \text{Li}_2\text{CO}_3 + 2 \text{Na}^+$ to produce the CO_2 -dependent electromotive force. As the CO_2 concentration increases, equilibrium shifts to the products and produces lesser voltage. The sensor voltage output is approximately 325 mV with 400 ppm CO_2 and 265 mV with 10000 ppm CO_2 ; overall, the output voltage is related to the partial pressure of carbon dioxide by the Nernst equation (Futurlec, 2004):

$$E_{cell} = E_{cell}^0 - \frac{RT}{2F} \cdot \ln P_{\text{CO}_2} \quad (\text{Equation 3.2})$$

where E_{cell} = produced voltage at actual conditions (V),

E_{cell}^0 = produced voltage at standard conditions (V),

$R = 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$, the ideal gas constant,

T = absolute temperature (K),

$F = 96,485 \text{ J V}^{-1} \text{ mol}^{-1}$, the Faraday constant, and
 P_{CO_2} = partial pressure of carbon dioxide (atm).



Figure 3.4: Carbon dioxide sensor (Futurlec, 2004)

Another major type of sensor for gaseous contaminants is a laser sensor, like the GasFinder from Boreal Laser (2009). Such a laser, as depicted in Figure 3.5, can measure several gases, including carbon dioxide, ammonia, methane, hydrogen sulfide, and hydrogen fluoride. Harper et al. (2010) used one of these lasers for measuring the ammonia and methane emissions downwind of a covered anaerobic digester in their study mentioned above.



Figure 3.5: GasFinder laser (Boreal Laser, 2009)

3.3 Theory of Permittivity

All types of dielectric molecules have some type of dipole, or sections with a partial positive and partial negative charge. These dipoles may be permanent due to unbalanced electronegative attractions or induced based on the London dispersion forces. In either case, this phenomenon is the chemical basis for studying permittivity (Lee et al., 2007).

Without an external electrical field present, the molecules will assume a random orientation to where there is no specific orientation for the dipoles. However, when an electric field is introduced, these molecules will line up—through a process called polarization—such that their dipoles maintain a common orientation. Figure 3.6 and Figure 3.7 illustrate the effect of an electric field. These molecules all have dipoles indicated by a red portion (partial positive charge) and a black portion (partial negative charge).

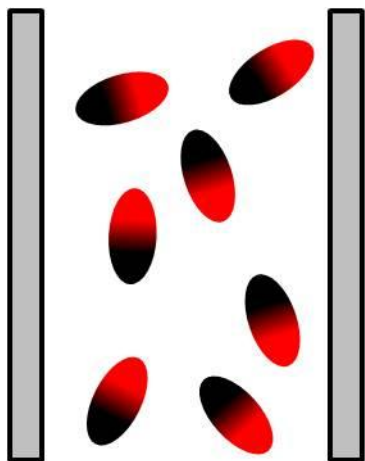


Figure 3.6: Molecules with no electric field

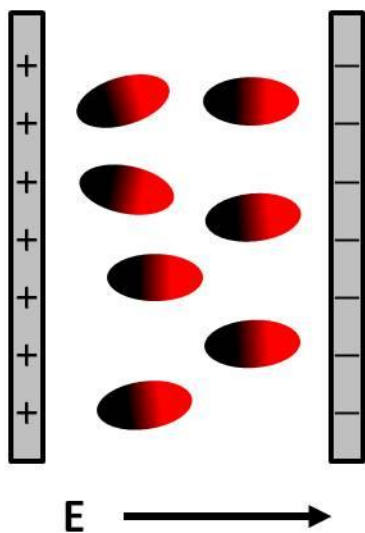


Figure 3.7: Molecules with an external electric field

As seen in Figure 3.7, the electric field causes the plate on the left to form a positive charge and the plate on the right to form a negative charge. Similarly, the molecules form an orientation to where the negative end of the dipole faces the positively charged plate and the positive end of the dipole faces the negatively charged plate.

Permittivity is highly dependent on the frequency of the applied electric field, and that dependence is the basis for several types of permittivity measurements. At relatively low frequencies, the alignment of dipoles causes a separation of charges, which stores energy and contributes to the real component of complex permittivity (Robinson et al., 2003). Furthermore, relatively high frequencies do not allow the molecules enough time to polarize, so there is not a sufficient separation of charges to store energy.

Based on the notation utilized by Nelson (1991) and Scholte et al. (2002), the following terms and relations will be utilized for the various components of permittivity:

$$\varepsilon_r = \varepsilon'_r - j \cdot \varepsilon''_r \quad (\text{Equation 3.3})$$

$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_0} \quad (\text{Equation 3.4})$$

$$\varepsilon''_r = \varepsilon''_d + \frac{\sigma_0}{\omega \cdot \varepsilon_0} \quad (\text{Equation 3.5})$$

where ε_r = relative complex permittivity (F m^{-1}),

ε'_r = real component of complex permittivity (F m^{-1}),

ε''_r = imaginary component of complex permittivity (F m^{-1}),

ε''_d = dielectric relaxation (F m^{-1}),

ε = complex permittivity (F m^{-1}),

$\varepsilon_0 = 8.8541878 \times 10^{-12} \text{ F m}^{-1}$, the permittivity of free space, and

$j = \sqrt{-1}$, the complex unit.

The real component of complex permittivity, ε'_r , can be described with two components that account for both the polarization of the electrodes and the frequency-dependent permittivity, as described by Scholte et al. (2002):

$$\varepsilon'_r = \varepsilon'_e + \varepsilon'_d \quad (\text{Equation 3.6})$$

where ϵ'_e = permittivity due to the polarization of electrodes (F m^{-1}), and ϵ'_d = frequency-dependent permittivity (F m^{-1}).

Furthermore, Robinson et al. (2003) gives the loss tangent according to:

$$\tan \delta = \frac{\epsilon''_r}{\epsilon'_r} = \frac{\epsilon''_d + \frac{\sigma_0}{\omega \cdot \epsilon_0}}{\epsilon'_r} \quad (\text{Equation 3.7})$$

where $\tan \delta$ = loss tangent (dimensionless),

σ_0 = DC electrical conductivity (S), and

$\omega = 2\pi f$, the angular frequency (rad s^{-1}).

3.4 Existing Permittivity Sensors

Various types of permittivity sensors have been used for decades in measuring the properties of water, biofuels, grains, and soil.

3.4.1 Water Permittivity Sensors

As described by Shultz (2009), permittivity sensors were utilized for measuring common contaminants in water, including nitrate salts and atrazine. Results for the study determined that the concentration and molar mass of the nitrate contaminants could be correlated with a coefficient of determination, R^2 , very close to one. However, the validation tests for atrazine were not nearly as reliable.

Utley et al. (2011) studied the utility of permittivity sensors in measuring suspended sediment concentrations. A simple linear regression was not possible, as the assumption of normally distributed variances was not met for the gain and phase data. Alternatively, using a partial least squares regression, the sediment could be measured at relatively higher concentrations (1000 mg L^{-1} or greater).

3.4.2 Biofuel Permittivity Sensors

Shultz (2009) also investigated how permittivity sensors could be used for measuring contaminants in biofuels. Water, glycerol, and glycerides often deter the performance of

biodiesel and were the focus of the study. Like the tests for nitrate concentration and molar mass, the R^2 values indicated a strong relationship.

3.4.3 Grain Permittivity Sensors

Dielectric properties of grain often give insight into their various properties, including moisture content and quality. Nelson (1991) found that both the real and imaginary components of complex permittivity in hard red winter wheat, yellow-dent field corn, winter barley, and soybeans increased with moisture content. Likewise, the permittivity measurements tended to be positively related with temperature, holding the frequency and moisture content fixed.

3.4.4 Soil Permittivity Sensors

Multiple properties of soil can be measured simultaneously with permittivity. Contrariwise to the permittivity data with grain, the gain of the signal—measured in decibels—tended to decrease with increasing volumetric water content at a broad range of frequencies (Lee et al., 2007). Alternatively, Scholte et al. (2002) studied the response at only 200 MHz. The real component of complex permittivity was nearly constant between 17% and 21% volumetric water content and linearly increased between 21% and 31%. Over the entire range of 17% and 31% moisture contents, the complex component of permittivity increased linearly.

Measuring the salinity at various concentrations was also compared at two specific frequencies by Lee et al. (2007). In almost all combinations of frequency (600 Hz and 200 MHz) and volumetric water content ($0.156 \text{ m}^3 \text{ m}^{-3}$ and $0.243 \text{ m}^3 \text{ m}^{-3}$), the regression analysis for the gain was stronger than that of the phase angle. The highest coefficient of determination was 0.989, showing that permittivity can be a good predictor for soil salinity.

3.5 Data Analysis and Statistical Methods

The process of regression analysis establishes a mathematical model that does the best job of fitting data. During this process, the predictor values (typically the x values) are modeled against the response variables (typically the y values). From this model, the predictor values can be used to forecast the values of the corresponding response variables.

Comparing the effectiveness of different models is frequently done with the coefficient of determination, or R^2 value. In a practical sense, the R^2 value represents the portion of variation in the responses that is accounted for by the model. Ranging between zero and one, the R^2 value can be multiplied by 100% to give a percentage of how much variation is explained. Mendenhall and Sincich (2012) outline the process for calculating the R^2 values in Equation 3.8, Equation 3.9, and Equation 3.10:

$$R^2 = 1 - \frac{SSE}{SS_{yy}} \quad (\text{Equation 3.8})$$

$$SSE = \sum (y_i - \hat{y}_i)^2 \quad (\text{Equation 3.9})$$

$$SS_{yy} = \sum (y_i - \bar{y})^2 \quad (\text{Equation 3.10})$$

where R^2 = coefficient of determination,

SSE = sum of squared errors,

SS_{yy} = sum of squared y deviations,

y_i = actual response,

\hat{y}_i = response predicted from the model, and

\bar{y} = average values of all responses.

A second parameter for comparing the effectiveness of models is the root mean square error, or RMSE. In contrast with the R^2 value, a smaller RMSE value is indicative of a better model. The RMSE value is not bounded between zero and one like the R^2 value is, so there is not a clear threshold for an acceptable value. Instead, the RMSE values are compared between various tests and regression methods. Ott and Longnecker (2001) give this parameter as:

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2}} \quad (\text{Equation 3.11})$$

where $RMSE$ = root mean square error,

y_i = actual response,

\hat{y}_i = response predicted from the model, and

n = number of observations.

3.5.1 Stepwise Regression

When many predictor variables exist, a variable screening method is needed to determine which predictor variables are truly significant. Stepwise regression is an example of such a procedure that begins with an empty model and adds significant predictors one-by-one, checking that all predictors are significant. Mendenhall and Sincich (2012) outline the process as follows:

1. All possible one-variable, linear models are fit to the data. The predictor variable whose corresponding t -value for the coefficient is the largest in magnitude begins the model, provided that it is significant at a specified α_{in} .
2. Each possible multivariable model containing the existing predictors plus one additional predictor from all remaining predictors is fit to the data. Of these remaining predictors, the one whose corresponding t -value has the largest magnitude is added to the model, provided that it is significant at a specified α_{in} .
3. The significance of each predictor is checked with compared with a specified α_{out} . Should a predictor not be significant, it is removed from the model.
4. Steps 2 and 3 are repeated until predictors can be neither added to nor removed from the model.

3.5.2 Wavelet Transform

Similar to a Fourier transform, the wavelet transform is a method of analyzing signals that preserves aspects of both the time and frequency domains. Wavelets must have a zero-average and are usually highly irregular. Scaling and shifting the original wavelets allows for matching the original signal. The simplest wavelet is the Haar wavelet, and several other wavelets are classified into families, which may include the Daubechies, Biorthogonal, Coiflets, and Symlets families (Misiti et al., 2006). Figure 3.8 shows the Haar wavelet plus an example from each of the four mentioned families.

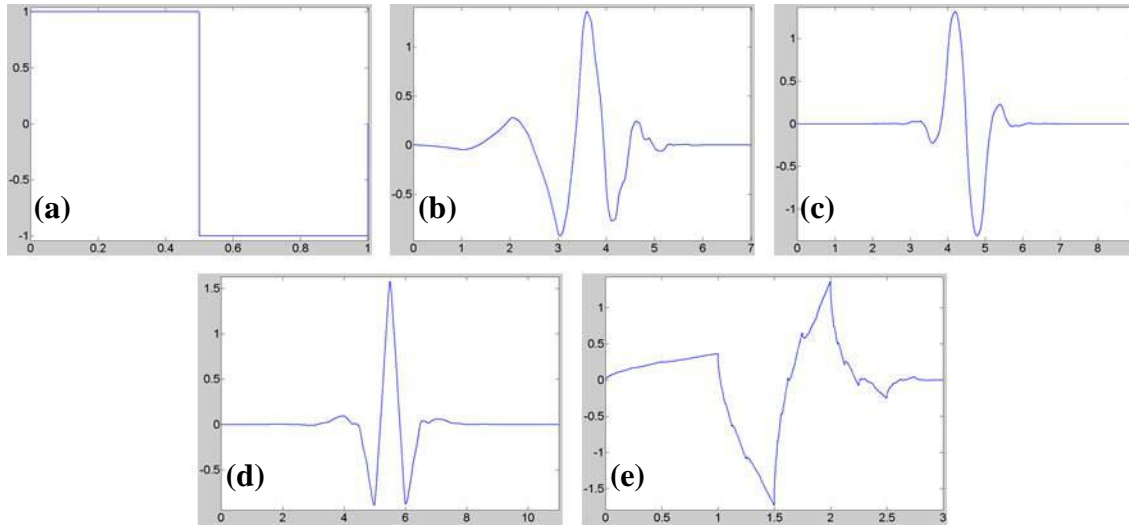


Figure 3.8: (a) The Haar wavelet, (b) a Daubechies wavelet, (c) a Biorthogonal wavelet, (d) a Coiflet wavelet, (e) a Symlet wavelet, adapted from Misiti et al. (2006)

The discrete wavelet transform utilizes a low-pass filter with its complementary high-pass filter to decompose a signal into its approximations (low frequency components) and details (high frequency components) (Misiti et al., 2006). For instance, Figure 3.9 shows an initial sample of 1024 data points. The first level of decomposition provides 1024 low frequency samples and 1024 high frequency samples, which are then reduced to 512 samples each via downsampling (Misiti et al., 2006). In Figure 3.9, these coefficients are identified as “ cA_1 ” and “ cD_1 ” for the first level approximations and details. Subsequent decompositions, which can be iterated repeatedly, take the samples from the approximation components and decompose them into the approximation and detail components of the next level. The subscript following each set of coefficients indicates the level of decomposition.

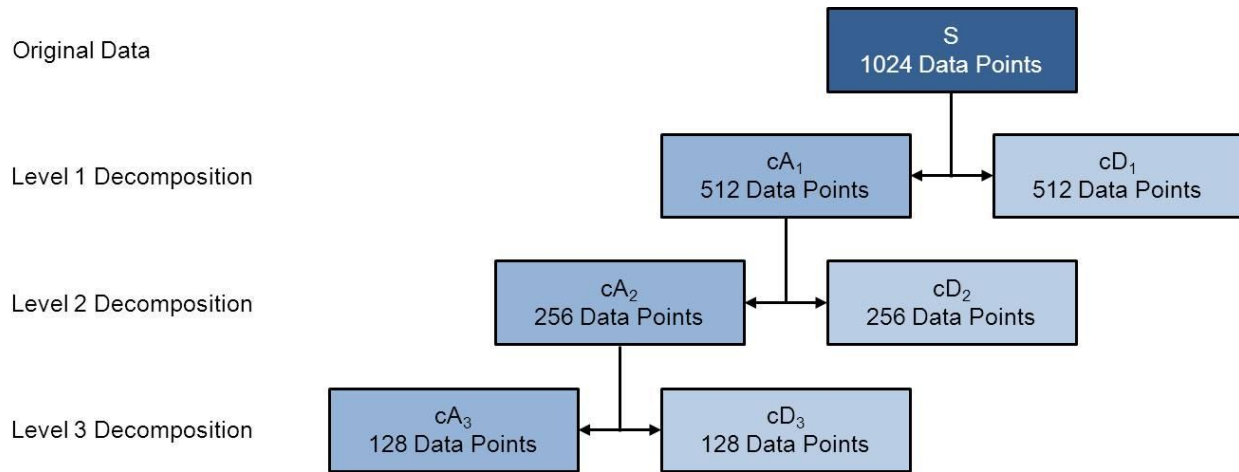


Figure 3.9: Wavelet decomposition, adapted from Misiti et al. (2006)

The approximation and detail coefficients can then be used as the predictor variables, so that a stepwise regression can determine which coefficients are significant when building a model.

3.5.3 Partial Least Squares Regression

Partial least squares regression (PLS) is an ideal method of regression when the number of variables surpasses the number of observations, as in the case of spectral data (Mevik, 2006). Similarly, the PLS procedure can be used where several predictors have multicollinear values, which is also important with spectral data (The MathWorks, 2012c).

Through constructing components—linear combinations of original predictor variables—PLS can achieve four objectives (The MathWorks, 2012c):

1. Ideal fitting of predictor variables to response variables, like in multiple linear regression.
2. Large variance of predictor variables through principal component analysis.
3. Large covariance between response variables and combinations of predictor variables.
4. A minimal number of terms needed to establish an accurate model, like in stepwise regression

Several software packages have algorithms for finding the appropriate matrices for scores and loading of predictor variables and response variables. The SAS Institute (2000) gives the model for PLS regression as:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} \quad (\text{Equation 3.12})$$

$$\mathbf{Y} = \mathbf{UQ}' + \mathbf{F} \quad (\text{Equation 3.13})$$

where \mathbf{X} = matrix of predictor variables,
 \mathbf{Y} = matrix of response variables,
 \mathbf{T} and \mathbf{U} = matrices of \mathbf{X} and \mathbf{Y} scores, respectively,
 \mathbf{P} and \mathbf{Q} = matrices of \mathbf{X} and \mathbf{Y} loadings, respectively, and
 \mathbf{E} and \mathbf{F} = matrices of \mathbf{X} and \mathbf{Y} error terms, respectively.

Another version of this method involves creating regression coefficients, which is utilized by the `plsregress` function in MATLAB[®] (The MathWorks, 2012d). This method uses the SIMPLS algorithm and excludes error terms so that the PLS regression can become:

$$\mathbf{Y} = \mathbf{XB} \quad (\text{Equation 3.14})$$

where \mathbf{X} = matrix of predictor variables,
 \mathbf{Y} = matrix of response variables, and
 \mathbf{B} = matrix of regression coefficients.

3.5.4 Neural Networks

Neural networks provide a powerful tool for data analysis. Simulating a biological neural network, an artificial neural network performs functions in parallel, rather than sequentially (Demuth and Beale, 2004). These networks can be used in a wide variety of applications, including pattern recognition and function fitting, which can be extended to analyzing spectral data.

A neural network can be represented as having an input matrix, a matrix of weights, and an output matrix that is compared with a target matrix. The overall connections are illustrated in Figure 3.10; while using the neural network for prediction, the x -variables represent the input while the y -variables are the target.

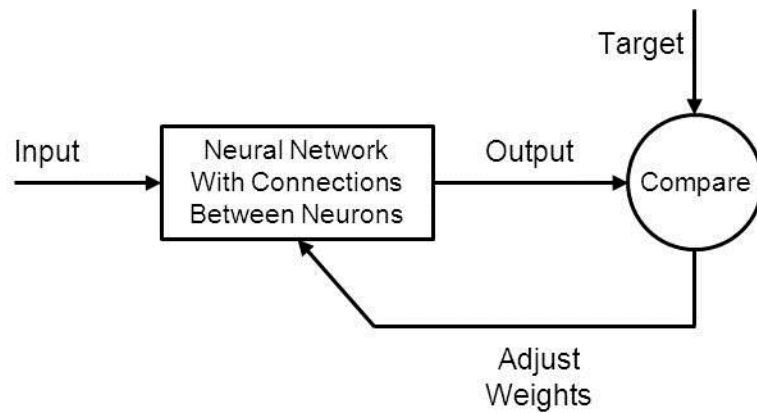


Figure 3.10: Example of a neural network layer, adapted from Demuth and Beale (2004)

Initially, there can be a large discrepancy between the output values and the target values. In a process called supervised training, the weights that connect all the neurons are adjusted to match the output to the target as close as possible. More iterations of supervised training will produce a more accurate result, with the major tradeoff being a longer processing time. Furthermore, this process can either be done with an entire set of inputs (batch training) or with inputs one by one (adaptive training) (Demuth and Beale, 2004).

CHAPTER 4 - Materials and Methods

4.1 Sensors

4.1.1 Probe Sensors

Two probe-style sensors were constructed to measure the frequency response data. As detailed in Appendix A, these two probes were constructed of several components. The most significant part of the sensor were six parallel sections of 0.51 mm thick type 302 stainless steel with an average gap of 2.05 mm between each plate. Stainless steel was chosen for its corrosion resistance and long life span. Alternating plates were electrically connected via 18-8 stainless steel machine screws while adjacent plates were electrically insulated with nylon 6/6 washers. One set of plates was attached to the primary wire of the coaxial cable while the other set was attached to the copper shielding.

The housing of each sensor is a 26.83 cm section of ¼" schedule 80 clear PVC pipe. Furthermore, an acetal through-wall fitting allows for an easy attachment to and removal from the top of the testing chamber. An O-ring on the threaded portion of this fitting gives an airtight seal between the testing chamber and the outside air.

Additionally, the probes both feature 1.6 m sections of both Omega Engineering[®] type-T (copper-constantan) thermocouple wire to record temperature readings and RG-58 50 Ω coaxial cable to transmit the frequency signal and response data.

The cross-sectional areas of the probes are slightly different, so “small probe” and “large probe” will be used to differentiate between the two. Appendix A gives the dimensions of both sensors. Figure 4.1 and Figure 4.3 show the two probe sensors while Figure 4.2 and Figure 4.4 show the frequency response spectra for the small probe and large probe, respectively, in ambient air.



Figure 4.1: Small probe sensor

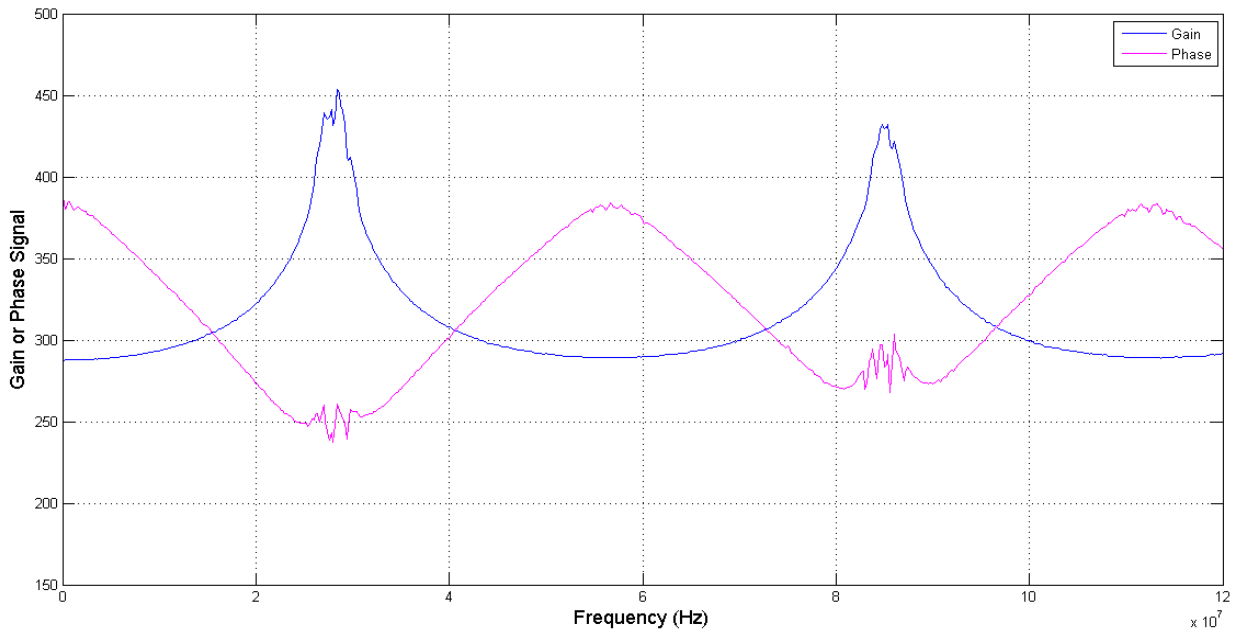


Figure 4.2: Frequency response data for the small probe in ambient air



Figure 4.3: Large probe sensor

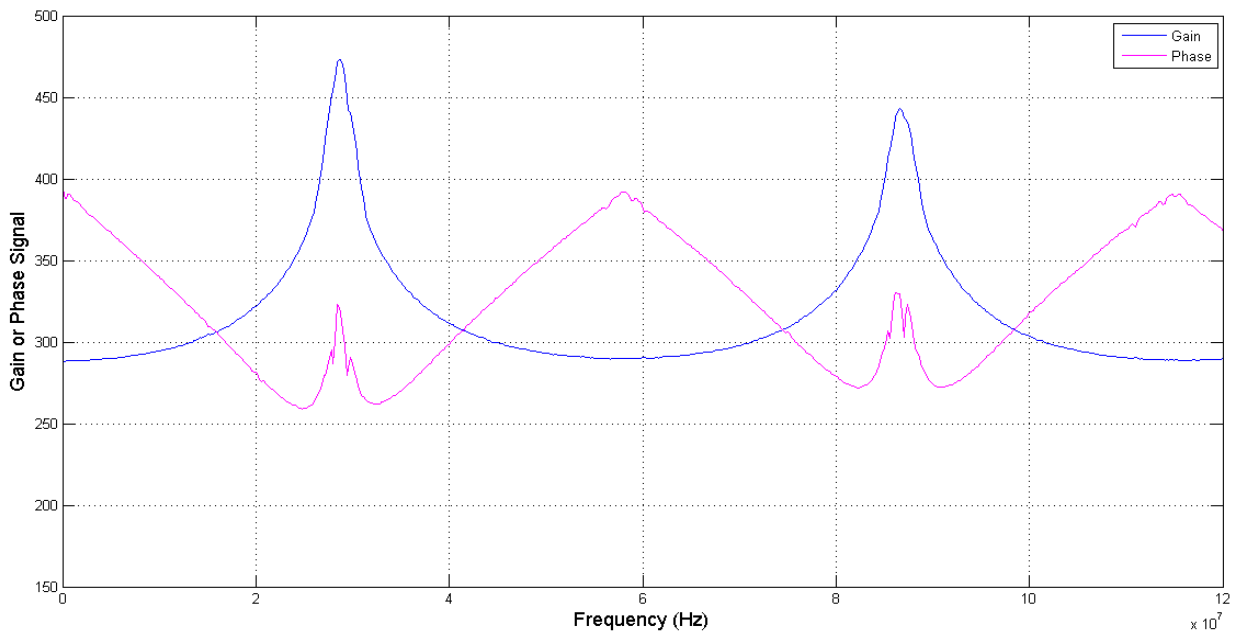


Figure 4.4: Frequency response data for the large probe in ambient air

4.1.2 Box Sensors

In Shultz (2009), two box-style sensors were constructed in a similar manner to the probe sensors. However, the housing for this pair of sensors was a plastic container. Six aluminum

plates, each 4.90 mm thick, were attached to this housing with an average gap of 3.5 mm between each plate. The pin of the BNC connector was electrically attached to alternating plates with 22 gauge wire; the BNC connector shell was connected in a similar manner to the remaining plates.

Like the probe sensors, since the two box sensors have different cross-sectional areas, “small box” and “large box” will distinguish between the two. Each plate on the small box is 5.10 cm wide by 3.45 cm tall, and each plate on the large box is 7.65 cm wide by 3.55 cm tall.

Figure 4.5 and Figure 4.6 show the small box sensor and its corresponding frequency response data in ambient air. Figure 4.7 and Figure 4.8 do the same for the large box sensor.



Figure 4.5: Small box sensor (Shultz, 2009)

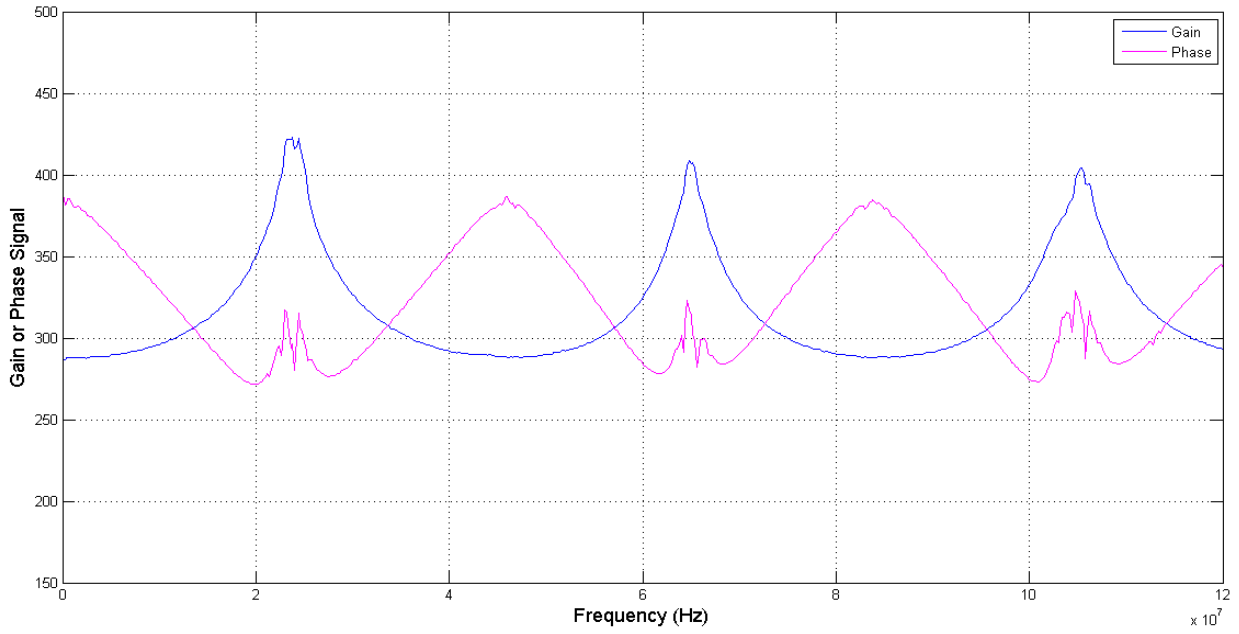


Figure 4.6: Frequency response data for the small box in ambient air



Figure 4.7: Large box sensor (Shultz, 2009)

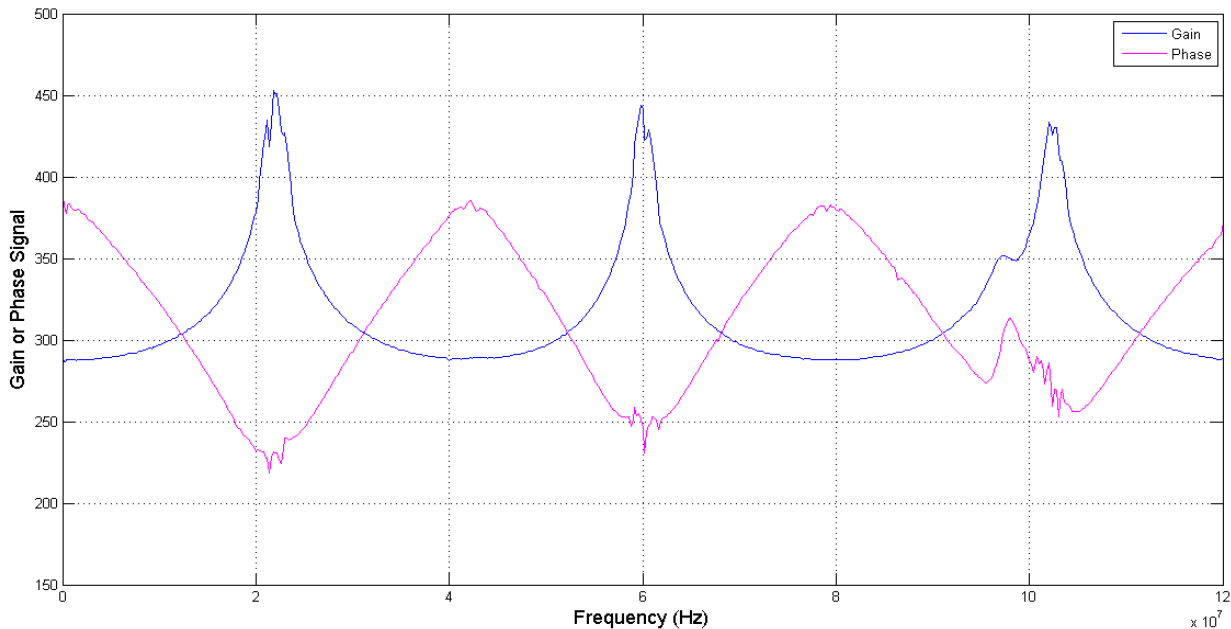


Figure 4.8: Frequency response data for the large box in ambient air

4.2 Testing Chamber

An Aqueon[®] size 10 glass aquarium was the primary component for all tests. The aquarium measures 25.4 cm in length by 50.2 cm in width by 30.2 cm in height. Therefore, the overall volume of the container is 38.5 L. Inside the chamber is an acrylic shelf that will hold the box sensors well away from any of the side walls. Furthermore, a 1.6 m section of coaxial cable, the same type used in the probe sensors, was included with the box to take measurements from the box sensors.

To contain the air under investigation, a rectangular section of acrylic serves as the top of the chamber. This acrylic has holes drilled to accommodate the probe sensors and shelf for the box sensors. All holes are sealed with O-rings to prevent gas transfer. Furthermore, this gas transfer is also prevented with a section of foam tape that lines the circumference of the aquarium.

For the injection of contaminant samples, a 10 mm, 3-layer Hamilton GC septum was embedded into the top of the chamber. While the septum fits snugly into the top of the chamber, a secondary piece of acrylic is glued below the septum to hold it in place.

Furthermore, as described with the sensitivity tests, the temperature inside the testing chamber is an important factor. Thus, two temperature probes from a Traceable[®] High Accuracy Refrigerator Thermometer was installed in the chamber. One probe was placed near the probe sensor and the other was placed near the box sensor shelf. This thermometer gives readings to 0.01°C and was calibrated with instruments traceable to the NIST.



Figure 4.9: Testing chamber

4.3 Control Box

4.3.1 System Hardware

Constructed in previous work, the control box used in this project is needed to both produce sinusoidal electromagnetic waves and read the corresponding gain and phase data (Tang, 2009). This gain and phase data is an alternative to the ϵ_r' and ϵ_r'' terms discussed with the theory of permittivity. An AD9854 signal generator evaluation board was installed to generate the sinusoidal signal and transmit it through the coaxial cable of the sensor. After the signal is returned from the sensor, an AD8302 evaluation board measures the differences in gain

and phase between the signal sent to and received from the sensor. Furthermore, a kitCON-167 microcontroller is in charge of several processes, including instructing the AD9854 to generate the proper frequency, reading the data from the AD8302, and transmitting the data through the RS-232 serial port. Figure 4.10 gives a block diagram of the control box.

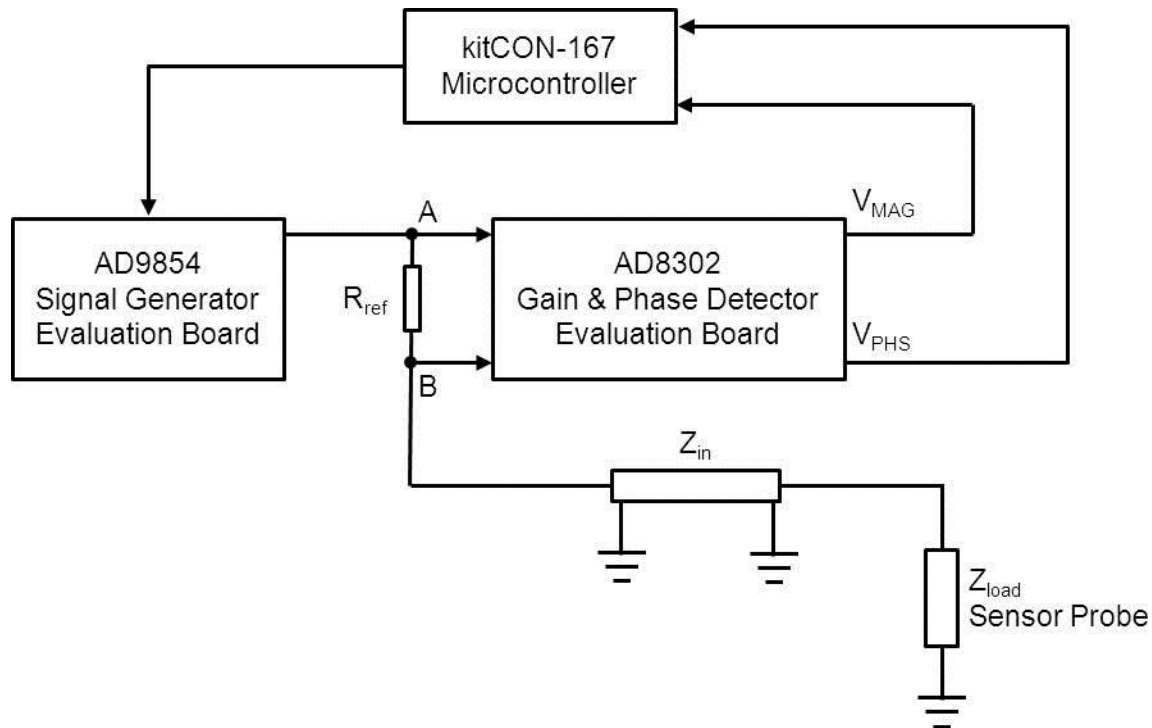


Figure 4.10: Block diagram of the control box, adapted from Tang (2009)

The measurements of gain and phase data are based on a voltage divider circuit. Comparing the sinusoidal signals at points A and B in Figure 4.10 can be represented as the transfer function shown in Figure 4.11. Furthermore, the gain and phase can be calculated from the signals at A and B using Equation 4.1 and Equation 4.2.

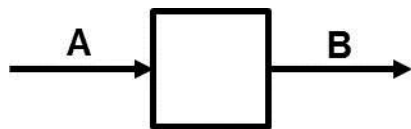


Figure 4.11: Transfer function for gain and phase measurements

$$Gain = 20 \log \left| \frac{\text{Amplitude}_B}{\text{Amplitude}_A} \right| \text{ (dB)} \quad \text{(Equation 4.1)}$$

$$Phase = \arctan \frac{B}{A} \text{ (degrees)} \quad \text{(Equation 4.2)}$$

Based on the datasheet of the AD8302 board (Analog Devices, Inc., 2002), the voltages V_{MAG} and V_{PHS} will be produced depending on the gain and phase measurements, respectively, as shown in Figure 4.12. These two measurements are then transmitted to the kitCON-167 microcontroller, which uses an analog-to-digital converter to convert the voltage signals into discrete readings (Tang, 2009).

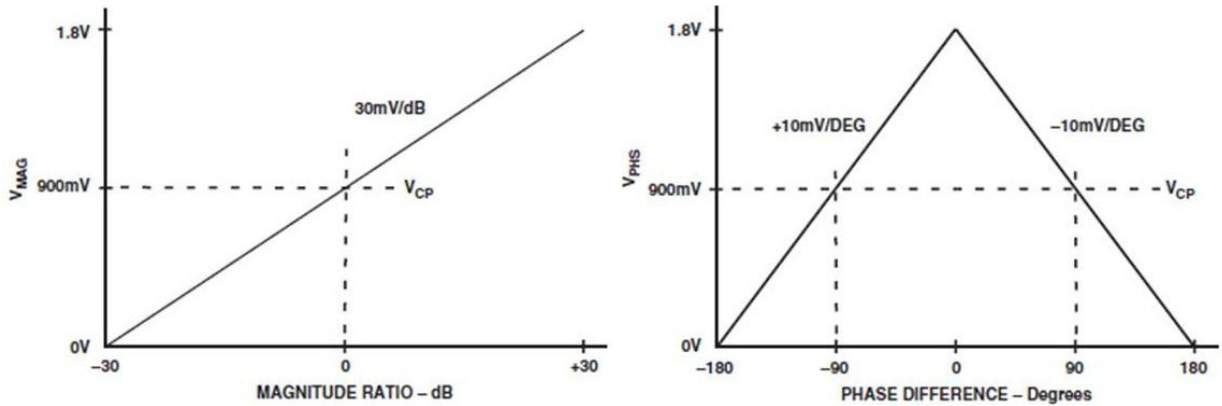


Figure 4.12: Ideal transfer gain and phase characteristics of AD8302 (Analog Devices, Inc., 2002)

Since the system does not have its own power supply, an Elenco[®] Precision XP-580 was used to provide a nominal 9 VDC for the control box. Before supplying power to the control box, a multimeter was used to check that the voltage was truly between 8.5 VDC and 9.0 VDC—enough to power the control box without risking damage to the internal circuitry. As shown in Figure 4.13, banana cables connect the power supply to the control box.



Figure 4.13: Control box with power supply

4.3.2 System Software

To control the system, a C program was compiled in the TASKING[®] Embedded Development Environment and downloaded to the kitCON-167 microcontroller. The program is designed to control the signal generator to produce three iterations of the entire spectrum of frequencies up to 120 MHz. Also, the program sends the data through the RS-232 serial port; each line of data includes a frequency index, the frequency divided by 10,000 Hz, the gain signal, and the phase signal. These lines of text data are easily transmitted through a serial cable and read with a terminal emulation program.

4.4 Sensitivity Tests

With these sensors, a wide variety of factors can corrupt the readings. Before applying the sensors to measure air contaminants, numerous factors were individually studied for their impact on the sensor readings. Temperature is clearly a factor that should be considered, as the dielectric constants of materials tend to be very dependent on temperature, just as many other chemical and physical properties. The boundary effect can also cause problems when the signal from the sensor interferes with the surroundings. As such, the position of the sensor was changed in two dimensions—depth and radial distance. Also, the external surface was changed between a slate tabletop, a metal tabletop, and a ceramic hot plate top. Tests were also run with a metal thermometer to test for any interference problems. Finally, the conformation of the coaxial cable on the sensor was changed to identify any effects.

Distilled water in a 1 L Nalgene[®] plastic container was the medium for testing the large probe sensor. A clamp on a metal ring stand held the sensor in place for the duration of the data acquisition. Figure 4.14 shows an example of the setup for testing the sensitivity of the sensors.

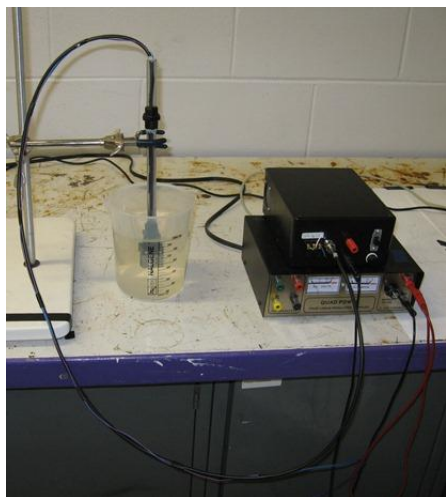


Figure 4.14: Setup for sensitivity tests

After completing the tests for a given factor, the resulting data were processed and plotted using a MATLAB[®] program unique to each factor. As an example, Appendix B gives the program for analyzing the effect of the height of the sensor.

4.5 Experimental Procedure

4.5.1 Sample Preparation

As the viscosity of glycerol is roughly 1580 times that of water ($\mu = 1.412 \text{ Pa s}$), a syringe or micropipette would be insufficient to introduce glycerol samples into the testing chamber. Furthermore, since the vapor pressure of glycerol is very low ($p^* = 0.001316 \text{ atm}$), quantifying the concentration of glycerol is not practical. Instead, the tests were treated as detection only, differentiating between ambient air and air with glycerol vapor present. Introducing 1 mL samples of glycerol (Fisher Scientific[®], Certified ACS, 99.8%) with a disposable pipette ensures that a vapor-liquid equilibrium is established.

Ethanol and ammonia have significantly higher vapor pressures than glycerol, as given by the Antoine Equation (Felder and Rousseau, 2005):

$$\log_{10} p^* = A - \frac{B}{T + C} \quad (\text{Equation 4.3})$$

where p^* = vapor pressure (torr or atm),

T = temperature ($^{\circ}\text{C}$), and

A, B, C = Antoine constants for a specific compound and temperature range.

Solving Equation 4.3 at 25°C gives a vapor pressure for ethanol and ammonia of $p^* = 0.07762$ atm and $p^* = 9.985$ atm, respectively. This means that several amounts of either contaminant can be introduced into a container, and the concentration can be determined. As the tests are run under roughly standard conditions, the ideal gas law gives a very good approximation of the amount of gas in a volume at a given temperature and pressure:

$$p \cdot V = n \cdot R \cdot T = \frac{m \cdot R \cdot T}{MW} \quad (\text{Equation 4.4})$$

where p = absolute pressure (atm),

V = volume of gas (L),

n = number of moles of gas (mol),

$R = 0.08206$ L atm K^{-1} mol^{-1} , the ideal gas constant,

T = absolute temperature (K),

m = mass of gas (g), and

MW = molecular weight of the gas (g mol^{-1}).

To find the mass of 1 ppm ethanol ($MW = 46.07$ g mol^{-1}), Equation 4.4 can be solved for m . Since the container is 38.5 L, 1 ppm occupies 3.85×10^{-5} L. Also, the local conditions were assumed to be 0.996 atm pressure and 299.45 K (26.3°C):

$$\begin{aligned} m &= \frac{p \cdot V \cdot MW}{R \cdot T} = \frac{0.996 \text{ atm} \cdot 3.85 \times 10^{-5} \text{ L} \cdot 46.07 \frac{\text{g}}{\text{mol}}}{\frac{0.08206 \text{ L} \cdot \text{atm}}{\text{K} \cdot \text{mol}} \cdot 299.45 \text{ K}} \\ &= 7.22 \times 10^{-5} \text{ g } C_2H_5OH \end{aligned} \quad (\text{Equation 4.5})$$

The ethanol (Fisher Scientific[®], anhydrous, denatured) has a density of 0.789 g mL⁻¹, so Equation 4.6 converts the mass of ethanol to the volume required:

$$7.22 \times 10^{-5} \text{ g } C_2H_5OH \cdot \frac{1 \text{ mL } C_2H_5OH}{0.789 \text{ g } C_2H_5OH} \quad (\text{Equation 4.6})$$

$$= 9.16 \times 10^{-5} \text{ mL } C_2H_5OH = 0.09156 \text{ } \mu\text{L } C_2H_5OH$$

With the sensor measurements, thirteen different concentrations of ethanol were used for the regression. Table 4.1 gives the list of concentrations along with the corresponding volume of ethanol introduced to the chamber and the ratio of the concentration to the ILO standard of 5,000 ppm. Each volume was rounded to the nearest microliter since the micropipette measures in whole microliter increments.

Table 4.1: Concentrations of ethanol used

Ethanol Concentration (ppm)	Ethanol Volume (μL)	Ratio to ILO Standard
0	0	0.00
98	9	0.02
295	27	0.06
699	64	0.14
1,005	92	0.20
1,507	138	0.30
2,501	229	0.50
3,014	276	0.60
3,997	366	0.80
5,002	458	1.00
7,503	687	1.50
10,004	916	2.00
12,505	1,145	2.50

To find the mass of 1 ppm ammonia ($MW = 17.031 \text{ g mol}^{-1}$), Equation 4.5 can be modified, changing only the molar mass term:

$$m = \frac{p \cdot V \cdot MW}{R \cdot T} = \frac{0.996 \text{ atm} \cdot 3.85 \times 10^{-5} \text{ L} \cdot 17.031 \frac{\text{g}}{\text{mol}}}{\frac{0.08206 \text{ L} \cdot \text{atm}}{\text{K} \cdot \text{mol}} \cdot 299.45 \text{ K}} \quad (\text{Equation 4.7})$$

$$= 2.66 \times 10^{-5} \text{ g } NH_3$$

Based on the manufacturer's data, the ammonium hydroxide solution (Fisher Scientific[®], Certified ACS Plus) has an assay of 29.47% (by mass) and a density of 0.9 g mL⁻¹. Equation 4.8 converts the mass of ammonia in Equation 4.7 to the volume of ammonium hydroxide required:

$$2.66 \times 10^{-5} \text{ g } NH_3 \cdot \frac{100 \text{ g solution}}{29.47 \text{ g } NH_3} \cdot \frac{1 \text{ mL solution}}{0.9 \text{ g solution}} \quad (\text{Equation 4.8})$$

$$= 1.00 \times 10^{-4} \text{ mL solution} = 0.100 \text{ } \mu\text{L solution}$$

Like the ethanol tests, Table 4.2 summarizes the thirteen concentrations of ammonia that were introduced into the testing chambers based on the corresponding volume of ammonium hydroxide:

Table 4.2: Concentrations of ammonia used

Ammonia Concentration (ppm)	Ammonium Hydroxide Volume (μL)	Ratio to OSHA Standard
0	0	0.00
50	5	0.17
100	10	0.33
200	20	0.67
300	30	1.00
450	45	1.50
600	60	2.00
750	75	2.50
900	90	3.00
1,200	120	4.00
1,500	150	5.00
2,400	240	8.00
3,000	300	10.00

4.5.2 Trial Procedure

Before each trial, the testing chamber was cleaned with distilled water to remove any residual chemicals. Similarly, the sensors were rinsed with distilled water and dried under a hot air blower. Five minutes after the sensors were replaced, the chamber was sealed. For ethanol and ammonia tests, the injection of each sample was carried out with a Hamilton GASTIGHT[®] series 1700 syringe for any sample less than 20 μL . Larger samples were introduced with a Fisherbrand[®] Finnpiette[®] II adjustable-volume micropipette.

After the sample was introduced to the container, enough time was allocated to allow for the system to reach equilibrium. A preliminary test was carried out with 30 μL ammonia in the testing chamber and results taken every five minutes. Based on this test, whose results are shown in Figure 4.15 and Figure 4.16, a 70 minute equilibrium time was chosen.

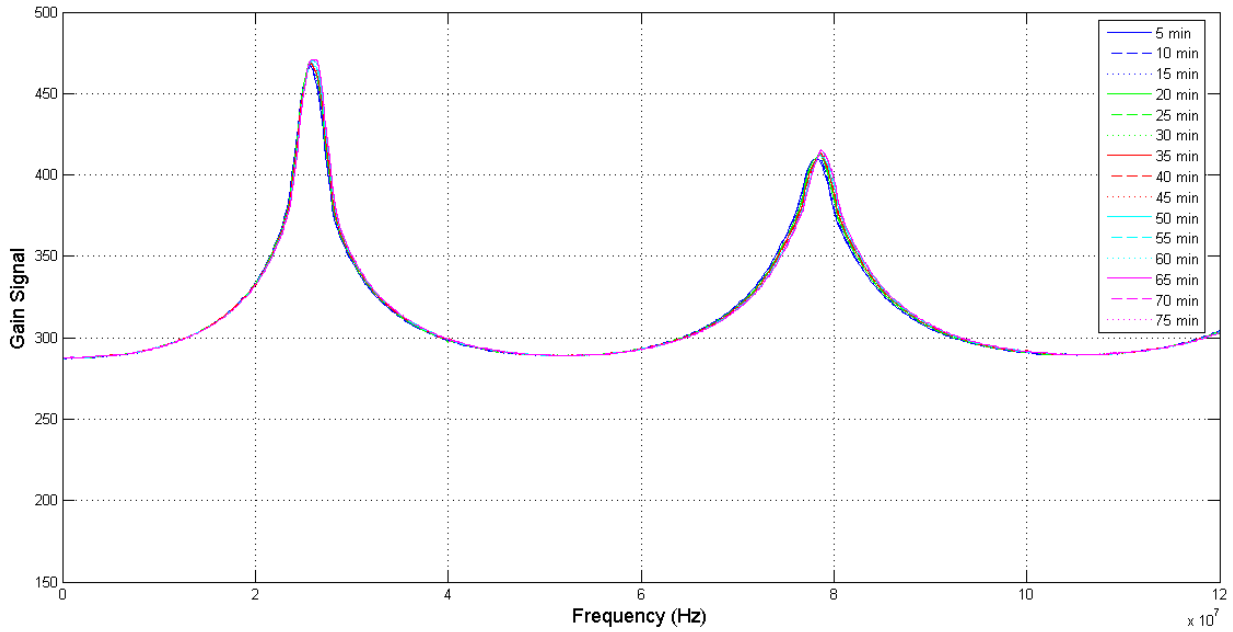


Figure 4.15: Gain data for determining equilibrium time

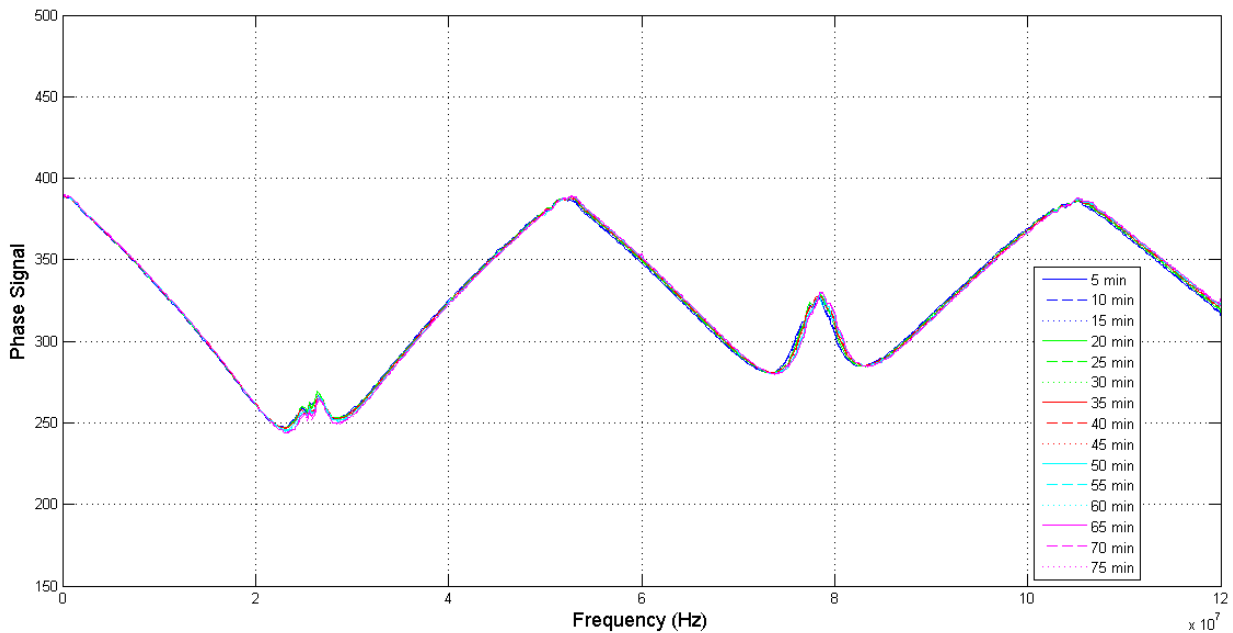


Figure 4.16: Phase data for determining equilibrium time

4.5.3 Data Acquisition

All results were captured through the control box and its accompanying serial output. With an RS-232 serial cable, all data were captured with the HyperTerminal[®] program and then saved as a comma delimited text file. From Microsoft Excel[®], this text file was converted into a XLS spreadsheet and then processed using one of several functions written in either MATLAB[®] or SAS[®].

4.6 Data Analysis

4.6.1 Qualitative Detection of Contaminant Presence in Air

As glycerol was only tested on a detection basis, a least squares means test at each frequency was used to distinguish between ambient air and air with glycerol vapor present. Either the gain or phase data was introduced into the first of two programs for this process. The program, written in SAS[®] 9.2 (The SAS Institute, 2008) and given in the first part of Appendix C, utilizes the ‘proc glm’ command that returns the average response at each combination of frequency and presence of contaminants. Additionally, the ‘lsmeans’ command with the ‘slice’ option returns p -values for testing the differences between readings with and without contaminants at a given frequency.

From this output, a spreadsheet is made in Microsoft Excel[®] to include the frequency in the first column and p -value (sorted in ascending order) in the second column. This spreadsheet is then processed with the MATLAB[®]-based Holm Comparison program in the second half of Appendix C. The program is based on the Holm-Bonferroni method for making multiple comparisons, which is a more conservative process than a direct comparison of p -values. Beginning with the most significant (smallest) p -value, the p -value is compared with α divided by the number of comparisons left. If the p -value is less than α divided by the number of comparisons left, the difference is declared to be significant and the process continues with the next lowest p -value. In case the difference is not significant, the process stops since no other frequencies will be significant (Jager, 2012).

Figure 4.17 demonstrates the process for deciding whether or not the readings at a given frequency are significant. In this flowchart, k represents the number of comparisons left.

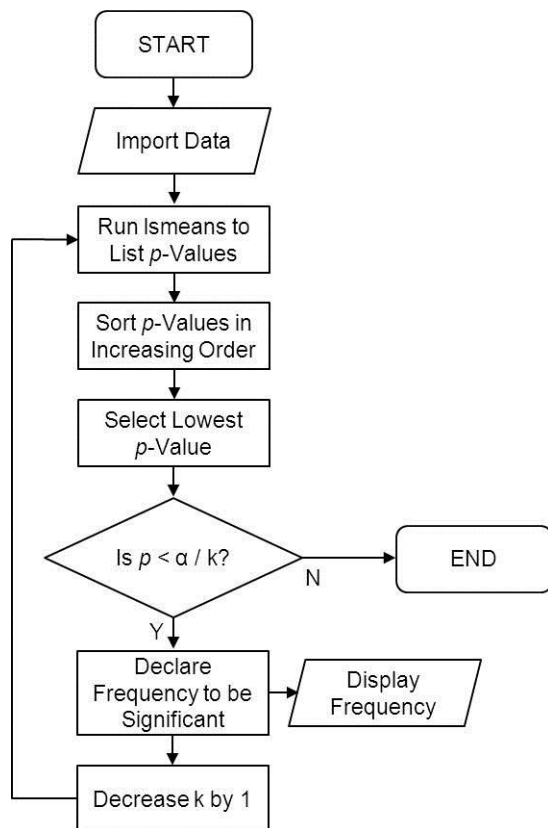


Figure 4.17: Flowchart of detection program

Additionally, a neural network was used to analyze the utility of measuring the contaminants on a truly qualitative basis, determining whether or not a contaminant was present. Using a procedure similar to that described in Section 4.6.2.4 with the code given in Appendix D, the neural network classifies all iterations from both a training dataset (that was used to train the neural network) and a validation dataset as either having contaminants or no contaminants. Instead of having the concentrations being in parts per million, all nonzero concentrations are identified as one. Should the neural network output be below 0.5, the predicted output decides there are no contaminants present; otherwise, the neural network determines there are contaminants present. Tables in Chapter 5 list the success rate detecting each contaminant.

4.6.2 Quantitative Detection of Contaminant Presence in Air

For ethanol and ammonia, where multiple concentrations were measured, the data were analyzed with stepwise regression, partial least squares regression, neural network analysis, and wavelet transformation, the same four regression methods used by Shultz (2009).

For each combination of sensor and contaminant, a summary spreadsheet details the individual spreadsheets where the frequency response data are stored, the amount of contaminant introduced for each sample, and the temperature at which each of the three iterations were taken. The Import Data function, whose code is given in Appendix I, takes each of the individual spreadsheets, separates the gain and phase data, creates a term for the interaction between gain and phase, and then returns matrices for use in other MATLAB[®] programs.

The data originally represents the entire spectrum of frequencies that the control box is capable of producing and measuring (up to 120 MHz). For the regression analysis, the data are divided into a range of frequencies for the specific peaks, depending on the sensor.

Table 4.3: Frequency ranges used in regression analysis

	Small Probe, Large Probe	Small Box	Large Box
Peak 1	2.10 x 10 ⁷ Hz to 3.78 x 10 ⁷ Hz	1.30 x 10 ⁷ Hz to 3.48 x 10 ⁷ Hz	1.60 x 10 ⁷ Hz to 2.78 x 10 ⁷ Hz
Peak 2	8.08 x 10 ⁷ Hz to 9.76 x 10 ⁷ Hz	5.84 x 10 ⁷ Hz to 7.12 x 10 ⁷ Hz	5.52 x 10 ⁷ Hz to 6.50 x 10 ⁷ Hz
Peak 3		9.74 x 10 ⁷ Hz to 1.13 x 10 ⁸ Hz	9.42 x 10 ⁷ Hz to 1.08 x 10 ⁸ Hz

The four programs for performing the various regression techniques all followed the same general pattern. First, the program imported the data from Microsoft Excel[®] spreadsheets and trimmed the data based on the specific sensor and data (gain only, phase only, gain and phase together, or the interaction between gain and phase). Using data from three iterations of all measurements at each concentration, a model is fitted through the predictor variables and response variables. The predictors of the data (the training dataset) are then run through the model to estimate the response. If any prediction yields a negative concentration, the prediction is changed to zero. Comparing these estimated responses to the actual responses provides an R² and RMSE parameters. Finally, a second set of three iterations of all measurements at each concentration (the validation dataset) is also used with the model to give R² and RMSE values.

All of the programs for analyzing the data were written in the MATLAB[®] R2012a (The MathWorks, 2012a) and the various additional toolboxes.

4.6.2.1 Stepwise Regression

After importing and trimming the data for the specific regression, the ‘stepwisefit’ function in the Statistics Toolbox™ (The MathWorks, 2012e) is the main function for creating the stepwise model. The ‘stepwisefit’ function returns a vector of coefficient estimates and a logical vector indicating which coefficients are significant. In these regressions, α_{in} and α_{out} were 0.05 and 0.10, respectively, since these are default values in the ‘stepwisefit’ function.

Once the ‘stepwisefit’ model is created, a list of significant frequencies is exported to a Microsoft Excel® spreadsheet. Next, this model is used to predict the contaminant concentration. This prediction is calculated by summing the products of the predictor variables, the coefficient vector, and the logical vector. By including this logical vector, only the significant predictors have an effect on the predictions. Once all these products are totaled, the intercept is added to complete the model.

After the model is completed, prediction, R^2 , and RMSE values are calculated for the training and validation datasets. Should no model be created, meaning that none of the predictors were found to be significant, the program does not attempt to create a list of significant frequencies and displays “N/A – no model made” for R^2 and RMSE values.

Figure 4.18 gives a flowchart of the process while Appendix E gives the full code.

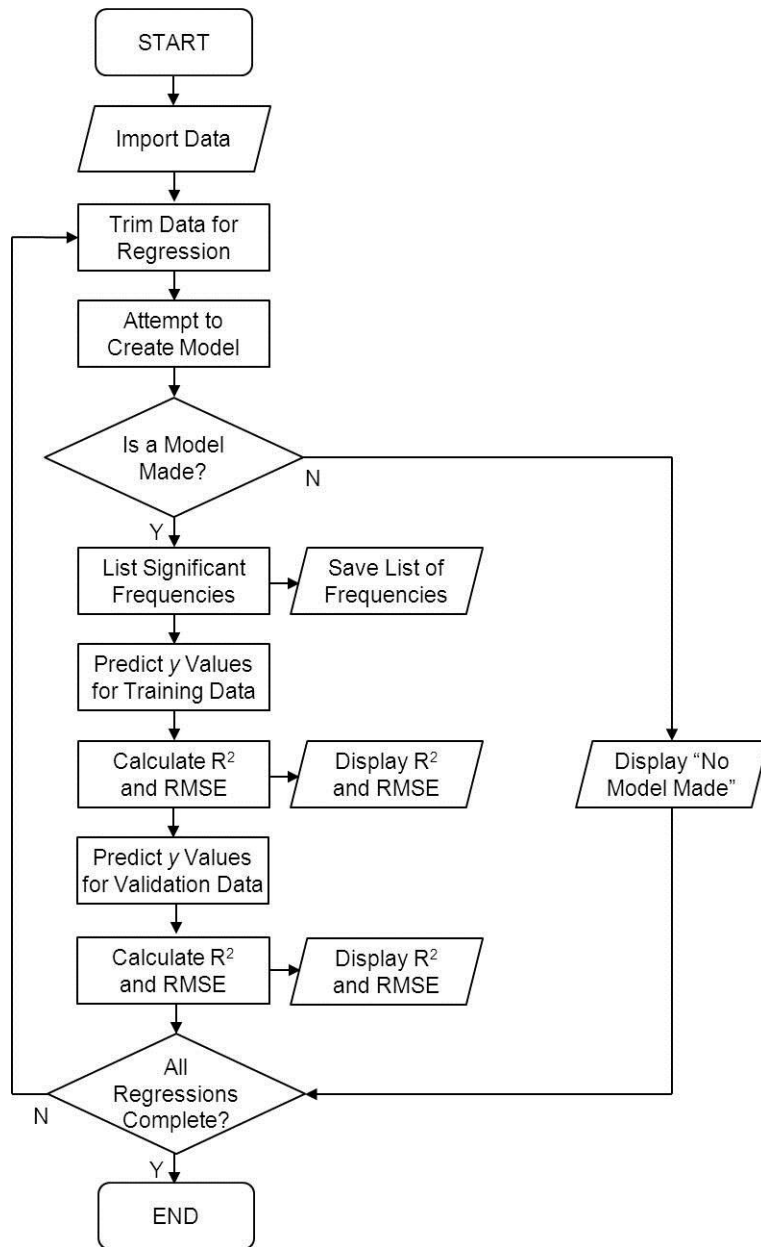


Figure 4.18: Flowchart of stepwise regression program

4.6.2.2 Wavelet Transform

The program for performing the wavelet regression is very similar to the stepwise regression program, with two major differences. First, all predictors are processed using the 'wavedec' function from the Wavelet Toolbox™ (The MathWorks, 2012f). A third level Haar wavelet was used for each regression. Also, the significant coefficients are exported to a Microsoft Excel® spreadsheet, rather than the significant frequencies. Figure 4.19 outlines the procedure while the full program is in Appendix F.

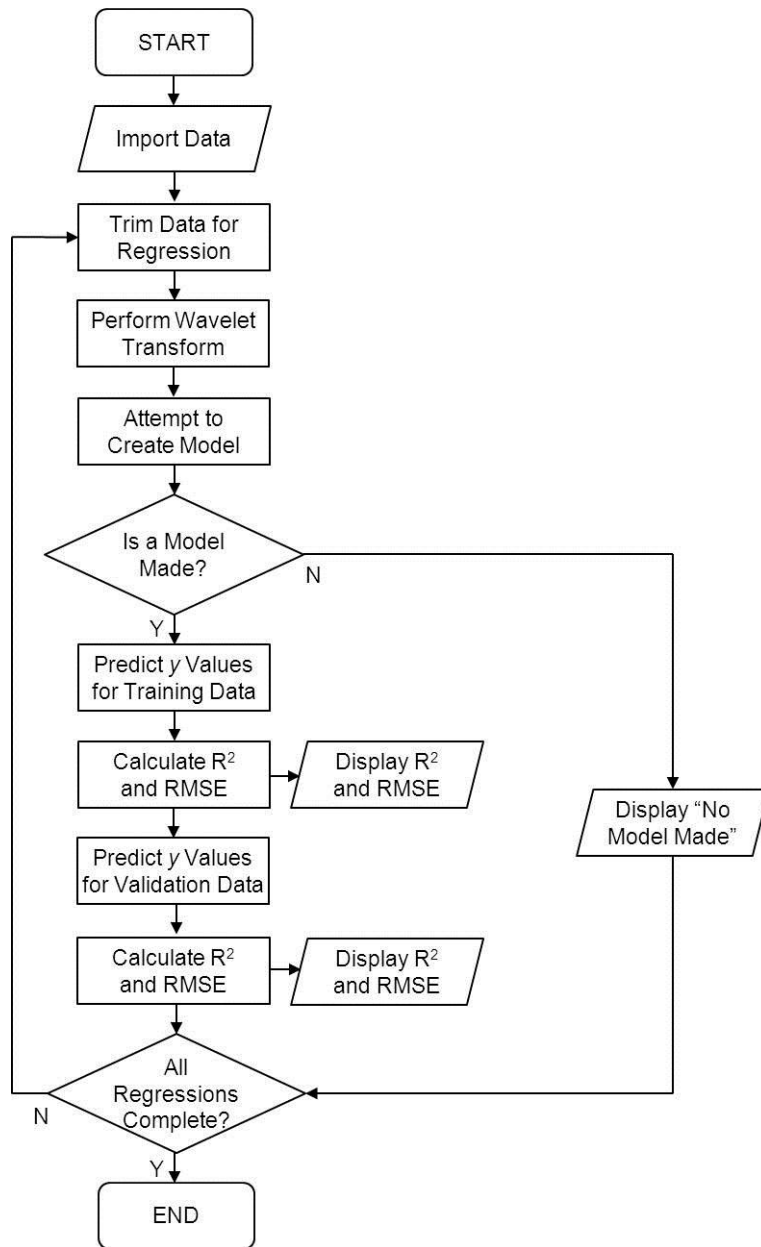


Figure 4.19: Flowchart of wavelet regression program

4.6.2.3 Partial Least Squares Regression

Following the importation and trimming of data, the ‘plsregress’ function in the Wavelet Toolbox™ (The MathWorks, 2012f) is the function used to return the PLS coefficients from the training dataset. The function requires a parameter that indicates the number of components to use in the regression. This is included as another input for the program, and was chosen to be four, six, and eight for separate executions of the regression.

For simplicity, the beta coefficients \mathbf{B} in Equation 3.14 were used for the model, rather than the matrices of scores and loadings in Equation 3.12 and Equation 3.13. This model is then used to predict the actual concentrations from the training and validation datasets. Illustrated in Figure 4.20, the complete code is included in Appendix G.

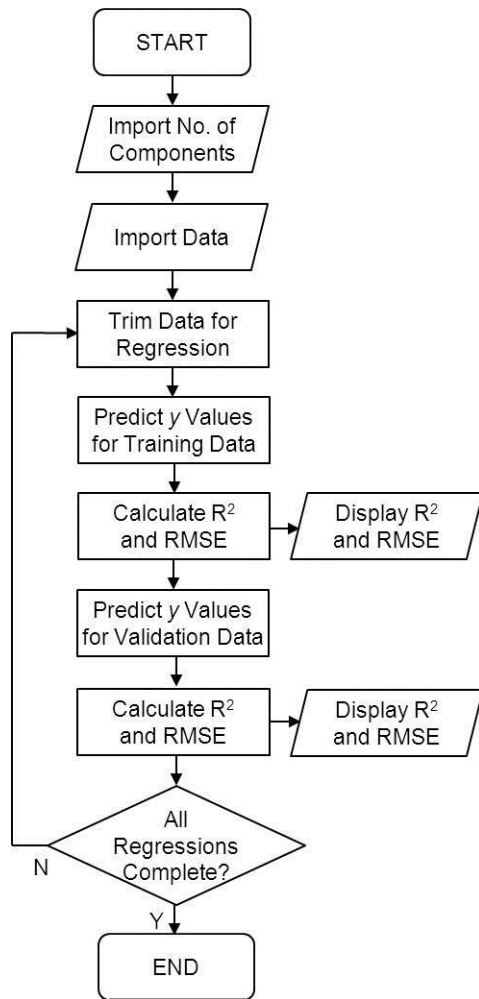


Figure 4.20: Flowchart of PLS regression program

4.6.2.4 Neural Networks

The program for neural network regression begins by reading the data, and then transposes the data so each row corresponds to a specific predictor. For regressions with the gain or phase data joined and with the gain and phase interaction, only every third predictor is used;

this reduces the amount of data used so that the neural network program can run properly. Otherwise, the computer runs out of memory, and the program terminates.

From the Neural Network Toolbox™ (The MathWorks, 2012b), the neural network is initialized with a hidden layer size of ten, a typical value for such applications. The data are then processed with the 'removeconstantrows' and 'mapminmax' functions to remove any data that is constant for all contaminant concentrations and normalizes the data in each row to a [-1, 1] range. All data are divided in a random order such that 60% of samples are to train the network, 20% to validate the network, and 20% to test the network.

After these parameters are set, the network is trained with the Levenberg-Marquardt backpropagation algorithm. This algorithm is relatively fast and very suitable for supervised training (Demuth and Beale, 2004). Once the training is complete, the network is ready to accept a matrix of predictor variables to model the response variables, which is done with both the training and validation datasets.

Figure 4.21 gives the framework for the entire process, and Appendix H gives the full code for the program.

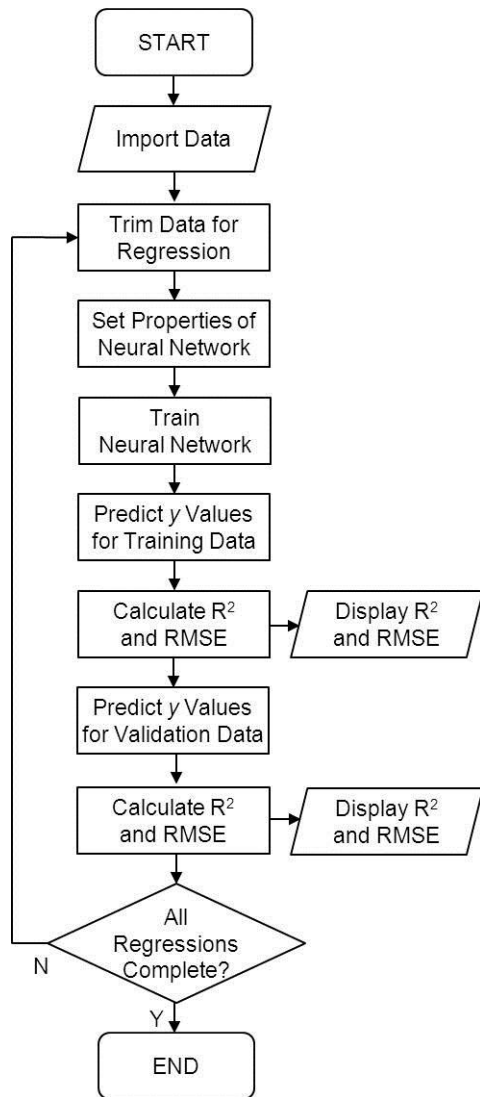


Figure 4.21: Flowchart of neural network regression program

CHAPTER 5 - Experimental Results and Discussion

5.1 Sensitivity Tests

Using the procedure described in Section 4.4, several factors were investigated to determine which had significant effects on the sensor measurements. The major factors examined included the temperature of the sample, the surface on which the sampling chamber was placed, the sensor position relative to the bottom and sides of the chamber, the presence of a thermometer, and the cable positioning. Each of the tests in this section was carried out in distilled water with the large probe sensor.

5.1.1 Temperature

Varying the temperature had a minor effect on frequency and magnitude of different peaks showing in the frequency responses. As shown in Figure 5.1 and Figure 5.2, there is a tendency for peaks at higher temperatures to not be as strong for both the gain and phase signal. However, since the room temperature where the air samples were taken could not be kept constant, the temperature for each iteration was included as another predictor for regression.

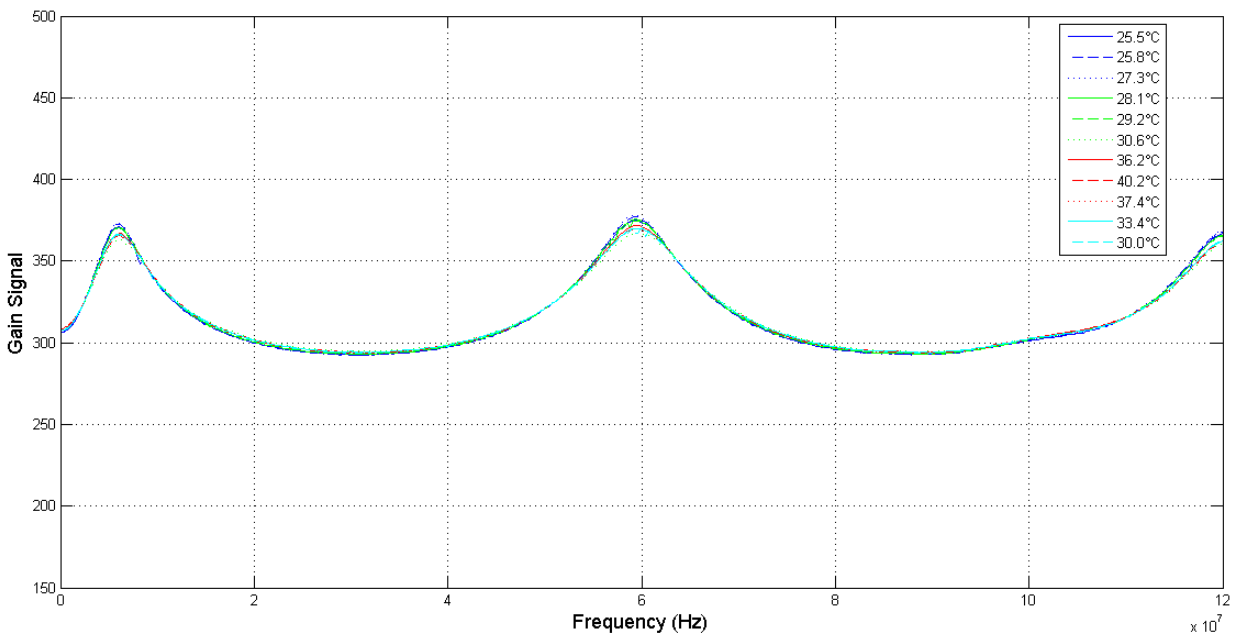


Figure 5.1: Gain data for various temperatures in distilled water

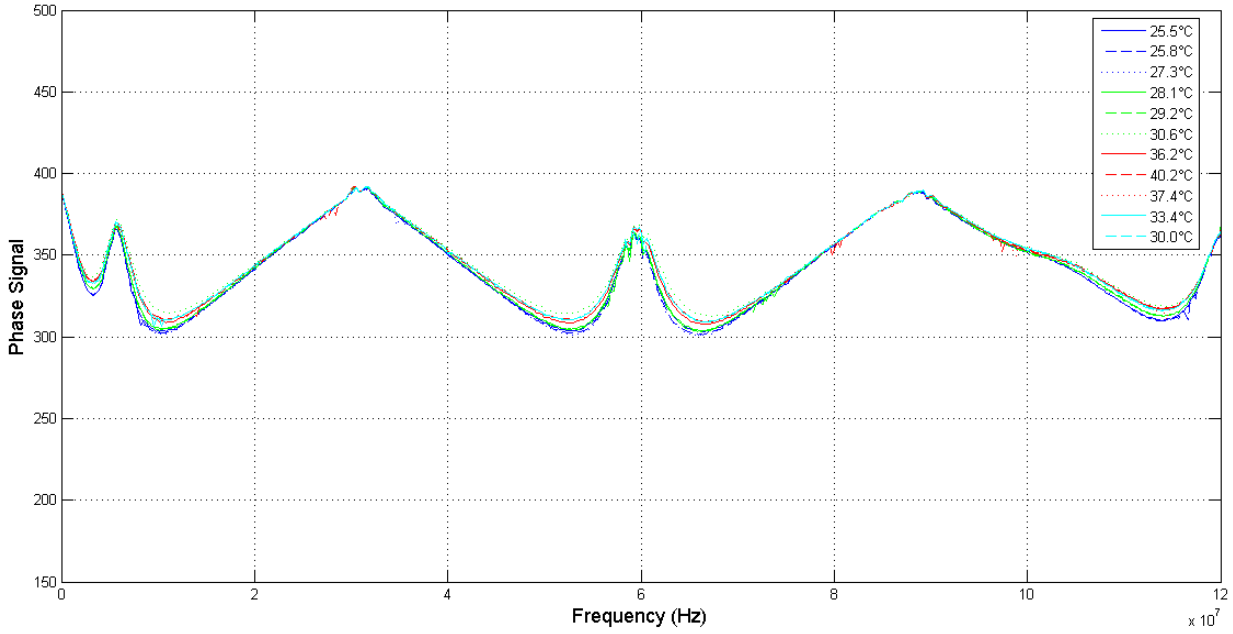


Figure 5.2: Phase data for various temperatures in distilled water

5.1.2 Testing Surface

Placing the sampling chamber on a variety of surfaces did not have a significant effect on the sensor measurements. Figure 5.3 and Figure 5.4 show very minute differences between tests carried out on a slate tabletop, a metal tabletop, and atop a ceramic hot plate.

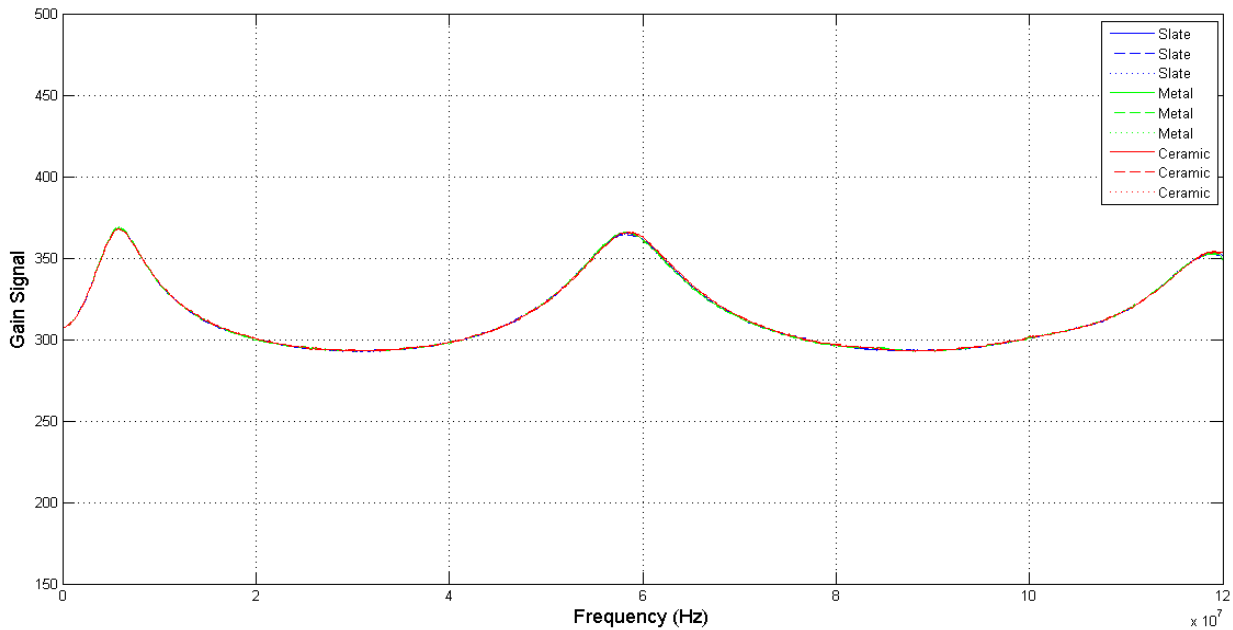


Figure 5.3: Gain data on various table surfaces in distilled water

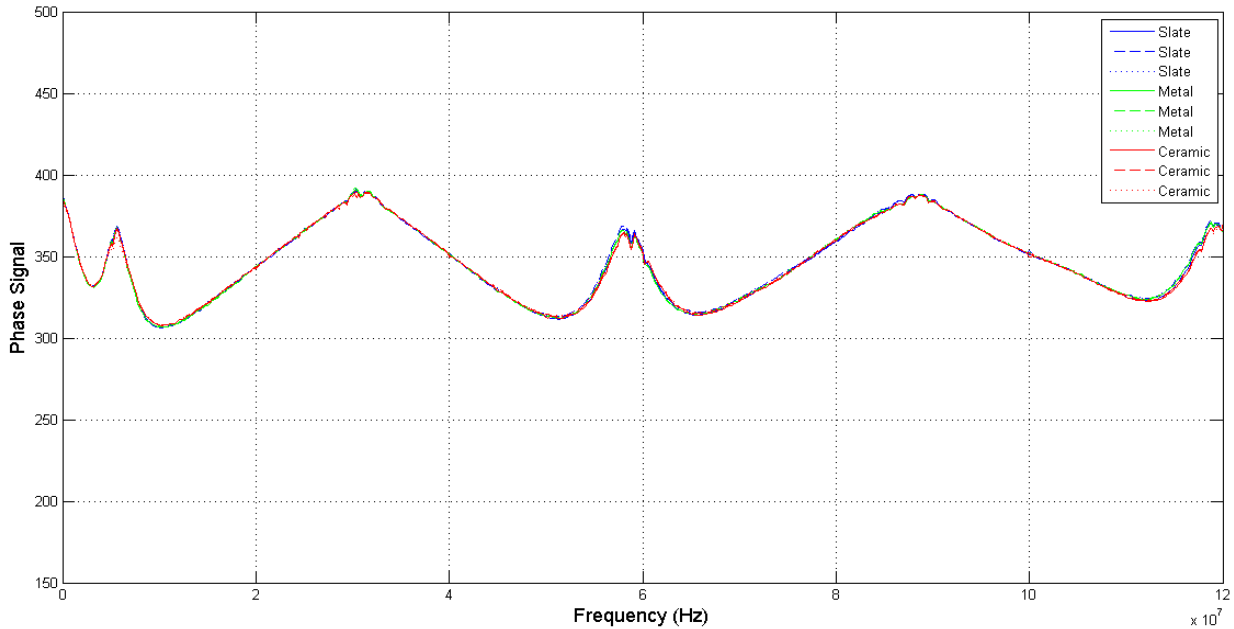


Figure 5.4: Phase data on various table surfaces in distilled water

5.1.3 Sensor Height

Based on Figure 5.5 and Figure 5.6, the height of the sensor relative to the bottom of the container had very little effect. At 1 cm, 3 cm, 5 cm, and 7 cm above the container bottom, both the gain and phase signals have very negligible differences. However, at 9 cm above the container bottom, there is a significant difference; this is because about half the area of the metal plates was exposed to the air.

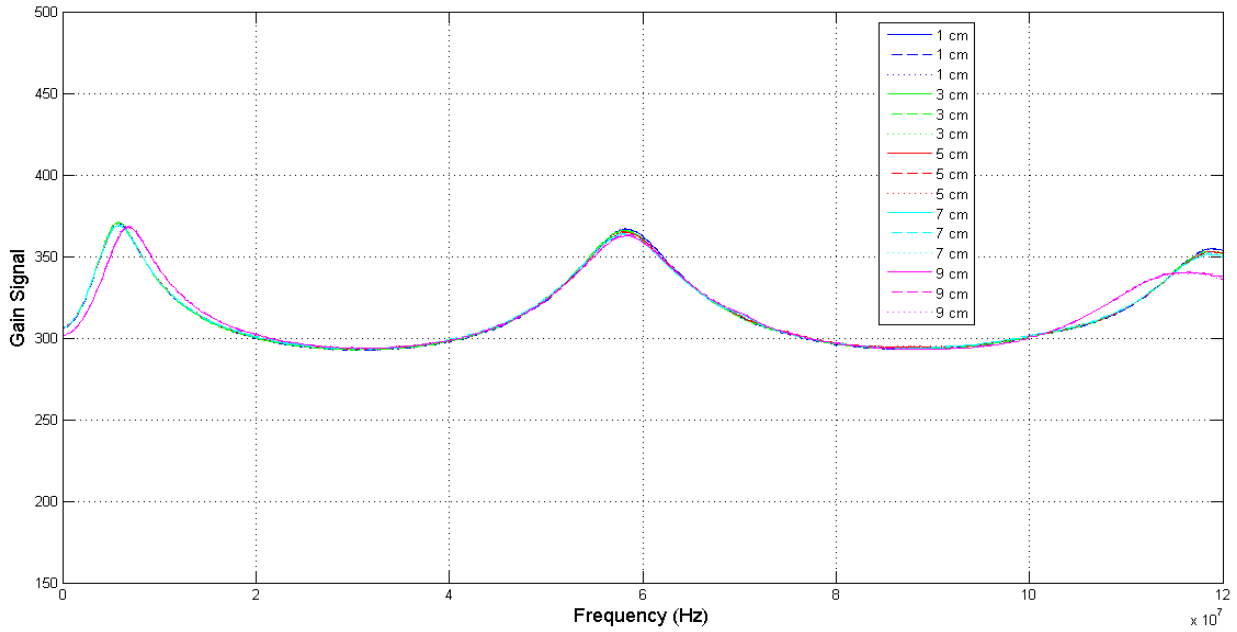


Figure 5.5: Gain data for various sensor heights in distilled water

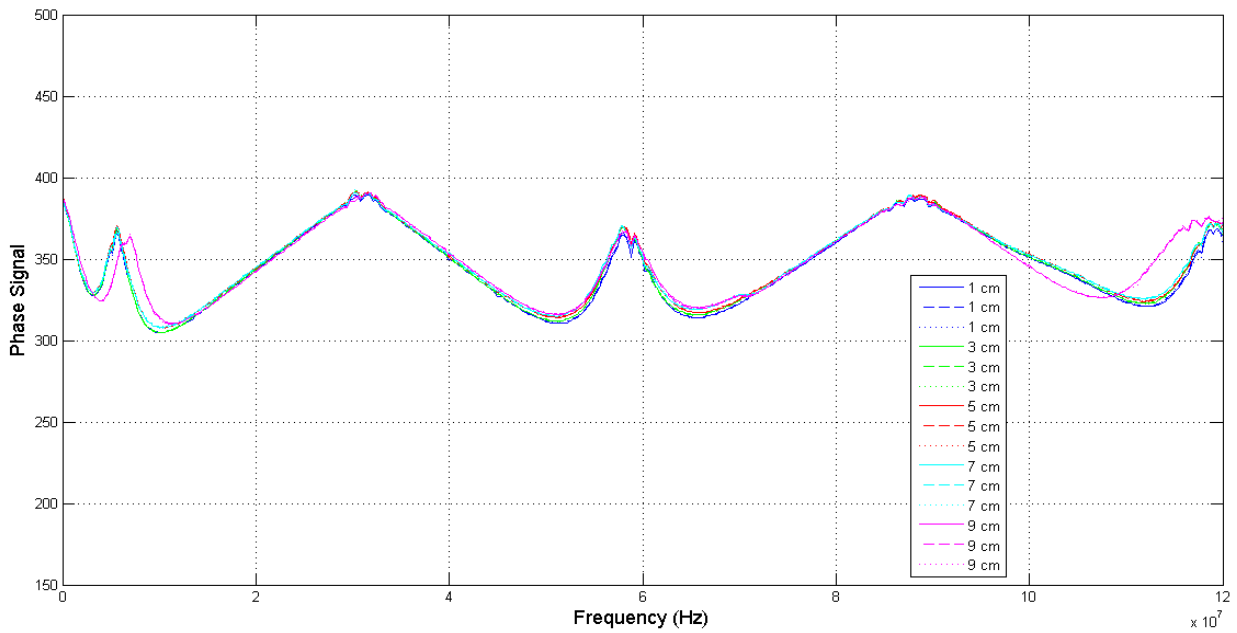


Figure 5.6: Phase data for various sensor heights in distilled water

5.1.4 Sensor Radial Distance

Similar to the sensor height, the boundary effect was minimal when compared at different radial distances. Gain and phase readings, shown in Figure 5.7 and Figure 5.8, are virtually

identical when the sensor is at the center of the container, near the edge of the container, or halfway between the center and edge.

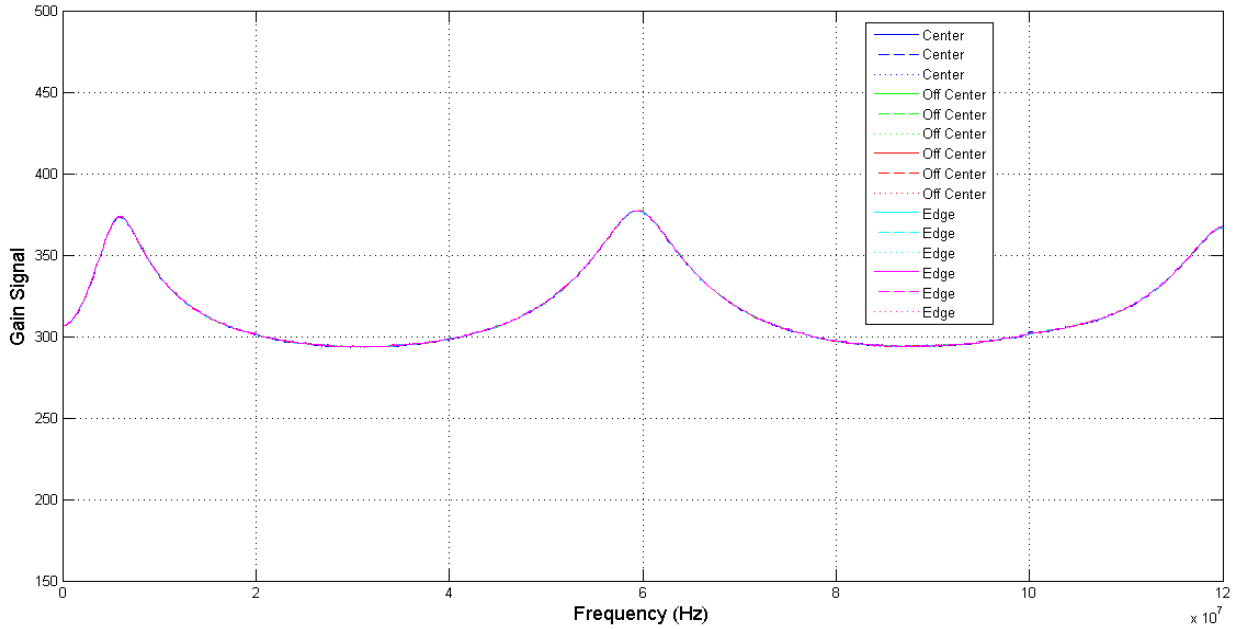


Figure 5.7: Gain data for various sensor radial positions in distilled water

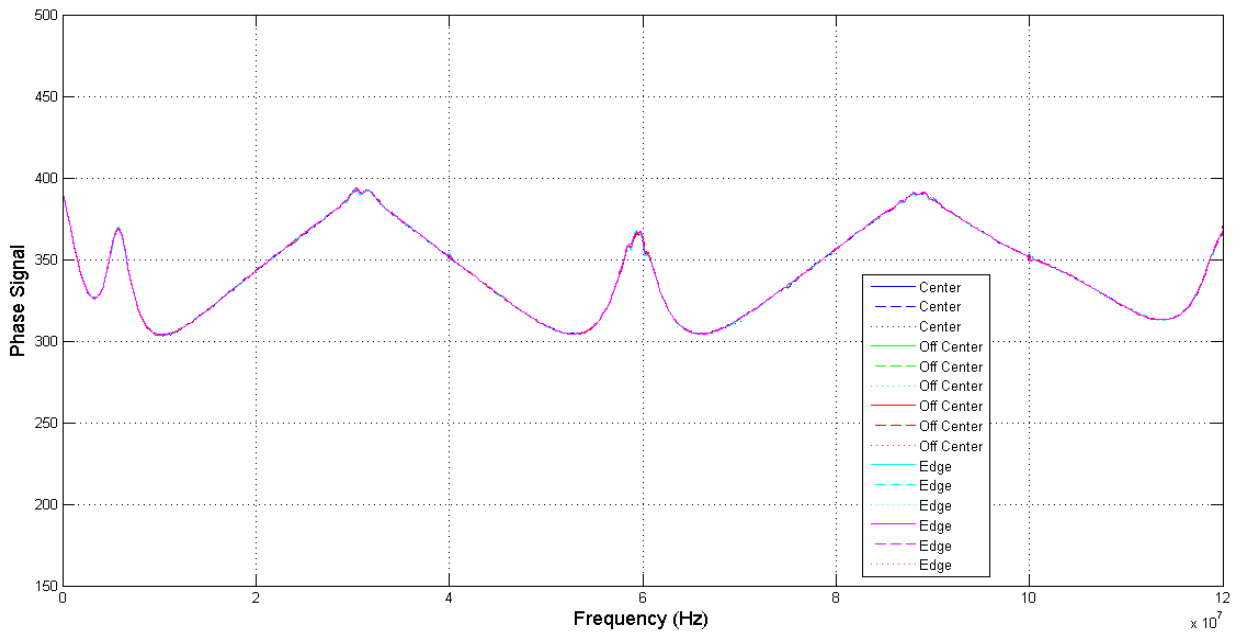


Figure 5.8: Phase data for various sensor radial positions in distilled water

5.1.5 Presence of a Thermometer

From the spectra in Figure 5.9 and Figure 5.10, a stainless steel thermometer was shown to be insignificant in affecting the sensor readings, regardless of the proximity to the sensor.

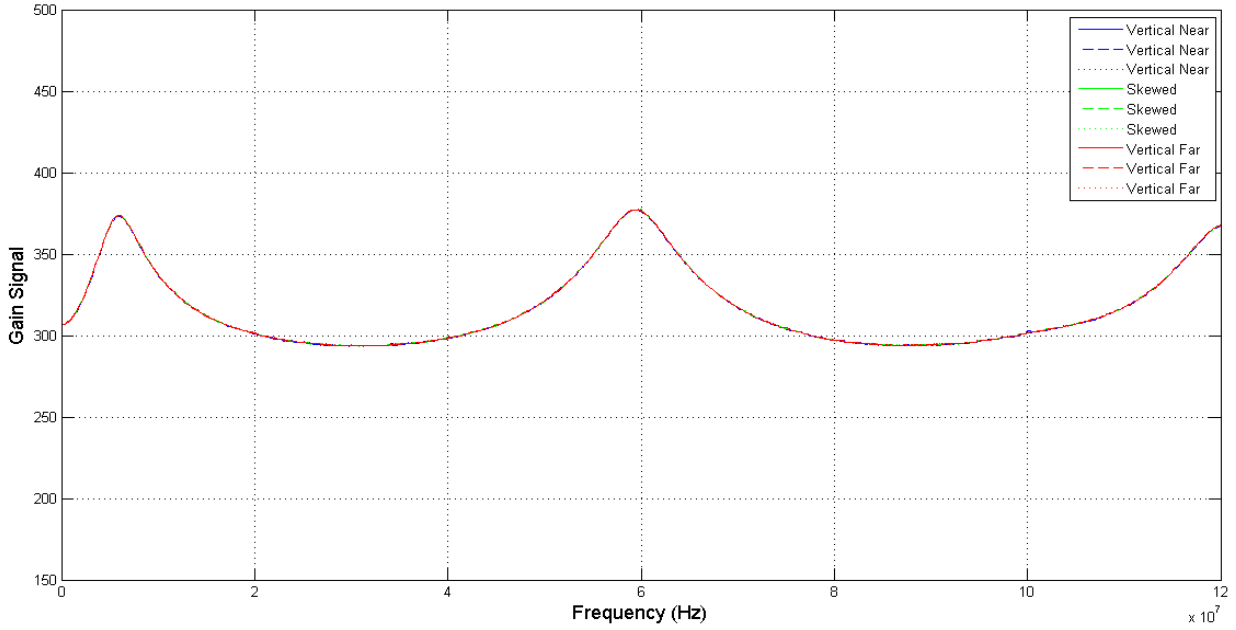


Figure 5.9: Gain data for various thermometer positions in distilled water

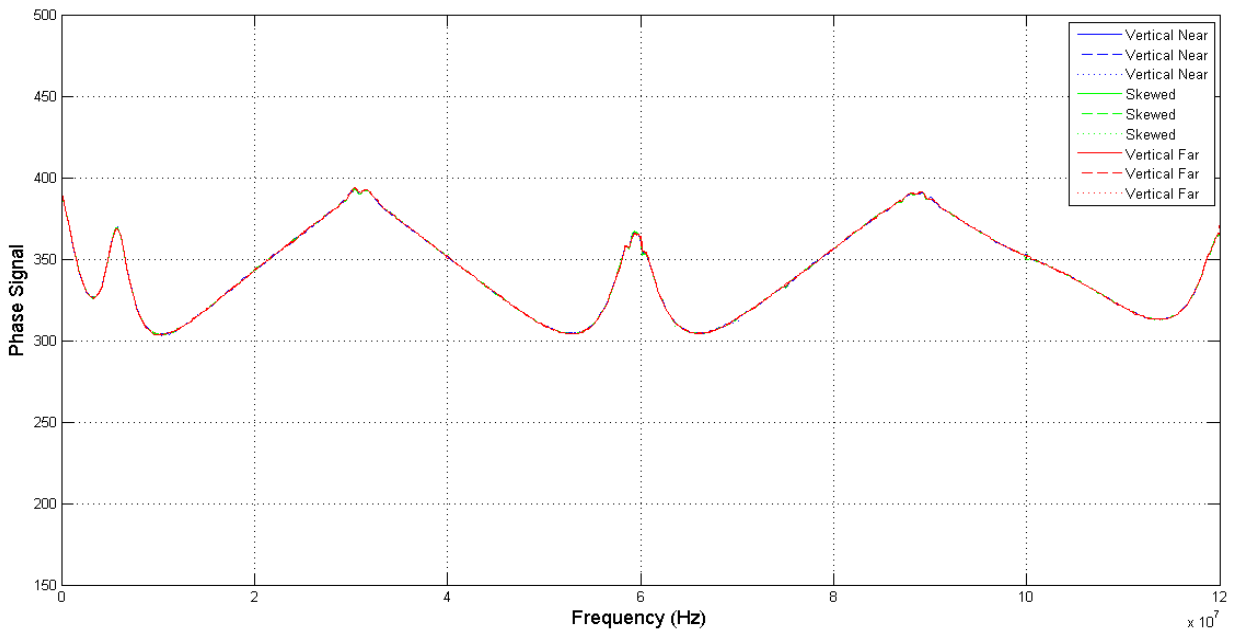


Figure 5.10: Phase data for various thermometer positions in distilled water

5.1.6 Cable Position

The position of the coaxial cable did have an impact on readings, as shown in Figure 5.11 and Figure 5.12. This position likely shifted the BNC connector, impacting the gain and phase data. In tests with air contaminants, the cable position was held fixed to marginalize this effect.

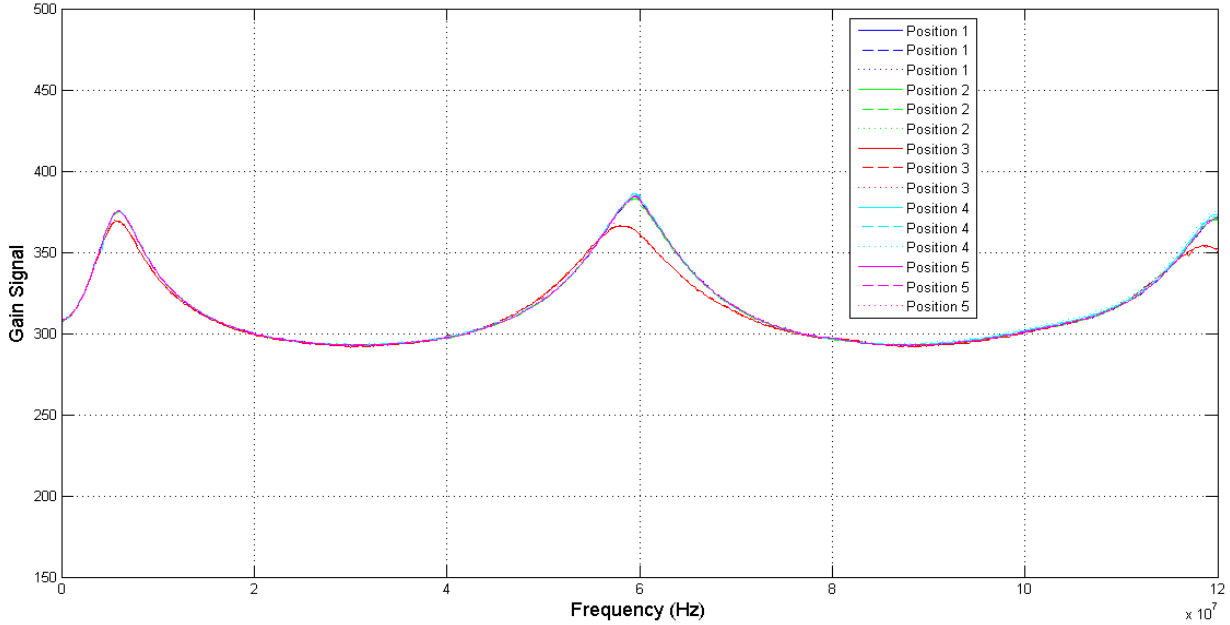


Figure 5.11: Gain data for various cable positions in distilled water

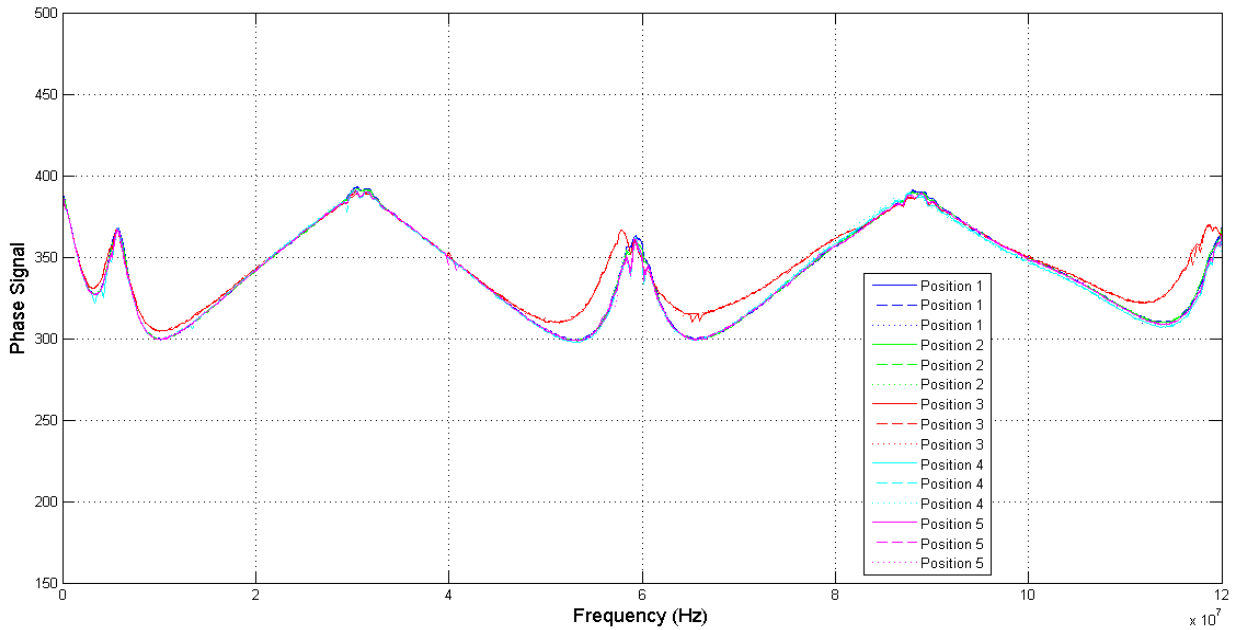


Figure 5.12: Phase data for various cable positions in distilled water

5.2 Glycerol Tests

Using the code in Appendix C, the gain and phase data were compared between ambient air and air with glycerol vapor present. A total of 12 iterations through all 609 frequencies, six without glycerol and six with glycerol, were used in the analysis. Least squares means and a comparison using the Holm-Bonferroni method, discussed in Section 4.6.1, determined which frequencies had significantly different measurements between the two types of samples.

Figure 5.13 and Figure 5.14 show distributions of significant frequencies for gain and phase data, respectively. These distributions divided the entire spectrum of 21 groups of 29 frequencies each. Appendix J identifies the complete list of frequencies that were identified as having a significant difference.

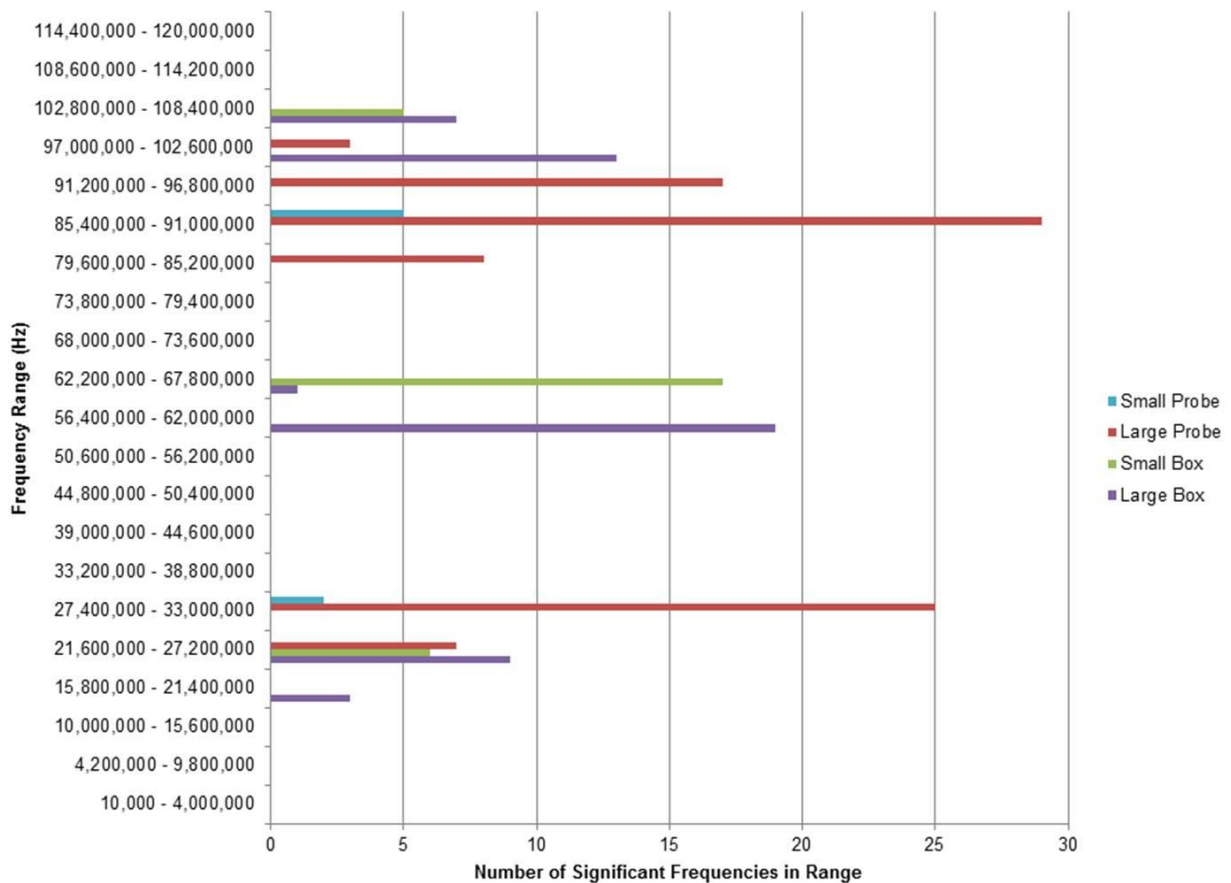


Figure 5.13: Distribution of significant frequencies with glycerol, gain data

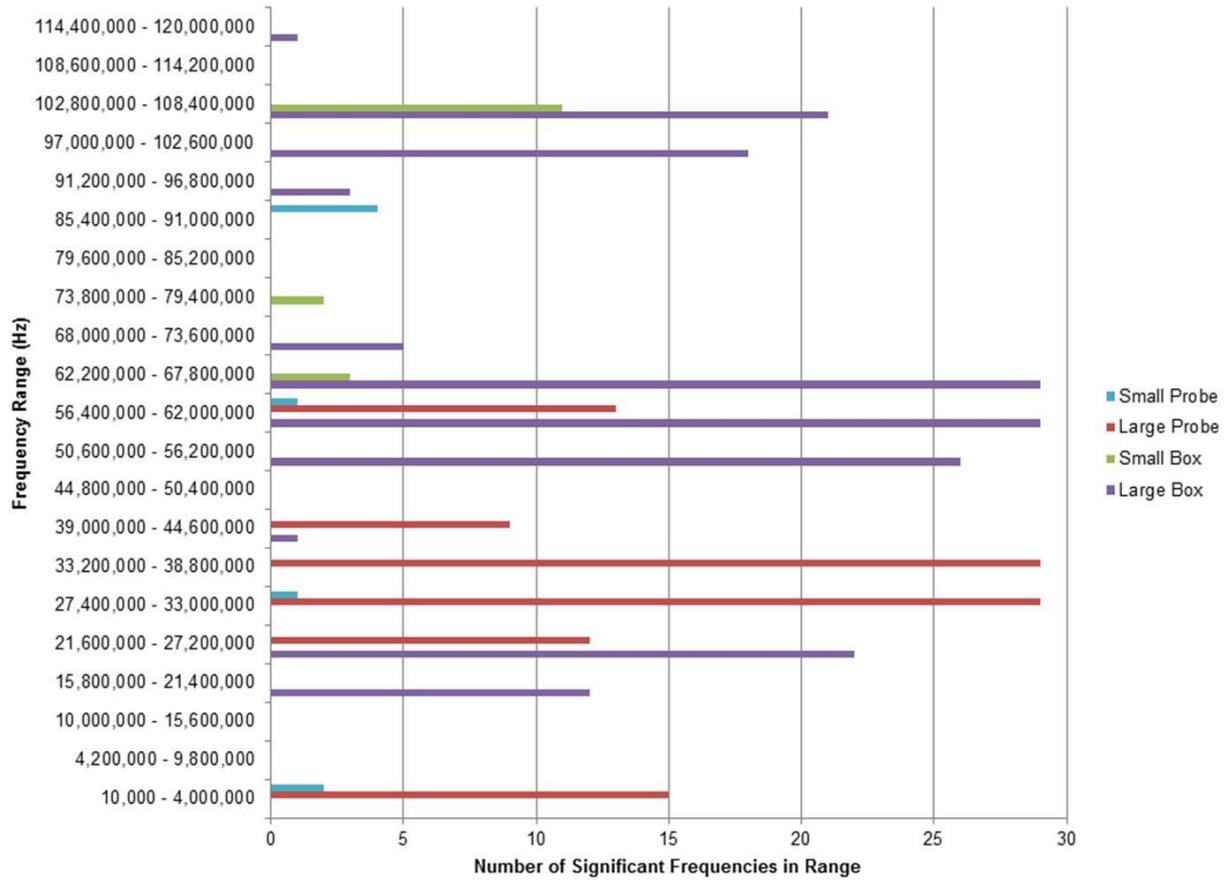


Figure 5.14: Distribution of significant frequencies with glycerol, phase data

The gain data showed the most significantly different frequencies with the large probe sensor (89) and the least with the small probe sensor (7). Most of the significantly different frequencies corresponded to the major peaks in the frequency response spectra in Figure 4.2, Figure 4.4, Figure 4.6, and Figure 4.8. Both the small and large probe sensors each had two major blocks of frequencies that had significantly different measurements, just as the small and large box sensors had three major blocks apiece.

Phase data showed that, like the gain data, the large probe and large box sensors had many more frequencies where the measurements between ambient air and air with glycerol present were significantly different (107 and 167, respectively), compared with the small probe and small box sensors (8 and 16, respectively). Like was seen with the gain data, the frequencies with significantly different measurements corresponded to the peaks in the ambient air spectra.

Using the neural network detection method correctly identified the vast majority of cases with either ambient air or air with glycerol present. A total of 24 iterations of the entire frequency range were used in the analysis, with six belonging to each combination of ambient air/air with glycerol and training/validation. As shown in Table 5.1, most cases were correctly identified as either being ambient air or air with glycerol present. In total, 184 of the 192 iterations were classified correctly, producing a 95.8% success rate.

Table 5.1: Success rates of neural network for detecting glycerol

	Ambient Air Training	Ambient Air Validation	Air with Glycerol Training	Air with Glycerol Validation
Small Probe Gain	5 of 6 (83%)	5 of 6 (83%)	6 of 6 (100%)	6 of 6 (100%)
Small Probe Phase	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)
Large Probe Gain	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)	4 of 6 (67%)
Large Probe Phase	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)
Small Box Gain	6 of 6 (100%)	5 of 6 (83%)	5 of 6 (83%)	6 of 6 (100%)
Small Box Phase	6 of 6 (100%)	5 of 6 (83%)	6 of 6 (100%)	4 of 6 (67%)
Large Box Gain	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)
Large Box Phase	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)	6 of 6 (100%)

5.3 Ethanol Tests

5.3.1 Qualitative Detection of Ethanol Presence in Air

Following the same procedure in Section 5.2 for glycerol, the frequency response measurements were compared between ambient air and air with ethanol at the International Labour Organization standard of 5,000 ppm (Stellman, 1998). Here, Figure 5.15 and Figure 5.16 illustrate the distributions of frequencies where the measurements were significantly different while Appendix J includes the full list of frequencies.

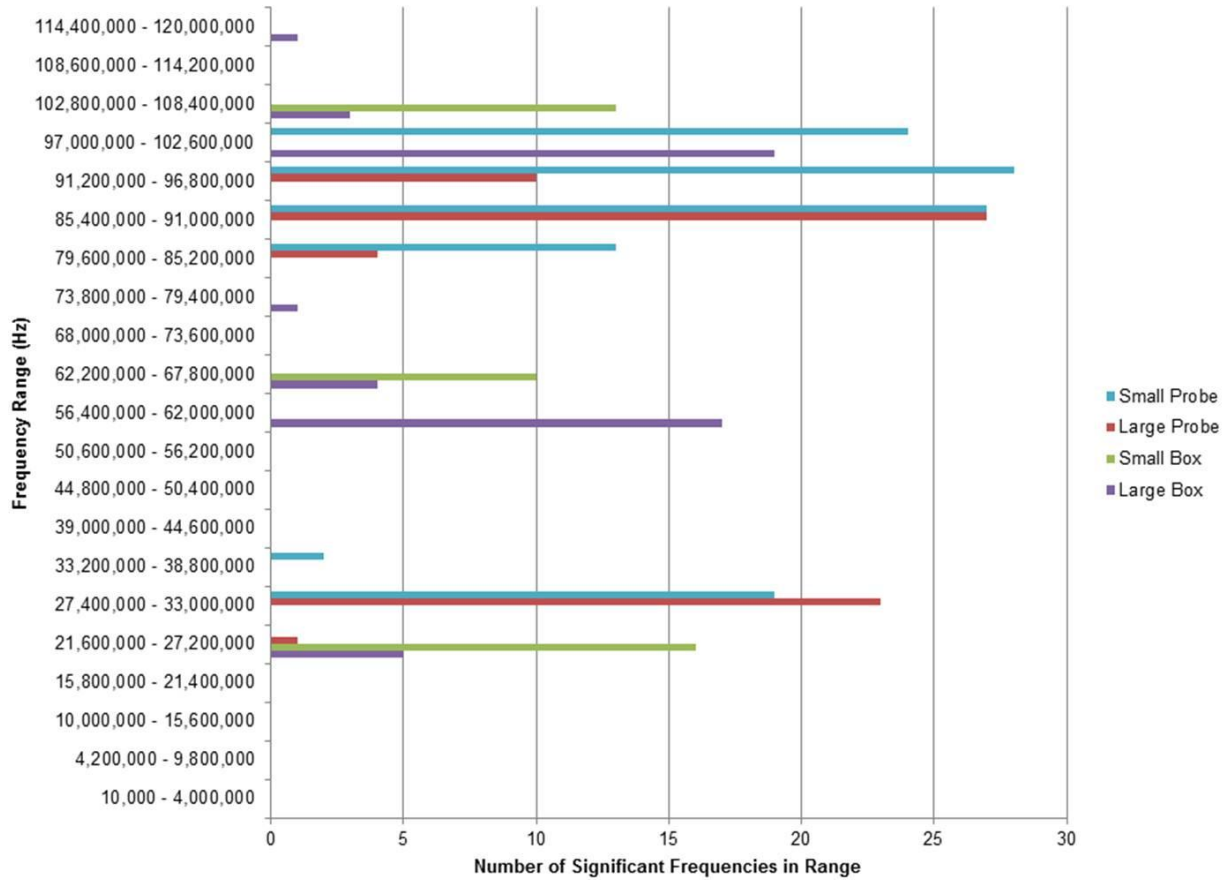


Figure 5.15: Distribution of significant frequencies with ethanol, gain data

One of the biggest differences between the gain data with glycerol and ethanol is the number of significant frequencies with the small probe sensor (7 compared to 113). In fact, this count is about twice as much as for the large probe sensor (65), where the small and large box sensor had 39 and 50 frequencies with significant differences, respectively. Again, these frequencies tended to coincide with the major peaks seen with the frequency response spectra. Additionally, for the small probe, many significant differences were in the 90,000,000 Hz to 100,000,000 Hz range, a slightly higher frequency than the main peak in ambient air.

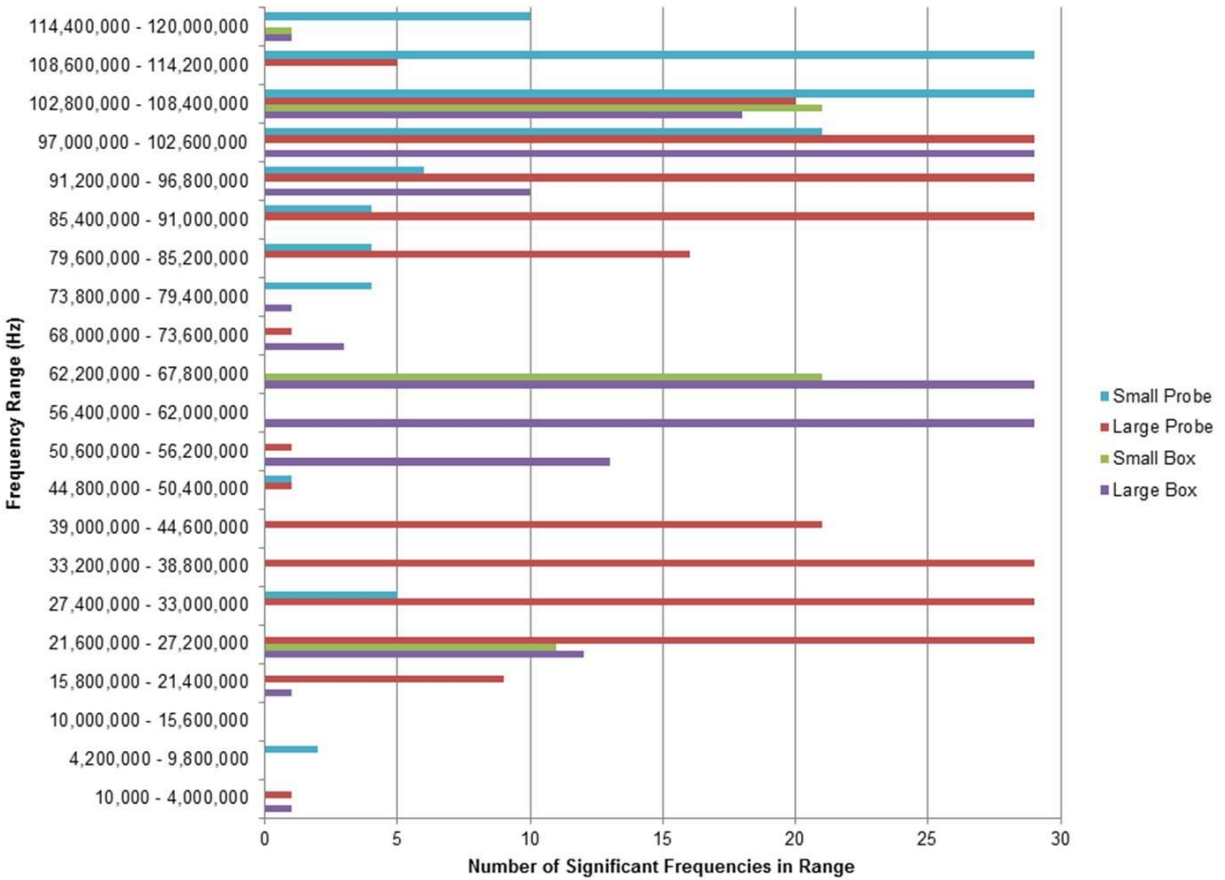


Figure 5.16: Distribution of significant frequencies with ethanol, phase data

Phase data showed many more contiguous blocks of frequencies with significant differences in measurements. The major blocks were from 98,600,000 Hz to 115,600,000 Hz for the small probe sensor; 20,400,000 Hz to 42,600,000 Hz and from 82,400,000 Hz to 104,000,000 Hz for the large probe sensor; and from 53,800,000 Hz to 68,200,000 Hz and from 95,400,000 Hz and from 106,200,000 Hz. Each frequency in these blocks had measurements that were significantly different between the ambient air and air with 5,000 ppm ethanol. Also, unlike the gain data, these contiguous blocks do not match up exactly with the major peaks seen with ambient air. Rather, especially for the large probe and large box sensors, the blocks of frequencies seemed to be centered near the peaks from ambient air, yet there is a much larger spread of frequencies.

With the neural network detection procedure, two complete runs, or six iterations, were used for each of the concentrations of ethanol listed in Table 4.1. Half of the iterations at each concentration formed the training dataset; the other half, the validation dataset. In six of the 48 cases of ambient air, the neural network declared that there was ethanol present, resulting in a false positive. For the 576 cases with ethanol present, 32 false negatives were found; most of these occurred at the lowest concentration (98 ppm), yet about one-third occurred at 5,002 ppm or above. Overall, the detection success rate was 93.9%.

Table 5.2: Success rates of neural network for detecting ethanol

	Ambient Air Training	Ambient Air Validation	Air with Ethanol Training	Air with Ethanol Validation
Small Probe Gain	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	31 of 36 (86%)
Small Probe Phase	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	31 of 36 (86%)
Large Probe Gain	3 of 3 (100%)	2 of 3 (67%)	36 of 36 (100%)	36 of 36 (100%)
Large Probe Phase	3 of 3 (100%)	2 of 3 (67%)	36 of 36 (100%)	36 of 36 (100%)
Small Box Gain	3 of 3 (100%)	2 of 3 (67%)	35 of 36 (97%)	32 of 36 (89%)
Small Box Phase	3 of 3 (100%)	3 of 3 (100%)	35 of 36 (97%)	26 of 36 (72%)
Large Box Gain	3 of 3 (100%)	2 of 3 (67%)	34 of 36 (94%)	35 of 36 (97%)
Large Box Phase	2 of 3 (67%)	2 of 3 (67%)	35 of 36 (97%)	34 of 36 (94%)

5.3.2 Quantitative Detection of Ethanol Presence in Air

The following tables detail the R^2 and RMSE values for regressions performed on each sensor, divided into the specific frequency range listed in Table 4.3. Each column represents the specific regression method used (stepwise regression; wavelet transformation followed by stepwise regression; four-, six-, or eight-component partial least squares regression; neural network regression). The rows are the data used in the regression (gain data only; phase data only; gain and phase data together; the interaction between gain and phase data) with ‘T’ standing for the training dataset and ‘V’ for the validation dataset. For the stepwise regressions, cell with three dashes indicate that a model could not be made from the input data.

Appendix K and Appendix L list the frequencies that were used in the stepwise models for the full sets of gain and phase data, respectively.

5.3.2.1 Small Probe Sensor

With stepwise regression and wavelet transformation followed by stepwise regression, the training model fit extremely well ($R^2 = 1.000$) using all frequencies; however, the model did not fit the validation dataset with much consistency ($R^2 < 0.250$). PLS regression showed a significant improvement including more components, reaching a maximum R^2 of 0.988 and 0.606 for the training and validation datasets, respectively. The neural network regression did not show a significant improvement over PLS regression.

Trimming the data to the specific peaks tended to yield a smaller R^2 values for the training datasets, but more reasonable R^2 values for the validations datasets with the stepwise regression and wavelet transformation with stepwise regression. Of these frequency ranges, the R^2 values range from 0.369 to 1.000 and from 0.001 to 0.447 for the training and validation datasets, respectively.

Table 5.3: R² values for ethanol with small probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.791	0.934	0.987	0.529
	Gain – V	0.006	0.061	0.472	0.538	0.525	0.016
	Phase – T	1.000	1.000	0.738	0.940	0.988	0.691
	Phase – V	0.081	0.003	0.547	0.504	0.452	0.036
	Joined – T	1.000	1.000	0.707	0.946	0.987	0.439
	Joined – V	0.159	0.202	0.519	0.536	0.606	0.128
	Interaction – T	1.000	1.000	0.730	0.925	0.975	0.671
	Interaction – V	0.080	0.227	0.534	0.508	0.514	0.252
Peak 1	Gain – T	0.427	0.714	0.608	0.714	0.887	0.712
	Gain – V	0.177	0.239	0.353	0.430	0.227	0.323
	Phase – T	0.369	0.985	0.638	0.835	0.958	0.460
	Phase – V	0.076	0.136	0.333	0.116	0.006	0.182
	Joined – T	0.785	0.710	0.649	0.751	0.924	0.780
	Joined – V	0.060	0.245	0.395	0.444	0.261	0.111
	Interaction – T	0.662	1.000	0.562	0.693	0.831	0.617
	Interaction – V	0.252	0.001	0.417	0.427	0.121	0.117
Peak 2	Gain – T	0.370	0.449	0.576	0.762	0.951	0.406
	Gain – V	0.215	0.236	0.442	0.436	0.334	0.038
	Phase – T	0.416	0.669	0.541	0.787	0.918	0.742
	Phase – V	0.386	0.034	0.447	0.304	0.170	0.425
	Joined – T	0.370	0.619	0.475	0.737	0.924	0.427
	Joined – V	0.215	0.123	0.412	0.429	0.285	0.309
	Interaction – T	0.370	0.891	0.522	0.766	0.904	0.885
	Interaction – V	0.215	0.039	0.426	0.408	0.148	0.240

Table 5.4: RMSE values for ethanol with small probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	1523.0	974.6	453.4	2135.5
	Gain – V	4962.4	5456.9	1982.9	1902.6	1890.2	2857.6
	Phase – T	0.0	0.0	1581.1	916.9	427.7	2008.3
	Phase – V	4615.6	3729.8	1847.3	2245.5	2396.5	4405.5
	Joined – T	0.0	0.0	1710.8	850.4	452.0	2491.7
	Joined – V	4238.6	3820.8	2124.9	1891.8	1747.1	3484.3
	Interaction – T	0.0	0.0	1596.0	1018.1	614.9	1814.8
	Interaction – V	3799.3	3662.6	1966.9	2340.3	2216.3	2514.0
Peak 1	Gain – T	1897.9	1673.9	1814.5	1689.9	1205.1	1729.9
	Gain – V	2016.1	2471.3	1884.9	1898.1	2515.4	1909.6
	Phase – T	1863.2	486.0	1645.7	1360.1	787.3	3721.6
	Phase – V	2611.9	3825.4	1802.8	1992.6	2854.9	3879.4
	Joined – T	1501.3	1669.1	1711.7	1644.8	982.3	1993.4
	Joined – V	2301.8	2071.3	1812.6	1901.8	2045.5	3082.7
	Interaction – T	1720.5	0.0	1697.8	1698.6	1416.3	2215.6
	Interaction – V	2304.4	4791.3	1702.4	1988.0	2609.9	2798.9
Peak 2	Gain – T	1895.9	1789.6	1895.6	1630.3	830.6	3309.3
	Gain – V	2434.9	2142.7	2645.6	2655.8	4062.7	3184.0
	Phase – T	1941.7	1837.6	1849.1	1531.8	1046.9	1753.2
	Phase – V	1966.2	3286.2	2081.7	2396.6	2607.0	2229.1
	Joined – T	1895.9	1847.2	1957.4	1590.1	1012.1	2305.9
	Joined – V	2434.9	2431.2	2267.0	2350.4	2863.5	2228.4
	Interaction – T	1895.9	1195.9	1923.8	1581.2	1125.5	1257.9
	Interaction – V	2434.9	4186.6	1962.7	2178.3	2405.8	3195.7

5.3.2.2 Large Probe Sensor

Using stepwise regression and wavelet transformation followed by stepwise regression when looking at the entire set of frequencies did not fit the model perfectly to the training datasets, as the maximum R^2 was 0.979. Yet, the R^2 value for these validation datasets had an average of 0.159; this means that the models are still over-fitting the training data. Again, the models created by PLS regression were improved by including more components, and the neural network regression tended to give lower R^2 values than even the four-component PLS models.

Selecting the individual peaks gave results similar to the small probe. The training dataset had an R^2 value ranging from 0.297 to 0.965 while the validation dataset had an R^2 value from 0.020 to 0.795.

Table 5.5: R^2 values for ethanol with large probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.671	0.976	0.788	0.931	0.975	0.780
	Gain – V	0.336	0.058	0.576	0.710	0.778	0.373
	Phase – T	0.966	0.974	0.851	0.952	0.980	0.020
	Phase – V	0.188	0.177	0.657	0.739	0.744	0.000
	Joined – T	0.974	0.979	0.813	0.943	0.979	0.645
	Joined – V	0.156	0.018	0.645	0.778	0.798	0.256
	Interaction – T	0.974	0.977	0.839	0.954	0.979	0.561
	Interaction – V	0.172	0.164	0.676	0.752	0.764	0.097
Peak 1	Gain – T	0.712	0.873	0.702	0.801	0.872	0.637
	Gain – V	0.435	0.088	0.471	0.348	0.357	0.251
	Phase – T	---	0.608	0.733	0.924	0.965	0.783
	Phase – V	---	0.020	0.516	0.566	0.565	0.260
	Joined – T	0.712	0.945	0.690	0.814	0.931	0.764
	Joined – V	0.435	0.423	0.571	0.569	0.513	0.303
	Interaction – T	0.712	0.935	0.691	0.819	0.905	0.755
	Interaction – V	0.435	0.542	0.564	0.634	0.618	0.429
Peak 2	Gain – T	0.857	0.641	0.742	0.828	0.943	0.606
	Gain – V	0.736	0.448	0.653	0.730	0.756	0.222
	Phase – T	0.944	0.297	0.744	0.855	0.933	0.700
	Phase – V	0.338	0.094	0.539	0.677	0.709	0.630
	Joined – T	0.857	0.837	0.725	0.889	0.954	0.511
	Joined – V	0.736	0.260	0.576	0.771	0.795	0.104
	Interaction – T	0.857	0.829	0.672	0.859	0.920	0.527
	Interaction – V	0.736	0.225	0.572	0.709	0.684	0.403

Table 5.6: RMSE values for ethanol with large probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1799.3	602.5	1600.6	988.7	610.6	1632.8
	Gain – V	2168.7	3909.1	1823.6	1603.0	1483.4	2214.6
	Phase – T	709.6	619.2	1393.7	844.6	561.1	3564.9
	Phase – V	4276.1	3570.2	1850.1	1689.3	1805.4	2321.7
	Joined – T	631.3	561.1	1528.4	915.6	560.7	2156.8
	Joined – V	3154.1	3130.1	1687.4	1453.8	1460.9	2698.5
	Interaction – T	633.6	594.4	1431.8	814.0	569.5	2470.8
	Interaction – V	3720.7	3303.2	1918.8	1589.2	1647.4	3348.9
Peak 1	Gain – T	1760.6	1291.4	1786.8	1536.2	1318.1	1569.2
	Gain – V	1944.7	2796.6	1988.5	2338.5	2414.3	1341.4
	Phase – T	---	1778.7	1683.5	1039.3	719.1	1511.6
	Phase – V	---	2182.9	1670.0	1888.2	1866.8	1942.2
	Joined – T	1760.6	905.1	1666.5	1538.1	998.7	1754.3
	Joined – V	1944.7	2476.3	1464.3	1895.2	2016.8	2428.1
	Interaction – T	1760.6	963.1	1732.8	1486.9	1160.0	1603.3
	Interaction – V	1944.7	2245.1	1636.2	1709.8	1894.8	1667.6
Peak 2	Gain – T	1344.4	1762.1	1653.3	1463.5	901.5	2015.1
	Gain – V	1869.9	1785.0	1530.5	1776.5	1942.5	1987.1
	Phase – T	890.6	1815.5	1659.7	1363.5	972.0	1261.1
	Phase – V	3267.3	918.5	2482.0	2233.7	2333.1	1344.7
	Joined – T	1344.4	1383.1	1737.0	1211.6	821.3	2122.9
	Joined – V	1869.9	2548.8	2084.8	1872.6	1975.2	1970.2
	Interaction – T	1344.4	1401.8	1840.9	1331.5	1042.7	1919.4
	Interaction – V	1869.9	2814.9	2193.8	2169.7	2590.4	1030.4

5.3.2.3 Small Box Sensor

Like the probe-style sensors, the stepwise and wavelet regression methods over-fit the data for all frequencies, with all training $R^2 = 1.000$ and all validation $R^2 \leq 0.055$. PLS regression shows lower R^2 values for the training dataset (0.768 to 0.996), but with improved R^2 values for the validation dataset (0.009 to 0.385). Neural network regression only had minor success in one case, but all training RMSE values were well above those of the other methods.

Trimming the data to the specific peaks gave generally strong coefficients of determination for the training data, but the models again tend to over-fit the data. From the validation datasets, the highest and average R^2 values were 0.564 and 0.079, respectively.

Table 5.7: R² values for ethanol with small box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.800	0.959	0.993	0.212
	Gain – V	0.012	0.001	0.009	0.059	0.031	0.167
	Phase – T	1.000	1.000	0.838	0.938	0.987	0.171
	Phase – V	0.025	0.048	0.385	0.286	0.268	0.001
	Joined – T	1.000	1.000	0.887	0.969	0.996	0.666
	Joined – V	0.001	0.014	0.257	0.296	0.306	0.000
	Interaction – T	1.000	1.000	0.768	0.966	0.994	0.176
	Interaction – V	0.055	0.019	0.161	0.221	0.174	0.002
Peak 1	Gain – T	0.684	0.584	0.540	0.804	0.950	0.363
	Gain – V	0.018	0.032	0.000	0.053	0.048	0.001
	Phase – T	0.624	0.732	0.680	0.896	0.960	0.151
	Phase – V	0.015	0.008	0.173	0.087	0.060	0.011
	Joined – T	1.000	1.000	0.627	0.869	0.973	0.575
	Joined – V	0.117	0.000	0.096	0.162	0.212	0.002
	Interaction – T	1.000	1.000	0.566	0.918	0.963	0.066
	Interaction – V	0.045	0.010	0.054	0.070	0.009	0.021
Peak 2	Gain – T	0.401	0.428	0.600	0.797	0.889	0.245
	Gain – V	0.003	0.027	0.002	0.000	0.036	0.080
	Phase – T	---	0.666	0.654	0.831	0.932	0.000
	Phase – V	---	0.042	0.000	0.001	0.002	0.037
	Joined – T	0.930	0.979	0.654	0.907	0.965	0.648
	Joined – V	0.002	0.002	0.026	0.081	0.121	0.002
	Interaction – T	0.764	1.000	0.612	0.876	0.896	0.301
	Interaction – V	0.016	0.000	0.022	0.026	0.014	0.016
Peak 3	Gain – T	0.685	0.913	0.569	0.902	0.970	0.063
	Gain – V	0.037	0.084	0.248	0.292	0.273	0.020
	Phase – T	0.473	0.487	0.672	0.851	0.940	0.483
	Phase – V	0.078	0.008	0.178	0.037	0.011	0.000
	Joined – T	0.855	1.000	0.638	0.884	0.956	0.474
	Joined – V	0.012	0.050	0.521	0.564	0.374	0.011
	Interaction – T	0.829	1.000	0.727	0.884	0.945	0.364
	Interaction – V	0.124	0.063	0.241	0.259	0.180	0.074

Table 5.8: RMSE values for ethanol with small box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	1529.3	769.7	324.0	3443.7
	Gain – V	3672.0	3499.1	1509.8	1774.8	1806.4	3237.4
	Phase – T	0.0	0.0	1389.2	956.3	453.9	3222.4
	Phase – V	7333.3	4100.0	1858.0	2027.5	1987.3	4246.5
	Joined – T	0.0	0.0	1201.4	685.6	264.3	2162.0
	Joined – V	2619.8	4211.3	2369.8	2217.2	2131.2	2333.2
	Interaction – T	0.0	0.0	1674.1	713.3	308.0	2647.1
	Interaction – V	2731.7	5185.0	1698.3	1887.9	2025.3	4085.6
Peak 1	Gain – T	1787.9	1912.4	1864.8	1539.1	826.1	3029.8
	Gain – V	3497.1	3492.9	2213.1	2449.8	3097.1	4390.9
	Phase – T	1789.7	1674.8	1687.7	1172.0	755.4	2866.0
	Phase – V	3508.0	3393.1	2657.9	2609.9	2497.0	3451.1
	Joined – T	0.0	0.0	1816.8	1312.1	627.0	2464.0
	Joined – V	2794.9	5051.8	3042.2	2688.6	2766.3	2034.3
	Interaction – T	0.0	0.0	1844.7	1051.5	733.5	3748.9
	Interaction – V	4732.6	4874.1	3030.9	3064.7	2935.0	3155.9
Peak 2	Gain – T	1811.9	1939.1	1849.1	1576.0	1216.3	2212.1
	Gain – V	1486.6	2489.0	1804.8	2598.6	2784.2	2451.1
	Phase – T	---	1752.9	1843.3	1427.5	969.5	2394.5
	Phase – V	---	4880.0	2480.7	3165.9	3872.0	3081.5
	Joined – T	985.5	569.5	1840.2	1115.4	714.8	2304.4
	Joined – V	3962.7	5254.1	3010.6	3090.5	2949.3	6103.7
	Interaction – T	1510.0	5.5	1876.6	1209.3	1153.2	3278.9
	Interaction – V	2115.3	4906.4	2406.3	3545.1	4261.1	2646.0
Peak 3	Gain – T	1691.8	1070.2	1886.0	1129.7	663.1	2839.7
	Gain – V	3026.6	3380.9	1351.9	1906.1	2273.8	3028.7
	Phase – T	1906.6	1892.9	1572.9	1401.0	919.3	3077.2
	Phase – V	1904.0	2247.2	1469.9	2299.5	2780.0	2590.2
	Joined – T	1324.7	0.0	1708.8	1224.9	790.6	2277.6
	Joined – V	2740.5	3787.6	2014.0	1909.1	2144.9	2464.7
	Interaction – T	1377.8	0.0	1692.3	1245.8	883.5	2584.0
	Interaction – V	2471.0	3401.6	2167.7	2278.4	2845.5	2833.3

5.3.2.4 Large Box Sensor

Just as the small box sensor had several R^2 values above 0.980, the large box sensor had many strong coefficients of determination when looking at all frequencies. However, the validation datasets had a much stronger fit than those of the small box sensor, with the R^2 up to

0.784, with an average of 0.564. Also, the neural network regression showed a strong R^2 value (0.808 for training and 0.660 for validation) for gain and phase data combined.

Table 5.9: R^2 values for ethanol with large box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.691	0.932	0.985	0.466
	Gain – V	0.454	0.144	0.625	0.687	0.748	0.011
	Phase – T	1.000	1.000	0.843	0.956	0.988	0.357
	Phase – V	0.521	0.274	0.729	0.743	0.771	0.036
	Joined – T	1.000	1.000	0.850	0.950	0.989	0.808
	Joined – V	0.547	0.632	0.759	0.740	0.784	0.660
	Interaction – T	1.000	1.000	0.799	0.940	0.985	0.584
	Interaction – V	0.632	0.555	0.710	0.738	0.777	0.258
Peak 1	Gain – T	0.237	0.452	0.636	0.811	0.947	0.249
	Gain – V	0.057	0.066	0.242	0.130	0.067	0.016
	Phase – T	0.610	0.783	0.693	0.885	0.942	0.800
	Phase – V	0.212	0.218	0.494	0.356	0.257	0.275
	Joined – T	0.911	0.799	0.715	0.900	0.976	0.540
	Joined – V	0.133	0.214	0.354	0.307	0.285	0.251
	Interaction – T	0.920	0.879	0.631	0.847	0.911	0.636
	Interaction – V	0.155	0.185	0.440	0.400	0.308	0.284
Peak 2	Gain – T	0.247	0.365	0.327	0.567	0.772	0.141
	Gain – V	0.110	0.038	0.224	0.090	0.023	0.001
	Phase – T	0.130	0.685	0.703	0.802	0.910	0.758
	Phase – V	0.136	0.297	0.591	0.683	0.668	0.211
	Joined – T	0.247	0.822	0.638	0.762	0.941	0.720
	Joined – V	0.110	0.414	0.529	0.675	0.658	0.279
	Interaction – T	0.247	0.866	0.553	0.675	0.859	0.791
	Interaction – V	0.110	0.078	0.486	0.555	0.503	0.546
Peak 3	Gain – T	0.676	0.666	0.576	0.739	0.895	0.669
	Gain – V	0.435	0.450	0.362	0.437	0.388	0.401
	Phase – T	0.873	0.829	0.531	0.717	0.891	0.457
	Phase – V	0.630	0.360	0.447	0.566	0.610	0.445
	Joined – T	0.820	0.999	0.601	0.795	0.903	0.837
	Joined – V	0.331	0.460	0.533	0.628	0.653	0.526
	Interaction – T	0.820	0.996	0.547	0.764	0.888	0.778
	Interaction – V	0.331	0.115	0.409	0.503	0.401	0.451

Isolating the individual peaks did not show an improvement compared to the full range of frequencies, with the smallest and maximum training RMSE values of 132.9 and 3639.0, respectively. This is far greater than that of using the entire frequency range, and the validation statistics follow a similar pattern.

Table 5.10: RMSE values for ethanol with large box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	1735.2	954.9	479.4	2625.7
	Gain – V	4381.9	3367.8	1920.7	1704.9	1499.6	4431.1
	Phase – T	0.0	0.0	1417.6	810.1	427.1	1884.9
	Phase – V	2601.8	3854.2	1713.8	1757.4	1590.6	3264.3
	Joined – T	0.0	0.0	1355.0	856.1	416.9	1740.4
	Joined – V	2217.5	2196.6	1541.9	1704.8	1479.6	2322.1
	Interaction – T	0.0	0.0	1540.1	936.7	486.1	2003.2
	Interaction – V	2141.3	2662.7	1709.6	1798.7	1677.3	2715.2
Peak 1	Gain – T	1654.3	1901.2	1867.4	1531.0	889.4	3639.0
	Gain – V	1748.4	3614.8	3134.6	5037.4	7597.4	5061.6
	Phase – T	1937.3	1603.3	1766.4	1230.1	905.8	1782.0
	Phase – V	3111.2	4242.7	2002.3	2707.1	3254.3	3277.7
	Joined – T	1094.2	1579.6	1785.1	1165.7	587.3	2218.5
	Joined – V	4850.2	4221.0	2925.0	3534.0	3801.2	2851.0
	Interaction – T	1047.6	1285.0	1871.5	1410.4	1125.7	2009.5
	Interaction – V	4270.5	3590.6	2442.3	2989.1	3250.2	2169.5
Peak 2	Gain – T	1714.9	1784.8	1831.4	1871.2	1628.4	1925.1
	Gain – V	1813.1	1973.8	1706.7	2476.3	3083.2	2544.6
	Phase – T	1337.0	1706.2	1752.4	1500.7	1115.8	1271.6
	Phase – V	1340.5	3077.3	2003.9	1754.1	2014.4	2853.6
	Joined – T	1714.9	1437.9	1534.4	1641.7	904.2	1769.4
	Joined – V	1813.1	3096.5	1819.1	1801.1	2209.5	3952.5
	Interaction – T	1714.9	1278.3	1712.4	1788.9	1351.0	1639.0
	Interaction – V	1813.1	5048.8	1925.1	2088.2	2478.4	2359.5
Peak 3	Gain – T	1741.4	1840.4	1662.8	1644.1	1170.7	2314.8
	Gain – V	2254.1	2429.4	2265.9	2377.2	3101.8	3050.8
	Phase – T	1215.0	1488.5	1834.6	1738.7	1178.4	2188.9
	Phase – V	2511.1	2677.6	1985.8	1947.1	2258.9	2273.9
	Joined – T	1418.9	132.9	1841.5	1542.7	1147.5	1502.2
	Joined – V	2872.5	3169.3	1930.0	1956.2	2003.8	2453.2
	Interaction – T	1418.9	254.1	1811.2	1660.0	1217.1	1209.2
	Interaction – V	2872.5	4651.0	2036.2	2182.7	2898.8	1959.3

5.4 Ammonia Tests

5.4.1 Qualitative Detection of Ammonia Presence in Air

With the same procedure for detecting glycerol in Section 5.2, each of the gain and phase measurements were compared with ambient air and air with 300 ppm ammonia, a standard set forth by the Occupational Safety and Health Administration (2003). An overview of the distribution is given in Figure 5.17 and Figure 5.18 with the full list given in Appendix J.

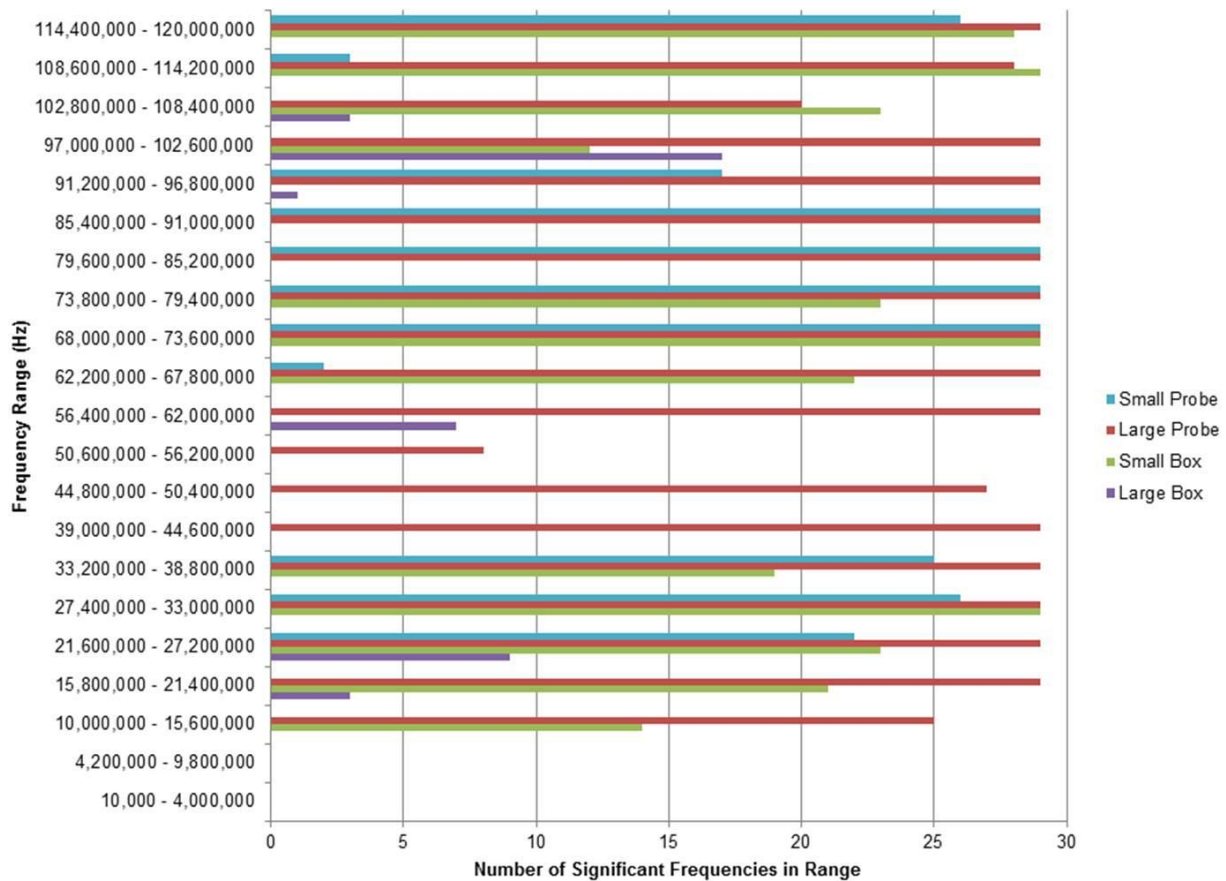


Figure 5.17: Distribution of significant frequencies with ammonia, gain data

This distribution shows that there were far more significantly different gain measurements found with ammonia than with either glycerol or ethanol. The small probe sensor had significantly different measurements at all frequencies between 19,800,000 Hz and 26,800,000 Hz, 60,400,000 Hz and 72,200,000 Hz, and 76,600,000 Hz and 110,600,000 Hz. With significant differences at 514 frequencies, the large probe sensor had major contiguous blocks from 11,000,000 Hz to 50,000,000 Hz and from 55,800,000 Hz to 106,000,000 Hz. The

small box sensor showed several, albeit discontinuous, blocks of frequencies with significant differences, whereas the large box sensor had significantly different measurements at 40 different frequencies in five smaller, contiguous blocks.

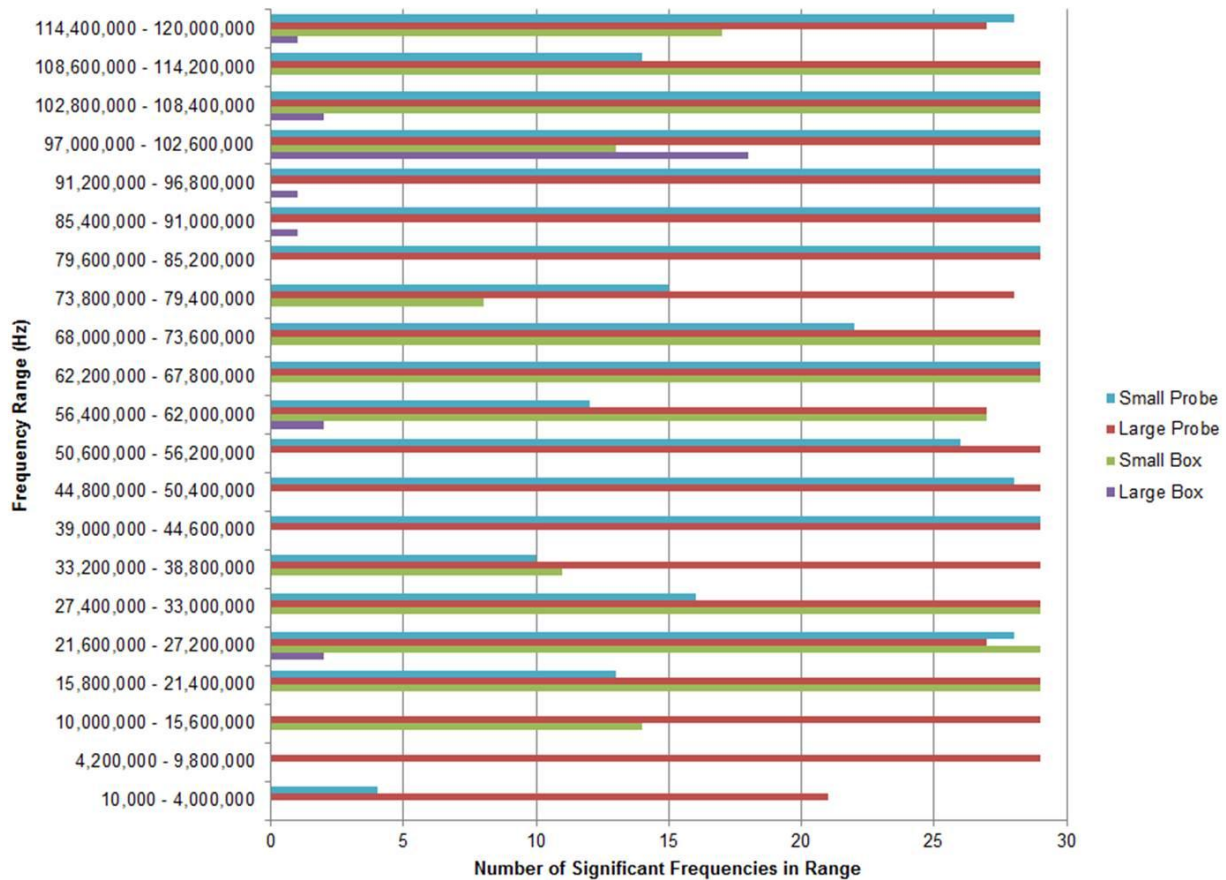


Figure 5.18: Distribution of significant frequencies with ammonia, phase data

Phase data had many significant differences in measurements with and without ammonia as well. The small probe, large probe, small box, and large box sensors had significant differences at 419, 594, 293, and 27 frequencies, respectively. Most of these frequencies are in large, contiguous blocks with only a few insignificant measurements scattered throughout. The largest contiguous block is with the large probe sensor, which had significant readings between 78,600,000 Hz and 116,000,000 Hz. In fact, with the exception of 58,600,000 Hz, 58,800,000 Hz, and 78,400,000 Hz, all frequencies between 25,600,000 Hz and 116,000,000 Hz have significantly different phase readings between ambient air and air with 300 ppm ammonia.

Like detecting ethanol, six iterations for each concentration in Table 4.2 were used in the detection method for ammonia. Only one false positive was detected, and ammonia was successfully detected in 559 of 576 cases. The false negatives tended to occur with at the lower and mid-range concentrations (50 ppm to 900 ppm), although there was not a specific pattern of false negatives. Nonetheless, ammonia was detected with a 97.1% overall success rate.

Table 5.11: Success rates of neural network for detecting ammonia

	Ambient Air Training	Ambient Air Validation	Air with Ammonia Training	Air with Ammonia Validation
Small Probe Gain	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	36 of 36 (100%)
Small Probe Phase	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	36 of 36 (100%)
Large Probe Gain	3 of 3 (100%)	3 of 3 (100%)	35 of 36 (97%)	36 of 36 (100%)
Large Probe Phase	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	36 of 36 (100%)
Small Box Gain	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	34 of 36 (94%)
Small Box Phase	3 of 3 (100%)	3 of 3 (100%)	36 of 36 (100%)	36 of 36 (100%)
Large Box Gain	3 of 3 (100%)	3 of 3 (100%)	34 of 36 (94%)	30 of 36 (83%)
Large Box Phase	3 of 3 (100%)	2 of 3 (67%)	33 of 36 (92%)	31 of 36 (86%)

5.4.2 Quantitative Detection of Ammonia Presence in Air

The frequency response data from the various ammonia concentrations is run through the same four regression programs, with the partial least squares program used with four, six, and eight components. As described fully in Section 5.3.2, each column corresponds to the specific regression methods and each row is the data used in the regression. Again, ‘T’ stands for the training dataset; ‘V’, the validation dataset.

Appendix M and Appendix N give the frequencies used in the stepwise regression model for the gain and phase data, respectively.

5.4.2.1 Small Probe Sensor

Considering the entire range of frequencies, stepwise regression and wavelet transformation followed by stepwise regression had very high R^2 values for the training dataset, and R^2 values ranging from 0.048 to 0.648 for the validation dataset. PLS regression tended to improve with an increasing number of components, and neural network regression gave RMSE values in the 380.8 to 734.2 range.

Analyzing the two peaks individually tended to reduce the training R^2 values, but with the advantage of improving the validation R^2 values. A very significant data subset was the gain and phase data together for the second peak; neural network regression gave training and validation R^2 values of 0.991 and 0.817, respectively.

Table 5.12: R^2 values for ammonia with small probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.533	0.750	0.961	0.487
	Gain – V	0.304	0.330	0.117	0.185	0.334	0.122
	Phase – T	1.000	1.000	0.453	0.901	0.970	0.798
	Phase – V	0.357	0.048	0.086	0.098	0.137	0.311
	Joined – T	1.000	1.000	0.426	0.873	0.973	0.561
	Joined – V	0.468	0.363	0.127	0.104	0.197	0.295
	Interaction – T	0.718	1.000	0.423	0.785	0.963	0.563
	Interaction – V	0.648	0.369	0.134	0.054	0.106	0.094
Peak 1	Gain – T	0.904	0.790	0.438	0.761	0.917	0.739
	Gain – V	0.387	0.389	0.207	0.146	0.175	0.384
	Phase – T	0.983	0.803	0.621	0.856	0.943	0.963
	Phase – V	0.500	0.038	0.078	0.303	0.416	0.417
	Joined – T	1.000	1.000	0.441	0.794	0.929	0.846
	Joined – V	0.407	0.248	0.028	0.053	0.161	0.344
	Interaction – T	0.947	1.000	0.453	0.746	0.853	0.946
	Interaction – V	0.383	0.272	0.065	0.186	0.184	0.522
Peak 2	Gain – T	0.718	0.627	0.260	0.856	0.954	0.782
	Gain – V	0.690	0.125	0.295	0.457	0.562	0.581
	Phase – T	0.763	0.184	0.526	0.760	0.917	0.943
	Phase – V	0.689	0.173	0.017	0.028	0.004	0.686
	Joined – T	0.763	0.930	0.437	0.735	0.926	0.991
	Joined – V	0.689	0.001	0.091	0.102	0.008	0.817
	Interaction – T	0.718	0.688	0.421	0.838	0.907	0.988
	Interaction – V	0.648	0.009	0.006	0.003	0.006	0.800

Table 5.13: RMSE values for ammonia with small probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	415.3	390.9	174.3	629.0
	Gain – V	845.2	654.9	626.4	605.0	582.1	710.5
	Phase – T	0.0	0.0	381.0	273.5	155.9	380.8
	Phase – V	758.2	880.4	566.5	832.9	792.3	615.4
	Joined – T	0.0	0.0	342.3	304.0	148.0	440.7
	Joined – V	586.4	851.7	442.1	748.1	751.8	734.2
	Interaction – T	332.5	0.0	365.3	373.7	174.0	534.4
	Interaction – V	397.4	842.5	499.4	783.9	905.5	592.7
Peak 1	Gain – T	253.9	339.2	434.5	369.3	246.6	340.8
	Gain – V	745.1	680.3	406.4	531.7	535.4	554.8
	Phase – T	118.6	335.1	432.6	315.4	210.0	169.8
	Phase – V	717.6	761.2	436.9	554.9	582.5	624.6
	Joined – T	0.0	0.0	406.2	360.3	231.9	321.4
	Joined – V	713.8	955.9	460.7	545.3	570.8	883.3
	Interaction – T	196.2	0.0	415.4	382.1	322.0	205.9
	Interaction – V	694.3	970.0	459.9	526.6	640.8	555.6
Peak 2	Gain – T	385.4	413.7	300.4	303.8	191.6	304.6
	Gain – V	353.0	589.2	318.3	458.6	390.6	378.9
	Phase – T	366.3	342.9	408.2	383.5	246.4	194.3
	Phase – V	360.0	385.3	554.2	775.9	710.2	413.9
	Joined – T	366.3	228.2	342.6	392.6	236.8	84.2
	Joined – V	360.0	667.4	389.9	712.8	775.7	363.5
	Interaction – T	332.5	421.2	447.6	328.2	263.0	96.2
	Interaction – V	397.4	710.0	674.3	717.3	740.9	409.7

5.4.2.2 Large Probe Sensor

Like many other sensor-contaminant combinations, the training R^2 values for the stepwise regression and wavelet transformation followed by stepwise regression were 1.000 when analyzing the entire range of frequencies. However, when validated, these models generally had lower RMSE values than with ethanol. With a few exceptions, the RMSE values tended to decrease as more components were used in the PLS regressions, and neural network regression gave R^2 values that, for the majority of regressions, were lower than any other methods.

The strongest R^2 values of any regression were found using the neural network regression with the gain and phase data combined from the second peak; the training and validation R^2 values were 1.000 and 0.996, respectively, with both RMSE values below 55. All combinations of data from this peak had R^2 values had training and validation R^2 values at or above 0.968.

Table 5.14: R^2 values for ammonia with large probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.538	0.591	0.937	0.612
	Gain – V	0.309	0.151	0.521	0.581	0.862	0.545
	Phase – T	1.000	1.000	0.611	0.627	0.884	0.952
	Phase – V	0.648	0.625	0.592	0.586	0.844	0.898
	Joined – T	1.000	1.000	0.599	0.640	0.892	0.319
	Joined – V	0.588	0.403	0.592	0.606	0.851	0.359
	Interaction – T	1.000	1.000	0.598	0.741	0.889	0.288
	Interaction – V	0.651	0.444	0.615	0.694	0.861	0.287
Peak 1	Gain – T	---	0.705	0.431	0.852	0.913	0.774
	Gain – V	---	0.637	0.529	0.837	0.861	0.735
	Phase – T	0.480	0.916	0.554	0.786	0.934	0.589
	Phase – V	0.561	0.549	0.600	0.680	0.888	0.601
	Joined – T	0.480	0.974	0.563	0.789	0.890	0.993
	Joined – V	0.561	0.581	0.606	0.701	0.837	0.964
	Interaction – T	0.480	1.000	0.550	0.870	0.917	0.628
	Interaction – V	0.561	0.396	0.575	0.811	0.862	0.986
Peak 2	Gain – T	0.377	0.450	0.394	0.610	0.843	0.996
	Gain – V	0.359	0.311	0.391	0.488	0.523	0.992
	Phase – T	0.884	0.918	0.467	0.660	0.883	0.980
	Phase – V	0.541	0.635	0.460	0.606	0.782	0.980
	Joined – T	0.986	0.975	0.499	0.616	0.802	1.000
	Joined – V	0.635	0.506	0.494	0.591	0.770	0.996
	Interaction – T	0.358	0.960	0.489	0.753	0.953	0.968
	Interaction – V	0.355	0.223	0.485	0.720	0.899	0.974

Table 5.15: RMSE values for ammonia with large probe sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	458.8	440.7	215.2	460.1
	Gain – V	1023.6	676.9	470.8	444.0	290.0	492.6
	Phase – T	0.0	0.0	424.2	438.5	287.3	184.4
	Phase – V	539.7	566.7	429.5	461.1	303.7	242.0
	Joined – T	0.0	0.0	443.0	436.8	279.0	695.6
	Joined – V	543.3	671.7	432.3	448.8	296.3	622.0
	Interaction – T	0.0	0.0	440.9	386.4	287.5	315.3
	Interaction – V	466.5	612.6	415.2	413.4	289.0	337.6
Peak 1	Gain – T	---	388.3	455.0	309.9	244.0	420.4
	Gain – V	---	453.7	368.0	294.9	267.2	444.2
	Phase – T	456.7	253.3	446.0	356.8	227.1	464.5
	Phase – V	379.0	551.1	396.9	418.7	263.7	400.1
	Joined – T	456.7	140.9	443.3	345.1	282.0	75.1
	Joined – V	379.0	474.9	389.5	386.9	305.8	161.0
	Interaction – T	456.7	0.0	455.4	285.2	251.2	690.0
	Interaction – V	379.0	736.8	424.1	311.5	275.8	95.4
Peak 2	Gain – T	446.0	446.5	449.7	446.6	323.8	55.4
	Gain – V	448.2	565.7	444.5	511.1	529.4	79.1
	Phase – T	274.4	250.9	458.7	409.4	286.2	125.2
	Phase – V	478.2	471.6	458.3	430.6	350.4	121.7
	Joined – T	106.1	142.0	456.5	394.8	365.2	20.1
	Joined – V	496.3	469.1	452.6	402.9	365.3	54.9
	Interaction – T	441.3	174.7	459.0	396.2	193.3	155.9
	Interaction – V	438.5	659.7	453.7	410.5	249.7	143.9

These regressions show a clear difference between the regression methods that tended to over-fit the training data and those that did not. Using the gain data from all frequencies, it is clear that the wavelet transformation followed by stepwise regression over-fit the data while the eight-component PLS regression did not. As shown in Figure 5.19, the wavelet transformation followed by stepwise regression model overestimates the lower concentrations (750 ppm and below) and underestimates the higher concentrations (900 ppm and above) when analyzing the validation dataset. Figure 5.20 illustrates how the PLS regression results in training data points that do not follow a near-exact trend line, but the validation data points have a much smaller deviation from the model established with the training data.

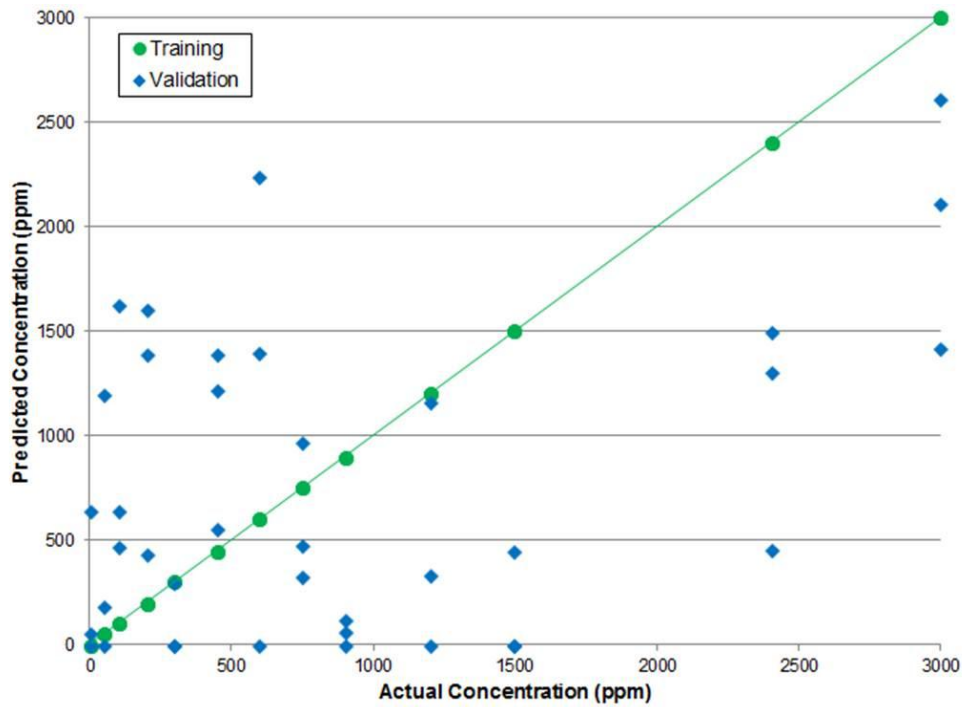


Figure 5.19: Comparison of actual and predicted ammonia concentrations using wavelet transformation followed by stepwise regression on data from large probe sensor

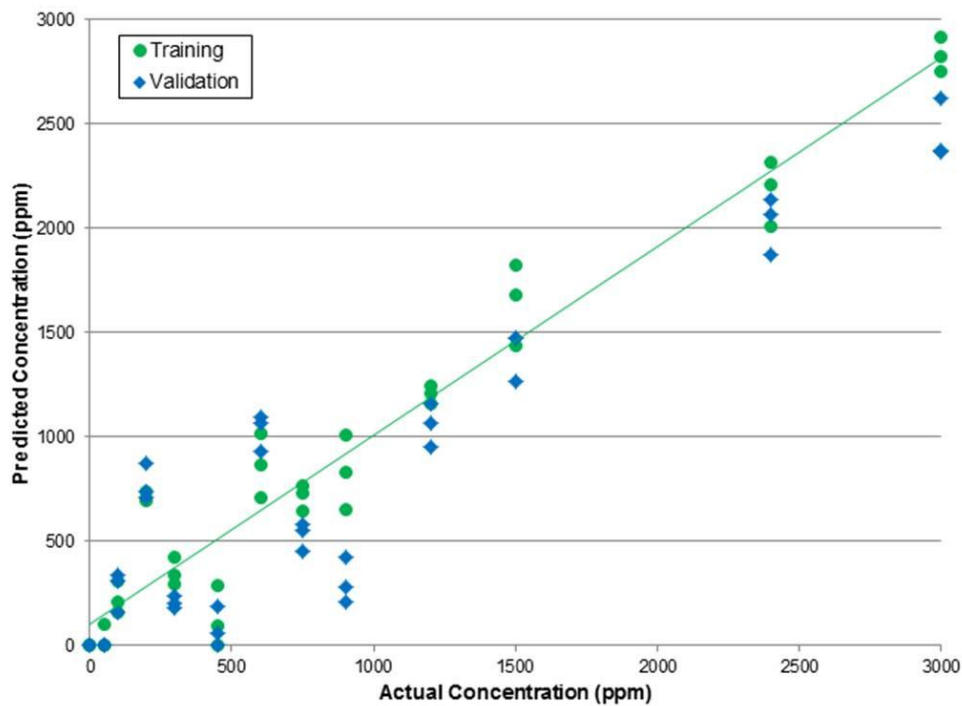


Figure 5.20: Comparison of actual and predicted ammonia concentrations with PLS regression on data from large probe sensor

5.4.2.3 Small Box Sensor

When analyzing the full range of frequencies, the training models tend to over-fit the data, as most of the training R^2 values were above 0.900 and only one validation R^2 value was above 0.100. Like the training results with the ethanol tests, the stepwise and wavelet transformation followed by stepwise regression all had R^2 values of 1.000, the PLS regressions increase the R^2 values with more components, and neural network regression tended to have the lowest R^2 values and highest RMSE values of any regression technique.

Limiting the data to the specific peak produced very mixed results. For stepwise regression, the R^2 values had a maximum of 0.339 for training and 0.081 for validation, with no models being made for any peaks with phase only data. Among the other regression methods, the training R^2 values had minimum, average, and maximum values of 0.181, 0.755, and 1.000, respectively; the validation R^2 values, 0.000, 0.036, and 0.252, respectively.

Table 5.16: R² values for ammonia with small box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	1.000	1.000	0.719	0.892	0.987	0.188
	Gain – V	0.022	0.004	0.050	0.031	0.034	0.009
	Phase – T	1.000	1.000	0.729	0.926	0.981	0.546
	Phase – V	0.174	0.002	0.005	0.021	0.011	0.041
	Joined – T	1.000	1.000	0.749	0.926	0.987	0.693
	Joined – V	0.003	0.015	0.015	0.040	0.032	0.067
	Interaction – T	1.000	1.000	0.656	0.845	0.976	0.503
	Interaction – V	0.000	0.047	0.027	0.031	0.034	0.029
Peak 1	Gain – T	0.188	0.666	0.644	0.859	0.986	0.602
	Gain – V	0.000	0.004	0.076	0.112	0.063	0.138
	Phase – T	---	0.942	0.621	0.902	0.964	0.807
	Phase – V	---	0.000	0.003	0.006	0.008	0.031
	Joined – T	0.188	1.000	0.671	0.895	0.974	0.193
	Joined – V	0.000	0.001	0.034	0.004	0.003	0.031
	Interaction – T	0.188	1.000	0.608	0.890	0.972	0.279
	Interaction – V	0.000	0.039	0.007	0.013	0.028	0.252
Peak 2	Gain – T	0.135	0.181	0.561	0.824	0.940	0.592
	Gain – V	0.001	0.001	0.007	0.043	0.008	0.020
	Phase – T	---	0.273	0.530	0.822	0.912	0.925
	Phase – V	---	0.016	0.040	0.000	0.002	0.009
	Joined – T	0.135	0.870	0.572	0.858	0.986	0.903
	Joined – V	0.001	0.004	0.052	0.008	0.001	0.054
	Interaction – T	0.135	1.000	0.340	0.789	0.962	0.914
	Interaction – V	0.001	0.023	0.089	0.015	0.005	0.000
Peak 3	Gain – T	0.339	0.737	0.505	0.776	0.965	0.932
	Gain – V	0.081	0.028	0.063	0.008	0.001	0.058
	Phase – T	---	0.802	0.572	0.754	0.897	0.977
	Phase – V	---	0.040	0.015	0.095	0.098	0.000
	Joined – T	0.339	0.701	0.493	0.799	0.930	0.704
	Joined – V	0.081	0.117	0.046	0.007	0.074	0.033
	Interaction – T	0.339	0.700	0.421	0.637	0.850	0.931
	Interaction – V	0.081	0.060	0.087	0.005	0.114	0.008

Table 5.17: RMSE values for ammonia with small box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.0	0.0	401.9	271.7	103.3	816.6
	Gain – V	899.8	867.0	861.0	901.0	876.0	1071.3
	Phase – T	0.0	0.0	401.0	231.1	124.7	453.3
	Phase – V	601.4	920.9	724.1	741.7	727.7	555.3
	Joined – T	0.0	0.0	391.1	229.6	103.5	430.1
	Joined – V	797.6	943.4	825.9	777.0	790.4	725.7
	Interaction – T	0.0	0.0	416.2	326.9	139.1	646.8
	Interaction – V	865.2	979.5	651.3	739.2	631.1	815.0
Peak 1	Gain – T	359.6	396.8	418.8	316.1	105.9	619.2
	Gain – V	289.0	710.0	550.2	520.4	648.2	655.0
	Phase – T	---	212.9	422.5	259.1	167.9	362.1
	Phase – V	---	793.7	582.5	567.0	557.4	515.2
	Joined – T	359.6	0.0	412.1	265.6	144.5	623.3
	Joined – V	289.0	997.4	647.8	673.5	636.2	482.7
	Interaction – T	359.6	0.0	420.6	276.0	149.5	640.4
	Interaction – V	289.0	878.6	519.9	508.7	447.6	479.3
Peak 2	Gain – T	305.8	354.4	428.2	344.0	214.6	579.0
	Gain – V	297.3	302.8	427.3	526.6	631.7	740.3
	Phase – T	---	409.7	459.3	348.6	259.0	234.5
	Phase – V	---	565.2	785.3	1122.5	1282.3	739.5
	Joined – T	305.8	296.5	455.0	320.5	107.8	284.4
	Joined – V	297.3	1158.2	946.3	1222.8	1283.5	595.1
	Interaction – T	305.8	16.2	435.9	372.0	170.1	274.4
	Interaction – V	297.3	1496.3	706.7	987.9	1353.0	632.5
Peak 3	Gain – T	414.9	389.1	457.8	354.1	167.0	220.5
	Gain – V	559.3	563.3	685.7	699.8	808.3	953.9
	Phase – T	---	360.9	450.8	395.2	276.8	131.3
	Phase – V	---	976.1	809.4	861.7	949.8	756.7
	Joined – T	414.9	405.9	459.0	368.6	229.3	427.2
	Joined – V	559.3	809.5	683.2	845.3	1142.8	369.8
	Interaction – T	414.9	416.0	454.3	442.4	328.8	234.8
	Interaction – V	559.3	832.2	610.6	751.5	531.7	673.7

5.4.2.4 Large Box Sensor

The regressions for the entire range of frequencies have R^2 values that followed the same pattern as many of the other sensor-contaminant combinations. Stepwise regression and wavelet transformation followed by stepwise regression all had R^2 values above 0.900 with training data,

yet only one R^2 value above 0.200 with validation data. Increasing the number of PLS components from four to six to eight generally improved the average R^2 values from 0.433 to 0.817 to 0.924 with training data, and from 0.333 to 0.382 to 0.348 with validation data. Neural network regression had the best R^2 values with the interaction data, with 0.664 for the training data and 0.427 for the validation data.

Regression on the individual peaks had very mixed results, just as it did with the small box sensor with ammonia. No stepwise models were created for any of the data for the individual peaks. R^2 values for the wavelet transformation followed by stepwise regression had an average of 0.609 and 0.169 for training and validation data, respectively. Training RMSE values with PLS regressions tended to decrease with more components used; validation RMSE values followed the opposite trend. Lastly, a great improvement was with the neural network on the third peak; the average R^2 values were 0.922 and 0.743 for training and validation data, respectively.

Table 5.18: R² values for ammonia with large box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	0.914	1.000	0.412	0.764	0.938	0.525
	Gain – V	0.186	0.169	0.280	0.277	0.343	0.171
	Phase – T	0.981	1.000	0.419	0.788	0.904	0.032
	Phase – V	0.007	0.026	0.344	0.422	0.322	0.103
	Joined – T	1.000	1.000	0.428	0.890	0.944	0.195
	Joined – V	0.187	0.292	0.348	0.397	0.372	0.032
	Interaction – T	1.000	1.000	0.473	0.825	0.911	0.664
	Interaction – V	0.108	0.127	0.361	0.430	0.354	0.427
Peak 1	Gain – T	---	0.688	0.476	0.778	0.952	0.326
	Gain – V	---	0.191	0.217	0.228	0.211	0.198
	Phase – T	---	0.118	0.381	0.789	0.909	0.560
	Phase – V	---	0.208	0.152	0.242	0.253	0.495
	Joined – T	---	0.847	0.593	0.770	0.950	0.035
	Joined – V	---	0.292	0.153	0.440	0.400	0.128
	Interaction – T	---	0.884	0.290	0.660	0.853	0.063
	Interaction – V	---	0.142	0.170	0.205	0.200	0.008
Peak 2	Gain – T	---	0.313	0.507	0.723	0.889	0.666
	Gain – V	---	0.122	0.403	0.455	0.407	0.177
	Phase – T	---	0.742	0.508	0.683	0.871	0.435
	Phase – V	---	0.050	0.342	0.238	0.183	0.272
	Joined – T	---	0.204	0.417	0.675	0.883	0.815
	Joined – V	---	0.114	0.419	0.275	0.243	0.480
	Interaction – T	---	0.969	0.455	0.673	0.846	0.719
	Interaction – V	---	0.257	0.450	0.420	0.316	0.222
Peak 3	Gain – T	---	0.727	0.229	0.709	0.839	0.901
	Gain – V	---	0.096	0.236	0.312	0.350	0.702
	Phase – T	---	0.241	0.397	0.616	0.842	0.949
	Phase – V	---	0.191	0.304	0.491	0.337	0.753
	Joined – T	---	0.764	0.435	0.669	0.817	0.941
	Joined – V	---	0.137	0.317	0.400	0.381	0.725
	Interaction – T	---	0.811	0.431	0.599	0.813	0.897
	Interaction – V	---	0.228	0.301	0.420	0.384	0.790

Table 5.19: RMSE values for ammonia with large box sensor

Data in Model		Stepwise	Wavelet	PLS 4	PLS 6	PLS 8	Neural Network
All Frequencies	Gain – T	250.1	0.0	431.0	361.5	213.2	476.6
	Gain – V	1691.6	1472.9	759.9	2049.1	2205.3	949.1
	Phase – T	123.6	0.0	420.9	314.2	262.1	385.3
	Phase – V	965.1	1710.5	599.9	982.8	1948.0	1003.1
	Joined – T	0.0	0.0	444.6	269.0	205.2	644.8
	Joined – V	1011.6	954.1	699.3	1646.6	1966.7	619.4
	Interaction – T	0.0	0.0	425.7	302.5	251.5	545.5
	Interaction – V	1119.8	1231.8	756.1	1302.9	1979.5	734.1
Peak 1	Gain – T	---	392.1	444.7	363.2	190.1	586.5
	Gain – V	---	991.4	1494.7	1561.7	1690.4	455.4
	Phase – T	---	296.4	406.2	347.0	258.8	363.5
	Phase – V	---	443.0	745.7	1388.6	1718.2	483.7
	Joined – T	---	305.2	342.2	358.7	193.5	545.2
	Joined – V	---	1126.4	433.9	624.7	1279.1	673.2
	Interaction – T	---	283.5	378.0	409.0	306.0	314.9
	Interaction – V	---	1529.1	614.0	1637.3	1577.5	330.0
Peak 2	Gain – T	---	373.7	439.8	402.4	269.3	389.0
	Gain – V	---	630.2	507.0	835.4	581.8	647.2
	Phase – T	---	387.4	444.7	415.2	290.9	345.2
	Phase – V	---	2694.0	1167.4	2006.0	2784.9	483.4
	Joined – T	---	370.6	452.6	409.7	289.8	455.9
	Joined – V	---	670.1	583.4	2000.5	2461.3	726.1
	Interaction – T	---	158.6	442.7	410.1	312.9	459.1
	Interaction – V	---	1385.1	564.7	973.0	1350.1	766.3
Peak 3	Gain – T	---	384.9	386.6	392.7	330.7	253.0
	Gain – V	---	896.0	643.0	1456.1	1261.0	502.7
	Phase – T	---	393.6	423.6	404.7	320.1	200.9
	Phase – V	---	383.9	537.7	556.5	1912.8	349.3
	Joined – T	---	383.5	433.5	410.9	337.8	212.5
	Joined – V	---	1280.1	712.0	1094.8	1588.8	465.4
	Interaction – T	---	333.9	425.5	397.8	345.5	262.3
	Interaction – V	---	1284.1	714.8	677.9	1399.5	444.5

5.5 Combined Tests

In addition to analyzing the three contaminants individually, the data from all tests were pooled together and run through two different neural network analyses. The first was to recognize the presence of contaminants with the second to identify which contaminant was present.

5.5.1 Recognizing Presence of a Contaminant

This procedure was identical to that described in Section 4.6.1 and used in Sections 5.2, 5.3.1, and 5.4.1. Here, the data from all three contaminants were combined and then run through the neural network. After combining these data, there are 12, 36, 36, and six iterations with ambient air, air with ammonia, air with ethanol, and air with glycerol, respectively, for both the training and validation datasets.

In general, the pooled data did not show a major difference from using data with the individual contaminants. However, there were many false positives, especially with the large box sensor. Overall, the success rates were 73.4%, 98.1%, 96.7%, and 94.8% for measuring ambient air, air with ammonia, air with ethanol, and air with glycerol, respectively. Table 5.20 details the complete data, with ‘T’ and ‘V’ for the training and validation data, respectively.

Table 5.20: Success rates of neural network for detecting contaminants with pooled data

	Ambient Air	Air with Ammonia	Air with Ethanol	Air with Glycerol
Small Probe Gain – T	8 of 12 (67%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)
Small Probe Gain – V	7 of 12 (58%)	36 of 36 (100%)	34 of 36 (94%)	6 of 6 (100%)
Small Probe Phase – T	11 of 12 (92%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)
Small Probe Phase – V	10 of 12 (83%)	36 of 36 (100%)	30 of 36 (83%)	6 of 6 (100%)
Large Probe Gain – T	12 of 12 (100%)	36 of 36 (100%)	35 of 36 (97%)	6 of 6 (100%)
Large Probe Gain – V	10 of 12 (83%)	35 of 36 (97%)	36 of 36 (100%)	5 of 6 (83%)
Large Probe Phase – T	11 of 12 (92%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)
Large Probe Phase – V	12 of 12 (100%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)

Continuation of Table 5.20

	Ambient Air	Air with Ammonia	Air with Ethanol	Air with Glycerol
Small Box Gain – T	9 of 12 (75%)	36 of 36 (100%)	36 of 36 (100%)	5 of 6 (83%)
Small Box Gain – V	9 of 12 (75%)	31 of 36 (86%)	35 of 36 (97%)	5 of 6 (83%)
Small Box Phase – T	9 of 12 (75%)	36 of 36 (100%)	35 of 36 (97%)	6 of 6 (100%)
Small Box Phase – V	8 of 12 (67%)	33 of 36 (92%)	29 of 36 (81%)	4 of 6 (67%)
Large Box Gain – T	10 of 12 (83%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)
Large Box Gain – V	1 of 12 (8%)	34 of 36 (94%)	35 of 36 (97%)	6 of 6 (100%)
Large Box Phase – T	12 of 12 (100%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)
Large Box Phase – V	2 of 12 (17%)	36 of 36 (100%)	36 of 36 (100%)	6 of 6 (100%)

5.5.2 Identifying the Contaminant Present

Instead of using ‘0’ as a target value for ambient air and ‘1’ for contaminated air, target values were assigned to the contaminants to order them by increasing dipole moments. Thus, ‘0’ was used for ambient air, ‘1’ for ammonia, ‘2’ for ethanol, and ‘3’ for glycerol. After training the neural network, the outputs were classified as being ambient air or having one of the contaminants. An output of less than 0.5 was ambient air, 0.5 to 1.5 was ammonia, 1.5 to 2.5 was ethanol, and 2.5 or above was glycerol.

While the success overall success rates for identifying the contaminants were not as strong as those for determining the presence of contaminants. Many of the tests with ambient air and air with glycerol were not identified successfully. However, especially with the small and large probe sensors, the air with ammonia and air with ethanol were correctly identified in several cases. Table 5.21 overviews the success rates of these tests; Appendix O, Appendix P, Appendix Q, and Appendix R list each individual test with the actual conditions, the output value from the neural network analysis, and the resulting classification.

Table 5.21: Success rates of neural network for identifying contaminants with pooled data

	Ambient Air	Air with Ammonia	Air with Ethanol	Air with Glycerol
Small Probe Gain – T	7 of 12 (58%)	36 of 36 (100%)	32 of 36 (89%)	4 of 6 (67%)
Small Probe Gain – V	5 of 12 (42%)	32 of 36 (89%)	24 of 36 (67%)	4 of 6 (67%)
Small Probe Phase – T	7 of 12 (58%)	30 of 36 (83%)	27 of 36 (75%)	4 of 6 (67%)
Small Probe Phase – V	6 of 12 (50%)	25 of 36 (69%)	15 of 36 (42%)	4 of 6 (67%)
Large Probe Gain – T	5 of 12 (42%)	34 of 36 (94%)	32 of 36 (89%)	3 of 6 (50%)
Large Probe Gain – V	4 of 12 (33%)	27 of 36 (75%)	25 of 36 (69%)	0 of 6 (0%)
Large Probe Phase – T	3 of 12 (25%)	28 of 36 (78%)	31 of 36 (86%)	1 of 6 (17%)
Large Probe Phase – V	3 of 12 (25%)	18 of 36 (50%)	29 of 36 (81%)	0 of 6 (0%)
Small Box Gain – T	9 of 12 (75%)	28 of 36 (78%)	33 of 36 (92%)	4 of 6 (67%)
Small Box Gain – V	4 of 12 (33%)	12 of 36 (33%)	22 of 36 (61%)	0 of 6 (0%)
Small Box Phase – T	11 of 12 (92%)	33 of 36 (92%)	35 of 36 (97%)	1 of 6 (17%)
Small Box Phase – V	5 of 12 (42%)	22 of 36 (61%)	23 of 36 (64%)	0 of 6 (0%)
Large Box Gain – T	12 of 12 (100%)	29 of 36 (81%)	32 of 36 (89%)	2 of 6 (33%)
Large Box Gain – V	4 of 12 (33%)	22 of 36 (61%)	21 of 36 (58%)	3 of 6 (50%)
Large Box Phase – T	9 of 12 (75%)	30 of 36 (83%)	29 of 36 (81%)	5 of 6 (83%)
Large Box Phase – V	2 of 12 (17%)	22 of 36 (61%)	26 of 36 (72%)	2 of 6 (33%)

CHAPTER 6 - Conclusions

6.1 Review of Research Objectives

6.1.1 Construction of Sensors

With stainless steel plates as the main sensing component, two probe-style sensors were constructed to go along with two box-style sensors constructed in previous work. Materials for the two probe-style sensors were under \$100. These sensors all withstood the repeated testing, cleaning, and drying for measuring the contaminants at the various concentrations.

6.1.2 Investigate Sensor Sensitivity

From the sensor sensitivity tests, sample temperature and cable position were found to be variables that had a significant impact on the frequency response. Conversely, the sensor readings were robust to the changes in the sensor vertical position, sensor radial position, the presence of a thermometer, and the surface on which the sample was placed.

6.1.3 Code for Analyzing Differences

A SAS[®] program is run on the initial data to give p -values for differences between measurements with and without contaminants. These p -values are then passed through a MATLAB[®] program that uses the Holm-Bonferroni method to determine which differences are truly significant. Appendix C contains both programs in their entirety. For applying a quantitative method for qualitative detection, Appendix D shows the neural network program that was used for this analysis.

6.1.4 Identifying Frequencies with Significantly Different Measurements

Each contaminant had a unique set of frequencies where measurements were significantly different between ambient air and air with contaminants. Appendix J gives the full list of frequencies where the differences were statistically different.

6.1.5 Code for Regression Analysis

MATLAB[®] programs for performing each of the various types of regressions were written and are included in Appendix E through Appendix H. Furthermore, two additional programs were written to import and trim the data (Appendix I).

6.1.6 Calculating R^2 and RMSE Values

Table 5.3 through Table 5.10 and Table 5.12 through Table 5.19 give all the R^2 and RMSE values for regressions from each contaminant, sensor, dataset, and regression method used. In several cases, the R^2 values were 1.000 for the training datasets, but the corresponding validation R^2 values were very low. However, the best regression had R^2 values of 1.000 and 0.996 for the training and validation datasets; this was using neural network regression with the second peak of the gain and phase data combined from the large probe sensor with ammonia.

6.2 Summary of Results

Overall, the qualitative detection methods tend to outperform the quantitative regression methods. In general, the probe style sensors tended to outperform the box style sensors, and the larger plate areas tended to outperform the smaller plate areas. Partial least squares regression did the best job of constructing a model without over-fitting, and including more components with PLS improved the regression. Most cases did not show that restricting the regressions to individual peaks was advantageous over using the entire frequency range.

6.2.1 Glycerol Tests

Run only on a detection basis, measurements at several frequencies showed a significant difference between ambient air and air with glycerol present. Gain data showed a total of 7, 89, 28, and 52 frequencies with significantly different measurements for the small probe, large probe, small box, and large box, respectively. Phase data had 8, 107, 16, and 167 significantly different measurements for each sensor. Neural network analysis had an overall success rate of 95.8% for detecting glycerol, with the large box having no errors.

6.2.2 Ethanol Tests

Distinguishing between ambient air and 5,000 ppm ethanol showed that 113, 65, 39, and 50 frequencies showed different gain measurements for the small probe, large probe, small box, and large box, respectively. For 115, 249, 54, and 147 measurements with the respective sensors, the phase data showed significant differences. Detection with the neural network analysis had an overall success rate of 93.9% accuracy despite each sensor having some errors.

Regression analysis on the range of concentrations showed that many models tended to over-fit the training data, producing strong training R^2 values and poor validation R^2 values. However, partial least squares regression tended to correct this problem, producing lower training and higher validation R^2 values.

6.2.3 Ammonia Tests

Detection tests comparing ambient air with 300 ppm ammonia had significant differences at many more frequencies than either glycerol or ethanol had. In total 237, 514, 272, and 40 frequencies had significantly different gain measurements for the small probe, large probe, small box, and large box sensors, respectively. Likewise, measurements at 419, 594, 293, and 27 frequencies were found with the phase data from the four sensors. Neural network-based detection was successful in 97.1% of all cases, with the small probe being 100% accurate and the large probe only having one error.

Running the regression methods on the series of concentrations had some instances of over-fitting, similar to the regressions with ethanol. These problems were not as common as the regressions with ethanol, but still some validation R^2 values were below 0.010. However, especially with the large probe sensor, several validation R^2 values were near or above 0.900, with the maximum being 0.996.

6.2.4 Combined Tests

Using the qualitative detection tests with data pooled from all contaminants did not show a significant change between using the tests with each contaminant individually. The major problem with these tests was the incidence of false positives. Yet, contaminated air was identified correctly in well over 90% of the cases. Identifying the specific contaminant showed some promise, especially with ammonia and ethanol.

CHAPTER 7 - Recommendations for Future Research

Despite the work done on the project to date, these areas still need further work:

1. Improvement of the sensor construction. Ideally, each of the metal plates would be cut out with a laser cutter to give much higher precision and uniformity. Also, the top of the sensors could be improved by creating a single plastic piece that accommodates each metal plate.
2. Analysis of the sensor parameters. So many factors of the sensor construction can impact the gain and phase readings. The effects of the number of plates, plate size, plate material, and gap between plates are clearly important variables.
3. Expansion of the frequency range. A larger frequency range, on the order of 10^0 to 10^9 Hz, should be utilized to give a more widespread, accurate spectrum.
4. Establish the minimum detection limits for each contaminant.
5. Ensure that the distribution of contaminants is uniform. This could be done with a fan that distributes the air but is turned off when reading the permittivity data.
6. Run detection tests at each level of ethanol and ammonia. While these tests were run at the concentrations set forth in the standards, running the detection tests at each frequency level would give a more complete picture of the sensors' ability to measure the contaminant.
7. Investigation of the interaction between contaminants. This work only investigated a single contaminant at one time. In practice, several contaminants might be produced in a given process. It is possible that such a combination of contaminants shows a significant difference in frequency response. Furthermore, humidity could also play a major role in permittivity measurements.
8. Run tests in the field. As these tests were all run in a sealed chamber, it is unlikely that the steady-state conditions would produce the same frequency response as in transient conditions. Air movement, contaminant variations, and temperature inconsistencies would all likely have a great impact on the measurements.

References

- Analog Devices, Inc. 2002. *LF-2.7 GHz RF/IF Gain and Phase Detector AD8302*. Retrieved 7 August 2012 from <http://www.analog.com/static/imported-files/data_sheets/AD8302.pdf>.
- BGI, Inc. 2003. *Audit procedure for the Andersen RAAS speciation sampler*. Retrieved 10 March 2012 from <<http://www.bgiusa.com/cal/andersen.pdf>>.
- Boreal Laser, Inc. 2009. *GasFinder specifications*. Retrieved 7 March 2012 from <<http://www.boreal-laser.com/gasfinder-specifications>>.
- Clean Air Trust. 2008. *Impact of E85 on Health*. Retrieved 26 June 2012 from <<http://www.cleanairtrust.org/E85-Health-Impact.html>>.
- Demuth, H. and M. Beale. 2004. *Neural network toolbox*. Natick, MA: The MathWorks.
- European Biodiesel Board. 2012. *Statistics: the EU biodiesel industry*. Retrieved 19 February 2012 from <<http://www.ebb-eu.org/stats.php>>.
- Felder, R.M., and R.W. Rousseau. 2005. *Elementary Principles of Chemical Processes*. Hoboken, MA: John Wiley & Sons, Inc.
- Figaro USA, Inc. 2005. *Product listing of target gases*. Retrieved 7 March 2012 from <<http://figarosensor.com/gaslist.html>>.
- Futurlec. 2004. *CO₂ Sensor – Carbon Dioxide Sensor*. Retrieved 2 May 2012 from <http://www.futurlec.com/CO2_Sensor.shtml>.
- Hargreaves, J.A. and C.S. Tucker. 2004. *Managing ammonia in fish ponds*. Retrieved 9 March 2012 from <<http://www.ca.uky.edu/wkrec/ManagingAmmonia.pdf>>.

Harper, L.A., T.K. Flesch, K.H. Weaver, and J.D. Wilson. 2010. The effect of biofuel production on swine farm methane and ammonia emissions. *Journal of Environmental Quality* 39(6): 1984-1992.

Jager, A.L. 16 July 2012. Personal communication with the author.

Lee, K.H., N. Zhang, W.B. Kuhn, and G.J. Kluitenberg. 2007. A frequency-response permittivity sensor for simultaneous measurement of multiple soil properties: Part I. The frequency-response method. *Transactions of the ASABE* 50(6): 2315-2326.

McElroy, M.B. 2009. *Energy: Perspectives, Problems, and Prospects*. New York: Oxford University Press.

Mendenhall, W., and T. Sincich. 2012. *Regression analysis*. Boston, MA: Prentice Hall.

Mevik, B.H. 2006. The pls package. *R News* 6(3): 12-17.

Misiti, M., Y. Misiti, G. Oppenheim, and J. Poggi. 2006. *Wavelet toolbox*. Natick, MA: The MathWorks.

National Biodiesel Board. 2007a. *Biodiesel backgrounder*. Retrieved 19 February 2012 from <http://biodiesel.org/pdf_files/fuelfactsheets/backgrounder.pdf>.

National Biodiesel Board. 2007b. *Biodiesel technical information*. Retrieved 14 March 2012 from <http://www.biodiesel.org/pdf_files/Biodiesel_Technical_Manual.pdf>.

Nelson, S.O. 1991. Dielectric Properties of Agricultural Products: Measurements and Applications. *IEEE Transactions on Electrical Insulation* 26(5): 845-869.

Omron Corporation. 2008. *Clean sensing system: ZN series*. Retrieved 11 March 2012 from <http://www.ia.omron.com/data_pdf/data_sheet/zn-a_zn-j_zn-pd_ca_csm1570.pdf>.

Ott, R.L., and M. Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis*. Pacific Grove, CA: Duxbury Thomson Learning.

Peterson, C.L., D.L. Reece, B.L. Hammond, J. Thompson, and S.M. Beck. 1997. Processing, characterization, and performance of eight fuels from lipids. *Applied Engineering in Agriculture* 13(1): 71-79.

Peterson, C.L., J.L. Cook, J.C. Thompson, and J.S. Taberski. 2002. Continuous flow biodiesel production. *Applied Engineering in Agriculture* 18(1): 5-11.

Robinson, D.A., S.B. Jones, J.M. Wraith, D. Or, and S.P. Friedman. 2003. A Review of Advances in Dielectric and Electrical Conductivity Measurement in Soils Using Time Domain Reflectometry. *Vadose Zone Journal* 2(4): 444-475.

Scholte, J.W., J.Q. Shang, and R.K. Rowe. 2002. Improved Complex Permittivity Measurement and Data Processing Technique for Soil-Water Systems. *Geotechnical Testing Journal* 25(2): 187-198.

Shultz, S. 2009. Calibration of permittivity sensors to measure contaminants in water and biodiesel fuel. M.S. Thesis. Department of Biological and Agricultural Engineering, Kansas State University.

Stellman, J.M. (ed.) 1998. *Encyclopaedia of Occupational Health and Safety: Guides, Indexes, Directory Volume 4*. Geneva: International Labour Organization.

Tang, N. 2009. A real-time control system for a frequency response-based permittivity sensor. M.S. Thesis. Department of Biological and Agricultural Engineering, Kansas State University.

The MathWorks, Inc. 2004. *MATLAB mathematics*. Natick, MA: The MathWorks, Inc.

The MathWorks, Inc. 2012a. MATLAB® R2012a [Software]. Natick, MA: The MathWorks, Inc.

The MathWorks, Inc. 2012b. Neural Network Toolbox™, version 7.0.3 [Software]. Natick, MA: The MathWorks, Inc.

The MathWorks, Inc. 2012c. *Partial least squares*. Retrieved 12 June 2012 from <<http://www.mathworks.com/help/toolbox/stats/briu0ni.html>>.

The MathWorks, Inc. 2012d. *Partial least-squares regression – MATLAB*. Retrieved 31 May 2012 from <<http://www.mathworks.com/help/toolbox/stats/plsregress.html>>.

The MathWorks, Inc. 2012e. Statistics Toolbox™, version 8.0 [Software]. Natick, MA: The MathWorks, Inc.

The MathWorks, Inc. 2012f. Wavelet Toolbox™, version 4.9 [Software]. Natick, MA: The MathWorks, Inc.

The SAS Institute, Inc. 2000. *Examples using the PLS procedure*. Retrieved 5 June 2012 from <<http://support.sas.com/rnd/app/papers/plsex.html>>.

The SAS Institute, Inc. 2008. SAS® 9.2 [Software]. Cary, NC: The SAS Institute, Inc.

Topp, G.C., S. Zegelin, and I. White. 2000. Impacts of the Real and Imaginary Components of Relative Permittivity on Time Domain Reflectometry Measurements in Soils. *Soil Science Society of America Journal* 64: 1244-1252.

U.S. OSHA. 2003. *Safety and health topics: Ammonia*. Retrieved 9 March 2012 from <http://www.osha.gov/dts/chemicalsampling/data/CH_218300.html>.

Uteley, B.C., T.M. Wynn, N. Zhang, and T.E. Teany. 2011. Evaluation of a permittivity sensor for continuous monitoring of suspended sediment concentration. *Transaction of the ASABE* 54(4): 1299-1309.

Appendix A - Design of Sensors

The following figures illustrate the detail of the stainless steel plates for the two probes. All dimensions are given in inches and millimeters.

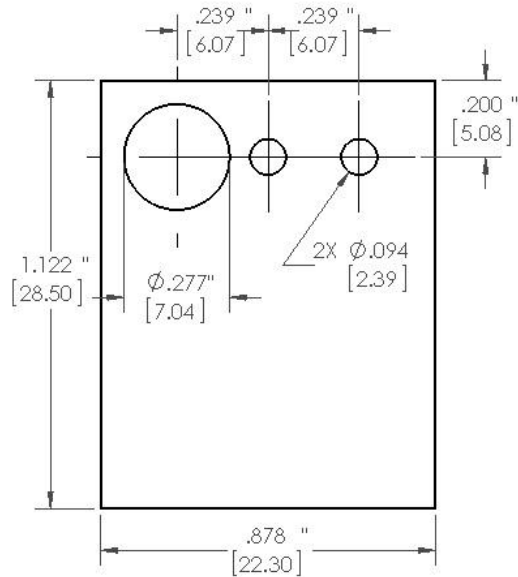


Figure A.1: Plate for the small probe

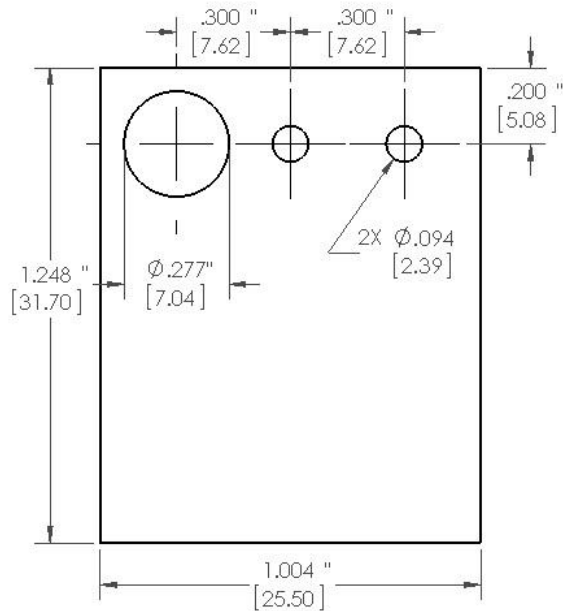


Figure A.2: Plate for the large probe

Appendix B - Code for Sensor Sensitivity Tests

```
1 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2 % Boundary Height Function %
3 % %
4 % This program takes Excel spreadsheets from the tests to check %
5 % for the sensitivity to the height of the sensor, then plots %
6 % either the gain or phase data. %
7 % %
8 % May 14, 2012 %
9 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
10
11 function Boundary_Height(GP)
12
13 %Input:
14 %GP - Graph gain or phase (Gain - 1, Phase - 2)
15
16 clc;
17
18 %Read data from Excel spreadsheet and check for proper size
19 data0 = Average_Iterations('051412-12');
20 data1 = Average_Iterations('051412-13');
21 data2 = Average_Iterations('051412-14');
22 data3 = Average_Iterations('051412-15');
23 data4 = Average_Iterations('051412-16');
24 data5 = Average_Iterations('051412-17');
25 data6 = Average_Iterations('051412-18');
26 data7 = Average_Iterations('051412-19');
27 data8 = Average_Iterations('051412-20');
28 data9 = Average_Iterations('051412-21');
29 data10 = Average_Iterations('051412-22');
30 data11 = Average_Iterations('051412-23');
31 data12 = Average_Iterations('051412-24');
32 data13 = Average_Iterations('051412-25');
33 data14 = Average_Iterations('051412-26');
34 [m n] = size(data0);
35
36 if((m ~= 609) || (n ~= 3))
37     error('The data dimensions are not valid.')
38 end
39
40 Frequency = data0(:, 1);
41
42 %Plot the averaged data versus frequency on a linear scale
43 hold off
44
45 if(GP == 1)
46     plot(Frequency, data0(:, 2), 'b-', Frequency, data1(:, 2), 'b--',
47          Frequency, data2(:, 2), 'b:', Frequency, data3(:, 2), 'g-', Frequency,
48          data4(:, 2), 'g--', Frequency, data5(:, 2), 'g:', Frequency, data6(:,
49          2), 'r-', Frequency, data7(:, 2), 'r--', Frequency, data8(:, 2), 'r:',
50          Frequency, data9(:, 2), 'c-', Frequency, data10(:, 2), 'c--',
51          Frequency, data11(:, 2), 'c:', Frequency, data12(:, 2), 'm-',
52          Frequency, data13(:, 2), 'm--', Frequency, data14(:, 2), 'm:')
```

```

47     ylabel('Gain Signal', 'FontSize', 13)
48     legend('1 cm', '1 cm', '1 cm', '3 cm', '3 cm', '3 cm', '5 cm', '5
cm', '5 cm', '7 cm', '7 cm', '7 cm', '9 cm', '9 cm', '9 cm')
49 end
50
51 if(GP == 2)
52     plot(Frequency, data0(:, 3), 'b-', Frequency, data1(:, 3), 'b--',
Frequency, data2(:, 3), 'b:', Frequency, data3(:, 3), 'g-', Frequency,
data4(:, 3), 'g--', Frequency, data5(:, 3), 'g:', Frequency, data6(:,
3), 'r-', Frequency, data7(:, 3), 'r--', Frequency, data8(:, 3), 'r:',
Frequency, data9(:, 3), 'c-', Frequency, data10(:, 3), 'c--',
Frequency, data11(:, 3), 'c:', Frequency, data12(:, 3), 'm-',
Frequency, data13(:, 3), 'm--', Frequency, data14(:, 3), 'm:')
53     ylabel('Phase Signal', 'FontSize', 13)
54     legend('1 cm', '1 cm', '1 cm', '3 cm', '3 cm', '3 cm', '5 cm', '5
cm', '5 cm', '7 cm', '7 cm', '7 cm', '9 cm', '9 cm', '9 cm')
55 end
56
57 xlabel('Frequency (Hz)', 'FontSize', 13)
58 axis([0 120000000 150 500])
59 grid on
60 hold on
61
62
63 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
64 % Average Iterations Function %
65 % %
66 % This program takes an Excel spreadsheet called 'FileName', %
67 % processes the gain and phase data, then returns a matrix with %
68 % the frequency, average gain, and average phase data. %
69 % %
70 % March 5, 2012 %
71 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
72
73 function [NewData] = Average_Iterations(File1)
74
75 %Inputs:
76     %File1 - XLS file where data is stored
77
78 %Outputs:
79     %NewData - Matrix with frequency, gain, and phase
80
81 clc;
82
83 %Read data from Excel spreadsheets and check for proper size
84 data1 = xlsread(strcat(File1, '.xls'));
85 [m1 n1] = size(data1);
86
87 if((m1 ~= 1910) || (n1 ~= 4))
88     error('The data dimensions are not valid.')
89 end
90
91 NewData = zeros(609, 3);
92

```



```
93 %Copy the data from the original file to the new matrix
94 for i = 1 : 609
95     %Frequency
96     NewData(i, 1) = 10000 * data1(i + 27, 2);
97
98     %Gain
99     NewData(i, 2) = (data1(i + 27, 3) + data1(i + 663, 3) + data1(i +
100     1299, 3)) / 3;
101
101     %Phase
102     NewData(i, 3) = (data1(i + 27, 4) + data1(i + 663, 4) + data1(i +
103     1299, 4)) / 3;
103 end
```

Appendix C - Code for Detection Tests

SAS[®] code for returning *p*-values:

```
1  proc glm data=responses;
2      class Glycerol Freq;
3      model Gain = Glycerol Freq Glycerol*Freq;
4      lsmeans Glycerol*Freq / slice=Freq;
5  run;
```

MATLAB[®] code for the Holm-Bonferroni comparison test:

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % Holm Comparison Function %
3  % %
4  % This program takes Excel spreadsheet corresponding to the %
5  % Sensor, Data, and Contaminant parameters, and performs the %
6  % Holm-Bonferroni comparison for each frequency. The Excel %
7  % spreadsheet must have the frequency and p-values in the two %
8  % columns, sorted by increasing p-values. %
9  % %
10 % July 16, 2012 %
11 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
12
13 function Holm_Comparison(Sensor, Data, Contaminant, Alpha)
14
15 %Inputs:
16 %Sensor - Sensor for data ('SP' 'LP' 'SB' or 'LB')
17 %Data - Type of data to plot (Gain - 'G', Phase - 'P')
18 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
19 'G', Ethanol - 'E')
20 %Alpha - Level of significance
21
22
23 %Import Excel file where data is stored
24 data = xlsread(strcat('Comparison', Sensor, Data, Contaminant, '.xls'));
25 [m n] = size(data);
26
27 if(n ~= 2)
28     error('The input data dimensions are not valid.')
29 end
30
31 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
32 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
33     error('The Sensor parameter is not valid.')
34 end
```

```

34
35     if((strcmp(Data, 'G') == 0) && (strcmp(Data, 'P') == 0))
36         error('The Data parameter is not valid.')
37     end
38
39     if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
&& (strcmp(Contaminant, 'E') == 0))
40         error('The Contaminant parameter is not valid.')
41     end
42
43     %Perform Holm-Bonferroni comparison
44     Significant = 1;
45     Iteration = 1;
46     while((Significant == 1) && (Iteration < m))
47         if(data(Iteration, 2) < (Alpha / (m - Iteration + 1)))
48             fprintf('%i\n', data(Iteration, 1))
49             Iteration = Iteration + 1;
50         else
51             fprintf('Data found insignificant at %i \n', data(Iteration, 1))
52             Significant = 0;
53         end
54     end

```

Appendix D - Neural Network Detection Program

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % Neural Detection Function %
3  % %
4  % This program runs a neural network-based comparison to detect %
5  % the presence of contaminants, using ambient air with an %
6  % output of 0 and contaminated air with an output of 1. %
7  % %
8  % Adapted from Demuth and Beale (2004) %
9  % %
10 % July 22, 2012 %
11 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
12
13 function Neural_Comparison(Sensor, Contaminant)
14
15 %Inputs:
16 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
17 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
18 %'G', Ethanol - 'E')
19
20 clc;
21
22 %Check for proper Sensor and Contaminant parameters
23 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
24 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
25     error('The Sensor parameter is not valid.')
26 end
27
28 if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
29 && (strcmp(Contaminant, 'E') == 0))
30     error('The Contaminant parameter is not valid.')
31 end
32
33 %Read data from Import Data function
34 [XT YT XV YV] = Import_Data(Sensor, Contaminant);
35
36 %Run the neural network comparison for entire region of frequencies
37 Compare_NN(XT, YT, XV, YV, Sensor, 'G')
38 Compare_NN(XT, YT, XV, YV, Sensor, 'P')
39
40 function Compare_NN(XT, YT, XV, YV, Sensor, Data)
41
42 if(strcmp(Data, 'G'))
43     str1 = 'gain';
44 elseif(strcmp(Data, 'P'))
45     str1 = 'phase';
46 end
47
48 %Trim data to specific range and create neural network
49 [XT1 YT1 XV1 YV1] = Trim_Data(XT, YT, XV, YV, Sensor, 0, Data);
50 for i = 1 : length(YT1)
51     if(YT1(i, 1) ~= 0)
```

```

49     YT1(i, 1) = 1;
50     end
51     if(YV1(i, 1) ~= 0)
52         YV1(i, 1) = 1;
53     end
54 end
55 XT1 = XT1'; YT1 = YT1'; XV1 = XV1';
56
57 net = fitnet(10);
58 net.inputs{1}.processFcns = {'removeconstantrows', 'mapminmax'};
59 net.outputs{2}.processFcns = {'removeconstantrows', 'mapminmax'};
60
61 net.divideFcn = 'dividerand';
62 net.divideMode = 'sample';
63 net.divideParam.trainRatio = 0.60;
64 net.divideParam.valRatio = 0.20;
65 net.divideParam.testRatio = 0.20;
66 net.trainFcn = 'trainlm';
67 net.performFcn = 'mse';
68
69 [net, ~] = train(net, XT1, YT1);
70
71 %Use neural network to predict presence of contaminants in training
dataset
72 fprintf('Using %s data:\n', str1)
73 yfit = (net(XT1))';
74 for i = 1 : length(yfit)
75     if(yfit(i, 1) >= 0.5)
76         fprintf(' Contaminants detected at %i ppm - Training\n', YT(i, 1))
77     else
78         fprintf(' No contaminants detected at %i ppm - Training\n', YT(i,
1))
79     end
80 end
81
82 %Use neural network to predict presence of contaminants in validation
dataset
83 yfit = (net(XV1))';
84 for i = 1 : length(yfit)
85     if(yfit(i, 1) >= 0.5)
86         fprintf(' Contaminants detected at %i ppm - Validation\n', YV(i,
1))
87     else
88         fprintf(' No contaminants detected at %i ppm - Validation\n',
YV(i, 1))
89     end
90 end
91 fprintf('\n')

```

Appendix E - Stepwise Regression Program

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % Stepwise Regression Function %
3  % %
4  % This program runs a stepwise regression on the data from the %
5  % entered sensor and contaminant, including all possible %
6  % combinations of data and frequency subsets. %
7  % %
8  % July 6, 2012 %
9  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
10
11 function Stepwise_Regression(Sensor, Contaminant)
12
13 %Inputs:
14 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
15 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
16 %'G', Ethanol - 'E')
17
18 clc;
19
20 %Check for proper Sensor and Contaminant parameters
21 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
22 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
23 error('The Sensor parameter is not valid.')
24 end
25
26 if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
27 && (strcmp(Contaminant, 'E') == 0))
28 error('The Contaminant parameter is not valid.')
29 end
30
31 %Read data from Import Data function
32 [XT YT XV YV] = Import_Data(Sensor, Contaminant);
33
34 %Run Stepwise for entire region of frequencies
35 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 0, 'G')
36 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 0, 'P')
37 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 0, 'J')
38 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 0, 'I')
39
40 %Run Stepwise for first region of frequencies
41 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 1, 'G')
42 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 1, 'P')
43 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 1, 'J')
44 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 1, 'I')
45
46 %Run Stepwise for second region of frequencies
47 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 2, 'G')
48 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 2, 'P')
49 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 2, 'J')
50 Calculate_Stepwise(XT, YT, XV, YV, Sensor, 2, 'I')
```

```

48 %Check third region for small box and large box sensors
49 if((strcmp(Sensor, 'SB') == 1) || (strcmp(Sensor, 'LB') == 1))
50     Calculate_Stepwise(XT, YT, XV, YV, Sensor, 3, 'G')
51     Calculate_Stepwise(XT, YT, XV, YV, Sensor, 3, 'P')
52     Calculate_Stepwise(XT, YT, XV, YV, Sensor, 3, 'J')
53     Calculate_Stepwise(XT, YT, XV, YV, Sensor, 3, 'I')
54 end
55
56 function Calculate_Stepwise(XT, YT, XV, YV, Sensor, Peak, Data)
57
58 if(strcmp(Data, 'G'))
59     str1 = 'gain';
60     StartFreq = [1; Frequency_List(Sensor, Peak)];
61 elseif(strcmp(Data, 'P'))
62     str1 = 'phase';
63     StartFreq = [1; Frequency_List(Sensor, Peak)];
64 elseif(strcmp(Data, 'J'))
65     str1 = 'gain and phase';
66     StartFreq = [1; Frequency_List(Sensor, Peak); Frequency_List(Sensor,
Peak)];
67 elseif(strcmp(Data, 'I'))
68     str1 = 'interaction';
69     StartFreq = [1; Frequency_List(Sensor, Peak); Frequency_List(Sensor,
Peak); Frequency_List(Sensor, Peak)];
70 end
71
72 if(Peak == 1)
73     str2 = 'first peak';
74 elseif(Peak == 2)
75     str2 = 'second peak';
76 elseif(Peak == 3)
77     str2 = 'third peak';
78 else
79     str2 = 'all frequencies';
80 end
81
82 %Trim data to specific range and perform stepwise regression
83 [XT1 YT1 XV1 YV1] = Trim_Data(XT, YT, XV, YV, Sensor, Peak, Data);
84 [beta, ~, ~, inmodel, stats] = stepwisefit(XT1, YT1, 'display', 'off');
85
86 %List the frequencies used in the model
87 if(sum(inmodel) ~= 0)
88     TrimmedFreq = StartFreq.*inmodel';
89     TrimmedFreq = TrimmedFreq(TrimmedFreq ~= 0);
90     xlswrite(strcat('Freq', Sensor, num2str(Peak), Data, '.xls'),
TrimmedFreq);
91 end
92
93 %Fit training dataset to the model then calculate the R^2 and RMSE
values
94 if(sum(inmodel) ~= 0)
95     [m, ~] = size(XT1);
96     yfit = zeros(m, 1);
97     for i = 1 : m
98         yfit(i, 1) = (XT1(i, :)) * (beta.*inmodel');
99     end

```

```

100     yfit = yfit + stats.intercept;
101     for i = 1 : length(yfit)
102         if(yfit(i, 1) < 0)
103             yfit(i, 1) = 0;
104         end
105     end
106     [~, m, b] = regression(YT1', yfit');
107     SSE = sum((yfit - (b + m.*YT1)).^2);
108     SSyy = sum((yfit - mean(yfit)).^2);
109     Rsq = 1 - SSE/SSyy;
110     RMSE = sqrt(SSE / (length(YT1) - 2));
111     fprintf('Training R^2 for %s data, %s:   %f\n', str1, str2, Rsq)
112     fprintf('Training RMSE for %s data, %s:  %f\n', str1, str2, RMSE)
113 else
114     fprintf('Training R^2 for %s data, %s:   N/A, no model made\n',
str1, str2)
115     fprintf('Training RMSE for %s data, %s:  N/A, no model made\n',
str1, str2)
116 end
117
118 %Fit validation dataset to the model then calculate the R^2 and RMSE
values
119 if(sum(inmodel) ~= 0)
120     [m, ~] = size(XV1);
121     yfit = zeros(m, 1);
122     for i = 1 : m
123         yfit(i, 1) = (XV1(i, :)) * (beta.*inmodel');
124     end
125     yfit = yfit + stats.intercept;
126     for i = 1 : length(yfit)
127         if(yfit(i, 1) < 0)
128             yfit(i, 1) = 0;
129         end
130     end
131     [~, m, b] = regression(YV1', yfit');
132     SSE = sum((yfit - (b + m.*YV1)).^2);
133     SSyy = sum((yfit - mean(yfit)).^2);
134     Rsq = 1 - SSE/SSyy;
135     RMSE = sqrt(SSE / (length(YT1) - 2));
136     fprintf('Validation R^2 for %s data, %s:   %f\n', str1, str2, Rsq)
137     fprintf('Validation RMSE for %s data, %s:  %f\n', str1, str2, RMSE)
138 else
139     fprintf('Validation R^2 for %s data, %s:   N/A, no model made\n',
str1, str2)
140     fprintf('Validation RMSE for %s data, %s:  N/A, no model made\n',
str1, str2)
141 end

```


Appendix F - Wavelet Regression Program

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % Wavelet Regression Function %
3  % %
4  % This program runs a wavelet transformation followed by a %
5  % stepwise regression on the data from the entered sensor and %
6  % contaminant, including all possible combinations of data and %
7  % frequency subsets. %
8  % %
9  % July 11, 2012 %
10 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
11
12 function Wavelet_Regression(Sensor, Contaminant)
13
14 %Inputs:
15 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
16 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
17 %'G', Ethanol - 'E')
18
19 clc;
20
21 %Check for proper Sensor and Contaminant parameters
22 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
23 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
24 error('The Sensor parameter is not valid.')
25 end
26
27 if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
28 && (strcmp(Contaminant, 'E') == 0))
29 error('The Contaminant parameter is not valid.')
30 end
31
32 %Read data from Import Data function
33 [XT YT XV YV] = Import_Data(Sensor, Contaminant);
34
35 %Run Wavelet and Stepwise for entire region of frequencies
36 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 0, 'G')
37 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 0, 'P')
38 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 0, 'J')
39 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 0, 'I')
40
41 %Run Wavelet and Stepwise for first region of frequencies
42 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 1, 'G')
43 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 1, 'P')
44 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 1, 'J')
45 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 1, 'I')
46
47 %Run Wavelet and Stepwise for second region of frequencies
48 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 2, 'G')
49 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 2, 'P')
50 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 2, 'J')
51 Calculate_Wavelet(XT, YT, XV, YV, Sensor, 2, 'I')
```

```

49
50 %Check third region for small box and large box sensors
51 if((strcmp(Sensor, 'SB') == 1) || (strcmp(Sensor, 'LB') == 1))
52     Calculate_Wavelet(XT, YT, XV, YV, Sensor, 3, 'G')
53     Calculate_Wavelet(XT, YT, XV, YV, Sensor, 3, 'P')
54     Calculate_Wavelet(XT, YT, XV, YV, Sensor, 3, 'J')
55     Calculate_Wavelet(XT, YT, XV, YV, Sensor, 3, 'I')
56 end
57
58 function Calculate_Wavelet(XT, YT, XV, YV, Sensor, Peak, Data)
59
60 if(strcmp(Data, 'G'))
61     str1 = 'gain';
62 elseif(strcmp(Data, 'P'))
63     str1 = 'phase';
64 elseif(strcmp(Data, 'J'))
65     str1 = 'gain and phase';
66 elseif(strcmp(Data, 'I'))
67     str1 = 'interaction';
68 end
69
70 if(Peak == 1)
71     str2 = 'first peak';
72 elseif(Peak == 2)
73     str2 = 'second peak';
74 elseif(Peak == 3)
75     str2 = 'third peak';
76 else
77     str2 = 'all frequencies';
78 end
79
80 %Trim data to specific range and perform wavelet decomposition
81 [XT1 YT1 XV1 YV1] = Trim_Data(XT, YT, XV, YV, Sensor, Peak, Data);
82 [m1, ~] = size(XV1);
83 for i = 1 : m1
84     [CT(i, :), ~] = wavedec(XT1(i, :), 3, 'haar');
85     [CV(i, :), ~] = wavedec(XV1(i, :), 3, 'haar');
86 end
87
88 %Perform stepwise regression
89 [beta, ~, ~, inmodel, stats] = stepwisefit(CT, YT1, 'display', 'off');
90
91 %Identify coefficients used in the model
92 if(sum(inmodel ~= 0))
93     Coeffs = CT.*inmodel;
94     xlswrite(strcat('Index', Sensor, num2str(Peak), Data, '.xls'),
95     Coeffs);
96 end
97
98 %Fit training dataset to the model then calculate the R^2 and RMSE
99 values
100 if(sum(inmodel) ~= 0)
101     [m, ~] = size(CT);
102     yfit = zeros(m, 1);
103     for i = 1 : m

```

```

102     yfit(i, 1) = (CT(i, :)) * (beta.*inmodel');
103 end
104 yfit = yfit + stats.intercept;
105 for i = 1 : length(yfit)
106     if(yfit(i, 1) < 0)
107         yfit(i, 1) = 0;
108     end
109 end
110 [~, m, b] = regression(YT1', yfit');
111 SSE = sum((yfit - (b + m.*YT1)).^2);
112 SSyy = sum((yfit - mean(yfit)).^2);
113 Rsq = 1 - SSE/SSyy;
114 RMSE = sqrt(SSE / (length(YT1) - 2));
115 fprintf('Training R^2 for %s data, %s: %f\n', str1, str2, Rsq)
116 fprintf('Training RMSE for %s data, %s: %f\n', str1, str2, RMSE)
117 else
118     fprintf('Training R^2 for %s data, %s: N/A, no model made\n',
str1, str2)
119     fprintf('Training RMSE for %s data, %s: N/A, no model made\n',
str1, str2)
120 end
121
122 %Fit validation dataset to the model then calculate the R^2 and RMSE
values
123 if(sum(inmodel) ~= 0)
124     [m, ~] = size(CV);
125     yfit = zeros(m, 1);
126     for i = 1 : m
127         yfit(i, 1) = (CV(i, :)) * (beta.*inmodel');
128     end
129     yfit = yfit + stats.intercept;
130     for i = 1 : length(yfit)
131         if(yfit(i, 1) < 0)
132             yfit(i, 1) = 0;
133         end
134     end
135     [~, m, b] = regression(YV1', yfit');
136     SSE = sum((yfit - (b + m.*YV1)).^2);
137     SSyy = sum((yfit - mean(yfit)).^2);
138     Rsq = 1 - SSE/SSyy;
139     RMSE = sqrt(SSE / (length(YT1) - 2));
140     fprintf('Validation R^2 for %s data, %s: %f\n', str1, str2, Rsq)
141     fprintf('Validation RMSE for %s data, %s: %f\n', str1, str2, RMSE)
142 else
143     fprintf('Validation R^2 for %s data, %s: N/A, no model made\n',
str1, str2)
144     fprintf('Validation RMSE for %s data, %s: N/A, no model made\n',
str1, str2)
145 end

```

Appendix G - Partial Least Squares Regression Program

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % PLS Regression Function %
3  % %
4  % This program runs a partial least squares regression on the %
5  % data from the entered sensor and contaminant, including all %
6  % possible combinations of data and frequency subsets. %
7  % %
8  % July 10, 2012 %
9  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
10
11 function PLS_Regression(Sensor, Contaminant, Components)
12
13 %Inputs:
14 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
15 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
16 %'G', Ethanol - 'E')
17 %Components - Number of components to be used in the regression
18
19 clc;
20
21 %Check for proper Sensor and Contaminant parameters
22 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
23 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
24 error('The Sensor parameter is not valid.')
25 end
26
27 if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
28 && (strcmp(Contaminant, 'E') == 0))
29 error('The Contaminant parameter is not valid.')
30 end
31
32 %Read data from Import Data function
33 [XT YT XV YV] = Import_Data(Sensor, Contaminant);
34
35 %Run PLS for entire region of frequencies
36 Calculate_PLS(XT, YT, XV, YV, Sensor, 0, 'G', Components)
37 Calculate_PLS(XT, YT, XV, YV, Sensor, 0, 'P', Components)
38 Calculate_PLS(XT, YT, XV, YV, Sensor, 0, 'J', Components)
39 Calculate_PLS(XT, YT, XV, YV, Sensor, 0, 'I', Components)
40
41 %Run PLS for first region of frequencies
42 Calculate_PLS(XT, YT, XV, YV, Sensor, 1, 'G', Components)
43 Calculate_PLS(XT, YT, XV, YV, Sensor, 1, 'P', Components)
44 Calculate_PLS(XT, YT, XV, YV, Sensor, 1, 'J', Components)
45 Calculate_PLS(XT, YT, XV, YV, Sensor, 1, 'I', Components)
46
47 %Run PLS for second region of frequencies
48 Calculate_PLS(XT, YT, XV, YV, Sensor, 2, 'G', Components)
49 Calculate_PLS(XT, YT, XV, YV, Sensor, 2, 'P', Components)
50 Calculate_PLS(XT, YT, XV, YV, Sensor, 2, 'J', Components)
51 Calculate_PLS(XT, YT, XV, YV, Sensor, 2, 'I', Components)
```

```

49
50 %Check third region for small box and large box sensors
51 if((strcmp(Sensor, 'SB') == 1) || (strcmp(Sensor, 'LB') == 1))
52     Calculate_PLS(XT, YT, XV, YV, Sensor, 3, 'G', Components)
53     Calculate_PLS(XT, YT, XV, YV, Sensor, 3, 'P', Components)
54     Calculate_PLS(XT, YT, XV, YV, Sensor, 3, 'J', Components)
55     Calculate_PLS(XT, YT, XV, YV, Sensor, 3, 'I', Components)
56 end
57
58 function Calculate_PLS(XT, YT, XV, YV, Sensor, Peak, Data, Components)
59
60 if(strcmp(Data, 'G'))
61     str1 = 'gain';
62 elseif(strcmp(Data, 'P'))
63     str1 = 'phase';
64 elseif(strcmp(Data, 'J'))
65     str1 = 'gain and phase';
66 elseif(strcmp(Data, 'I'))
67     str1 = 'interaction';
68 end
69
70 if(Peak == 1)
71     str2 = 'first peak';
72 elseif(Peak == 2)
73     str2 = 'second peak';
74 elseif(Peak == 3)
75     str2 = 'third peak';
76 else
77     str2 = 'all frequencies';
78 end
79
80 %Trim data to specific range and perform PLS regression
81 [XT1 YT1 XV1 YV1] = Trim_Data(XT, YT, XV, YV, Sensor, Peak, Data);
82 [~, ~, ~, ~, beta] = plsregress(XT1, YT1, Components);
83
84 %Fit training dataset to the model then calculate the R^2 and RMSE
values
85 yfit = [ones(size(XT1, 1), 1) XT1] * beta;
86 for i = 1 : length(yfit)
87     if(yfit(i, 1) < 0)
88         yfit(i, 1) = 0;
89     end
90 end
91 [~, m, b] = regression(YT1', yfit');
92 SSE = sum((yfit - (b + m.*YT1)).^2);
93 SSyy = sum((yfit - mean(yfit)).^2);
94 Rsq = 1 - SSE/SSyy;
95 RMSE = sqrt(SSE / (length(YT1) - 2));
96 fprintf('Training R^2 for %s data, %s:    %f\n', str1, str2, Rsq)
97 fprintf('Training RMSE for %s data, %s:    %f\n', str1, str2, RMSE)
98
99 %Fit validation dataset to the model then calculate the R^2 and RMSE
values
100 yfit = [ones(size(XT1, 1), 1) XV1] * beta;
101 for i = 1 : length(yfit)

```

```
102     if(yfit(i, 1) < 0)
103         yfit(i, 1) = 0;
104     end
105 end
106 [~, m, b] = regression(YV1', yfit');
107 SSE = sum((yfit - (b + m.*YV1)).^2);
108 SSyy = sum((yfit - mean(yfit)).^2);
109 Rsq = 1 - SSE/SSyy;
110 RMSE = sqrt(SSE / (length(YT1) - 2));
111 fprintf('Validation R^2 for %s data, %s: %f\n', str1, str2, Rsq)
112 fprintf('Validation RMSE for %s data, %s: %f\n', str1, str2, RMSE)
```

Appendix H - Neural Network Regression Program

```
1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  % Neural Regression Function %
3  % %
4  % This program runs a regression on the neural network based on %
5  % data from the entered sensor and contaminant, including all %
6  % possible combinations of data and frequency subsets. %
7  % %
8  % Adapted from The MathWorks (2012b) %
9  % %
10 % July 11, 2012 %
11 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
12
13 function Neural_Regression(Sensor, Contaminant)
14
15 %Inputs:
16 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
17 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
18 %'G', Ethanol - 'E')
19
20 clc;
21
22 %Check for proper Sensor and Contaminant parameters
23 if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
24 (strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
25     error('The Sensor parameter is not valid.')
26 end
27
28 if((strcmp(Contaminant, 'A') == 0) && (strcmp(Contaminant, 'G') == 0)
29 && (strcmp(Contaminant, 'E') == 0))
30     error('The Contaminant parameter is not valid.')
31 end
32
33 %Read data from Import Data function
34 [XT YT XV YV] = Import_Data(Sensor, Contaminant);
35
36 %Run neural network regression for entire region of frequencies
37 Calculate_NN(XT, YT, XV, YV, Sensor, 0, 'G')
38 Calculate_NN(XT, YT, XV, YV, Sensor, 0, 'P')
39 Calculate_NN(XT, YT, XV, YV, Sensor, 0, 'J')
40 Calculate_NN(XT, YT, XV, YV, Sensor, 0, 'I')
41
42 %Run neural network regression for first region of frequencies
43 Calculate_NN(XT, YT, XV, YV, Sensor, 1, 'G')
44 Calculate_NN(XT, YT, XV, YV, Sensor, 1, 'P')
45 Calculate_NN(XT, YT, XV, YV, Sensor, 1, 'J')
46 Calculate_NN(XT, YT, XV, YV, Sensor, 1, 'I')
47
48 %Run neural network regression for second region of frequencies
49 Calculate_NN(XT, YT, XV, YV, Sensor, 2, 'G')
50 Calculate_NN(XT, YT, XV, YV, Sensor, 2, 'P')
51 Calculate_NN(XT, YT, XV, YV, Sensor, 2, 'J')
```

```

49 Calculate_NN(XT, YT, XV, YV, Sensor, 2, 'I')
50
51 %Check third region for small box and large box sensors
52 if((strcmp(Sensor, 'SB') == 1) || (strcmp(Sensor, 'LB') == 1))
53     Calculate_NN(XT, YT, XV, YV, Sensor, 3, 'G')
54     Calculate_NN(XT, YT, XV, YV, Sensor, 3, 'P')
55     Calculate_NN(XT, YT, XV, YV, Sensor, 3, 'J')
56     Calculate_NN(XT, YT, XV, YV, Sensor, 3, 'I')
57 end
58
59 function Calculate_NN(XT, YT, XV, YV, Sensor, Peak, Data)
60
61 if(strcmp(Data, 'G'))
62     str1 = 'gain';
63 elseif(strcmp(Data, 'P'))
64     str1 = 'phase';
65 elseif(strcmp(Data, 'J'))
66     str1 = 'gain and phase';
67 elseif(strcmp(Data, 'I'))
68     str1 = 'interaction';
69 end
70
71 if(Peak == 1)
72     str2 = 'first peak';
73 elseif(Peak == 2)
74     str2 = 'second peak';
75 elseif(Peak == 3)
76     str2 = 'third peak';
77 else
78     str2 = 'all frequencies';
79 end
80
81 %Trim data to specific range and create neural network
82 [XT1 YT1 XV1 YV1] = Trim_Data(XT, YT, XV, YV, Sensor, Peak, Data);
83 if(strcmp(Data, 'J') || strcmp(Data, 'I'))
84     XT1 = XT1(:, 1:3:end); XV1 = XV1(:, 1:3:end);
85 end
86 XT1 = XT1'; YT1 = YT1'; XV1 = XV1'; YV1 = YV1';
87
88 net = fitnet(10);
89 net.inputs{1}.processFcns = {'removeconstantrows', 'mapminmax'};
90 net.outputs{2}.processFcns = {'removeconstantrows', 'mapminmax'};
91
92 net.divideFcn = 'dividerand';
93 net.divideMode = 'sample';
94 net.divideParam.trainRatio = 0.60;
95 net.divideParam.valRatio = 0.20;
96 net.divideParam.testRatio = 0.20;
97 net.trainFcn = 'trainlm';
98 net.performFcn = 'mse';
99
100 [net, ~] = train(net, XT1, YT1);
101
102 %Fit training dataset to the neural network then calculate the R^2 and
    RMSE values

```



```

103     yfit = (net(XT1))';
104     YT1 = YT1';
105     for i = 1 : length(yfit)
106         if(yfit(i, 1) < 0)
107             yfit(i, 1) = 0;
108         end
109     end
110     [~, m, b] = regression(YT1', yfit');
111     SSE = sum((yfit - (b + m.*YT1)).^2);
112     SSyy = sum((yfit - mean(yfit)).^2);
113     Rsq = 1 - SSE/SSyy;
114     RMSE = sqrt(SSE / (length(YT1) - 2));
115     fprintf('Training R^2 for %s data, %s: %f\n', str1, str2, Rsq)
116     fprintf('Training RMSE for %s data, %s: %f\n', str1, str2, RMSE)
117
118     %Fit validation dataset to the neural network then calculate the R^2
    and RMSE values
119     yfit = (net(XV1))';
120     YV1 = YV1';
121     for i = 1 : length(yfit)
122         if(yfit(i, 1) < 0)
123             yfit(i, 1) = 0;
124         end
125     end
126     [~, m, b] = regression(YV1', yfit');
127     SSE = sum((yfit - (b + m.*YV1)).^2);
128     SSyy = sum((yfit - mean(yfit)).^2);
129     Rsq = 1 - SSE/SSyy;
130     RMSE = sqrt(SSE / (length(YT1) - 2));
131     fprintf('Validation R^2 for %s data, %s: %f\n', str1, str2, Rsq)
132     fprintf('Validation RMSE for %s data, %s: %f\n', str1, str2, RMSE)

```

Appendix I - Supplementary Code

```
1 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2 % Import Data Function %
3 % %
4 % This program takes a summary Excel spreadsheet, reads the %
5 % data from the list of Excel spreadsheets, and returns %
6 % matrices for the training and validation datasets. %
7 % %
8 % June 8, 2012 %
9 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
10
11 function [XT YT XV YV] = Import_Data(Sensor, Contaminant)
12
13 %Inputs:
14 %Sensor - Sensor for regression ('SP' 'LP' 'SB' or 'LB')
15 %Contaminant - Contaminant for regression (Ammonia - 'A', Glycerol -
16 %'G', Ethanol - 'E')
17
18 %Outputs:
19 %XT - Matrix of training predictor variables
20 %YT - Matrix of training response variables
21 %XV - Matrix of validation predictor variables
22 %YV - Matrix of validation response variables
23
24
25 clc;
26
27 %Read data from Excel spreadsheets and check for proper size
28 data1 = xlsread(strcat('Training', Sensor, Contaminant, '.xls'));
29 data2 = xlsread(strcat('Validation', Sensor, Contaminant, '.xls'));
30 [m1 n1] = size(data1);
31 [m2 n2] = size(data2);
32
33 if((n1 ~= 6) || (n2 ~= 6) || (m1 ~= m2))
34 error('The data dimensions are not valid.')
35 end
36
37 XT = zeros(3 * m1, 1828);
38 YT = zeros(3 * m1, 1);
39 XV = zeros(3 * m2, 1828);
40 YV = zeros(3 * m2, 1);
41
42 %Copy the data from the original files to the training matrices
43 %Iterate over the list of files in the Training file
44 for i = 1 : m1
45 if(data1(i, 2) < 10)
46 data = xlsread(strcat('0', num2str(data1(i, 1)), '-0',
47 num2str(data1(i, 2)), '.xls'));
48 else
49 data = xlsread(strcat('0', num2str(data1(i, 1)), '-',
50 num2str(data1(i, 2)), '.xls'));
51 end
52 [m3 n3] = size(data);
```

```

49
50     if((m3 ~= 1910) || (n3 ~= 4))
51         error(strcat('The data dimensions of ', data1(i, 1), '.xls are not
valid.'))
52     end
53
54     %Response
55     YT(3 * i - 2, 1) = data1(i, 3);
56     YT(3 * i - 1, 1) = data1(i, 3);
57     YT(3 * i, 1) = data1(i, 3);
58
59     %Temperature
60     XT(3 * i - 2, 1) = data1(i, 4);
61     XT(3 * i - 1, 1) = data1(i, 5);
62     XT(3 * i, 1) = data1(i, 6);
63
64     for j = 1 : 609
65         %Gain
66         XT(3 * i - 2, j + 1) = data(j + 27, 3);
67         XT(3 * i - 1, j + 1) = data(j + 663, 3);
68         XT(3 * i, j + 1) = data(j + 1299, 3);
69
70         %Phase
71         XT(3 * i - 2, j + 610) = data(j + 27, 4);
72         XT(3 * i - 1, j + 610) = data(j + 663, 4);
73         XT(3 * i, j + 610) = data(j + 1299, 4);
74
75         %Interaction
76         XT(3 * i - 2, j + 1219) = data(j + 27, 3) * data(j + 27, 4);
77         XT(3 * i - 1, j + 1219) = data(j + 663, 3) * data(j + 663, 4);
78         XT(3 * i, j + 1219) = data(j + 1299, 3) * data(j + 1299, 4);
79     end
80 end
81
82 %Iterate over the list of files in the Validation file
83 for i = 1 : m2
84     if(data2(i, 2) < 10)
85         data = xlsread(strcat('0', num2str(data2(i, 1)), '-0',
num2str(data2(i, 2)), '.xls'));
86     else
87         data = xlsread(strcat('0', num2str(data2(i, 1)), '-',
num2str(data2(i, 2)), '.xls'));
88     end
89     [m3 n3] = size(data);
90
91     if((m3 ~= 1910) || (n3 ~= 4))
92         error(strcat('The data dimensions of ', data2(i, 1), '.xls are not
valid.'))
93     end
94
95     %Response
96     YV(3 * i - 2, 1) = data2(i, 3);
97     YV(3 * i - 1, 1) = data2(i, 3);
98     YV(3 * i, 1) = data2(i, 3);
99

```

```

100     %Temperature
101     XV(3 * i - 2, 1) = data2(i, 4);
102     XV(3 * i - 1, 1) = data2(i, 5);
103     XV(3 * i, 1) = data2(i, 6);
104
105     for j = 1 : 609
106         %Gain
107         XV(3 * i - 2, j + 1) = data(j + 27, 3);
108         XV(3 * i - 1, j + 1) = data(j + 663, 3);
109         XV(3 * i, j + 1) = data(j + 1299, 3);
110
111         %Phase
112         XV(3 * i - 2, j + 610) = data(j + 27, 4);
113         XV(3 * i - 1, j + 610) = data(j + 663, 4);
114         XV(3 * i, j + 610) = data(j + 1299, 4);
115
116         %Interaction
117         XV(3 * i - 2, j + 1219) = data(j + 27, 3) * data(j + 27, 4);
118         XV(3 * i - 1, j + 1219) = data(j + 663, 3) * data(j + 663, 4);
119         XV(3 * i, j + 1219) = data(j + 1299, 3) * data(j + 1299, 4);
120     end
121 end
122
123 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
124 % Trim Data Function
125 %
126 % This program takes the input training and validation data
127 % and returns a portion of the input data based on the Sensor,
128 % Peak, and Data parameters.
129 %
130 % June 11, 2012
131 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
132
133 function [XT YT XV YV] = Trim_Data(XT1, YT1, XV1, YV1, Sensor, Peak,
Data)
134
135 %Inputs:
136 %XT1 - Original matrix of training predictor variables
137 %YT1 - Original matrix of training response variables
138 %XV1 - Original matrix of validation predictor variables
139 %YV1 - Original matrix of validation response variables
140 %Sensor - Sensor for data ('SP' 'LP' 'SB' or 'LB')
141 %Peak - Peak to trim to (0 to use entire spectrum)
142 %Data - Type of data to plot (Gain - 'G', Phase - 'P', Join - 'J',
Interaction - 'I')
143
144 %Outputs:
145 %XT - Trimmed matrix of training predictor variables
146 %YT - Trimmed matrix of training response variables
147 %XV - Trimmed matrix of validation predictor variables
148 %YV - Trimmed matrix of validation response variables
149
150 %Check inputs for proper size and parameters
151 [m1 n1] = size(XT1);
152 [m2 n2] = size(YT1);

```

```

153     [m3 n3] = size(XV1);
154     [m4 n4] = size(YV1);
155
156     if((m1 ~= m2) || (m2 ~= m3) || (m3 ~= m4) || (n1 ~= 1828) || (n2 ~= 1)
|| (n3 ~= 1828) || (n4 ~= 1))
157         error('The input data dimensions are not valid.')
158     end
159
160     if((strcmp(Sensor, 'SP') == 0) && (strcmp(Sensor, 'LP') == 0) &&
(strcmp(Sensor, 'SB') == 0) && (strcmp(Sensor, 'LB') == 0))
161         error('The Sensor parameter is not valid.')
162     end
163
164     if((Peak ~= 0) && (Peak ~= 1) && (Peak ~= 2) && (Peak ~= 3))
165         error('The Peak parameter is not valid.')
166     end
167
168     if((strcmp(Data, 'G') == 0) && (strcmp(Data, 'P') == 0) &&
(strcmp(Data, 'J') == 0) && (strcmp(Data, 'I') == 0))
169         error('The Data parameter is not valid.')
170     end
171
172     %Split data into its temperature, gain, phase, and interaction
components
173     XTt = XT1(:, 1); XTg = XT1(:, 2:610); XTp = XT1(:, 611:1219); XTi =
XT1(:, 1220:1828);
174     XVt = XV1(:, 1); XVg = XV1(:, 2:610); XVp = XV1(:, 611:1219); XVi =
XV1(:, 1220:1828);
175
176     %Copy response variables
177     YT = YT1;
178     YV = YV1;
179
180     %Trim data based on Sensor and Peak parameter
181     if((strcmp(Sensor, 'SP') == 1) || (strcmp(Sensor, 'LP') == 1))
182         if(Peak == 1)
183             XTg = XTg(:, 114:198); XTp = XTp(:, 114:198); XTi = XTi(:,
114:198);
184             XVg = XVg(:, 114:198); XVp = XVp(:, 114:198); XVi = XVi(:,
114:198);
185         elseif(Peak == 2)
186             XTg = XTg(:, 413:497); XTp = XTp(:, 413:497); XTi = XTi(:,
413:497);
187             XVg = XVg(:, 413:497); XVp = XVp(:, 413:497); XVi = XVi(:,
413:497);
188         end
189     end
190
191     if(strcmp(Sensor, 'SB') == 1)
192         if(Peak == 1)
193             XTg = XTg(:, 74:183); XTp = XTp(:, 74:183); XTi = XTi(:, 74:183);
194             XVg = XVg(:, 74:183); XVp = XVp(:, 74:183); XVi = XVi(:, 74:183);
195         elseif(Peak == 2)
196             XTg = XTg(:, 301:365); XTp = XTp(:, 301:365); XTi = XTi(:,
301:365);

```

```

197     XVg = XVg(:, 301:365); XVp = XVp(:, 301:365); XVi = XVi(:,
198     301:365);
198     elseif(Peak == 3)
199         XTg = XTg(:, 496:574); XTp = XTp(:, 496:574); XTi = XTi(:,
200         496:574);
200         XVg = XVg(:, 496:574); XVp = XVp(:, 496:574); XVi = XVi(:,
201         496:574);
201     end
202 end
203
204 if(strcmp(Sensor, 'LB') == 1)
205     if(Peak == 1)
206         XTg = XTg(:, 89:148); XTp = XTp(:, 89:148); XTi = XTi(:, 89:148);
207         XVg = XVg(:, 89:148); XVp = XVp(:, 89:148); XVi = XVi(:, 89:148);
208     elseif(Peak == 2)
209         XTg = XTg(:, 285:334); XTp = XTp(:, 285:334); XTi = XTi(:,
210         285:334);
210         XVg = XVg(:, 285:334); XVp = XVp(:, 285:334); XVi = XVi(:,
211         285:334);
211     elseif(Peak == 3)
212         XTg = XTg(:, 480:549); XTp = XTp(:, 480:549); XTi = XTi(:,
213         480:549);
213         XVg = XVg(:, 480:549); XVp = XVp(:, 480:549); XVi = XVi(:,
214         480:549);
214     end
215 end
216
217 %Concatenate data based on the Data parameter
218 if(strcmp(Data, 'G') == 1)
219     XT = [XTt XTg]; XV = [XVt XVg];
220 end
221
222 if(strcmp(Data, 'P') == 1)
223     XT = [XTt XTp]; XV = [XVt XVp];
224 end
225
226 if(strcmp(Data, 'J') == 1)
227     XT = [XTt XTg XTp]; XV = [XVt XVg XVp];
228 end
229
230 if(strcmp(Data, 'I') == 1)
231     XT = [XTt XTg XTp XTi]; XV = [XVt XVg XVp XVi];
232 end

```

Appendix J - Significant Frequencies Found with Detection Method

Each shaded box represents that a significant difference between ambient air and air with contaminants for each combination of sensor, gain or phase, contaminant, and frequency.

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
10,000																								
20,000																								
30,000																								
40,000																								
60,000																								
80,000																								
120,000																								
170,000																								
250,000																								
350,000																								
500,000																								
700,000																								
1,000,000																								
1,200,000																								
1,400,000																								
1,600,000																								
1,800,000																								
2,000,000																								
2,200,000																								
2,400,000																								
2,600,000																								
2,800,000																								
3,000,000																								
3,200,000																								
3,400,000																								
3,600,000																								
3,800,000																								
4,000,000																								
4,200,000																								
4,400,000																								
4,600,000																								
4,800,000																								
5,000,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
5,200,000																								
5,400,000																								
5,600,000																								
5,800,000																								
6,000,000																								
6,200,000																								
6,400,000																								
6,600,000																								
6,800,000																								
7,000,000																								
7,200,000																								
7,400,000																								
7,600,000																								
7,800,000																								
8,000,000																								
8,200,000																								
8,400,000																								
8,600,000																								
8,800,000																								
9,000,000																								
9,200,000																								
9,400,000																								
9,600,000																								
9,800,000																								
10,000,000																								
10,200,000																								
10,400,000																								
10,600,000																								
10,800,000																								
11,000,000																								
11,200,000																								
11,400,000																								
11,600,000																								
11,800,000																								
12,000,000																								
12,200,000																								
12,400,000																								
12,600,000																								
12,800,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
13,000,000																								
13,200,000																								
13,400,000																								
13,600,000																								
13,800,000																								
14,000,000																								
14,200,000																								
14,400,000																								
14,600,000																								
14,800,000																								
15,000,000																								
15,200,000																								
15,400,000																								
15,600,000																								
15,800,000																								
16,000,000																								
16,200,000																								
16,400,000																								
16,600,000																								
16,800,000																								
17,000,000																								
17,200,000																								
17,400,000																								
17,600,000																								
17,800,000																								
18,000,000																								
18,200,000																								
18,400,000																								
18,600,000																								
18,800,000																								
19,000,000																								
19,200,000																								
19,400,000																								
19,600,000																								
19,800,000																								
20,000,000																								
20,200,000																								
20,400,000																								
20,600,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase			
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	
20,800,000																									
21,000,000																									
21,200,000																									
21,400,000																									
21,600,000																									
21,800,000																									
22,000,000																									
22,200,000																									
22,400,000																									
22,600,000																									
22,800,000																									
23,000,000																									
23,200,000																									
23,400,000																									
23,600,000																									
23,800,000																									
24,000,000																									
24,200,000																									
24,400,000																									
24,600,000																									
24,800,000																									
25,000,000																									
25,200,000																									
25,400,000																									
25,600,000																									
25,800,000																									
26,000,000																									
26,200,000																									
26,400,000																									
26,600,000																									
26,800,000																									
27,000,000																									
27,200,000																									
27,400,000																									
27,600,000																									
27,800,000																									
28,000,000																									
28,200,000																									
28,400,000																									

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
28,600,000																								
28,800,000																								
29,000,000																								
29,200,000																								
29,400,000																								
29,600,000																								
29,800,000																								
30,000,000																								
30,200,000																								
30,400,000																								
30,600,000																								
30,800,000																								
31,000,000																								
31,200,000																								
31,400,000																								
31,600,000																								
31,800,000																								
32,000,000																								
32,200,000																								
32,400,000																								
32,600,000																								
32,800,000																								
33,000,000																								
33,200,000																								
33,400,000																								
33,600,000																								
33,800,000																								
34,000,000																								
34,200,000																								
34,400,000																								
34,600,000																								
34,800,000																								
35,000,000																								
35,200,000																								
35,400,000																								
35,600,000																								
35,800,000																								
36,000,000																								
36,200,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
36,400,000																								
36,600,000																								
36,800,000																								
37,000,000																								
37,200,000																								
37,400,000																								
37,600,000																								
37,800,000																								
38,000,000																								
38,200,000																								
38,400,000																								
38,600,000																								
38,800,000																								
39,000,000																								
39,200,000																								
39,400,000																								
39,600,000																								
39,800,000																								
40,000,000																								
40,200,000																								
40,400,000																								
40,600,000																								
40,800,000																								
41,000,000																								
41,200,000																								
41,400,000																								
41,600,000																								
41,800,000																								
42,000,000																								
42,200,000																								
42,400,000																								
42,600,000																								
42,800,000																								
43,000,000																								
43,200,000																								
43,400,000																								
43,600,000																								
43,800,000																								
44,000,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase			
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	
44,200,000																									
44,400,000																									
44,600,000																									
44,800,000																									
45,000,000																									
45,200,000																									
45,400,000																									
45,600,000																									
45,800,000																									
46,000,000																									
46,200,000																									
46,400,000																									
46,600,000																									
46,800,000																									
47,000,000																									
47,200,000																									
47,400,000																									
47,600,000																									
47,800,000																									
48,000,000																									
48,200,000																									
48,400,000																									
48,600,000																									
48,800,000																									
49,000,000																									
49,200,000																									
49,400,000																									
49,600,000																									
49,800,000																									
50,000,000																									
50,200,000																									
50,400,000																									
50,600,000																									
50,800,000																									
51,000,000																									
51,200,000																									
51,400,000																									
51,600,000																									
51,800,000																									

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
52,000,000																								
52,200,000																								
52,400,000																								
52,600,000																								
52,800,000																								
53,000,000																								
53,200,000																								
53,400,000																								
53,600,000																								
53,800,000																								
54,000,000																								
54,200,000																								
54,400,000																								
54,600,000																								
54,800,000																								
55,000,000																								
55,200,000																								
55,400,000																								
55,600,000																								
55,800,000																								
56,000,000																								
56,200,000																								
56,400,000																								
56,600,000																								
56,800,000																								
57,000,000																								
57,200,000																								
57,400,000																								
57,600,000																								
57,800,000																								
58,000,000																								
58,200,000																								
58,400,000																								
58,600,000																								
58,800,000																								
59,000,000																								
59,200,000																								
59,400,000																								
59,600,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
59,800,000																								
60,000,000																								
60,200,000																								
60,400,000																								
60,600,000																								
60,800,000																								
61,000,000																								
61,200,000																								
61,400,000																								
61,600,000																								
61,800,000																								
62,000,000																								
62,200,000																								
62,400,000																								
62,600,000																								
62,800,000																								
63,000,000																								
63,200,000																								
63,400,000																								
63,600,000																								
63,800,000																								
64,000,000																								
64,200,000																								
64,400,000																								
64,600,000																								
64,800,000																								
65,000,000																								
65,200,000																								
65,400,000																								
65,600,000																								
65,800,000																								
66,000,000																								
66,200,000																								
66,400,000																								
66,600,000																								
66,800,000																								
67,000,000																								
67,200,000																								
67,400,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
67,600,000																								
67,800,000																								
68,000,000																								
68,200,000																								
68,400,000																								
68,600,000																								
68,800,000																								
69,000,000																								
69,200,000																								
69,400,000																								
69,600,000																								
69,800,000																								
70,000,000																								
70,200,000																								
70,400,000																								
70,600,000																								
70,800,000																								
71,000,000																								
71,200,000																								
71,400,000																								
71,600,000																								
71,800,000																								
72,000,000																								
72,200,000																								
72,400,000																								
72,600,000																								
72,800,000																								
73,000,000																								
73,200,000																								
73,400,000																								
73,600,000																								
73,800,000																								
74,000,000																								
74,200,000																								
74,400,000																								
74,600,000																								
74,800,000																								
75,000,000																								
75,200,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
75,400,000																								
75,600,000																								
75,800,000																								
76,000,000																								
76,200,000																								
76,400,000																								
76,600,000																								
76,800,000																								
77,000,000																								
77,200,000																								
77,400,000																								
77,600,000																								
77,800,000																								
78,000,000																								
78,200,000																								
78,400,000																								
78,600,000																								
78,800,000																								
79,000,000																								
79,200,000																								
79,400,000																								
79,600,000																								
79,800,000																								
80,000,000																								
80,200,000																								
80,400,000																								
80,600,000																								
80,800,000																								
81,000,000																								
81,200,000																								
81,400,000																								
81,600,000																								
81,800,000																								
82,000,000																								
82,200,000																								
82,400,000																								
82,600,000																								
82,800,000																								
83,000,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase			
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	
83,200,000																									
83,400,000																									
83,600,000																									
83,800,000																									
84,000,000																									
84,200,000																									
84,400,000																									
84,600,000																									
84,800,000																									
85,000,000																									
85,200,000																									
85,400,000																									
85,600,000																									
85,800,000																									
86,000,000																									
86,200,000																									
86,400,000																									
86,600,000																									
86,800,000																									
87,000,000																									
87,200,000																									
87,400,000																									
87,600,000																									
87,800,000																									
88,000,000																									
88,200,000																									
88,400,000																									
88,600,000																									
88,800,000																									
89,000,000																									
89,200,000																									
89,400,000																									
89,600,000																									
89,800,000																									
90,000,000																									
90,200,000																									
90,400,000																									
90,600,000																									
90,800,000																									

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
91,000,000																								
91,200,000																								
91,400,000																								
91,600,000																								
91,800,000																								
92,000,000																								
92,200,000																								
92,400,000																								
92,600,000																								
92,800,000																								
93,000,000																								
93,200,000																								
93,400,000																								
93,600,000																								
93,800,000																								
94,000,000																								
94,200,000																								
94,400,000																								
94,600,000																								
94,800,000																								
95,000,000																								
95,200,000																								
95,400,000																								
95,600,000																								
95,800,000																								
96,000,000																								
96,200,000																								
96,400,000																								
96,600,000																								
96,800,000																								
97,000,000																								
97,200,000																								
97,400,000																								
97,600,000																								
97,800,000																								
98,000,000																								
98,200,000																								
98,400,000																								
98,600,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase			
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	
98,800,000																									
99,000,000																									
99,200,000																									
99,400,000																									
99,600,000																									
99,800,000																									
100,000,000																									
100,200,000																									
100,400,000																									
100,600,000																									
100,800,000																									
101,000,000																									
101,200,000																									
101,400,000																									
101,600,000																									
101,800,000																									
102,000,000																									
102,200,000																									
102,400,000																									
102,600,000																									
102,800,000																									
103,000,000																									
103,200,000																									
103,400,000																									
103,600,000																									
103,800,000																									
104,000,000																									
104,200,000																									
104,400,000																									
104,600,000																									
104,800,000																									
105,000,000																									
105,200,000																									
105,400,000																									
105,600,000																									
105,800,000																									
106,000,000																									
106,200,000																									
106,400,000																									

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase		
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia
106,600,000																								
106,800,000																								
107,000,000																								
107,200,000																								
107,400,000																								
107,600,000																								
107,800,000																								
108,000,000																								
108,200,000																								
108,400,000																								
108,600,000																								
108,800,000																								
109,000,000																								
109,200,000																								
109,400,000																								
109,600,000																								
109,800,000																								
110,000,000																								
110,200,000																								
110,400,000																								
110,600,000																								
110,800,000																								
111,000,000																								
111,200,000																								
111,400,000																								
111,600,000																								
111,800,000																								
112,000,000																								
112,200,000																								
112,400,000																								
112,600,000																								
112,800,000																								
113,000,000																								
113,200,000																								
113,400,000																								
113,600,000																								
113,800,000																								
114,000,000																								
114,200,000																								

Frequency (Hz)	Small Probe Gain			Large Probe Gain			Small Box Gain			Large Box Gain			Small Probe Phase			Large Probe Phase			Small Box Phase			Large Box Phase			
	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	Glycerol	Ethanol	Ammonia	
114,400,000																									
114,600,000																									
114,800,000																									
115,000,000																									
115,200,000																									
115,400,000																									
115,600,000																									
115,800,000																									
116,000,000																									
116,200,000																									
116,400,000																									
116,600,000																									
116,800,000																									
117,000,000																									
117,200,000																									
117,400,000																									
117,600,000																									
117,800,000																									
118,000,000																									
118,200,000																									
118,400,000																									
118,600,000																									
118,800,000																									
119,000,000																									
119,200,000																									
119,400,000																									
119,600,000																									
119,800,000																									
120,000,000																									

Appendix K - Frequencies Selected in Stepwise Regression for Ethanol (Gain)

Small Probe
14,000,000
25,000,000
27,800,000
39,200,000
40,200,000
41,400,000
41,800,000
42,200,000
44,800,000
52,000,000
52,200,000
53,400,000
55,400,000
58,800,000
60,800,000
61,200,000
62,600,000
65,000,000
65,600,000
71,800,000
75,800,000
77,200,000
80,400,000
81,000,000
96,800,000
98,800,000
101,800,000
104,800,000
105,000,000
107,200,000
108,200,000
108,600,000
110,000,000
115,600,000
119,000,000
119,800,000

Large Probe
28,600,000
40,000,000
100,400,000

Small Box
350,000
9,400,000
12,800,000
17,600,000
18,400,000
26,000,000
27,000,000
29,600,000
30,000,000
33,000,000
33,600,000
34,200,000
40,000,000
40,800,000
45,400,000
50,200,000
53,000,000
59,200,000
63,400,000
66,600,000
69,000,000
72,000,000
80,400,000
81,000,000
86,200,000
90,800,000
91,000,000
92,200,000
92,400,000
93,400,000
96,000,000
101,200,000
106,600,000
108,000,000
114,200,000
118,400,000

Large Box
120,000
5,000,000
7,200,000
15,200,000
24,200,000
26,800,000
31,600,000
34,600,000
36,400,000
39,800,000
46,000,000
48,000,000
48,400,000
53,600,000
54,000,000
54,600,000
55,200,000
55,400,000
56,400,000
56,600,000
63,400,000
65,200,000
71,000,000
77,000,000
81,200,000
81,400,000
85,800,000
87,600,000
88,200,000
99,200,000
99,600,000
107,200,000
108,800,000
112,200,000
114,400,000

Appendix L - Frequencies Selected in Stepwise Regression for Ethanol (Phase)

Small Probe	Large Probe	Small Box	Large Box
60,000	13,000,000	2,200,000	2,800,000
6,200,000	13,400,000	12,000,000	4,200,000
12,000,000	53,000,000	16,000,000	5,600,000
12,400,000	58,200,000	18,000,000	13,800,000
13,000,000	63,400,000	25,800,000	18,200,000
19,600,000	67,600,000	28,400,000	19,200,000
20,800,000	69,200,000	29,600,000	19,800,000
27,400,000	69,400,000	31,200,000	21,400,000
28,400,000	78,600,000	38,000,000	23,400,000
34,200,000	113,400,000	41,600,000	27,800,000
35,800,000	113,800,000	42,200,000	28,600,000
36,800,000	116,200,000	52,000,000	33,800,000
38,400,000		53,400,000	34,200,000
41,000,000		54,600,000	35,200,000
43,800,000		58,400,000	35,600,000
45,600,000		59,400,000	36,000,000
49,600,000		60,200,000	37,000,000
49,800,000		60,800,000	41,200,000
54,000,000		62,400,000	44,800,000
56,000,000		71,000,000	45,200,000
65,200,000		81,400,000	52,600,000
66,400,000		89,600,000	76,000,000
70,600,000		92,200,000	76,200,000
78,000,000		93,000,000	77,600,000
80,000,000		94,800,000	78,800,000
82,800,000		95,600,000	83,200,000
85,000,000		97,200,000	85,000,000
87,600,000		101,000,000	87,800,000
87,800,000		101,400,000	94,600,000
94,200,000		102,000,000	106,600,000
97,800,000		102,600,000	107,200,000
106,000,000		108,200,000	109,000,000
107,600,000		110,400,000	112,400,000
114,600,000		113,400,000	115,400,000
115,000,000		118,800,000	
116,200,000		119,600,000	

Appendix M - Frequencies Selected in Stepwise Regression for Ammonia (Gain)

Small Probe	Large Probe	Small Box	Large Box
10,000	10,000	15,000,000	15,600,000
2,200,000	30,000	15,800,000	61,600,000
4,400,000	500,000	29,600,000	68,800,000
8,600,000	1,400,000	29,800,000	71,000,000
13,000,000	2,600,000	35,000,000	86,800,000
25,000,000	2,800,000	35,200,000	89,000,000
25,400,000	3,200,000	39,200,000	91,800,000
27,000,000	3,400,000	42,200,000	98,400,000
32,400,000	3,800,000	42,400,000	119,600,000
33,400,000	4,400,000	46,200,000	119,800,000
39,200,000	4,600,000	46,600,000	
39,400,000	5,000,000	59,600,000	
41,000,000	5,400,000	62,200,000	
44,200,000	6,800,000	63,600,000	
47,800,000	7,800,000	71,200,000	
49,200,000	8,800,000	73,400,000	
51,000,000	9,200,000	75,400,000	
53,000,000	9,800,000	76,400,000	
54,400,000	10,000,000	79,000,000	
54,600,000	10,400,000	80,800,000	
55,000,000	10,800,000	84,400,000	
56,200,000	11,400,000	86,000,000	
56,800,000	12,000,000	86,200,000	
57,800,000	12,200,000	88,200,000	
62,800,000	12,400,000	90,000,000	
67,200,000	13,600,000	98,000,000	
71,200,000	15,200,000	98,400,000	
82,400,000	40,000,000	100,200,000	
83,000,000	47,800,000	101,600,000	
95,200,000	49,000,000	102,600,000	
101,200,000	52,200,000	108,800,000	
105,200,000	54,000,000	109,600,000	
106,600,000	56,000,000	112,600,000	
112,600,000	58,400,000	112,800,000	
116,800,000	59,800,000	113,800,000	
	62,800,000	115,600,000	
	73,400,000		

Appendix N - Frequencies Selected in Stepwise Regression for Ammonia (Phase)

Small Probe	Large Probe	Small Box	Large Box
3,200,000	10,000	10,000	6,000,000
5,200,000	170,000	1,000,000	7,000,000
6,000,000	2,000,000	2,400,000	7,600,000
6,200,000	2,600,000	3,200,000	33,400,000
7,000,000	4,000,000	3,600,000	35,000,000
7,200,000	4,200,000	4,200,000	41,600,000
9,800,000	5,800,000	5,000,000	43,200,000
12,600,000	8,000,000	5,400,000	44,400,000
13,800,000	8,400,000	5,800,000	45,400,000
15,000,000	9,000,000	6,400,000	47,400,000
15,400,000	12,200,000	8,600,000	78,000,000
15,800,000	12,400,000	9,800,000	80,000,000
35,200,000	13,600,000	36,200,000	80,600,000
35,800,000	13,800,000	39,800,000	83,000,000
36,000,000	15,400,000	42,600,000	
37,600,000	18,200,000	42,800,000	
42,400,000	19,200,000	43,600,000	
46,600,000	20,200,000	48,400,000	
53,200,000	20,600,000	49,600,000	
54,200,000	21,600,000	50,400,000	
55,200,000	23,200,000	52,200,000	
57,200,000	27,600,000	53,000,000	
58,400,000	27,800,000	53,800,000	
60,000,000	28,000,000	54,400,000	
70,400,000	28,600,000	55,800,000	
75,000,000	28,800,000	72,200,000	
80,200,000	45,400,000	72,400,000	
82,600,000	48,200,000	75,800,000	
91,400,000	51,000,000	77,200,000	
100,000,000	55,800,000	78,400,000	
107,000,000	56,200,000	83,000,000	
112,000,000	72,800,000	83,400,000	
112,200,000	74,400,000	86,400,000	
112,800,000	75,400,000	86,800,000	
113,600,000	100,000,000	89,200,000	
	113,600,000	118,200,000	
	115,800,000		

Appendix O - Results with Neural Network Combined Test with Small Probe Sensor

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
No contaminants	Ammonia 1.03	Nothing 0.32	Nothing 0.44	Nothing -0.19
No contaminants	Ammonia 0.70	Ammonia 0.68	Nothing 0.04	Nothing -0.61
No contaminants	Ammonia 0.91	Ammonia 0.58	Nothing -0.04	Nothing 0.10
No contaminants	Nothing -0.05	Ethanol 1.56	Ammonia 0.67	Ethanol 2.13
No contaminants	Nothing -0.11	Ammonia 0.53	Ammonia 0.52	Ammonia 1.11
No contaminants	Nothing -0.10	Ammonia 1.42	Nothing 0.44	Ammonia 1.47
No contaminants	Nothing -0.05	Nothing -0.05	Ammonia 0.67	Ammonia 0.67
No contaminants	Nothing -0.11	Nothing -0.11	Ammonia 0.52	Ammonia 0.52
No contaminants	Nothing -0.10	Nothing -0.10	Nothing 0.44	Nothing 0.44
No contaminants	Nothing 0.05	Nothing 0.05	Nothing 0.45	Nothing 0.45
No contaminants	Ethanol 1.53	Ethanol 1.53	Ethanol 1.94	Ethanol 1.94
No contaminants	Ethanol 1.87	Ethanol 1.87	Nothing 0.48	Nothing 0.48
50 ppm Ammonia	Ammonia 1.07	Ammonia 1.39	Ammonia 0.97	Nothing 0.44
50 ppm Ammonia	Ammonia 1.07	Nothing 0.40	Ammonia 1.03	Ammonia 1.22
50 ppm Ammonia	Ammonia 1.09	Nothing 0.49	Ammonia 0.79	Ammonia 1.16
100 ppm Ammonia	Ammonia 1.09	Ammonia 1.03	Nothing 0.48	Ammonia 0.60
100 ppm Ammonia	Ammonia 0.99	Ammonia 0.99	Nothing 0.34	Ethanol 1.57
100 ppm Ammonia	Ammonia 1.01	Ammonia 1.01	Ammonia 0.66	Ammonia 1.44
200 ppm Ammonia	Ammonia 0.97	Ammonia 1.20	Ethanol 1.71	Nothing 0.45
200 ppm Ammonia	Ammonia 1.03	Ammonia 1.04	Ammonia 1.16	Ammonia 0.83
200 ppm Ammonia	Ammonia 1.03	Ammonia 1.21	Ammonia 1.15	Ammonia 0.92
300 ppm Ammonia	Ammonia 1.08	Ammonia 1.00	Ammonia 0.83	Nothing -0.67
300 ppm Ammonia	Ammonia 1.09	Ammonia 0.64	Ammonia 1.29	Ammonia 0.86
300 ppm Ammonia	Ammonia 1.10	Nothing 0.44	Ammonia 1.35	Ammonia 0.82
450 ppm Ammonia	Ammonia 1.15	Ammonia 1.26	Ammonia 1.00	Ammonia 1.01
450 ppm Ammonia	Ammonia 1.05	Ammonia 1.10	Ammonia 0.92	Ammonia 1.16
450 ppm Ammonia	Ammonia 1.03	Ammonia 1.00	Ammonia 0.74	Ammonia 1.15
600 ppm Ammonia	Ammonia 0.91	Ammonia 0.87	Ammonia 0.60	Ammonia 1.06
600 ppm Ammonia	Ammonia 1.05	Ammonia 0.54	Ammonia 1.21	Ammonia 1.38
600 ppm Ammonia	Ammonia 1.05	Ammonia 0.57	Ammonia 1.27	Ethanol 1.81
750 ppm Ammonia	Ammonia 1.14	Ammonia 1.06	Ammonia 0.78	Nothing -0.57
750 ppm Ammonia	Ammonia 1.06	Ammonia 1.06	Ammonia 0.59	Nothing 0.29
750 ppm Ammonia	Ammonia 1.08	Ammonia 1.02	Ammonia 1.05	Ammonia 1.44
900 ppm Ammonia	Ammonia 0.68	Ammonia 1.06	Nothing 0.49	Ammonia 0.73
900 ppm Ammonia	Ammonia 1.07	Ammonia 1.01	Ammonia 1.28	Ammonia 1.20
900 ppm Ammonia	Ammonia 1.01	Nothing 0.32	Ammonia 1.26	Ammonia 1.48
1200 ppm Ammonia	Ammonia 1.09	Ammonia 0.96	Ammonia 0.81	Nothing -0.15
1200 ppm Ammonia	Ammonia 1.01	Ammonia 0.84	Nothing 0.36	Ammonia 0.54
1200 ppm Ammonia	Ammonia 1.14	Ammonia 1.18	Ammonia 0.52	Ammonia 1.26
1500 ppm Ammonia	Ammonia 1.08	Ammonia 0.97	Ammonia 0.50	Ammonia 1.13
1500 ppm Ammonia	Ammonia 1.09	Ammonia 0.81	Ammonia 1.05	Ammonia 1.42
1500 ppm Ammonia	Ammonia 1.13	Ammonia 0.68	Ammonia 1.12	Nothing 0.00
2400 ppm Ammonia	Ammonia 1.03	Ammonia 1.09	Ammonia 0.84	Nothing 0.25

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
2400 ppm Ammonia	Ammonia 0.89	Ammonia 1.10	Ammonia 1.16	Ammonia 0.89
2400 ppm Ammonia	Ammonia 1.04	Ammonia 1.12	Ammonia 1.17	Ammonia 0.92
3000 ppm Ammonia	Ammonia 1.08	Ammonia 1.34	Ammonia 0.50	Nothing 0.47
3000 ppm Ammonia	Ammonia 1.07	Ammonia 0.62	Ammonia 0.52	Ammonia 1.32
3000 ppm Ammonia	Ammonia 1.11	Ammonia 0.93	Nothing 0.41	Ammonia 1.28
98 ppm Ethanol	Ethanol 1.87	Ethanol 1.94	Ethanol 2.16	Ethanol 2.07
98 ppm Ethanol	Ethanol 2.12	Ammonia 1.07	Ethanol 2.21	Ammonia 1.08
98 ppm Ethanol	Ethanol 1.86	Ammonia 1.09	Ammonia 1.13	Ammonia 1.26
295 ppm Ethanol	Ethanol 1.58	Ethanol 1.81	Ammonia 1.07	Ethanol 1.71
295 ppm Ethanol	Ethanol 1.71	Ammonia 1.31	Ethanol 2.02	Ammonia 1.22
295 ppm Ethanol	Ammonia 1.24	Ammonia 1.39	Ethanol 2.06	Ammonia 1.18
699 ppm Ethanol	Ethanol 1.75	Ethanol 1.57	Ethanol 2.37	Ethanol 2.09
699 ppm Ethanol	Ethanol 1.62	Ethanol 1.65	Ethanol 1.97	Ammonia 1.40
699 ppm Ethanol	Ethanol 1.83	Ammonia 1.21	Ethanol 1.51	Ammonia 0.85
1005 ppm Ethanol	Ethanol 1.82	Ethanol 2.42	Ethanol 2.22	Ethanol 2.16
1005 ppm Ethanol	Ethanol 2.01	Ammonia 1.40	Ethanol 2.13	Ammonia 1.48
1005 ppm Ethanol	Ethanol 1.79	Ammonia 1.41	Ammonia 1.41	Ammonia 1.23
1507 ppm Ethanol	Ethanol 1.74	Ethanol 2.01	Ethanol 2.39	Ethanol 1.78
1507 ppm Ethanol	Ethanol 1.72	Ammonia 1.46	Ethanol 1.87	Ammonia 1.21
1507 ppm Ethanol	Ethanol 1.83	Ammonia 1.36	Ethanol 2.09	Ethanol 1.62
2501 ppm Ethanol	Ethanol 1.81	Ethanol 1.69	Ethanol 2.12	Ethanol 1.62
2501 ppm Ethanol	Ethanol 2.00	Ethanol 1.56	Ethanol 1.82	Ammonia 1.42
2501 ppm Ethanol	Ethanol 2.03	Ethanol 1.78	Ammonia 1.43	Ammonia 0.90
3014 ppm Ethanol	Ethanol 1.72	Ethanol 1.66	Ethanol 2.06	Ammonia 1.19
3014 ppm Ethanol	Ethanol 1.79	Ethanol 1.82	Ethanol 1.67	Ethanol 1.80
3014 ppm Ethanol	Ethanol 1.74	Ammonia 1.06	Glycerol 2.52	Ammonia 1.05
3997 ppm Ethanol	Ethanol 1.76	Ethanol 2.19	Ammonia 1.27	Ammonia 1.44
3997 ppm Ethanol	Ammonia 1.23	Ethanol 1.55	Ethanol 2.12	Ethanol 1.50
3997 ppm Ethanol	Ethanol 1.86	Ammonia 0.95	Ethanol 2.00	Ammonia 1.03
5002 ppm Ethanol	Ethanol 2.41	Ethanol 2.08	Ethanol 2.23	Ethanol 1.90
5002 ppm Ethanol	Ethanol 1.96	Ethanol 1.99	Ethanol 2.26	Ammonia 1.18
5002 ppm Ethanol	Ethanol 1.98	Ethanol 1.59	Ethanol 1.51	Ammonia 0.85
7503 ppm Ethanol	Ethanol 2.07	Ethanol 1.85	Ethanol 1.55	Ethanol 1.70
7503 ppm Ethanol	Ethanol 2.43	Ethanol 2.15	Ethanol 1.74	Ethanol 1.66
7503 ppm Ethanol	Glycerol 2.59	Ethanol 2.21	Ethanol 2.28	Ethanol 1.85
10004 ppm Ethanol	Ethanol 1.80	Ethanol 1.91	Ethanol 2.11	Ethanol 1.75
10004 ppm Ethanol	Ethanol 2.02	Ethanol 1.68	Ammonia 1.36	Ethanol 1.53
10004 ppm Ethanol	Ethanol 2.01	Ethanol 2.44	Ammonia 1.31	Ammonia 1.30
12505 ppm Ethanol	Ammonia 1.36	Ethanol 2.28	Ammonia 1.02	Ammonia 1.33
12505 ppm Ethanol	Ethanol 2.08	Ammonia 1.19	Ethanol 2.04	Ammonia 0.89
12505 ppm Ethanol	Ethanol 1.82	Ethanol 1.78	Ethanol 1.66	Ammonia 1.02
Glycerol Present	Ethanol 1.57	Ethanol 1.57	Ammonia 1.24	Ammonia 1.24
Glycerol Present	Glycerol 2.64	Glycerol 2.64	Glycerol 2.90	Glycerol 2.90
Glycerol Present	Ethanol 2.23	Ethanol 2.23	Glycerol 2.79	Glycerol 2.79
Glycerol Present	Glycerol 2.60	Glycerol 2.60	Glycerol 2.60	Glycerol 2.60
Glycerol Present	Glycerol 2.66	Glycerol 2.66	Ethanol 2.45	Ethanol 2.45
Glycerol Present	Glycerol 2.75	Glycerol 2.75	Glycerol 2.55	Glycerol 2.55

Appendix P - Results with Neural Network Combined Test with Large Probe Sensor

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
No contaminants	Nothing 0.01	Nothing 0.43	Nothing 0.32	Nothing 0.02
No contaminants	Nothing 0.26	Ammonia 0.54	Nothing -0.01	Nothing 0.03
No contaminants	Nothing 0.40	Nothing 0.11	Nothing -0.14	Nothing 0.43
No contaminants	Nothing 0.40	Nothing 0.33	Ethanol 2.03	Ammonia 0.90
No contaminants	Ammonia 0.51	Ethanol 1.90	Ammonia 1.33	Ethanol 2.20
No contaminants	Ethanol 2.40	Ammonia 1.43	Ammonia 0.97	Ammonia 1.40
No contaminants	Ammonia 1.27	Ammonia 1.24	Ammonia 1.39	Ethanol 1.68
No contaminants	Ammonia 1.29	Ethanol 2.01	Ammonia 1.46	Ethanol 1.56
No contaminants	Ethanol 1.68	Ammonia 1.03	Ethanol 1.55	Ethanol 1.58
No contaminants	Nothing 0.33	Nothing 0.40	Ammonia 0.90	Ethanol 2.03
No contaminants	Ethanol 1.90	Ammonia 0.51	Ethanol 2.20	Ammonia 1.33
No contaminants	Ammonia 1.43	Ethanol 2.40	Ammonia 1.40	Ammonia 0.97
50 ppm Ammonia	Ammonia 0.99	Ammonia 0.77	Ethanol 1.63	Ethanol 1.57
50 ppm Ammonia	Ammonia 0.91	Nothing 0.48	Ethanol 1.82	Ethanol 1.65
50 ppm Ammonia	Ammonia 0.98	Nothing 0.13	Ethanol 1.59	Ethanol 1.55
100 ppm Ammonia	Ammonia 1.03	Ethanol 1.70	Ammonia 1.13	Ammonia 1.37
100 ppm Ammonia	Ethanol 1.69	Ammonia 0.99	Ammonia 1.29	Ethanol 1.59
100 ppm Ammonia	Ammonia 1.49	Ammonia 0.82	Ammonia 1.39	Ethanol 1.69
200 ppm Ammonia	Ammonia 1.19	Ammonia 0.87	Ammonia 0.90	Ammonia 1.25
200 ppm Ammonia	Ammonia 0.96	Ammonia 1.27	Ammonia 1.11	Ethanol 1.55
200 ppm Ammonia	Ammonia 0.99	Ammonia 0.83	Ammonia 1.07	Ammonia 0.96
300 ppm Ammonia	Ethanol 1.71	Ammonia 1.33	Glycerol 2.62	Ethanol 2.21
300 ppm Ammonia	Ammonia 1.10	Ammonia 0.73	Ethanol 2.46	Ethanol 2.26
300 ppm Ammonia	Ammonia 1.47	Ammonia 0.95	Ethanol 2.27	Glycerol 2.50
450 ppm Ammonia	Ammonia 0.99	Ammonia 1.21	Ammonia 1.27	Ethanol 1.72
450 ppm Ammonia	Ammonia 0.96	Ammonia 0.89	Ammonia 1.40	Ethanol 2.05
450 ppm Ammonia	Ammonia 0.96	Ammonia 0.51	Ammonia 1.44	Ethanol 1.74
600 ppm Ammonia	Ammonia 0.96	Ammonia 0.52	Ammonia 0.67	Ammonia 0.88
600 ppm Ammonia	Ammonia 0.71	Nothing 0.28	Ammonia 0.78	Ammonia 0.60
600 ppm Ammonia	Ammonia 0.89	Ammonia 0.66	Ammonia 0.99	Ammonia 1.10
750 ppm Ammonia	Ammonia 0.95	Ammonia 0.66	Ammonia 1.21	Ammonia 1.38
750 ppm Ammonia	Ammonia 0.96	Ammonia 0.56	Ethanol 1.78	Ethanol 1.91
750 ppm Ammonia	Ammonia 0.97	Ammonia 0.87	Ammonia 1.32	Ethanol 1.90
900 ppm Ammonia	Ammonia 1.27	Nothing 0.00	Ethanol 1.90	Nothing 0.43
900 ppm Ammonia	Ammonia 1.13	Nothing 0.14	Ammonia 0.63	Ammonia 0.64
900 ppm Ammonia	Ammonia 1.13	Nothing 0.40	Ammonia 0.73	Ammonia 0.80
1200 ppm Ammonia	Ammonia 1.04	Ammonia 0.75	Ammonia 1.04	Ammonia 0.65
1200 ppm Ammonia	Ammonia 0.86	Nothing 0.48	Ammonia 0.96	Ammonia 0.98
1200 ppm Ammonia	Ammonia 0.99	Ammonia 1.02	Ammonia 1.07	Ammonia 1.13
1500 ppm Ammonia	Ammonia 1.00	Ammonia 1.10	Ammonia 0.68	Nothing 0.44
1500 ppm Ammonia	Ammonia 1.00	Ammonia 1.21	Ammonia 0.56	Ammonia 0.80
1500 ppm Ammonia	Ammonia 1.16	Ammonia 1.06	Ammonia 0.69	Nothing 0.45
2400 ppm Ammonia	Ammonia 1.12	Ammonia 0.83	Ammonia 0.94	Ammonia 0.77

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
2400 ppm Ammonia	Ammonia 1.40	Ethanol 1.59	Ammonia 0.97	Ammonia 1.07
2400 ppm Ammonia	Ammonia 1.15	Ammonia 1.22	Ammonia 0.70	Ammonia 1.11
3000 ppm Ammonia	Ammonia 1.00	Ammonia 1.30	Ammonia 0.88	Ammonia 1.29
3000 ppm Ammonia	Ammonia 1.37	Ammonia 1.06	Ammonia 0.71	Ammonia 0.98
3000 ppm Ammonia	Ammonia 1.16	Ammonia 1.04	Ammonia 1.24	Ethanol 1.63
98 ppm Ethanol	Ethanol 1.94	Ethanol 1.71	Ethanol 2.48	Glycerol 2.54
98 ppm Ethanol	Ethanol 1.66	Ethanol 1.70	Ethanol 2.19	Ethanol 2.42
98 ppm Ethanol	Ethanol 1.98	Ethanol 2.37	Ethanol 1.96	Ethanol 2.16
295 ppm Ethanol	Ethanol 1.82	Ethanol 2.04	Ethanol 2.00	Ethanol 2.02
295 ppm Ethanol	Ammonia 1.06	Ethanol 2.02	Ethanol 1.95	Ethanol 2.22
295 ppm Ethanol	Ethanol 1.83	Glycerol 2.56	Ethanol 2.13	Ethanol 2.12
699 ppm Ethanol	Ethanol 1.80	Ethanol 1.99	Ethanol 2.24	Ethanol 2.05
699 ppm Ethanol	Ethanol 1.75	Ethanol 2.03	Ethanol 2.06	Ethanol 2.14
699 ppm Ethanol	Ethanol 2.04	Ethanol 2.07	Ethanol 2.11	Ethanol 2.03
1005 ppm Ethanol	Ammonia 0.56	Ammonia 0.95	Glycerol 2.66	Glycerol 3.15
1005 ppm Ethanol	Ammonia 0.67	Ammonia 1.15	Glycerol 2.61	Glycerol 3.39
1005 ppm Ethanol	Nothing 0.48	Ammonia 0.84	Glycerol 2.66	Glycerol 2.52
1507 ppm Ethanol	Ethanol 1.85	Glycerol 2.88	Ethanol 1.98	Ethanol 2.26
1507 ppm Ethanol	Ethanol 1.83	Ethanol 1.59	Ethanol 2.06	Ethanol 2.32
1507 ppm Ethanol	Ethanol 1.79	Ethanol 1.53	Ethanol 1.97	Ethanol 2.09
2501 ppm Ethanol	Ethanol 1.78	Glycerol 2.62	Ethanol 2.41	Ethanol 2.12
2501 ppm Ethanol	Ethanol 1.92	Ethanol 1.58	Ethanol 2.05	Ethanol 2.05
2501 ppm Ethanol	Ethanol 2.00	Ethanol 1.68	Ethanol 2.02	Ethanol 2.03
3014 ppm Ethanol	Ethanol 1.77	Ethanol 1.53	Ethanol 2.37	Glycerol 2.50
3014 ppm Ethanol	Ethanol 2.04	Ethanol 1.64	Ethanol 2.39	Glycerol 2.70
3014 ppm Ethanol	Ethanol 1.86	Ethanol 2.01	Ethanol 1.91	Ethanol 2.20
3997 ppm Ethanol	Ethanol 1.91	Ethanol 1.97	Ethanol 2.32	Ethanol 2.45
3997 ppm Ethanol	Ethanol 1.84	Ammonia 1.41	Ethanol 2.28	Ethanol 2.45
3997 ppm Ethanol	Ethanol 1.60	Ethanol 1.67	Ethanol 2.16	Ethanol 2.07
5002 ppm Ethanol	Ethanol 1.78	Glycerol 2.62	Ethanol 2.41	Ethanol 2.12
5002 ppm Ethanol	Ethanol 1.92	Ethanol 1.58	Ethanol 2.05	Ethanol 2.05
5002 ppm Ethanol	Ethanol 2.00	Ethanol 1.68	Ethanol 2.02	Ethanol 2.03
7503 ppm Ethanol	Ethanol 1.85	Ethanol 2.02	Ethanol 2.15	Ethanol 2.38
7503 ppm Ethanol	Ethanol 1.79	Glycerol 2.53	Ethanol 2.12	Ethanol 2.27
7503 ppm Ethanol	Ethanol 1.99	Ethanol 2.13	Ethanol 2.19	Ethanol 2.22
10004 ppm Ethanol	Ethanol 2.05	Ethanol 1.57	Ammonia 1.42	Glycerol 3.30
10004 ppm Ethanol	Ethanol 1.94	Ammonia 1.48	Ethanol 1.55	Ethanol 2.33
10004 ppm Ethanol	Ethanol 1.93	Ammonia 1.40	Ammonia 1.49	Ethanol 1.78
12505 ppm Ethanol	Ethanol 1.81	Ethanol 1.91	Ethanol 1.89	Ethanol 1.93
12505 ppm Ethanol	Ethanol 1.89	Ethanol 2.03	Ethanol 1.82	Ethanol 1.91
12505 ppm Ethanol	Ethanol 1.96	Ethanol 1.84	Ethanol 1.86	Ethanol 1.95
Glycerol Present	Glycerol 2.78	Ethanol 1.85	Ethanol 1.86	Ethanol 2.07
Glycerol Present	Ethanol 2.10	Ammonia 1.33	Glycerol 2.64	Ethanol 2.03
Glycerol Present	Ethanol 1.51	Ethanol 1.75	Ethanol 1.86	Ethanol 2.02
Glycerol Present	Glycerol 2.74	Ethanol 2.45	Ethanol 1.86	Ethanol 1.99
Glycerol Present	Glycerol 2.62	Ethanol 1.80	Ethanol 2.02	Ethanol 1.93
Glycerol Present	Ethanol 2.22	Ethanol 2.05	Ethanol 1.82	Ethanol 1.88

Appendix Q - Results with Neural Network Combined Test with Small Box Sensor

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
No contaminants	Nothing 0.01	Nothing -0.02	Nothing 0.02	Nothing 0.41
No contaminants	Nothing 0.43	Nothing 0.00	Ammonia 0.62	Nothing 0.40
No contaminants	Nothing 0.04	Nothing -0.40	Nothing 0.02	Nothing 0.04
No contaminants	Nothing 0.06	Ethanol 1.72	Nothing 0.13	Ammonia 0.61
No contaminants	Ammonia 1.49	Ethanol 1.51	Nothing 0.03	Ammonia 0.93
No contaminants	Nothing 0.00	Ammonia 1.41	Nothing 0.10	Ammonia 1.00
No contaminants	Glycerol 2.61	Ammonia 0.63	Nothing 0.12	Ammonia 1.04
No contaminants	Ammonia 1.17	Ammonia 1.25	Nothing 0.13	Nothing 0.33
No contaminants	Nothing 0.49	Ammonia 1.40	Nothing 0.26	Ethanol 1.50
No contaminants	Nothing 0.00	Nothing 0.36	Nothing 0.08	Nothing 0.40
No contaminants	Nothing -0.06	Ammonia 1.39	Nothing 0.19	Ammonia 0.99
No contaminants	Nothing 0.02	Ammonia 0.99	Nothing 0.15	Ammonia 1.44
50 ppm Ammonia	Ammonia 1.07	Ammonia 1.23	Ammonia 0.75	Ethanol 1.71
50 ppm Ammonia	Ammonia 1.09	Ammonia 0.60	Ammonia 1.01	Ethanol 1.57
50 ppm Ammonia	Ammonia 1.01	Nothing -0.08	Ammonia 1.00	Ethanol 1.58
100 ppm Ammonia	Ammonia 1.03	Ammonia 0.82	Ammonia 1.04	Ammonia 1.28
100 ppm Ammonia	Ethanol 2.10	Ammonia 0.68	Ammonia 1.01	Ethanol 2.24
100 ppm Ammonia	Ammonia 1.39	Ammonia 0.77	Ammonia 1.01	Ethanol 1.91
200 ppm Ammonia	Ammonia 1.02	Nothing -0.26	Ammonia 1.01	Ethanol 1.99
200 ppm Ammonia	Nothing 0.27	Nothing -0.31	Ammonia 1.05	Ammonia 0.96
200 ppm Ammonia	Nothing 0.04	Nothing -0.03	Ammonia 1.16	Ammonia 1.31
300 ppm Ammonia	Ammonia 0.51	Nothing 0.05	Ammonia 1.03	Ammonia 0.75
300 ppm Ammonia	Nothing 0.37	Nothing -0.26	Ammonia 1.22	Ammonia 0.68
300 ppm Ammonia	Ammonia 0.99	Nothing 0.17	Ammonia 1.12	Ammonia 0.90
450 ppm Ammonia	Ammonia 1.03	Nothing 0.08	Ethanol 1.54	Ethanol 1.68
450 ppm Ammonia	Ammonia 1.11	Nothing -0.31	Ammonia 1.00	Ammonia 1.19
450 ppm Ammonia	Ammonia 1.00	Nothing -0.68	Ammonia 1.27	Ammonia 1.09
600 ppm Ammonia	Nothing -0.48	Nothing -0.09	Ammonia 0.99	Ammonia 1.08
600 ppm Ammonia	Ammonia 1.00	Nothing -0.35	Ammonia 0.99	Ammonia 0.84
600 ppm Ammonia	Ammonia 1.00	Nothing -0.75	Ammonia 1.03	Ammonia 1.08
750 ppm Ammonia	Ethanol 1.70	Nothing 0.41	Ethanol 1.82	Ammonia 0.96
750 ppm Ammonia	Ammonia 0.90	Nothing -0.09	Ammonia 0.99	Ammonia 1.25
750 ppm Ammonia	Ammonia 1.03	Nothing -0.49	Ammonia 0.98	Ammonia 0.77
900 ppm Ammonia	Ammonia 1.08	Nothing 0.09	Ethanol 1.89	Ethanol 1.74
900 ppm Ammonia	Ammonia 0.77	Ammonia 1.17	Ammonia 1.02	Ethanol 2.34
900 ppm Ammonia	Ammonia 1.02	Ethanol 1.55	Ammonia 0.53	Ammonia 1.27
1200 ppm Ammonia	Ethanol 1.56	Ammonia 1.30	Ammonia 1.06	Ammonia 0.89
1200 ppm Ammonia	Ammonia 1.01	Nothing -0.23	Ammonia 1.01	Ammonia 0.73
1200 ppm Ammonia	Ammonia 1.01	Nothing 0.36	Ammonia 0.76	Ammonia 0.66
1500 ppm Ammonia	Ammonia 1.07	Ammonia 0.66	Ammonia 0.98	Nothing 0.16
1500 ppm Ammonia	Ammonia 0.81	Nothing -0.07	Ammonia 1.03	Ammonia 1.04
1500 ppm Ammonia	Ammonia 1.01	Nothing 0.37	Ammonia 1.05	Nothing 0.27
2400 ppm Ammonia	Ammonia 1.02	Ammonia 0.50	Ammonia 0.98	Ammonia 1.30

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
2400 ppm Ammonia	Ammonia 1.05	Nothing -0.15	Ammonia 0.98	Ethanol 1.68
2400 ppm Ammonia	Ethanol 2.01	Ammonia 0.50	Ammonia 1.42	Ethanol 1.66
3000 ppm Ammonia	Ammonia 1.00	Ammonia 0.63	Ammonia 1.02	Ammonia 0.93
3000 ppm Ammonia	Ammonia 0.60	Nothing 0.36	Ammonia 1.29	Ammonia 0.72
3000 ppm Ammonia	Ammonia 0.99	Ammonia 0.98	Ammonia 1.01	Nothing -0.17
98 ppm Ethanol	Ethanol 2.02	Ammonia 1.12	Ethanol 1.72	Ammonia 0.78
98 ppm Ethanol	Ethanol 1.87	Ethanol 1.67	Ethanol 2.19	Ammonia 1.24
98 ppm Ethanol	Ethanol 2.01	Ammonia 1.46	Ethanol 2.05	Ethanol 1.69
295 ppm Ethanol	Ethanol 2.01	Ammonia 0.95	Ethanol 1.94	Ethanol 1.94
295 ppm Ethanol	Ethanol 2.01	Ammonia 1.16	Ethanol 2.31	Ethanol 1.56
295 ppm Ethanol	Ethanol 2.00	Ammonia 1.00	Ethanol 2.41	Ethanol 1.83
699 ppm Ethanol	Ethanol 2.06	Ethanol 2.15	Ethanol 2.10	Ethanol 1.92
699 ppm Ethanol	Ethanol 2.03	Ethanol 1.64	Ethanol 1.74	Glycerol 3.01
699 ppm Ethanol	Ethanol 2.22	Ethanol 1.97	Ethanol 2.11	Ethanol 1.87
1005 ppm Ethanol	Ethanol 2.06	Ammonia 1.14	Ethanol 1.86	Ethanol 1.98
1005 ppm Ethanol	Ethanol 2.25	Ethanol 1.60	Ammonia 1.40	Ethanol 2.05
1005 ppm Ethanol	Ethanol 1.87	Ethanol 1.68	Ethanol 1.89	Ethanol 2.05
1507 ppm Ethanol	Ethanol 2.02	Ethanol 1.89	Ethanol 2.02	Glycerol 2.90
1507 ppm Ethanol	Ethanol 2.40	Ethanol 1.82	Ethanol 2.06	Ethanol 1.62
1507 ppm Ethanol	Ethanol 2.30	Ethanol 1.61	Ethanol 2.22	Ethanol 1.91
2501 ppm Ethanol	Ethanol 2.01	Ethanol 1.76	Ethanol 2.01	Ethanol 1.91
2501 ppm Ethanol	Ethanol 2.03	Nothing 0.20	Ethanol 2.10	Ethanol 1.69
2501 ppm Ethanol	Ethanol 2.03	Ethanol 2.19	Ethanol 2.21	Ammonia 1.43
3014 ppm Ethanol	Ethanol 1.84	Ethanol 1.71	Ethanol 1.97	Ethanol 1.92
3014 ppm Ethanol	Ammonia 1.13	Ammonia 1.26	Ethanol 2.17	Ethanol 2.39
3014 ppm Ethanol	Glycerol 2.76	Ethanol 2.45	Ethanol 2.17	Ethanol 1.90
3997 ppm Ethanol	Ethanol 2.00	Ethanol 2.21	Ethanol 1.74	Ethanol 1.60
3997 ppm Ethanol	Glycerol 2.90	Ethanol 2.08	Ethanol 2.29	Ammonia 1.48
3997 ppm Ethanol	Ethanol 2.00	Ammonia 0.96	Ethanol 2.27	Ethanol 2.12
5002 ppm Ethanol	Ethanol 2.04	Nothing 0.25	Ethanol 1.93	Ammonia 1.43
5002 ppm Ethanol	Ethanol 2.01	Ethanol 1.98	Ethanol 2.15	Ethanol 1.63
5002 ppm Ethanol	Ethanol 2.07	Ethanol 1.59	Ethanol 2.11	Ammonia 1.39
7503 ppm Ethanol	Ethanol 1.98	Ammonia 0.78	Ethanol 1.91	Ethanol 1.89
7503 ppm Ethanol	Ethanol 1.99	Ammonia 1.18	Ethanol 2.21	Ammonia 1.20
7503 ppm Ethanol	Ethanol 2.00	Ethanol 2.11	Ethanol 2.26	Ammonia 1.38
10004 ppm Ethanol	Ethanol 2.04	Ammonia 1.02	Ethanol 1.89	Ammonia 1.15
10004 ppm Ethanol	Ethanol 2.02	Ethanol 1.92	Ethanol 2.11	Ethanol 1.80
10004 ppm Ethanol	Ethanol 1.99	Ethanol 1.95	Ethanol 2.15	Ethanol 1.75
12505 ppm Ethanol	Ethanol 2.07	Ethanol 2.12	Ethanol 2.03	Ammonia 1.41
12505 ppm Ethanol	Ethanol 2.07	Ammonia 1.12	Ethanol 1.89	Ammonia 1.43
12505 ppm Ethanol	Ethanol 2.01	Ethanol 2.16	Ethanol 1.59	Ethanol 1.50
Glycerol Present	Ammonia 1.25	Ammonia 1.00	Glycerol 2.80	Ammonia 0.82
Glycerol Present	Glycerol 3.02	Ammonia 1.40	Ammonia 0.75	Ethanol 1.83
Glycerol Present	Glycerol 2.97	Ethanol 1.52	Ammonia 0.96	Ammonia 1.46
Glycerol Present	Glycerol 3.02	Ethanol 1.70	Ammonia 1.34	Ammonia 1.05
Glycerol Present	Glycerol 3.13	Ammonia 0.95	Ethanol 1.85	Ammonia 1.49
Glycerol Present	Ethanol 1.75	Nothing 0.12	Ethanol 1.68	Ammonia 1.38

Appendix R - Results with Neural Network Combined Test with Large Box Sensor

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
No contaminants	Nothing 0.01	Nothing 0.23	Nothing 0.03	Nothing -0.36
No contaminants	Nothing 0.01	Nothing -0.02	Nothing 0.08	Ethanol 1.72
No contaminants	Nothing 0.01	Nothing 0.26	Nothing 0.06	Nothing -1.07
No contaminants	Nothing 0.00	Ethanol 1.77	Nothing 0.27	Ethanol 1.58
No contaminants	Nothing 0.00	Ethanol 1.50	Ethanol 2.09	Ammonia 1.39
No contaminants	Nothing 0.00	Ammonia 0.92	Nothing 0.11	Ethanol 1.57
No contaminants	Nothing 0.01	Ammonia 1.20	Nothing 0.09	Ammonia 0.66
No contaminants	Nothing 0.31	Nothing 0.38	Nothing 0.06	Ammonia 0.64
No contaminants	Nothing 0.00	Ammonia 0.77	Ammonia 0.77	Ammonia 1.34
No contaminants	Nothing 0.00	Ethanol 1.77	Nothing 0.27	Ethanol 1.58
No contaminants	Nothing 0.00	Ethanol 1.50	Ethanol 2.09	Ammonia 1.39
No contaminants	Nothing 0.00	Ammonia 0.92	Nothing 0.11	Ethanol 1.57
50 ppm Ammonia	Ammonia 1.00	Ammonia 0.98	Ammonia 1.00	Ethanol 1.91
50 ppm Ammonia	Ammonia 1.00	Ammonia 1.16	Ammonia 1.12	Ethanol 1.50
50 ppm Ammonia	Ammonia 1.00	Ammonia 1.08	Ammonia 1.11	Ethanol 1.68
100 ppm Ammonia	Ammonia 1.00	Ammonia 0.55	Ethanol 1.70	Ammonia 1.04
100 ppm Ammonia	Ammonia 1.00	Nothing 0.49	Ethanol 1.71	Ethanol 1.52
100 ppm Ammonia	Ammonia 0.89	Ammonia 0.69	Ammonia 0.98	Ammonia 1.03
200 ppm Ammonia	Ammonia 1.00	Ethanol 1.79	Ammonia 1.25	Ammonia 0.94
200 ppm Ammonia	Ammonia 1.00	Ammonia 1.02	Ammonia 0.96	Nothing 0.44
200 ppm Ammonia	Ethanol 1.52	Ammonia 1.34	Ammonia 0.55	Ammonia 0.84
300 ppm Ammonia	Ammonia 1.01	Nothing -0.02	Ammonia 0.99	Ammonia 1.42
300 ppm Ammonia	Ammonia 0.53	Ammonia 0.55	Ammonia 0.97	Ammonia 1.25
300 ppm Ammonia	Ammonia 0.72	Nothing 0.32	Ammonia 0.68	Ammonia 0.92
450 ppm Ammonia	Ammonia 1.00	Ammonia 0.57	Ammonia 1.00	Ammonia 1.28
450 ppm Ammonia	Ammonia 1.01	Ammonia 0.85	Ammonia 1.01	Ammonia 0.89
450 ppm Ammonia	Ethanol 1.51	Ammonia 1.06	Ammonia 1.01	Ammonia 0.73
600 ppm Ammonia	Ammonia 1.00	Nothing 0.27	Ammonia 0.99	Ammonia 1.49
600 ppm Ammonia	Nothing 0.26	Nothing 0.30	Ammonia 0.90	Nothing 0.46
600 ppm Ammonia	Ammonia 0.88	Nothing 0.26	Ammonia 0.99	Nothing 0.37
750 ppm Ammonia	Ammonia 1.01	Nothing 0.24	Ammonia 0.77	Ethanol 1.69
750 ppm Ammonia	Ammonia 0.73	Nothing 0.27	Ammonia 1.06	Ammonia 1.31
750 ppm Ammonia	Nothing 0.29	Nothing 0.09	Ammonia 1.00	Ammonia 1.40
900 ppm Ammonia	Ammonia 1.23	Ammonia 1.41	Ammonia 0.62	Nothing 0.49
900 ppm Ammonia	Ammonia 1.01	Ammonia 0.70	Ammonia 0.98	Ammonia 0.75
900 ppm Ammonia	Ammonia 1.01	Nothing 0.13	Ammonia 0.98	Ammonia 0.65
1200 ppm Ammonia	Ethanol 1.70	Ammonia 1.39	Ammonia 1.08	Ammonia 1.32
1200 ppm Ammonia	Ammonia 0.67	Ammonia 0.58	Ethanol 2.00	Ethanol 2.01
1200 ppm Ammonia	Ammonia 1.01	Ammonia 0.89	Ethanol 2.27	Ammonia 1.17
1500 ppm Ammonia	Ammonia 1.01	Ammonia 0.53	Nothing 0.49	Ammonia 0.65
1500 ppm Ammonia	Nothing 0.36	Ammonia 0.76	Ammonia 1.00	Nothing 0.28
1500 ppm Ammonia	Ammonia 1.01	Ammonia 0.83	Ammonia 0.98	Nothing 0.30
2400 ppm Ammonia	Ethanol 2.19	Glycerol 2.52	Ammonia 0.95	Ammonia 0.53

Actual Conditions	Prediction with Gain Data Training	Prediction with Gain Data Validation	Prediction with Phase Data Training	Prediction with Phase Data Validation
2400 ppm Ammonia	Ammonia 1.01	Glycerol 2.67	Ammonia 1.10	Nothing 0.44
2400 ppm Ammonia	Ammonia 1.01	Ethanol 1.98	Ethanol 1.62	Ammonia 0.91
3000 ppm Ammonia	Ammonia 1.00	Ammonia 0.83	Ammonia 0.61	Nothing 0.45
3000 ppm Ammonia	Ammonia 0.83	Ammonia 1.24	Ammonia 0.99	Ammonia 0.61
3000 ppm Ammonia	Ammonia 0.72	Ammonia 0.92	Ammonia 0.99	Ammonia 0.51
98 ppm Ethanol	Ethanol 2.00	Ethanol 1.81	Ethanol 1.96	Ammonia 0.80
98 ppm Ethanol	Ammonia 1.28	Ammonia 0.98	Ethanol 1.97	Ethanol 1.67
98 ppm Ethanol	Ethanol 2.00	Ethanol 1.54	Ammonia 1.08	Ammonia 1.34
295 ppm Ethanol	Ethanol 2.00	Ethanol 1.95	Ethanol 1.98	Glycerol 2.88
295 ppm Ethanol	Ethanol 2.08	Ethanol 1.67	Ethanol 1.97	Ethanol 2.26
295 ppm Ethanol	Ethanol 2.00	Glycerol 2.52	Glycerol 2.65	Ethanol 2.02
699 ppm Ethanol	Ethanol 2.00	Ethanol 2.21	Ethanol 1.94	Ethanol 2.15
699 ppm Ethanol	Ethanol 2.00	Ethanol 1.64	Ammonia 1.23	Ammonia 1.24
699 ppm Ethanol	Ethanol 1.62	Ammonia 0.85	Ethanol 2.27	Ethanol 1.89
1005 ppm Ethanol	Ethanol 2.00	Ammonia 1.42	Ethanol 1.72	Nothing -0.71
1005 ppm Ethanol	Ethanol 2.00	Ammonia 0.93	Ethanol 1.98	Ethanol 1.78
1005 ppm Ethanol	Ethanol 1.84	Ammonia 1.39	Ethanol 1.95	Ethanol 2.14
1507 ppm Ethanol	Ethanol 2.00	Ethanol 2.47	Ethanol 1.69	Ethanol 2.13
1507 ppm Ethanol	Ethanol 1.94	Ethanol 1.78	Ethanol 1.95	Ammonia 1.30
1507 ppm Ethanol	Ethanol 2.00	Ammonia 1.34	Glycerol 2.78	Ethanol 1.88
2501 ppm Ethanol	Ethanol 2.17	Ethanol 1.50	Ethanol 2.00	Ethanol 2.12
2501 ppm Ethanol	Ethanol 2.00	Ethanol 1.55	Ethanol 1.96	Glycerol 2.62
2501 ppm Ethanol	Ethanol 2.00	Ammonia 1.40	Ethanol 1.95	Ethanol 1.90
3014 ppm Ethanol	Ethanol 2.00	Ammonia 1.43	Ethanol 2.01	Ammonia 1.40
3014 ppm Ethanol	Ethanol 2.00	Glycerol 2.71	Ethanol 1.71	Ethanol 2.34
3014 ppm Ethanol	Ethanol 2.00	Ethanol 1.52	Ethanol 1.99	Ethanol 2.42
3997 ppm Ethanol	Ethanol 2.00	Ethanol 1.61	Ammonia 1.11	Ethanol 1.51
3997 ppm Ethanol	Ethanol 2.00	Ethanol 2.23	Ethanol 1.99	Ethanol 1.64
3997 ppm Ethanol	Ethanol 1.60	Ethanol 1.54	Ethanol 1.99	Ethanol 2.27
5002 ppm Ethanol	Ethanol 2.00	Ethanol 2.33	Ethanol 1.95	Ethanol 1.72
5002 ppm Ethanol	Ethanol 2.00	Ethanol 1.65	Ethanol 1.93	Ethanol 2.18
5002 ppm Ethanol	Ethanol 1.72	Glycerol 2.60	Ethanol 2.01	Ethanol 2.01
7503 ppm Ethanol	Ethanol 2.21	Ethanol 1.79	Ethanol 2.13	Ammonia 1.14
7503 ppm Ethanol	Ammonia 0.96	Ammonia 1.47	Ethanol 1.98	Ethanol 1.91
7503 ppm Ethanol	Ethanol 1.68	Ethanol 1.77	Ethanol 1.58	Ethanol 2.32
10004 ppm Ethanol	Ethanol 2.00	Ethanol 1.87	Ammonia 1.28	Ethanol 1.50
10004 ppm Ethanol	Ammonia 1.40	Ammonia 1.41	Ethanol 1.75	Ammonia 1.27
10004 ppm Ethanol	Ammonia 0.74	Ammonia 1.38	Glycerol 2.71	Ethanol 1.50
12505 ppm Ethanol	Ethanol 2.00	Ethanol 1.79	Ethanol 1.95	Ethanol 1.95
12505 ppm Ethanol	Ethanol 2.00	Ammonia 1.38	Ethanol 1.98	Ethanol 2.32
12505 ppm Ethanol	Ethanol 2.00	Ethanol 2.13	Ethanol 1.98	Ethanol 1.89
Glycerol Present	Ammonia 1.22	Ammonia 1.34	Ethanol 2.13	Ethanol 1.84
Glycerol Present	Ammonia 1.10	Ammonia 1.33	Glycerol 2.99	Ethanol 2.08
Glycerol Present	Ammonia 1.16	Nothing 0.33	Glycerol 2.98	Ammonia 1.43
Glycerol Present	Glycerol 2.69	Glycerol 3.07	Glycerol 2.87	Glycerol 2.65
Glycerol Present	Glycerol 3.00	Glycerol 2.66	Glycerol 2.96	Glycerol 2.94
Glycerol Present	Ethanol 2.29	Glycerol 2.61	Glycerol 2.98	Ethanol 2.48