Increased Classification Rates of Chemical Vapor Detectors Using Novel Sensor Types and Optimized Sensing Geometries

Thesis by

Marc D. Woodka

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2009
(Submitted September 11, 2008)

Acknowledgements

There are a number of people to whom I owe much thanks, many of whom I would not be here without. In chronological order: to start, I thank my parents for their love and support throughout the years. This is a thing often taken for granted, but it puts those who grew up with it at a tremendous advantage in life. Without this, I certainly would not be here today.

I also thank Prof. Georges Belfort and my graduate student mentor Dr. Gautam Baruah, who welcomed me as an undergraduate student researcher at Rensselaer Polytechnic Institute. They demonstrated to me an excitement and enthusiasm for research, and made my decision to attend graduate school at Caltech an easy one.

Thanks to my research advisor, Professor Nate Lewis, who allowed me the time and freedom to guide my research and overcome the hurdles at various steps throughout my graduate career – the extra time required to study for the qualifying exams, to learn various programming languages, to figure things out, etc. Also, thanks for your well-timed motivation and ideas, and finally, thanks for letting me represent the nose project at various points throughout the past few years – it mixed things up for me, and provided me with a greater appreciation for the grand scheme of things.

Many thanks are due to Dr. Bruce Brunschwig, effectively my second research advisor. During my first year, when I realized I needed to learn to program in LabVIEW, I looked to Bruce for assistance, and he immediately gave me several hours of his time and guidance. This was when I first recognized Bruce as an invaluable resource, and I have utilized him since for a number of meaningful discussions, from which I have learned much. He was also a very critical reviewer of all chapters in this thesis, which made the chapters come out better in the end. Thanks for the discussions, ideas, insights, and careful readings!

When I first joined the Lewis group, Brian Sisk and Ting Gao were the graduate student and post-doc, respectively, on the nose project. Thanks to Brian for introducing me to the lab and mentoring me in Matlab, and thanks to Ting for leaving behind some unfinished work, which became my first publication and formed Chapter 2 of my thesis. Also, when my research was just beginning, Mike Burl from JPL suggested using non-negative least squares for mixture analysis, which has since been the best performing algorithm for the task. I hopefully would have stumbled across it on my own, but we'll never know… Thanks for the discussion and suggestion!

Thanks to everybody else in the Lewis group: David Knapp for knowing and sharing with others how things work, and for keeping things running smoothly; Joe Nemanick for introducing me to facilities such as the machine shop; Tony Fitch and Don Walker, for listening and trying to understand what was going on in my world of research, and offering suggestions and comments; Edgardo Garcia, for several collaborations and discussions, and for coming back from Puerto Rico; Stephen Maldonado, for collaborations and useful discussions; and Anna Folinsky and Heather McCaig, for interactions and discussions throughout the years. Also, thanks to everybody else from the Lewis group who has made my time here enjoyable and entertaining: Pat for getting me into diving; Craig for the visits and discussions; Cinco (Davy) for the trips to Grandma's up at Big Bear and good conversation; Pat, Steve, Tony, Don, Craig and Edgardo for the fun outside of lab. Everybody else in the Lewis group not listed here: thanks for making the Lewis group a great place to spend 5 years.

Thanks to the Chemical Engineering support group: Shane for the many well-researched adventures; Armin for the laid-back attitude and the good times (Tucson was a blast); Andrew for the well-researched culinary options in the area (Hometown Buffet); and Ubaldo and Win for the good company at these culinary options. Also, thanks to Ameri for his time and patience during our first year, when he took the time to share with me his deep understanding of all things ChemE.

Thanks to my committee, Professors Rick Flagan, Mark Davis, and Pietro Perona, for their interest in my research project and for taking the time out of their schedule to read my thesis and sit on my committee. Special thanks to Rick Flagan for suggesting FEMLAB to solve my modeling issues during my candidacy – it took me over a year of fighting with an implicit finite-difference Crank-Nicholson scheme before I finally gave in, but when I did FEMLAB worked wonderfully.

Finally, I would like to thank Andrea for her love and support through the years, and her patience in my finishing up here at Caltech. The last few years have been wonderful, and I look forward to our new life on the east coast.

Abstract

Broadly responsive vapor sensor arrays, or so-called "electronic noses," have been explored and/or used for many years as a means to detect the vapors present in the headspace of a variety of targets, such as coffees, wines, vinegars, oils, explosives, and nerve agents, and for disease diagnosis. Electronic nose sensing modalities often exhibit a response that is linear with concentration, and additive with respect to multiple vapors. Ideally, one could simply train an array towards the pure vapors of interest, and use that pure vapor training to identify either pure vapors or vapor mixtures during field-testing. This, however, has proven difficult, and has limited the utility of this vapor detection approach for a number of applications.

This thesis utilized a low-cost, low-power sensing modality, insulator – carbon black composite chemiresistors, and exploited their linear response properties to enhance the classification rates of both pure vapors and vapor mixtures, based on pure vapor training. Sensors utilizing non-polymeric, small organic molecules as the insulating component were demonstrated to offer enhanced separation between pure vapor response clusters, and lower detection limits, relative to the traditional use of polymers as the insulating phase. These sensors were then used in a sensing geometry that induced a space- and time, or spatiotemporal (ST) dependence, to the sensor response, which increased the amount of chemical information extracted from the sensor response. This ST response information allowed for the correct classification of vapor mixtures consisting of up to 5 components, with training on only pure vapors.

A mass uptake model for the ST response of the sensors was developed, and vapor detection and mixture analysis was simulated for chamber geometries and vapor delivery flow rates spanning several orders of magnitude. The data were first used to define an optimized ST sensing regime for mixture analysis, based on two dimensionless Peclet number analogs. The data were then used to identify the inherent properties of the pure vapor training data that allowed for mixture analysis to be performed at high levels, specifically that the minimum resolution factor between all binary vapor combinations in the training library was sufficiently high.

Finally, the utility of the ST response was demonstrated to offer enhanced pure-vapor classification rates, relative to the traditional steady state approach typically employed with broadly responsive array-based sensing. These enhanced classification rates were demonstrated using a number of classification algorithms, including a bioinspired algorithm based on Fisher's Linear Discriminant. In summary, the results demonstrated herein quantify, in different ways, what is required for classification optimization, and in doing so increase the utility of this approach to vapor detection for a number of applications.

Contents

List of Figures

List of Tables

Chapter 1

Introduction: Electronic Noses and Polymer-Carbon Black Composite Vapor Sensors

1.1. Electronic Nose Introduction

No man-made sensor system combines the sensitivity, low power, rapid response, selectivity, and ability to track an odorant to its source that is characteristic of the olfactory system of a canine. In mammals, G-protein-coupled receptors (GPCRs) are a broad class of trans-membrane receptors that are used in many physiological processes, such as visual transduction, hormonal regulation, and stimulation and inhibition of various processes.[1] The mammalian genome possess at least 1,000 olfactory receptor genes[2] that are part of the broader GPCRs class. Hence, olfactory receptors constitute the largest family of GPCRs in mammals. These $\approx$1,000 genes can potentially encode up to $\approx$1,000 different functional odor receptors. However, mammals are able to detect over 10,000 different odors. Thus, the receptors must be broadly responsive in their response properties. In this architecture, a given receptor will be triggered by more than one odorant, and an odorant, in turn, will produce a response from more than one receptor.[3] Olfactory receptors are triggered in the olfactory epithelium, and then send a response through the olfactory bulb to the brain, for processing and odorant identification.

This biological process of olfaction lays the foundation for artificially created "electronic noses." Electronic noses began to develop into their modern form in the mid-1980s. Prior to this, broadly responsive sensor arrays had been investigated to a limited extent; however, bottlenecks in electronics and computational capability limited progress.[4,5] New technologies have been developed, and implementations of various pattern recognition algorithms, the workhorse of any electronic nose configuration, have flourished in the past 15 years.

Figure 1.1 schematically depicts a generic electronic nose architecture. In an electronic nose, an array of multiple sensors is used. Each sensor is broadly responsive, responding to many odorants, as is the case in the mammalian olfactory system. In practice, a gaseous odorant is exposed to the sensor array, generating a time-varying sensor signal, $S_{i,j}(t)$. Each of these sensor signals is then processed, and a single metric response descriptor is generated for each sensor, creating an array response. During a training phase, the array is exposed numerous times to odorants that will be used later to challenge the array in identification tasks. The training process creates an odorant database. During subsequent testing of unidentified odorants, the array response is compared against the training library, and a prediction is made as to the identity and concentration of the odorant. Various levels of processing are required, and ultimately some form of pattern recognition is used to identify the odorant, mimicking the steps involved in odor identification in the mammalian olfactory system.

Electronic noses implement this generic architecture in various ways. Surface acoustic wave devices,[6,7] tin oxide sensors,[8,9] conducting organic polymers,[10,11] polymer-coated quartz crystal microbalances,[12,13] polymer-coated micromachined cantilevers,[14] dye-impregnated polymers coated onto optical fibers or beads,[15,16] and polymer composite chemically sensitive resistors[17,18] comprise only a few of the broadly responsive sensors that have been employed in the construction of electronic noses. Pattern recognition algorithms that have been used include linear, statistically based methods such as partial least squares, principal components analysis (PCA), Fisher's linear discriminant (FLD), k-nearest neighbors, and soft independent modeling of class analogy; as well as non-linear, non-statistical methods that include various implementations of artificial neural networks.[8,19-21] This thesis will detail the recent developments of one sensing modality that has been developed at Caltech: insulator-carbon back composite chemiresistors.

1.2. Vapor Sensing With Insulator – Conductor Composites

1.2.1. Phase Equilibrium

When two phases come into contact, equilibrium will eventually be established for all species present. This is true regardless of the types of phases involved. The equilibration process between a vapor and a solid sorption based material forms the basis of sorption-based sensors.

Consider a vapor in equilibrium with an ideal solution of the vapor in a solid sorbent. In ideal solutions at equilibrium, Raoult's law is followed for all components.[22,23] Thus:

$$P_A = x_s P_A^o, \tag{1}$$

where P_A is the pressure exerted by species A in the vapor phase, x_s is the mole fraction of the vapor in the solid sorbent, and P_A^o is the vapor pressure of the pure liquid at the same temperature. Rearranging eq (1) gives:

$$x_s = \frac{P_A}{P_A^o}. \tag{2}$$

The mole fraction, and consequently, the molarity, molality or weight percent of the vapor in the solid is a function of the ratio of the pressure exerted on the solid by species *A* divided by the vapor pressure of pure *A*. Therefore, for an ideal solution, the solubility of a vapor in any sensor material is a function only of the properties of the vapor, and not the solid.

For non-ideal solutions, eq (1) can be written in terms of activities:[22,23]

$$P_A = a_{A,s} P_A^o = \gamma_A x_s P_A^o, \tag{3}$$

where $a_{A,s}$ is the activity of vapor species *A* in the solid, γ_A is the activity coefficient of species *A* in the solid, and the vapor is assumed to be an ideal gas at ambient pressure. For real solutions, Henry's law is a better approximation for the solute then Raoult's law, hence:

$$P_A \approx k_H x_s, \tag{4}$$

where k_H is the Henry's law constant, equal to $\gamma_A P_A^o$ (eq (3)). Combining eq (3) - (4) yields:

$$x_s = \frac{1}{\gamma_A}\frac{P_A}{P_A^o} = \frac{P_A}{k_H}. \tag{5}$$

The interactions between the solid and vapor that are responsible for the differences in solubilities of a given gas in different solids, and the differences in Henry's law constants, or more generally the non-ideality, are therefore taken into account by differences in the activity coefficient for each gas/solid combination. The equilibrium mole fraction of the species in the solid is therefore dictated by the ratio of the pressure exerted by the species on the solid to the vapor pressure of the species, divided by the activity coefficient (eq (5)).

1.2.2. Insulator – Conductor Systems and Percolation Theory

When a non-conducting substance (i.e., an electrical insulator) is mixed with a conductive substance, the electrical properties of the resulting blend will exhibit several regimes of electrical conductivity as the mass fraction of the insulator (conductor) is varied. At high fractions of insulator, the composite is an insulator with a resistance near that of the pure insulator. At high fractions of conductor, the composite is a conductor. In both regimes, the resistance is approximately a linear function of the mass fraction of the conductor. The regime at which the composite begins to change from an insulator to a conductor is known as the percolation regime. In this regime, the conductance has a highly non-linear dependence on the mass fraction of conductor. At the percolation threshold, the first continuous conducting pathway occurs within the composite. At this point, the resistance of the composite depends exponentially on the loading of the conductor, and small increases in the loading of the conductor will then produce large increases in the conductivity of the composite.

Quantitatively, the resistance, R, of an insulator-conductor composite can be expressed as:

$$R = \frac{(z-2)R_c R_m}{B_1 + B_2 + [(B_1+B_2)^2 + 2(z-2)R_c R_m]^{1/2}}, \tag{6}$$

where

$$B_1 = R_c[-1 + (z/2)(1 - \upsilon_c / f)] \tag{7}$$

and

$$B_2 = R_m[(z\upsilon_c /(2f)) - 1]. \tag{8}$$

Here, R_c and R_m are the resistances of the pure conducting and insulating substances, respectively; z is the coordination number for the conducting particles in the insulating matrix (which takes into account how the two substances pack); υ_c is the volume fraction of conducting material in the composite; and f is the total packing fraction of the composite ($f \leq 1$).[24,25] Figure 1.2 plots R vs. υ_c for a hypothetical insulator-conductor composite with $R_c = 10^{-5}$ Ω-cm, $R_m = 10^{10}$ Ω-cm, $z = 4$, and a total packing fraction of $f = 0.5$. The percolation threshold, υ_p, is denoted by the sharp drop in the resistance at a conducting volume fraction $\upsilon_c = \upsilon_p = 0.25$, i.e., the volume fraction of conducting material at the percolation threshold $\upsilon_p = 2f/z$.

1.2.3. Sensing Due to Phase Equilibration Using Percolative Chemiresistor Sensors

When an insulator-conductor composite is exposed to various vapor environments, the composite will come to equilibrium with each species (odorant) in the vapor phase. Assuming that the analyte diffuses throughout the composite, its volume will increase to accommodate the presence of the analyte (Figure 1.3). This increase in volume will decrease υ_c. Assuming that the sorbed material is non-conducting, the resistance of the composite will therefore increase.

The resistance across such an insulator-conductor composite will thus normally increase upon exposure to an odorant (although if exposure to an analyte causes contraction, or if the analyte is conductive, a decrease in resistance is predicted). A generic sensor mechanism is thus illustrated in Figure 1.3. Upon exposure to an odorant, the odorant partitions into the sensor material, causing a swelling and a concomitant increase in resistance. When the odorant is removed from the vapor phase, the odorant partitions out of the sensor, the sensor shrinks, and the electrical resistance therefore decreases. Using composites of various conductors and insulators, and monitoring the resistance across these composites, provides a low-power method for transducing chemical sensory information into an electrical signal.

Various materials have been used for the insulating and conducting phases of such sensors. The insulating phase is typically a low glass-transition temperature polymer or polymer blend, and the conducting phase is typically carbon black.[18,24] However, insulating phases have included ligands chemically attached to conductive gold nanoparticle cores,[26] and the conductive phase has included gold nanoparticles,[26] conductive polymeric materials,[17,27,28] and could include other materials such as colloidal Ag, colloidal Au, or colloidal TiO_2.

1.2.4. Polymer-Carbon Black Composite Chemiresistor Properties: Response Linearity and Additivity

The ability to adjust the relative amount of the conducting material in the insulator – conductor composites allow choice of the regime in which the sensor will operate: close to the percolation threshold, or in the linear regime having a high concentration of the conductive component. Operation in the linear regime, while less sensitive, is generally preferred. In this regime, the response is directly proportional to the concentration of analyte vapor. This relationship results in a linear correlation between the sensor response and the concentration of individual components of the vapor mixtures. This linear relationship can, of course, break down at high odorant concentrations.

Figure 1.4 shows the equilibrium sensor response for a poly(butadiene)-carbon black composite vapor sensor (20% mass fraction of carbon black) upon exposure to various odorants at different fractions of their vapor pressures.[29] The sensor exhibits a linear response with increasing concentration, until either the percolation threshold is reached for the sensor (not normally possible for high carbon black loadings) or until the sorption isotherm becomes nonlinear with vapor concentration.[4]

In the linear response regime, the sensor response to a mixture of odorants is simply the sum of the individual responses to the odorants that comprise the mixture. For example, Figure 1.5 displays the response of a prototypical sensor upon exposure to some test mixtures of odorants. In this experiment, benzene and chloroform were each exposed to the sensor at 2% of each of their vapor pressures (4600 ppm for chloroform and 2200 ppm for benzene) separately, and then in combination.[29] For each odorant at the given concentration, the sensor exhibited a characteristic increase in resistance, regardless of the other odorants present. The sensor response was both additive and independent of the order of exposure.

1.3. Outline of This Thesis

The goal of any electronic nose array-based arrangement is to robustly determine the identity and concentration of unknown vapors. Ideally, due to the linearity and additivity displayed in Figures

1.4 and 1.5, respectively, regardless of whether the unknown vapors are present in the pure form or as vapor mixtures, the array will only require training on the pure vapor components. This work addresses this task using a number of different approaches. Non-polymeric – carbon black composite chemiresistor sensors are introduced and demonstrated as a promising alternative to the traditional polymer – carbon black composite approach, wherein a higher concentration of functional groups present in the sorptive component of the sensor composite allows for enhanced vapor – sensor interactions, and an enhanced ability to discriminate between chemically similar vapors (Chapter 2) This work was started by a post-doc in our laboratory and left unfinished – my contribution was everything except the initial fabrication and analysis of the non polymer – carbon black composites (20% CB): this included the development of all figures, tables, analyses and discussions presented. Further, a means of increasing the amount of information extracted from sensor arrays is demonstrated by invoking a space- and time-, or spatiotemporal (ST) dependency, of the array's response. This ST approach takes advantage of the linear and additive response properties of the sensors, and is demonstrated to significantly improve the ability of the sensors to identify and quantify vapor mixtures with training on only the pure vapor components (Chapter 3). A model for the ST response of sensor arrays is developed and implemented to define an optimized ST mixture analysis regime, defined by two dimensionless numbers characterizing the competing mass transport processes across various dimensions of the linear sensor array vapor channel (Chapter 4). This same modeled response data is then analyzed in terms of various inherent properties of the pure vapor response data, and a method for predicting the ability to analyze vapor mixtures with only pure vapor training is introduced (Chapter 5). Finally, an increased ability to classify pure vapors, using the ST vapor detection approach, is demonstrated (Chapter 6). Various pattern recognition and classification techniques are introduced and implemented, and their benefits and downfalls are discussed. The results of this work should increase the practicality and usability of broadly responsive array-based vapor sensing.

1.4. References

(1) Stryer, L.; Bourne, H. R. *Annual Review of Cell Biology* **1986**, *2*, 391-419.

(2) Buck, L.; Axel, R. *Cell* **1991**, *65*, 175-187.

(3) Axel, R. *Scientific American* **1995**, *273*, 154-159.

(4) Gardner, J. W.; Bartlett, P. N. *Electronic Noses: Principles and Applications*; Oxford University Press: New York, NY, 1999.

(5) Gardner, J. W.; Bartlett, P. N. *Sensors and Actuators B* **1994**, *18*, 211-220.

(6) Patrash, S. J.; Zellers, E. T. *Analytical Chemistry* **1993**, *65*, 2055-2066.

(7) Rose-Pehrsson, S. L.; Grate, J. W.; Ballantine, D. S.; Jurs, P. C. *Analytical Chemistry* **1988**, *60*, 2801-2811.

(8) Getino, J.; Horrillo, M. C.; Gitierrez, J.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Sensors and Actuators B* **1997**, *43*, 200-205.

(9) Srivastava, R.; Dwivedi, R.; Srivastava, S. K. *Sensors and Actuators B* **1998**, *50*, 175-180.

(10) Gardner, J. W.; Bartlett, P. N. *Sensors and Actuators A* **1995**, *51*, 57-66.

(11) Harris, P. D.; Arnold, W. M.; Andrews, M. K.; Partridge, A. C. *Sensors and Actuators B* **1997**, *42*, 177-184.

(12) Fu, Y.; Finklea, H. O. *Analytical Chemistry* **2003**, *75*, 5387-5393.

(13) Grate, J. W.; Patrash, S. J.; Kaganove, S. N.; Abraham, M. H.; Wise, B. M.; Gallagher, N. B. *Analytical Chemistry* **2001**, *73*, 5247-5259.

(14) Baller, M. K.; Lang, H. P.; Fritz, J.; Gerber, C.; Gimzewski, J. K.; Drechsler, U.; Rothuizen, H.; Despont, M.; Vettiger, P.; Battiston, F. M.; Ramseyer, J. P.; Fornaro, P.; Meyer, E.; Guntherodt, H.-J. *Ultramicroscopy* **2000**, *82*, 1-9.

(15) Albert, K. J.; Walt, D. R. *Analytical Chemistry* **2001**, *73*, 2501-2508.

(16) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*.

(17) Freund, M. S.; Lewis, N. S. *Proceedings of the National Academy of Sciences, U.S.A.* **1995**, *92*, 2652-2656.

(18) Doleman, B. J.; Sanner, R. D.; Severin, E. J.; Grubbs, R. H.; Lewis, N. S. *Analytical Chemistry* **1998**, *70*, 2560-2564.

(19) Ciosek, P.; Wroblewsk, W. *Sensors and Actuators B* **2006**, *114*, 85-93.

(20) Pardo, M.; Sisk, B. C.; Sberveglieri, G.; Lewis, N. S. *Sensors and Actuators B* **2006**, *115*, 647-655.

(21) Vaid, T. P.; Burl, M. C.; Lewis, N. S. *Analytical Chemistry* **2001**, *73*, 321-331.

(22) Atkins, P.; de Paula, J. *Atkin's Physical Chemistry*, 8th ed.; Oxford University Press: New York, 2006.

(23) McQuarrie, D. A.; Simon, J. D. *Molecular Thermodynamics*; University Science Books: Sausalito, CA, 1999.

(24) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chemistry of Materials* **1996**, *8*, 2298-2312.

(25) Lundberg, B.; Sundqvist, B. *Journal of Applied Physics* **1986**, *60*, 1074-1079.

(26) Briglin, S. M.; Gao, T.; Lewis, N. S. *Langmuir* **2004**, *20*, 299-305.

(27) Sotzing, G. A.; Briglin, S. M.; Grubbs, R. H.; Lewis, N. S. *Analytical Chemistry* **2000**, *72*, 3181-3190.

(28) Sotzing, G. A.; Phend, J. N.; Grubbs, R. H.; Lewis, N. S. *Chemistry of Materials* **2000**, *12*, 593-595.

(29) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Analytical Chemistry* **2000**, *72*, 658-668.

Figure 1.1: Generic electronic nose architecture.[5] An unknown analyte, *j*, interacts with each sensor in the array (comprised of *n* total sensors), causing a change in some time-varying signal, $S_{i,j}(t)$. The signal is processed to create a single metric response descriptor for each sensor, $X_{i,j}$. The array response is then put through a pattern recognition algorithm, which compares the unknown array response to a library of responses. The algorithm predicts if analyte *j* is one of the analytes within the libray.

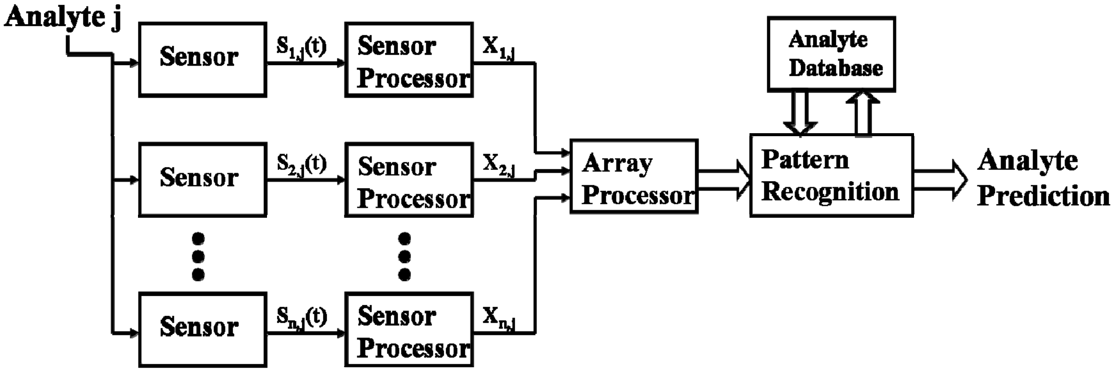


Figure 1.2: Resistance vs. the volume fraction of conducting material for a hypothetical insulator-conductor composite with $R_c = 10^{-5}$ Ω-cm, $R_m = 10^{10}$ Ω-cm, a coordination number (z) of 4, and a total packing fraction (f) of 0.5 (eq (6)). The volume fraction at which the sharp drop in resistance occurs is the percolation threshold for this hypothetical composite.

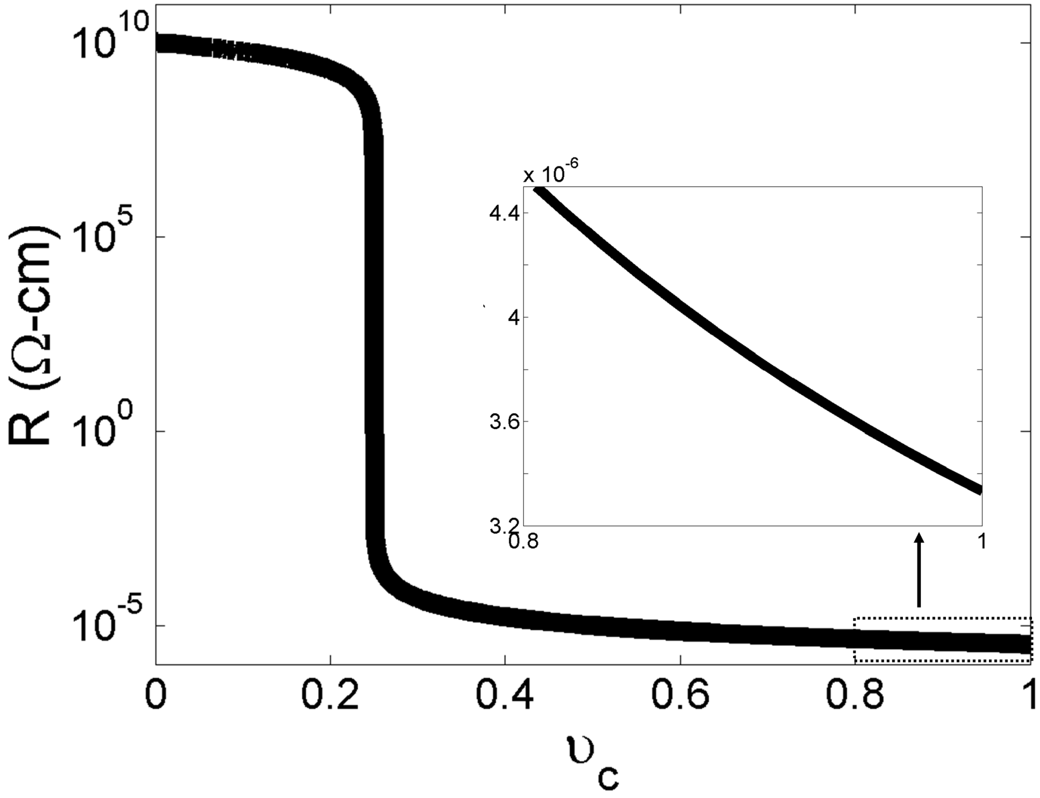


Figure 1.3: A schematic representation of the response mechanism of an insulator – conductor composite chemiresistor vapor sensor. In pure background air, current is passed through the material with some resistance, *R*. When an analyte is added to the background air, the analyte partitions into the sensor material, causing a swelling. This swelling, in turn, causes a decrease in υ_c (eq (6)), and an increase in the dc electrical resistance between the two electrical leads (Figure 1.2). When the analyte is removed from the background air, the analyte partitions out of the sensor, and the sensor returns to its initial state.

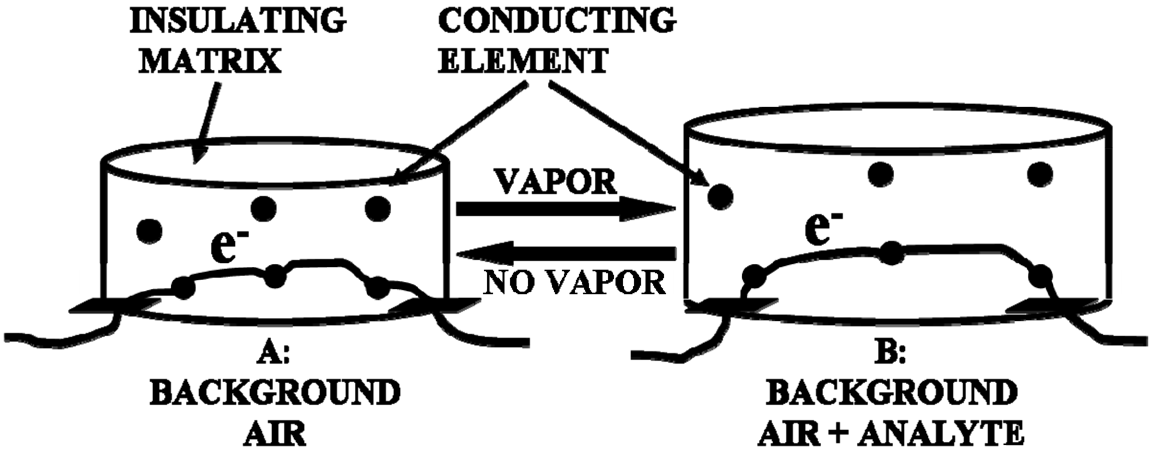


Figure 1.4: Average relative differential equilibrium resistance responses of a poly(butadiene)-carbon black composite vapor sensor (20% by weight carbon black) upon exposure to various analytes, as a function of analyte concentration.[29]

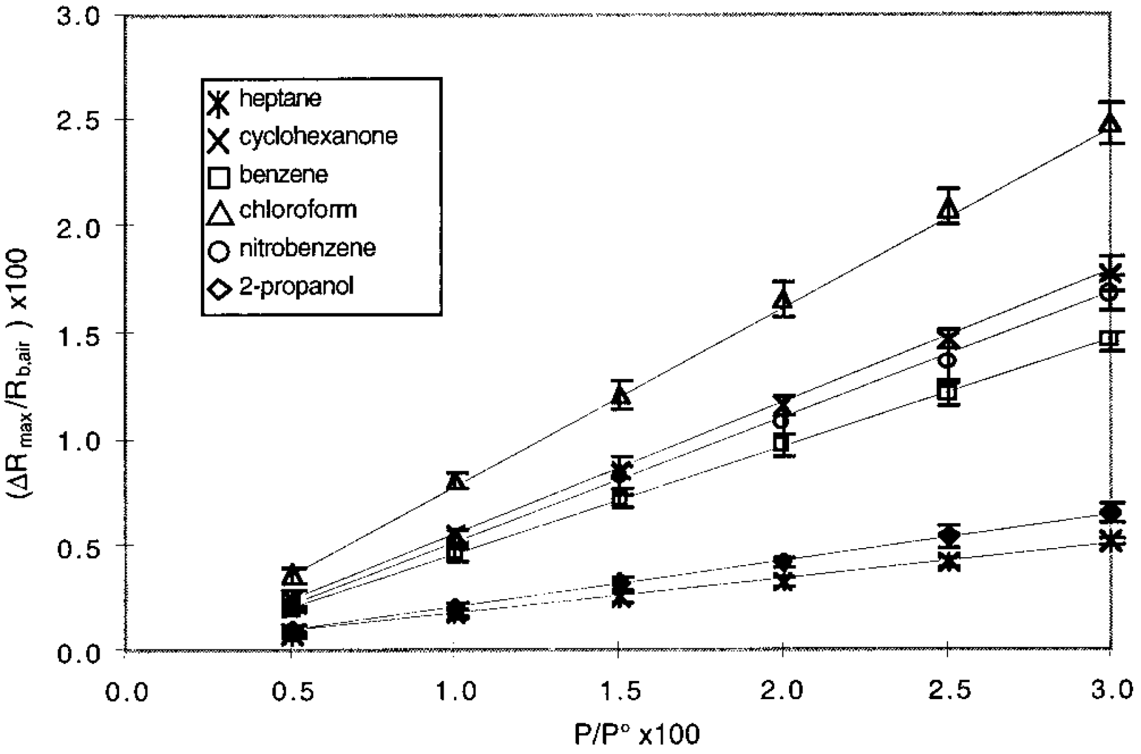
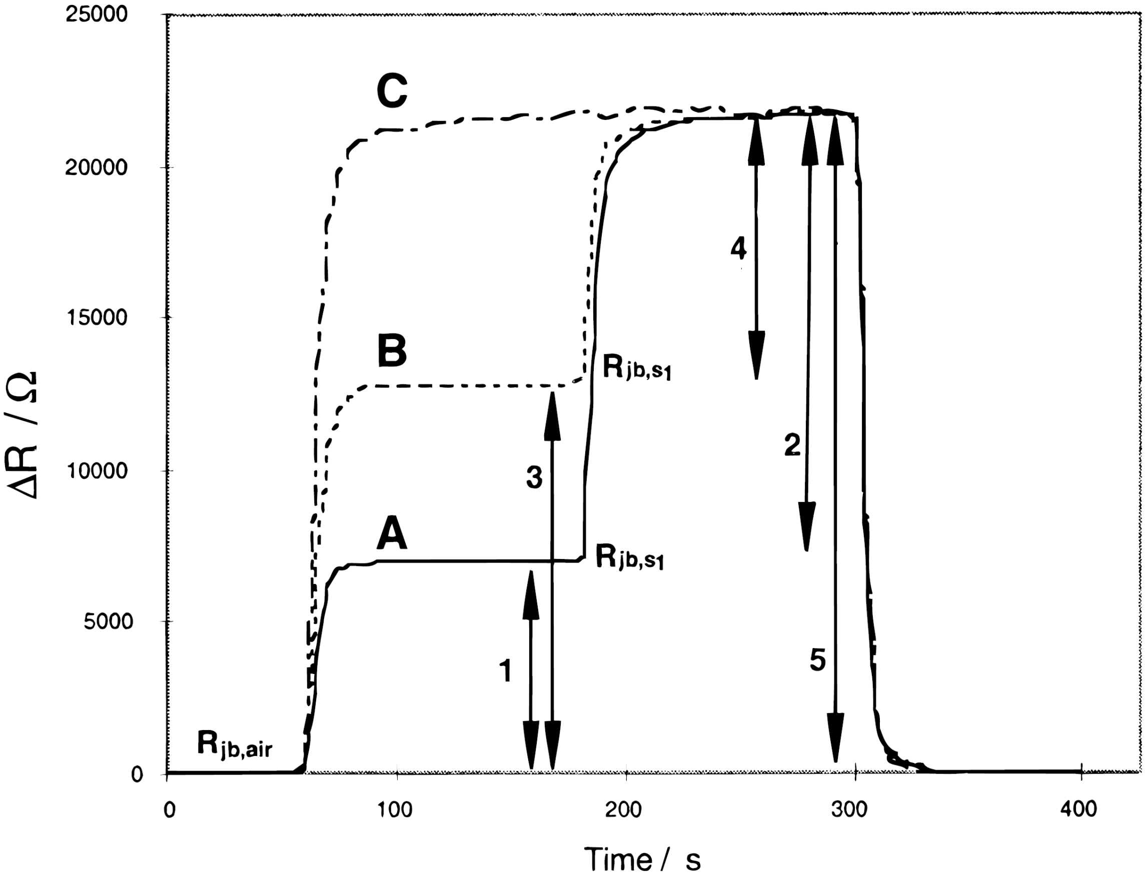


Figure 1.5: Differential resistance response for a poly(ethyelene-co-vinyl acetate)-carbon black composite vapor sensor (20% carbon black). A) Exposure to benzene at *P*/*P*° = 0.020, followed by a simultaneous exposure to benzene at *P*/*P*° =0.020 and chloroform at *P*/*P*° = 0.020. B) Exposure to chloroform at *P*/*P*° = 0.020 followed by a simultaneous exposure to chloroform at *P*/*P*° = 0.020 and benzene at *P*/*P*° = 0.020. C) Simultaneous exposure to benzene at *P*/*P*° = 0.020 and chloroform at *P*/*P*° = 0.020.[29]



Chapter 2

Chemiresistors for Array-Based Vapor Sensing Using Composites of Carbon Black with Low Volatility Organic Molecules[*]

2.1. Abstract

Chemically sensitive resistors have been fabricated from composites of carbon black and low volatility, non-polymeric, organic molecules such as propyl gallate, lauric acid, and dioctyl phthalate. Sorption of organic vapors into the non-conductive phase of such composites produced rapid and reversible changes in the relative differential resistance response of the sensing films. Arrays of these sensors, in which each sensing film was comprised of carbon black and a chemically distinct non-polymeric organic molecule or blend of organic molecules, produced characteristic response patterns upon exposure to a series of different organic test vapors. The use of non-polymeric sorption phases allowed fabrication of sensors having a high density of randomly oriented functional groups and provided excellent discrimination between analytes. By comparison to polymer – carbon black composite vapor sensors and sensor arrays, such sensors provided lower detection limits as well as enhanced clustering and enhanced resolution ability between test analytes.

2.2. Introduction

Array-based vapor sensing has attracted significant interest for its ability to detect and discriminate between various analyte vapors.[1] Surface acoustic wave devices,[2-4] tin oxide sensors,[5-7] conducting organic polymers,[8-10] polymer-coated quartz crystal microbalances,[11-13]

---

[*] This chapter is reproduced according to American Chemical Society copyright guidelines, from "Chemiresistors for Array-Based Vapor Sensing Using Composites of Carbon Black with Low Volatility Organic Molecules" by Ting Gao, Marc D. Woodka, Bruce S. Brunschwig, and Nathan S. Lewis, *Chemistry of Materials*, 18, (22), 5193-5202, 2006. Copyright 2006, American Chemical Society.

polymer-coated micromachined cantilevers,[14] thin film capacitors,[15] dye-impregnated polymers coated onto optical fibers or beads,[16-18] transition metal based dyes,[19,20] and polymer-composite chemically sensitive resistors[21-23] have all been explored in array-based sensing approaches. In this architecture, each sensor is not designed to respond selectively to a single analyte, but instead each analyte produces a distinct fingerprint response pattern from the array of broadly responsive sensors. Pattern recognition algorithms can then be used to obtain information on the identity, properties and concentration of the vapor exposed to the sensor array.[24-27]

One especially attractive signal transduction mode involves the use of chemically sensitive resistors as the sensor array elements.[21-23] Such sensors are inherently low power,[28,29] are compatible with VLSI processing,[7,30] can be deposited onto a variety of substrates including interdigitated electrodes,[31] glass,[32] ceramic,[33] or other insulating materials, and can be fabricated in a wide variety of form factors to optimize signal/noise ratios and produce desired physical sensor and sensor array configurations.[32] Significant attention in our laboratory has been devoted to the investigation of chemiresistive vapor detectors fabricated from composites of carbon black and insulating organic polymers,[21,22,32,34,35] in which the carbon black serves as the electrically conductive phase and the organic polymeric phase absorbs the organic vapor into the sensor.

The sensitivity of sorption-based detectors depends on the interactions between the analyte and the sorption material.[36] Vapor sensors with enhanced sensitivity to analytes having specific functional groups, such as amines or carboxylic acids, can be obtained through fabrication of sorption materials which target functional groups of the analyte of interest.[37,38] Increasing the density of the functional groups in the sorption material could further increase the amount of vapor sorption and therefore produce an additional increase in the sensitivity of such chemically resistive vapor detectors. In this work, we describe the properties of chemiresistive vapor sensors that are comprised of composites of conductive carbon black particles and an insulating organic material, wherein the sorption phase consists of simple, non-polymeric, low vapor pressure organic materials. Such sorption films have a relatively high density of functional groups and thereby could provide very effective sorption of organic analyte vapors. The random arrangement of the organic molecules in the sorption phase should produce a high vapor permeability and therefore lead to rapid sensor response times, and could produce highly reversible responses that show relatively little history effects or hysteresis in response to a wide range of organic analyte vapors. The use of non-polymeric materials opens up a wide range of sorption phases having desirable chemical functionality and physical properties that are in general not readily accessible in polymeric materials.

2.3. Experimental

2.3.1. Materials

The insulating materials used in fabricating the sensor films (Figure 2.1) and the plasticizer dioctyl phthalate, were used as received from either Aldrich Chemical Co. or Acros Organics Co. Reagent grade toluene, n-hexane, tetrahydrofuran (THF), ethanol, ethyl acetate, cyclohexane, n-heptane, n-octane, and isooctane were used as received from Aldrich Chemical Co. Black Pearls 2000 (BP 2000), a furnace carbon black material, was donated by Cabot Co. (Billerica, MA) and was used as received.

2.3.2. Detectors

Detector substrates were fabricated by evaporating 30 nm of chromium and 70 nm of gold onto glass microscope slides using 0.2 cm wide drafting tape as a mask. After evaporation, the mask was removed and the glass slides were cut into 1.0 cm × 2.5 cm pieces.

Sensor films consisted of suspensions of various amounts of carbon black and either pure organic material or mixtures thereof in 20 mL of either toluene or THF. Typically, the desired mass of organic sorption material was dissolved in 20 mL of solvent, and sufficient carbon black was then suspended in this solution to produce the desired mass fraction of organic material and carbon black, by weight of solids (Table 2.1). Prior to fabrication of the sensor films, the casting suspension was sonicated for > 30 min at room temperature. An airbrush (Iwata, Inc.) was used[39] to spray these suspensions across the 0.2 cm gap on the detector substrates until the resistance between the two leads was 10-100 kΩ, as measured by a Keithley model 2002 multimeter. After fabrication, all sensors were placed in a stream of dry air for at least 24 h prior to exposure to the test analytes.

2.3.3. Measurements

The instrumentation and apparatus for resistance measurements and for delivery of analyte vapors has been described previously.[23,34,35] The array of sensors was housed in an aluminum assembly that was connected by Teflon tubing to a computer-controlled, calibrated vapor generation and delivery system. To initiate an experiment, the detectors were placed into a flow chamber and an air flow of 5 L min^{-1} containing 1.10 ± 0.15 parts per thousand (ppth) of water vapor was introduced until the resistance of the detectors stabilized. An individual analyte exposure to the detectors consisted of a three-step process that was initiated with 70 s of airflow to achieve a smooth baseline resistance. Analyte vapor at a controlled concentration in flowing air was then introduced to the detectors for 80 s, followed by 60 s of airflow to insure that the baseline resistance value was restored before the next exposure.

Analytes consisted of five nonpolar hydrocarbons (cyclohexane, n-hexane, n-heptane, n-octane, and isooctane) as well as ethanol and ethyl acetate. In the primary set of data collection for composite sensors having high carbon black loadings, these seven analytes were presented in random order 200 times each to the detector array during a single run over 4 days, at a partial pressure in air such that $P/P^o = 0.0050$, where P is the partial pressure and P^o is the vapor pressure of the analyte at room temperature. In a separate run to evaluate the concentration dependence of the sensor response, concentrations of n-hexane and ethanol were varied at ten different values of P/P^o within the range $0.00020 < P/P^o < 0.00625$, with five exposures to each analyte/concentration combination, in randomized order. Each exposure consisted of 100 s of laboratory air, followed by 100 s of analyte, followed by 100 s of laboratory air, at a flow rate of 5 L min^{-1}.

An identical data run was used to evaluate the performance of the sensors with low carbon black loadings, with the seven analytes presented in random order 200 times each to the detector array during a single run over 4 days. Additionally, subsequent runs which were identical in their randomized analyte exposure order, exposure times and protocols were performed to assess the long term drift and stability of the sensors. The second run was initiated 2 days after the completion of the first run; the third run was initiated 2 days after the completion of the second run, and the fourth run was initiated 6 months after the completion of the third run. In these experiments, analytes were presented to the detector array at concentrations corresponding to $P/P^o = 0.0050$.

2.3.4. Data Processing

The response of a sensor to a particular analyte was expressed as $\Delta R_{max}/R_b$, where R_b is the baseline resistance of the sensor and ΔR_{max} is the steady-state resistance change upon exposing the sensor to analyte (after correcting for baseline drift). The value of ΔR_{max} was obtained from $R_{max}-R_b$, where R_{max} is the maximum resistance value observed during the analyte exposure, calculated by averaging over at least 3 consecutive resistance measurements (in most cases 4 or 5) in the steady-state portion of the response signal. The value of R_b was calculated by averaging over 5 resistance measurements before the exposure initiated. The ratiometric quantity $\Delta R_{max}/R_b$ was used as the response descriptor because it has been shown in similar detector films to be both relatively insensitive to the vapor introduction technique and to increase linearly with analyte concentration.[23,35] All data processing was performed using Matlab (The Mathworks, Natick, MA).

2.3.5. Quantification of Classification Performance

For quantification of the analyte classification performance, the responses from each of the datasets were sum-normalized. This process was performed using eq (1):

$$S'_{ij} = \frac{S_{ij}}{\sum_{j=1}^{n} S_{ij}}, \tag{1}$$

where S_{ij} refers to the $\Delta R_{max}/R_b$ sensor response signal of the j^{th} detector (out of n total detectors) to the i^{th} analyte exposure, and S'_{ij} represents the sum-normalized analog of S_{ij}. For sensors exhibiting a response that is linear with analyte concentration, this normalization procedure produces a unit vector in n -dimensional space defining a location in this space characteristic of each test analyte, regardless of analyte concentration.

The Fisher Linear Discriminant (FLD) algorithm was used on sum-normalized sensor response data to analyze the classification performance of the sensors. In the FLD approach, the responses of a training set were used to calculate a vector which projected response data onto the one-dimensional space that maximized the separation between two sets of data clusters.[40] For normalized data (eq (1)) produced by the responses of an n-detector array, this projection has the form:

$$D_i = \sum_{j=1}^{n-1} c_j S'_{ij} \ , \tag{2}$$

where c_j represents one of the $n - 1$ weighting factors from the hyperplane determined by the FLD algorithm. The value of D_i (hereafter referred to as the D-value) is a single, scalar metric that characterizes the position, along a vector normal to the hyperplane decision boundary, of the detector array data produced by an individual analyte exposure. The chosen hyperplane decision boundary is defined as the point in one-dimensional projected space for which a data point lying on this plane has an equal probability of belonging to either of the two data clusters.

The FLD algorithm maximizes the separation, or clustering, of the two distinct populations of D-values that arise from a single binary separation task. This clustering is measured by the resolution factor (*rf*) characteristic of a separation task, as given in eq (3):[27]

$$rf = \frac{\delta}{(\sigma_1^2 + \sigma_2^2)^{0.5}}, \tag{3}$$

where δ is the difference in the population means of the D-values, and σ_1 and σ_2 are the standard deviations of the two populations of D-values that correspond to the two analytes of the separation task. The FLD algorithm was used to evaluate the separation between each possible pairwise combination of analytes in the data set.

Because a supervised algorithm inherently introduces some bias into the analysis, a train/test scheme was employed. For each pair of analytes that comprised a single separation task, the first 100 exposures to each analyte (exposures 1-100, data set 1) were used to generate a training set and a set of coefficients (comprising a classification model) as described in eq (2). A decision boundary was then developed by defining the hyperplane at which an unknown analyte exposure would have an equal probability of belonging to either analyte population of the given binary separation task. All subsequent data were treated as test data, projected onto the optimized dimension for separation, and analyte identities were classified according to their positions relative to the fixed FLD decision boundary.

The signal to noise ratio (SNR) of a sensor for a given exposure was calculated as:

$$SNR = \frac{\Delta R_{\max}}{\sigma_{baseline}}, \tag{4}$$

where $\sigma_{baseline}$ represents the standard deviation in baseline resistance before analyte delivery, calculated using at least 5 data points.

The same analytes at $P/P^o = 0.0050$ have been previously exposed to carbon black-polymer composite chemiresistors. Such data have been analyzed in the same manner as that for the sensors under study, and is given for comparison.[21,22,32,34,35] Specifically, resolution factors and signal to noise ratios were compared for both types of sensors from previously recorded and reported data. For detection limit determination, carbon black – polymer composite sensors were also exposed simultaneously with carbon black – non polymer composite sensors to ensure equal vapor deliveries and representative analyses.

2.4. Results

2.4.1. Vapor Response Characteristics and Reproducibility

Carbon black loadings of ≈ 10, 25, 50, and 75% by weight were investigated, and 75% loadings exhibited a higher SNR, lower detection limit, and enhanced clustering relative to other loadings. Thus, results on sensor films made from 75% carbon black loadings are primarily reported herein. Additional results are described for a 6-month stability and drift study that was performed on sensors having various, lower carbon black loading levels. In each case, the carbon black loading was sufficient to insure that the chemiresistors were above their percolation threshold, i.e., in the highly conductive state of the composite in which the films displayed simple, ohmic resistance behavior between two electrically conductive contacting leads. Such composites consist of highly interconnected networks of conductive particles in a matrix of insulating organic material, but the structure of the organic material is difficult to elucidate directly from scanning electron microscopy, X-ray photoelectron spectroscopy or other spectroscopic methods due to the high mole fraction of carbon black in the composites.

Table 2.1 presents information on the high (75%) and low (25%) carbon black loaded polymer- and non polymer-based sensor arrays. The first exposure in Figure 2.2 shows the baseline-corrected resistance response of a non polymer- and polymer-carbon black composite sensor on exposure to n-hexane at $P/P^o = 0.0050$. Shown are tetracosanoic acid/dioctyl phthalate (75% carbon black, sensor A4) and poly(ethylene-co-vinyl acetate) (40% carbon black, sensor C2) films, which both exhibited the highest signal to noise for each of their respective sensor array types investigated. The resistance of the films increased when analyte vapor was present but rapidly (i.e., within seconds) returned to its original baseline resistance value after the vapor exposure had been discontinued. Non polymer-carbon black composite sensors consistently displayed signal to noise ratios and response magnitudes comparable to those obtained with the well-studied polymer-carbon black composite sensors evaluated in this work.

Figure 2.2 also displays the sensor response repeatability, showing six sensor responses, with 1, 35, 44, 62, and 71 hr, as well as random continuous exposure cycles to the test analytes, occurring between the second, third, fourth, fifth, and sixth displayed sensor response and the first displayed sensor response, respectively. As observed in Figure 2.2, in all cases, the sensor fully returned to the same response on exposure to n-hexane at $P/P^o = 0.0050$, as well as returned to the same baseline resistance on exposure to laboratory air. This was the case for the majority of exposures (> 95%), however hysteresis did occur randomly in a small percentage of exposures. Therefore, sensor responses were baseline corrected, forcing sensor readings to fully return to their initial baseline resistance; this ensured that $\Delta R_{max}/R_b$ was due solely to the sensor/analyte interaction and not due to sensor drift.

Table 2.2 presents the sensitivities and standard deviations of the responses measured for the different carbon black composite sensors exposed to the 7 test analytes studied in this work at $P/P^o = 0.0050$ in air. Sensitivities varied significantly across the analytes tested, and a given analyte produced different responses on different sensor films.

Different levels of variability were observed in the response of each of the sensors. Part of this variability in the response amplitude can be ascribed to sensor noise, which is inherent and unique to each of the sensors, as well as to variation in room temperature during the exposures. For example, a 1 ^{o}C change in room temperature produces a 4.5% change in the vapor pressure of n-hexane (the vapor pressures of n-hexane at 20 and 21 ^{o}C are 119.9 and 125.3 Torr, respectively).[41] Additionally, slight (though significant) drift was observed for several of the sensors, though this did not affect the ability to accurately model and predict based on sensor array response patterns.

Signal to noise ratios were calculated for each sensor on exposure to each of the test analytes. Table 2.3 details the means and standard deviations of the SNRs for each carbon black – non polymer composite sensor on exposure to the various test analytes each presented 200 times in random order at $P/P^o = 0.0050$ (sensors A1-A7). For comparison, Table 2.3 also presents the SNRs of the carbon black-polymer composite sensors on exposure to these analytes at the same partial pressure of $P/P^o = 0.0050$ (sensors C1-C9). The two sensor types exhibited similar SNR values, with different sensors performing better in different cases.

2.4.2. Concentration Dependence of Sensor Response

Figures 2.3a and 2.3b display the responses of several typical carbon black – non polymer composites as a function of the vapor phase concentration of n-hexane and ethanol, respectively. For the relatively low analyte concentrations used in this study, the sensor responses were well-described by a linear dependence on P/P^o, indicating operation above the percolation threshold. This relationship has been observed for carbon black-polymer composite sensors operating above the percolation threshold.[35]

Table 2.4a presents the limits of detection based on the $\Delta R_{max}/R_b$ vs. concentration data presented in Figure 2.3. Signal to noise ratios were calculated (eq (4)) for each of the sensors on exposure to hexane and ethanol at various fractions of their vapor pressure ($0.00020 < P/P^o < 0.00625$), and detection was taken to be the partial pressure at which SNR = 3. Limits of detection ranged from $P/P^o = 0.0002$ to $P/P^o = 0.00075$, with most values near 0.00035 or 0.0005. These thresholds were converted to parts per million for display. For comparison, Table 2.4b gives detection limits for several carbon black – polymer composites, exposed simultaneously with optimized carbon black – non polymer composite sensors to ensure a representative comparison. The limits of detection for the carbon black – polymer composite sensors were in accord with values reported previously.[42] The carbon black – non polymer composite sensors exhibited approximately comparable detection limits when compared to these well-studied and developed carbon black – polymer composite sensors.

2.4.3. Sensor Specificity

Figure 2.4 presents the mean responses, averaged over 200 randomly ordered exposures to each analyte, for each of the carbon black – non polymer composite films to the seven test analyte vapors at $P/P^o = 0.0050$. Large differences in sensitivity were observed between the responses of a given sensor upon exposure to the various test analytes. For example, quinacrine dihydrochloride dihydrate (sensor A2) displayed a strong positive response on exposure to a prototypical polar analyte, ethanol, while displaying a strong negative response to a prototypical nonpolar analyte, n-hexane. This can be attributed to insolubility of the latter compound with nonpolar solvents resulting from dielectric constant differences and molecular size. Additionally, a tetracosanoic acid/dioctyl phthalate – carbon black composite (sensor A7) exhibited an n-hexane/ethanol response ratio of 22, while a quinacrine dihydrochloride dihydrate/dioctyl phthalate – carbon black composite (sensor A6) displayed an n-hexane/ethanol response ratio of 0.3. For comparison, of the polymer – carbon black composite sensors investigated, the greatest response ratio of ethanol to n-hexane was produced by poly(ethylene-co-vinyl acetate) (sensor C2), with a ratio of 4, and the smallest ratio was achieved by poly(vinyl butyral) (sensor C8), with a ratio of 0.4 (Table 2.2). Clearly, the use of organic molecular sorption phases having a high density of hydrophilic or hydrophobic functional groups can produce sensor arrays that display large discrimination power between differing test pairs of analytes.

2.4.4. Sensor Array Response to Various Analytes

Principal components analysis[27] was used to visualize the differences in normalized autoscaled response patterns of a 7 element carbon black composite sensor array (Table 2.1, sensors A1-A7) exposed randomly 200 times to each of the seven test analytes at P/P^o = 0.0050. The points plotted in Figure 2.5 represent unique response patterns of the sensor array to each of the analytes presented. The response vectors are displayed with respect to the first three principal components of the data set, which contained 99% of the variance in detector response. Several major clusters are observed: ethanol, ethyl acetate, and c-hexane, as well as a clustering of the remaining alkanes. This remaining cluster of alkanes also displays a distinct pattern, which is shown inset in Figure 2.5. Even at the relatively low analyte concentrations used in this study, the sensor array readily distinguished extremely well between chemically similar analytes.

The classification performance of the sensor array was quantified by use of the Fisher Linear Discriminant algorithm for pairwise analyte classification. The figure of merit to determine the effectiveness of the FLD model is the resolution factor, *rf* (eq (3)), which quantifies the statistical separation between the two data clusters of interest. The first 100 normalized exposures to each analyte were used as a training set and the remaining 100 normalized exposures to each analyte, from the same set of data collection, was used as a test set. This train/test scheme was adopted to avoid bias resulting from possible overfitting of data.

Table 2.5a presents resolution factors for the carbon black – non polymer composite sensor array (sensors A1-A7). For comparison, Table 2.5b presents resolution factors for an array of carbon black – polymer composite sensors consisting of 9 sensor types (sensors C1-C9). This 9-sensor carbon black – polymer composite array was chosen from a non-exhaustive search seeking the best 9-sensor array that maximized the lowest resolution factors reported (> 15 9-sensor array combinations were investigated, and the "best" sensors based on experience, polycaprolactone, poly(ethylene-co-vinyl acetate), and poly(ethylene oxide), were always included). In terms of the ability to resolve between various analytes, the non-polymeric composite sensor array performed highly favorably relative to the well-developed and well-studied polymer-based sensor array, with significant increases in resolution in many previously difficult classification tasks. For example, in classifying n-hexane from c-hexane, n-heptane, n-octane, or i-octane, resolution factors of 2.5, 1.2, 1.7, and 3.5, respectively, were observed for the polymer composite-based sensor array. The use of a carbon black-non polymer composite sensor array increased these resolution factors to 6.1, 6.4, 9.9, and 6.2, respectively. A resolution factor of 1 implies 72% correct classification, 2 implies 92% correct classification, and 3 implies 98% correct classification. This new sensor type thus takes previous classification tasks, which performed at levels slightly above chance, and provided the ability to consistently and confidently correctly classify analytes.

2.4.5. Stability and Drift

A FLD model for each binary separation task, consisting of projection weights and a decision boundary, was constructed from sensor responses in the first data set of the first 100 exposures to each analyte. This model was then applied to 700 subsequent exposures spread over 4 sets that spanned six months of data collection. The exposures for each binary classification task were then projected onto the FLD vector characteristic for the given classification task, placing data into the one-dimensional space which initially maximized the resolution factor between the two analytes of interest. These analyte projections were compared to the originally modeled decision boundary for the given binary separation, and thereby assigned to be in one of the two analyte clusters. The classification rate was defined as the number of correct classifications divided by the number of classification attempts. Table 2.6 lists the performance factors for all combinations of binary separations for each set of data collection.

Binary classification rates were comparable throughout the first 3 data sets, which spanned one month. However, the fourth data set, collected 6 months after the initially trained model, yielded extremely low classification rates in many situations. In terms of the Fisher model, two explanations of this performance loss exist: 1) a new dimension for each binary analyte separation captures maximum resolution between analyte clusters, so that a new model, with different projection weights for each analyte, and a new decision boundary, needs to be created; or 2) the same model approximately captures maximum resolution between analyte clusters, but the clusters have drifted with respect to the original decision boundary. In the latter case, a calibration scheme has proven capable of restoring the classification performance of carbon black-polymeric composite sensors.[43] To reduce this type of drift, sensor responses were adjusted by a multiplicative calibration factor:

$$S_{a,t} = S_{c,t}\frac{S_{a,0}}{S_{c,0}}, \tag{5}$$

where $S_{\mathrm{a,t}}$ and $S_{\mathrm{c,t}}$ indicate the $\Delta R/R_b$ response signals for an analyte *a* and calibrant *c*, respectively, at some time *t* after training, and $S_{\mathrm{a,0}}$ and $S_{\mathrm{c,0}}$ are the initial responses to analyte *a* and calibrant *c*.[43]

Table 2.7 presents the classification rates for each binary separation, using each analyte as a calibrant, when the initial model (based on exposures 1-100, data set 1) was used on the final data set (200 exposures, recorded 6 months after the initial data set). The first three exposures from the final data set were used to calibrate the model according to eq (5), and were then followed by 47 test exposures. This cycle of calibrate/test was repeated 3 additional times, accounting for all 200 exposures of the final data set. Cases where reasonable performances were attained are shown in bold text. Of the 21 combinations of binary analyte classification tasks, 17 yielded classification rates of ≥ 0.90.

For binary classifications with low classification rates, the sensor array was still capable of resolving between analyte pairs in the dataset; however, a rigorous training period was again required to construct a new model for effective analyte separation. For example, the binary classification of n-hexane and n-heptane yielded a performance of 0.51 and had a resolution factor of 0.02 when the initial model was applied to the final data set. However, if the first 100 exposures of data set 4 were used to construct a new model, a resolution factor of 1.5 and a classification rate of 0.88 was achieved for the final 100 exposures of data set 4. These values were comparable to those obtained from training on the first 100 exposures and testing on the final 100 exposures of data set 1, with a classification rate of 0.82 (Table 2.6a). Thus, no sensor performance was lost, but the initial model describing the sensor response behavior changed significantly, resulting in the loss of predictive ability.

Figure 2.6a shows projections of 700 exposures, spread over 4 sets of data collection, for a FLD model constructed from the first 100 exposures in data set 1. Figure 2.6b shows these same projections, when a calibration scheme was adopted in which 3 exposures were first used as calibrant runs, followed by 47 test exposures, with the process repeated throughout the remaining 700 exposures of the data set. The projected dimension clearly maintained a reasonable level of separation between the two analytes (although this was no longer the optimal one dimensional space for resolution), however the analyte clusters drifted relative to the decision boundary. The calibration process shifted these projections back to the decision boundary, and classification performance was restored.

2.5. Discussion

The vapor sensing properties of the carbon black – non polymeric composite sensors and sensor arrays compared favorably in all aspects to the well-investigated carbon black – polymer composite sensing films. The non-polymer sensors provided improved analyte clustering and analyte resolution/classification capability, as well as a high level of signal to noise and low detection limit thresholds.

A measure of the performance of a sensor array is the resolution factor, which is a measure of the ability of a given sensor array to distinguish between and discriminate among various analytes. In this respect, the carbon black – non polymer composite sensors surpassed the performance of previous sensor classes, including our well-studied carbon black – polymer composite sensors (Table 2.5a,b). Significant improvements were observed, in particular, in the ability of the sensor array to distinguish between chemically similar alkanes, namely n-hexane, cyclohexane, n-heptane, n-octane, and isooctane.

The non-polymer sensors are well-suited to detect and exploit subtle differences between analytes, owing to a higher density and random arrangement of functional groups, as well as an enhanced signal to noise ratio for analyte detection. In typical carbon black – polymer composite sensors, functional groups are present at certain repeat units along the polymer backbone, and this structural motif places a limit on the functional group density as well as a limit on possible analyte-polymer interactions, due to steric hindrance. With the carbon black – non polymer composite sensor array, a higher functional group density, as well as random packing, can provide more specific sensor-analyte interactions which are able to better capture subtle differences in analyte properties. High signal to noise ratios provide the means of detecting and describing these subtle differences, which would likely be lost in the noise of other sensor types. These combinations allow carbon black – non polymer composite sensors to more precisely define the position of extremely similar analytes in sensor response space, which translates into enhanced clustering and resolution ability.

The carbon black – non polymer composite sensors also exhibited lower detection limits relative to typical carbon black – polymer composite sensors (Table 2.4a-b). Thus, carbon black – non polymer composite sensors are more suitable for trace vapor detection, which broadens the potential areas of application of these sensors.

The low mass fraction carbon black – non polymer sensor array showed relatively little long-term drift over extended time periods. Specifically, for most binary separation tasks, the non-polymeric composite sensors provided good analyte classification levels for at least 6 months after an initial training phase. When the sensors were used 6 months after an initial training period, 11 of the 21 binary separation tasks were performed with correct classification rates of > 90% (Tables 2.6-7). When a simple calibration scheme, which involved only 3 calibration exposures per 50 exposures, was performed, the number of binary separation tasks with > 90% correct classification after six months increased to 17. The cases where performance was unacceptable even after calibration were the same as those reported for carbon black-polymer composite sensors, for example n-hexane vs. n-heptane or n-heptane vs. n-octane.[43]

Plasticizers such as dioctyl phthalate (a viscous liquid) have been added to polymers to lower their glass transition temperature and decrease the sensor response time to various vapors. The sensors studied herein showed response times that were rapid, both with and without the presence of dioctyl phthalate or similar plasticizers (Figure 2.2). This rapid time response is characteristic of the use of low molecular weight non-polymeric organic molecules as the sorbent phase.

For many diseases, specific volatile organic compounds such as amines and fatty acids are found in the breath and urine of infected individuals. For bio-sensing applications, it is desirable to have sensors with a high sensitivity to these species. A key feature of using molecularly based sorbent phases is the ability to tune the sensitivity towards different classes of chemicals. The ratio of the $\Delta R_{max}/R_b$ responses of two carbon black – non polymer composite sensors, tetracosanoic acid/dioctyl phthalate and quinacrine dihydrochloride dihydrate/dioctyl phthalate, on exposure to n-hexane and to ethanol, was 22 and 0.3, respectively. Additionally, the sensor consisting of pure quinacrine dihydrochloride dihydrate exhibited a strong positive response on exposure to polar analytes, and a strong negative response on exposure to nonpolar analytes. Such large differences for various other analytes could likely be produced by further development of this class of sensors.

2.6. Conclusions

Composites made from homogeneous or blended organic molecules and carbon black showed fast response times, good reversibility, high stability, and an excellent ability to discriminate and classify between both similar and dissimilar types of analytes. This type of composite sensor offers a higher density of functional groups, as well as a random orientation and random exposure of these functional groups within the sensing material due to the lack of a restricting polymer backbone. A 7-sensor array robustly resolved even extremely similar test analytes, such as n-hexane and n-heptane. Excellent signal-to-noise ratios were achieved with these carbon black – non polymer composite sensors, which provided lower limits of detection relative to the evaluated carbon black – polymer composite sensors.

2.7. References

(1) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R. *Chem. Rev.* **2000**, *100*, 2595.
(2) Ballantine, D. S.; Rose, S. L.; Grate, J. W.; Wohltjen, H. *Anal. Chem.* **1986**, *58*, 3058.
(3) Rose-Pehrsson, S.; Grate, J.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801.
(4) Patrash, S. J.; Zellers, E. T. *Anal. Chem.* **1993**, *65*, 2055.
(5) Srivastava, R.; Dwivedi, R.; Srivastava, S. K. *Sens. Actuators, B* **1998**, *50*, 175.
(6) Getino, J.; Horrillo, M. C.; Gutierrez, J.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Sens. Actuators, B* **1997**, *B43*, 200.
(7) Bednarczyk, D.; DeWeerth, S. P. *Sens. Actuators, B* **1995**, *B27*, 271.
(8) Huang, J.; Virji, S.; Weiller, B. H.; Kaner, R. B. *Chem.-A Euro. J.* **2004**, *10*, 1315.
(9) Partridge, A. C.; Jansen, M. L.; Arnold, W. M. *Mater. Sci.* **2000**, *12*, 37.
(10) Bartlett, P. N.; Archer, P. B. M.; Lingchung, S. K. *Sens. Actuators* **1989**, *19*, 125.
(11) Fu, Y.; Finklea, H. O. *Anal. Chem.*, **2003**; *75*; 5387.
(12) Haupt, K.; Noworyta, K.; Kutner, W. *Anal. Comm.* **1999**, *36*, 391.
(13) Mirmohseni, A.; Oladegaragoze, A. *Sens. Actuators, B* **2004**, *102*, 261.
(14) Lang, H. P.; Baller, M. K.; Berger, R.; Gerber, C.; Gimzewski, J. K.; Battiston, F. M.; Fornaro, P.; Ramseyer, J. P.; Meyer, E.; Guntherodt, H. J. *Anal. Chim. Acta* **1999**, *393*, 59.
(15) Willing, B.; Kohli, M.; Muralt, P.; Oehler, O. *Infrared Phys.* **1998**, *39*, 443.
(16) Albert, K. J.; Walt, D. R.; Gill, D. S.; Pearce, T. C. *Anal. Chem.* **2001**, 2501.
(17) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*, 697.
(18) Goodey, A.; Lavigne, J. J.; Savoy, S. M.; Rodriguez, M. D.; Curey, T.; Tsao, A.; Simmons, G.; Wright, J.; Yoo, S. J.; Sohn, Y.; Anslyn, E. V.; Shear, J. B.; Neikirk, D. P.; McDevitt, J. T. *J. Am. Chem. Soc.* **2001**, 2559.
(19) Daws, C. A.; Exstrom, C. L.; Sowa, J. R.; Mann, K. R. *Chem. Mater.* **1997**, 363.
(20) Rakow, N. A.; Suslick, K. S. *Nature* **2000**, *406*, 710.
(21) Burl, M. C.; Sisk, B. C.; Vaid, T. P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *87*, 130.
(22) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 2652.
(23) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(24) Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta* **1986**, *185*, 1.
(25) Burns, J. A.; Whitesides, G. M. *Chem. Rev.* **1993**, *93*, 2583.
(26) Kowalski, B. R.; Bender, C. F. *Anal. Chem.* **1972**, *44*, 1405.
(27) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.
(28) Korotchenkov, G. S.; Dmitriev, S. V.; Brynzari, V. I. *Sens. Actuators, B* **1999**, *54*, 202.
(29) Han, K. R.; Kim, C. S.; Kang, K. T.; Koo, H. J.; Il Kang, D.; He, J. W. *Sens. Actuators, B* **2002**, *81*, 182.
(30) Wilson, D. M.; Deweerth, S. P. *Sens. Mat.* **1998**, *10*, 169.
(31) Shurmer, H. V.; Corcoran, P.; Gardner, J. W. *Sens. Actuators, B* **1991**, *4*, 29.
(32) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *82*, 54.
(33) Shurmer, H. V.; Gardner, J. W.; Corcoran, P. *Sens. Actuators, B* **1990**, *1*, 256.
(34) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177.
(35) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658.
(36) Popovska-Pavlovska, F.; Raka, L. *J. Poly. Sci. B, Poly. Phys.* **2004**, *42*, 267.

(37) Tillman, E. S.; Koscho, M. E.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **2003**, *75*, 1748.
(38) Sotzing, G. A.; Phend, J. N.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **2000**, *12*, 593.
(39) Koscho, M. E.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem* **2002**, *74*, 1307.
(40) Fisher, R. A. *Ann. Eugenic* **1936**, 179.
(41) Weast, R. C. *CRC Handbook of Chemistry and Physics, 70th Ed.*; CRC Press, Inc.: Boca Raton, Florida, 1989/90.
(42) Doleman, B. J.; Lewis, N. S. *Sens. Actuators, B* **2001**, 41.
(43) Sisk, B. C.; Lewis, N. S. *Sens. Actuators, B* **2005**, *104*, 249.

Table 2.1: Sorption material used in carbon black-non-polymeric composite sensors for (A1-A7) 75% and (B1-B9) 25%, by mass, CB loadings. 20 ml of either THF or toluene was added to sorption and plasticizer materials, followed by addition of CB, followed by sonication for > 30 min. (C1-C9) Sorption material used in CB – polymer composite sensors, as reported previously.[43] Where noted, the plasticizers dioctyl phthalate (DP) and di(ethylene glycol) dibenzoate (DEGB) were used.

		amount (mg)		
sensor	sorption material	sorption	plasticizer	CB
A1	propyl gallate	50	0	150
A2	quinacrine dihydrochloride dihydrate	50	0	150
A3	lauric acid / DP	35	15	150
A4	tetracosane / DP	35	15	150
A5	tetracosanoic acid	50	0	150
A6	quinacrine dihydrochloride dihydrate / DP	35	15	150
A7	tetracosanoic acid / DP	35	15	150
B1	tetraoctylammonium bromide / DP	80	80	20
B2	lauric acid / DP	80	70	20
B3	tetracosanoic acid	80	0	30
B4	tetracosanoic acid / DP	80	50	20
B5	tetracosanoic acid / DP	100	60	40
B6	propyl gallate	160	0	40
B7	1,2,5,6,9,10-hexabromocyclododecane / DP	100	60	40
B8	quinacrine dihydrochloride dihydrate	160	0	40
B9	quinacrine dihydrochloride dihydrate / DP	100	60	40
C1	polycaprolactone / DEBG	80	80	40
C2	poly(ethylene-co-vinyl acetate) / DEBG	80	80	40
C3	poly(ethylene oxide) / DEBG	80	80	40
C4	poly(ethylene glycol) / DEBG	80	80	40
C5	poly(methyl vinyl ether-co-maleic anhydride) / DEBG	80	80	40
C6	poly(4-vinyl phenol) / DEBG	80	80	40
C7	polycarbonate / DEBG	80	80	40
C8	poly(vinyl butyral) / DEBG	80	80	40
C9	polystyrene (PVS) / DEBG	80	80	40

Table 2.2: Sensor response, $\Delta R/R_b$ (×10,000), of carbon black – non polymer composite and carbon black – polymer composite sensors (Table 2.1) to seven test analytes presented at a concentration of $P/P^o = 0.0050$. The sensors were subjected to 200 randomly ordered exposures to each analyte; means and standard deviations are given for each sensor (mean ± standard deviation).

sensor	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
A1	4.2±0.9	6.4±0.8	8.0±0.7	1.7±0.3	4.4±1.0	6.3±1.0	3.1±1.3
A2	-12.5±7.2	15.0±2.9	2.6±-5.1	-5.1±2.9	-14.7±9.3	-16.6±9.4	-21.3±10.8
A3	21.7±2.9	1.3±0.2	8.3±0.8	8.9±0.7	25.1±4.5	38.0±7.4	27.4±4.1
A4	11.9±2.1	1.0±0.2	4.5±0.7	5.6±1.6	13.2±3.8	19.9±6.1	15.9±4.0
A5	18.0±0.9	0.9±0.3	6.1±0.5	5.2±0.5	23.2±0.9	36.5±1.4	24.2±0.8
A6	2.4±0.3	6.9±1.0	2.4±0.4	1.6±0.2	2.2±0.4	2.7±0.5	2.7±0.4
A7	18.7±1.0	0.8±0.2	5.8±0.3	6.8±1.0	23.1±1.5	35.6±2.6	24.8±1.2
C1	3.2±0.2	6.1±0.2	13.3±0.3	4.9±0.2	2.8±0.2	2.8±0.1	3.3±0.1
C2	18.2±0.5	12.0±0.5	48.4±1.3	27.9±0.6	16.6±0.5	17.9±0.4	19.7±0.5
C3	4.2±0.4	3.2±0.4	5.8±0.2	5.6±0.2	4.2±0.2	4.7±0.2	5.2±0.2
C4	2.1±0.2	2.7±0.2	12.3±0.4	3.4±.2	1.7±0.2	1.6±0.2	1.8±0.2
C5	20.2±0.6	6.6±0.3	37.8±1.1	31.0±1.0	18.5±0.6	20.1±0.5	22.4±0.5
C6	18.7±0.6	11.9±0.5	49.6±1.9	28.7±1.1	16.9±0.6	18.2±0.5	20.0±0.5
C7	14.7±0.5	7.5±0.3	55.3±2.0	23.8±0.9	12.9±0.5	13.5±0.4	14.5±0.3
C8	0.6±0.1	1.5±0.1	5.7±0.2	0.5±0.1	0.5±0.1	0.4±0.1	0.2±0.1
C9	6.8±0.5	12.2±0.5	34.6±0.9	6.1±0.4	5.7±0.4	5.2±0.3	2.3±0.3

Table 2.3: Signal to noise ratios (SNR) of carbon black – non polymer composite and carbon black – polymer composite sensors (Table 2.1) to seven test analytes presented at a concentration of P/P^o = 0.0050. The sensors were subjected to 200 randomly ordered exposures to each analyte; means and standard deviations are given for each sensor (mean ± standard deviation).

	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
A1	90±62	142±89	99±49	45±31	73±41	65±31	38±32
A2	-136±109	109±65	25±22	-52±45	-151±145	-97±79	-230±172
A3	152±62	16±7	81±33	86±33	150±41	164±36	155±46
A4	100±49	13±6	46±19	54±30	97±51	131±44	101±40
A5	55±19	5±4	25±14	22±15	64±23	73±18	75±36
A6	25±10	68±34	24±11	18±8	23±10	27±10	29±13
A7	99±26	14±8	61±27	82±38	98±23	90±17	112±24
C1	102±40	102±40	505±190	215±81	134±54	138±46	143±58
C2	465±211	211±102	763±187	809±276	586±220	636±240	746±313
C3	32±12	30±10	107±45	61±22	39±14	43±16	56±23
C4	29±12	62±23	190±87	60±22	32±11	35±14	42±20
C5	104±45	53±21	193±76	182±75	133±51	146±56	198±84
C6	46±21	311±124	585±278	68±32	54±20	46±18	38±15
C7	238±77	146±57	1355±654	526±217	295±181	304±111	320±123
C8	30±12	87±38	206±80	24±8	34±15	33±13	15±8
C9	65±30	54±23	326±111	49±13	77±32	70±22	29±11

Table 2.4: Approximate limits of detection of (a) carbon black – non polymer composite sensors and (b) carbon black – polymer composite sensors (Table 2.1) for the detection of n-hexane and ethanol. The limit of detection is defined as the vapor concentration at which the SNR = 3.

a)	limit of detection (ppm)						
	A1	A2	A3	A4	A5	A6	A7
n-hexane	110	100	100	100	100	60	140
ethanol	50	50	50	40	40	40	40

b)	limit of detection (ppm)			
	C1	C2	C3	C8
n-hexane	160	140	140	120
ethanol	80	50	50	70

Table 2.5: Resolution factors displaying the ability of a) the carbon black – non polymer composite sensor array (sensors A1-A7), and b) the carbon black – polymer composite sensor array (sensors C1-C9) to distinguish between test analytes presented at $P/P^o = 0.0050$. In each case, for a given separation task, a Fisher linear discriminant model was trained on exposures 1-100, and exposures 101-200 were then tested using the model. Reported values are for testing exposures 101-200.

a)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	44.6	13.3	6.1	6.4	9.9	6.2
ethanol		N/A	27	36.5	47.5	51.7	50
ethyl acetate			N/A	14.3	15.4	20.6	14.5
c-hexane				N/A	8.2	10.1	6.9
n-heptane					N/A	4.2	3.7
n-octane						N/A	4.8
i-octane							N/A

b)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	10.7	6.1	2.5	1.2	1.7	3
ethanol		N/A	24.2	29.1	23.3	25.2	25.9
ethyl acetate			N/A	30.4	15.5	27.1	32.1
c-hexane				N/A	3.9	4.4	10.2
n-heptane					N/A	1.7	6.8
n-octane						N/A	6.7
i-octane							N/A

Table 2.6: Binary classification rates of an array comprised of sensors B1-B9 (Table 2.1) using FLD with a statistical decision boundary for classification, with binary models trained on the first 100 exposures to data set 1, and applied to data testing exposures from data set one and data sets collected at later times. Classification rates shown are for testing exposures collected from a) data set 1, b) data set 2, b) data set 3, and d) data set 4, where data sets 2, 3 and 4 were collected two days, six days, and six months, respectively, after the initial training data was collected for data set 1.

a)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	1	1	1	0.82	0.95	1
ethanol		N/A	1	1	1	1	1
ethyl acetate			N/A	1	1	1	1
c-hexane				N/A	1	1	0.92
n-heptane					N/A	0.84	1
n-octane						N/A	1
i-octane							N/A

b)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	1	1	1	0.73	0.79	1
ethanol		N/A	1	1	1	1	1
ethyl acetate			N/A	1	1	1	1
c-hexane				N/A	1	1	0.56
n-heptane					N/A	0.59	1
n-octane						N/A	1
i-octane							N/A

Table 2.6 (continued):

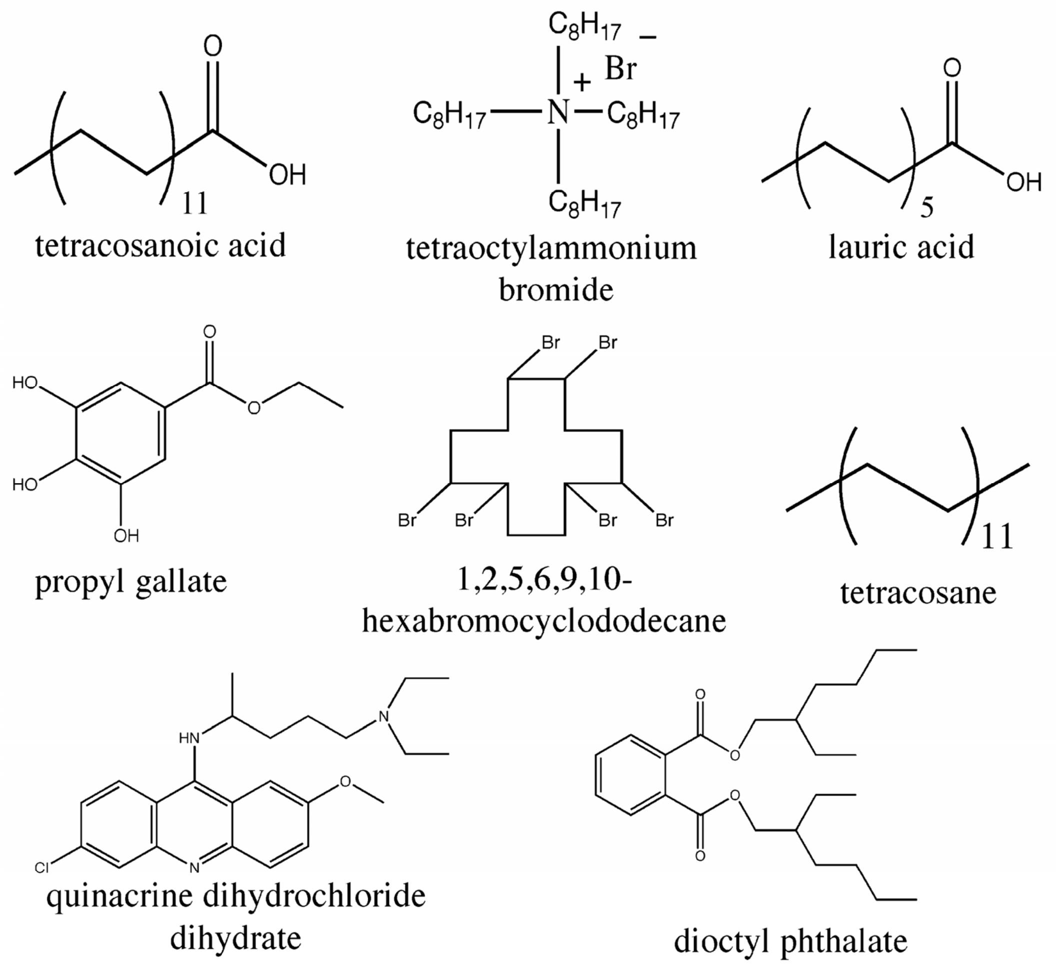
c)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	1	1	0.99	0.66	0.79	1
ethanol		N/A	1	1	1	1	1
ethyl acetate			N/A	1	1	0.99	1
c-hexane				N/A	0.99	0.99	0.54
n-heptane					N/A	0.64	1
n-octane						N/A	1
i-octane							N/A

d)	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane	N/A	0.94	0.98	0.51	0.51	0.5	0.59
ethanol		N/A	1	0.88	0.95	0.91	0.98
ethyl acetate			N/A	0.99	0.98	0.99	0.9
c-hexane				N/A	0.52	0.51	0.5
n-heptane					N/A	0.5	0.59
n-octane						N/A	0.62
i-octane							N/A

Table 2.7: Classification rates of an array of sensors B1-B9 when a FLD model was trained on 100 exposures from data set 1, and tested on 200 exposures from data set 4, six months later, with the use of various calibrations. Binary classification tasks capable of high performances (> 90% correct classification) with a 6 month period between the train and test phase are shown in bold.

	calibrant						
task	n-hexane	ethanol	ethyl acetate	c-hexane	n-heptane	n-octane	i-octane
n-hexane / ethanol	0.58	0.98	1	0.82	0.86	0.96	0.95
n-hexane / ethyl acetate	0.57	0.96	0.98	0.7	0.85	0.73	0.84
n-hexane / c-hexane	0.86	0.52	0.51	0.83	0.88	0.9	0.74
n-hexane / n-heptane	0.5	0.56	0.55	0.5	0.53	0.5	0.49
n-hexane / n-octane	0.49	0.57	0.56	0.51	0.53	0.55	0.51
n-hexane / i-octane	0.91	0.59	0.6	0.88	0.95	0.97	0.86
ethanol / ethyl acetate	0.51	1	1	0.75	0.84	0.86	0.76
ethanol / c-hexane	0.58	0.95	0.99	0.83	0.85	0.98	0.95
ethanol / n-heptane	0.59	0.9	0.99	0.83	0.86	0.98	0.97
ethanol / n-octane	0.57	0.89	0.99	0.84	0.85	0.97	0.96
ethanol / i-octane	0.57	0.99	1	0.85	0.86	0.99	0.98
ethyl acetate / c-hexane	0.57	0.86	0.98	0.73	0.84	0.73	0.83
ethyl acetate / n-heptane	0.58	0.76	0.97	0.71	0.85	0.74	0.85
ethyl acetate / n-octane	0.57	0.97	0.99	0.72	0.85	0.74	0.85
ethyl acetate / i-octane	0.53	0.53	0.89	0.72	0.81	0.72	0.82
c-hexane / n-heptane	0.86	0.7	0.68	0.82	0.86	0.91	0.78
c-hexane / n-octane	0.9	0.91	0.79	0.82	0.91	0.95	0.83
c-hexane / i-octane	0.48	0.5	0.5	0.58	0.48	0.54	0.57
n-heptane / n-octane	0.49	0.52	0.51	0.5	0.51	0.54	0.51
n-heptane / i-octane	0.89	0.89	0.8	0.88	0.93	0.97	0.9
n-octane / i-octane	0.89	0.9	0.88	0.87	0.91	0.96	0.91

Figure 2.1: Structures of materials used in this study. All of these materials, except dioctyl phthalate (liquid), are solids at room temperature.



tetracosanoic acid

tetraoctylammonium bromide

lauric acid

propyl gallate

1,2,5,6,9,10-hexabromocyclododecane

tetracosane

quinacrine dihydrochloride dihydrate

dioctyl phthalate

Figure 2.2: Response of sensor A4 (75% CB, non polymer, Table 2.1) sensor C2 (40% CB, polymer, Table 2.1) on exposure to n-hexane at $P/P^o = 0.0050$. A single exposure of sensor C2, and six exposures of sensor A4, are shown. Continuous random exposures to each of the test analytes occurred, with 1, 35, 44, 62, and 71 h occuring between the first response shown and the second, third, fourth, fifth, and sixth responses, respectively.

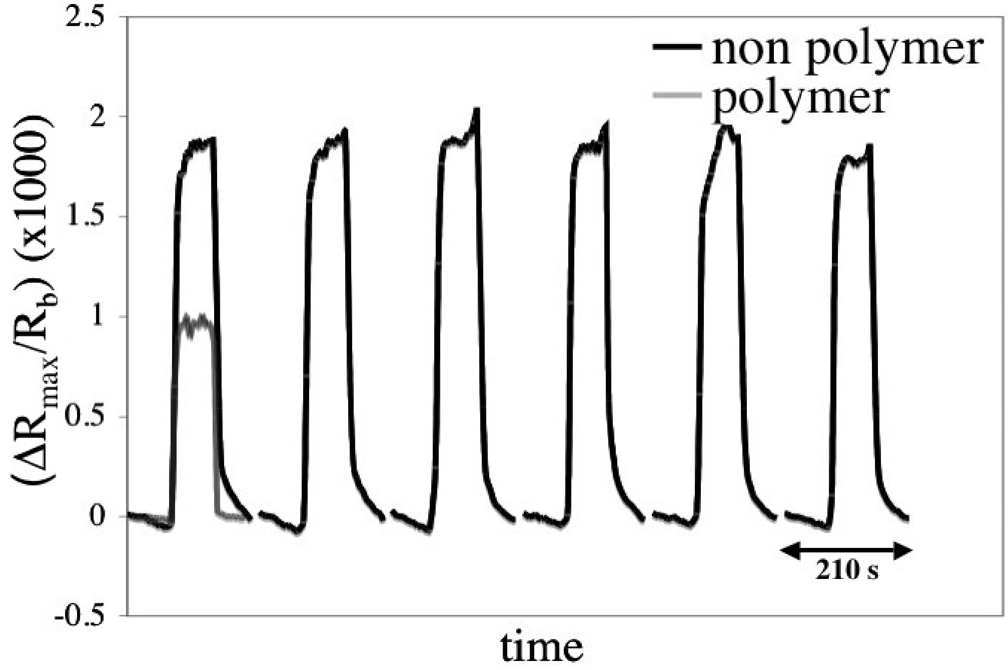


Figure 2.3: Responses of several non polymer-carbon black composite sensors (Table 2.1) to a) n-hexane and b) ethanol at various concentrations.

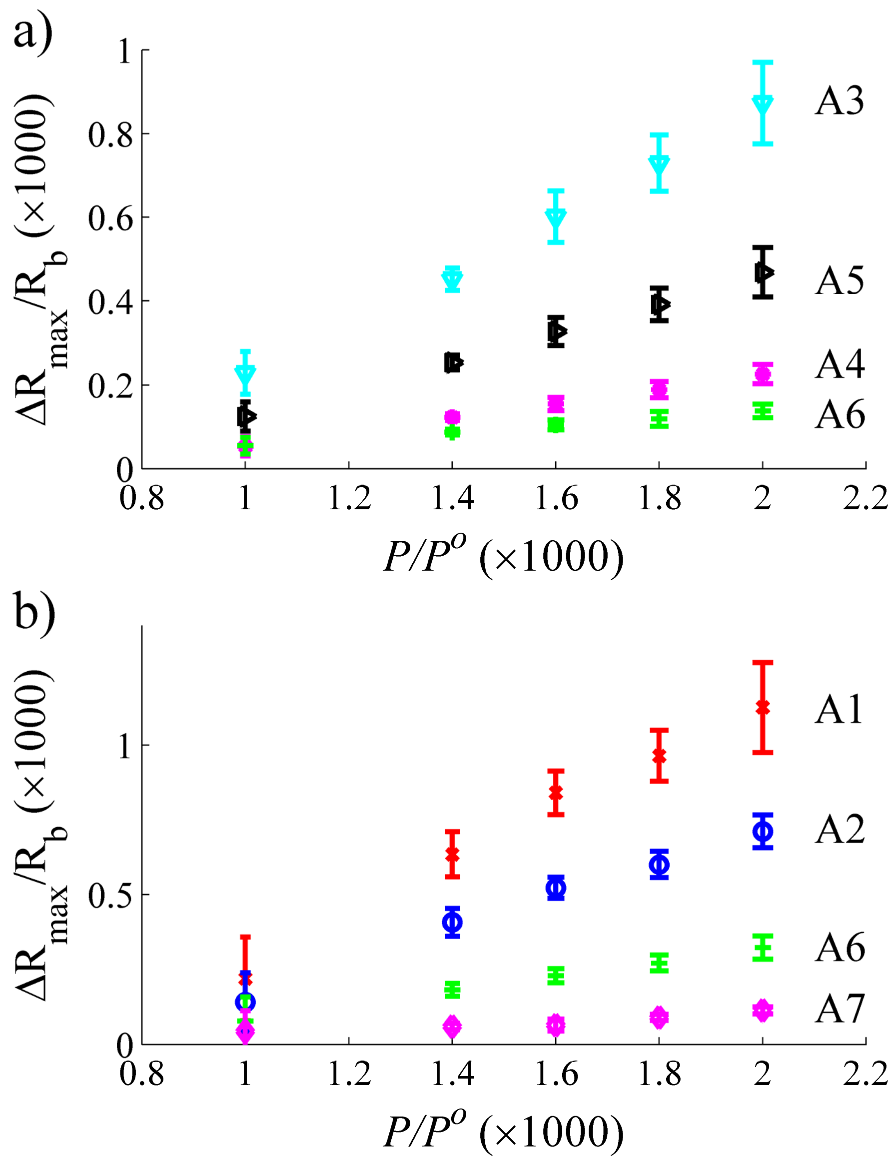


Figure 2.4: 3-D pattern depicting the average responses (Table 2.1) to the 7 test analytes at a concentration of $P/P^o = 0.0050$ in air.

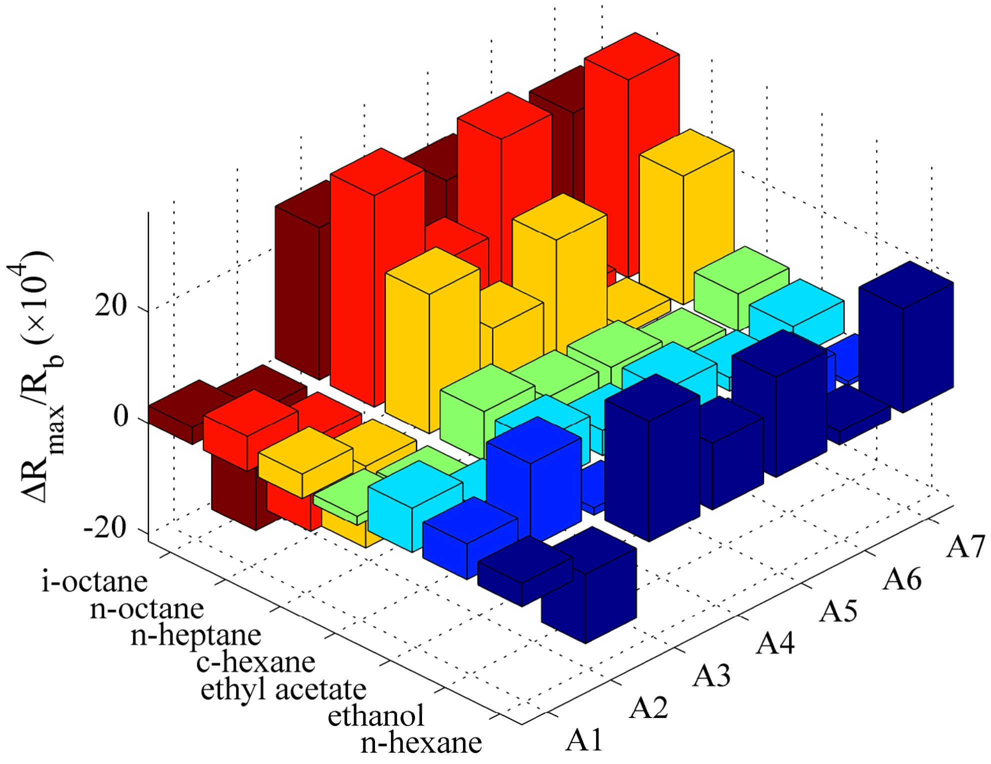


Figure 2.5: Principal components analysis showing principal components 1, 2 and 3 from normalized sensor array response data to all analytes on exposure to sensors A1-A7. The sub-plot details a second PCA performed on only the overlapping alkane clusters (highlighted). For visualization ease, only the first 50 exposures to each of the test analytes are analyzed and shown.

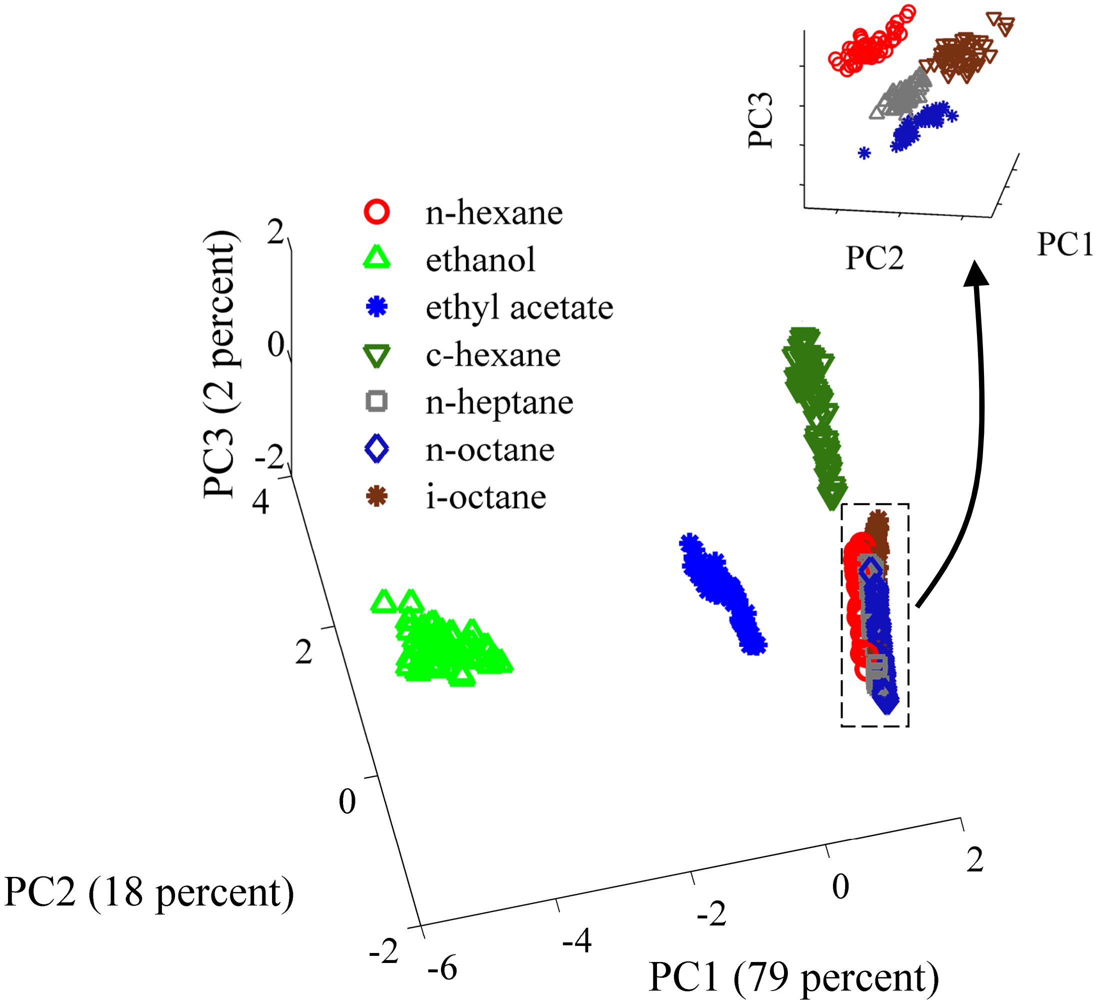
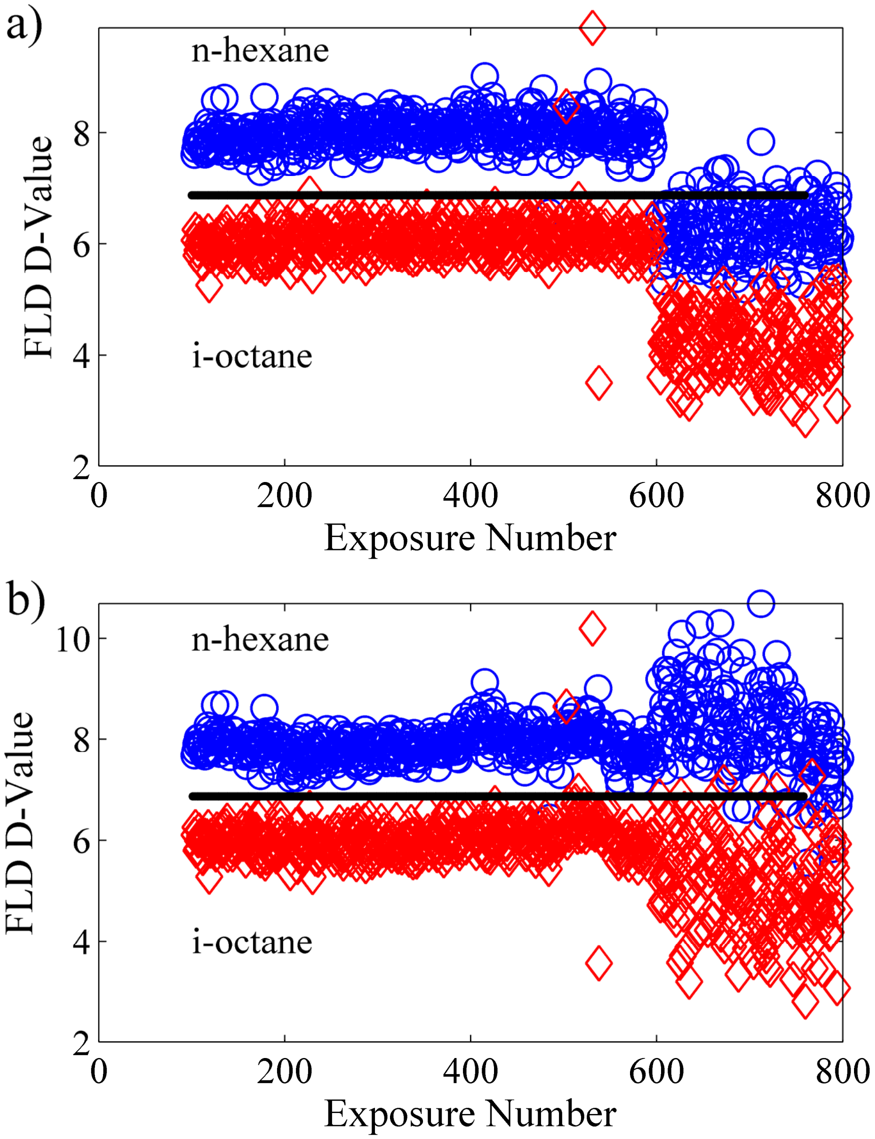


Figure 2.6: "Waterfall" plots detailing drift of FLD D-values (the single dimensional projection of the sensor array response which initially maximized the resolution factor for the classification task at hand) vs. exposure number for the n-hexane/i-octane binary separation task using an array of sensors B1-B9. The first 100 exposures of data were used to train the model. A decision boundary (solid line) based on these first 100 exposures is shown. Results are shown for a) no calibration and b) calibration using n-octane.



Chapter 3

Use of Spatiotemporal Response Information from Sorption-Based Sensor Arrays to Identify and Quantify the Composition of Analyte Mixtures[*]

3.1. Abstract

Linear sensor arrays made from small molecule/carbon black composite chemiresistors placed in a low headspace volume chamber, with vapor delivered at low flow rates, allowed for the extraction of chemical information that significantly increased the ability of the sensor arrays to identify vapor mixture components and to quantify their concentrations. Each sensor sorbed vapors from the gas stream, and thereby, as in gas chromatography, separated species having high vapor pressures from species having low vapor pressures. Instead of producing steady state (SS) sensor responses that were representative of the thermodynamic equilibrium partitioning of analyte between each sensor and the initial vapor phase, the sensor responses varied depending on the position of the sensor in the chamber and the time since the beginning of the analyte exposure. The concomitant spatiotemporal (ST) sensor array response therefore provided information that was a function of time as well as of the position of the sensor in the chamber. The responses to pure analytes and to multi-component analyte mixtures comprised of hexane, decane, ethyl acetate, chlorobenzene, ethanol, and/or butanol, were recorded along each of the sensor arrays. Use of a non-negative least squares (NNLSQ) method for analysis of the ST data enabled the correct identification and quantification of the composition of 2-, 3-, 4- and 5-component

---

[*] This chapter is reproduced according to American Chemical Society copyright guidelines, from "Use of Spatiotemporal Response Information from Sorption-Based Sensor Arrays to Identify and Quantify the Composition of Analyte Mixtures" by Marc D. Woodka, Bruce S. Brunschwig, and Nathan S. Lewis, *Langmuir*, 23, (26), 13232-13241, 2007. Copyright 2007, American Chemical Society.

mixtures from arrays using only 4 chemically different sorbent films. In contrast, when traditional time- and position-independent SS sensor response information was used, these same mixtures could not be identified or quantified robustly. The work has also demonstrated that for ST data, NNLSQ yielded significantly better results than analyses using extended disjoint principal components regression (EDPCR). The ability to correctly identify and quantify constituent components of vapor mixtures through the use of such ST information significantly expands the capabilities of such broadly cross-responsive arrays of sensors.

3.2. Introduction

Cross-responsive array-based vapor sensors have received significant attention in the recent literature. Such sensors include coated surface acoustic wave devices,[1-3] tin oxide sensors,[4,5] conducting organic polymers,[6-8] coated quartz crystal microbalances,[9,10] polymer-coated micromachined cantilevers,[11,12] dye-impregnated polymers coated onto optical fibers or beads,[13] polymer/carbon black composite chemiresistors,[14-16] and low volatility small molecule/carbon black composite chemiresistors.[17] Sensor arrays made from a variety of composite materials encompass a broad range of collective vapor/sensor interactions, producing a diversity of response values upon exposure to a given analyte. Arrays of such sensors, coupled with various pattern recognition approaches, are able to discriminate between different vapors.[18-20] Such arrangements have been termed “artificial” or “electronic” noses, due to their similarities to mammalian olfactory processes.[21,22]

Most reports dealing with cross-responsive sensor arrays have investigated the response of such arrays toward single analytes.[17,23,24] Alternatively, responses to complex mixtures have been used to “fingerprint” vapor mixtures rather than identify their constituents. Under this implementation, electronic noses have distinguished between different types of beers,[25] hops,[26] wines,[27] vinegars,[28] coffees,[29] and teas.[30] Zellers and co-workers have claimed that responses from an array consisting of 10 unique polymer-coated surface acoustic wave devices could, at best, provide robust information on simulated 3-component mixtures, where the simulated mixture responses assumed perfect additivity of pure vapor responses with superimposed Gaussian error as noise.[31]

Typical sensor array studies have placed each detector of the array in a nominally equivalent position relative to the analyte flow.[11,17,32] In this mode of operation, the partition coefficient between each sensor and vapor, K_{eq}, and the pressure of the vapor in the flow stream, determine the differences in sensor response. Di Francesco and co-workers have argued that an effective chamber design should assure that the transient time necessary to reach a stationary and uniform vapor concentration is much shorter than the sensor response time, so that all sensors are exposed at the same time to the same concentration of analyte. These workers have accordingly performed computational fluid dynamics (CFD) modeling studies in search of an optimal sensor geometry for this purpose.[33] Ali and co-workers have also sought to ensure that the entire incoming sample stream was distributed rapidly, simultaneously, and evenly over all of the sensors in an array. They have modeled vapor delivery in sensor geometries via CFD, and verified their results with flow visualization and measured sensor responses.[34] These studies are reflective of the dominant mode of sensor array implementation, in which each sensor is exposed, to the extent possible, to an identical, time-independent, stream of analyte vapor for the extraction of steady-state (SS) response information.

Exploitation of the spatiotemporal (ST) aspects of a non-uniform flow system may, however, yield additional information on the composition of analyte mixtures. For example, the flow dynamics of sniffing, as well as differences in the binding affinities of different odor receptors, are important for odor perception in mammalian olfaction. In humans, the vapor flow rate is different through the two nostrils of a given individual, because blood flow-induced occlusions in the nostrils cause the flow rate to vary with time and therefore vary which nostril has the higher flow rate. These varying flow patterns have been shown to affect odor perception.[35] Consistently, a sensor chamber modeled after a canine nasal cavity, having sensors placed throughout the cavity, has been shown to provide enhanced discrimination in various classification tasks relative to a single sensor array placed solely at the inlet of the cavity.[36] Additionally, the implementation of an olfactory microsystem which mimics the mammalian mucous layer with a polymeric layer has been shown to enhance discrimination between simple and complex odors.[37]

To measure the composition of certain vapor mixtures, metal-oxide semiconductor field-effect transistor (MOSFET) sensors have been placed along a sensor chamber that also contained thin films of palladium or platinum. These films catalytically partially decomposed certain vapors, so when mixtures of hydrogen, ammonia, acetylene, and/or ethanol were introduced into the sensor chamber, some of the constituents of the mixtures reacted, forming compounds not readily detected by the MOSFET sensors. For example, hydrogen flowing over the Pd catalyst reacted to form water, to which the sensors were insensitive. The initial sensors responded to the original vapors in the sampled stream, while later sensors responded to reduced concentrations of certain species. Using this approach, significant improvements were obtained in the identification of binary mixtures.[38-40] This method, however, is limited to mixtures having components that react at expensive catalytic surfaces to form products not easily detected at the sensors in the array.

More recently, arrays of polymer/carbon black composite sensors have been emplaced in a chamber with a low headspace volume while the analyte vapor was sampled at various flow rates. Depending on the vapor flow rate, pure test vapors and test vapor mixtures showed a concentration profile along the array as a function of time.[41,42] In this approach, the sensor material acted similarly to a stationary phase in a gas chromatographic (GC) column, with vapors partitioning into the sensor material as dictated by their solid/gas partition coefficient, K_{eq}. In this arrangement, the vapor species are not physically changed, unlike the situation involving the use of catalytic surfaces. Instead, the vapors are simply sorbed and retained by the sensor material. The progress of each vapor front down the sensor array is dictated by the flow rate, chamber geometry, and mass uptake by the upstream sensor films.

In this work, arrays of low volatility organic molecule/carbon black composite vapor sensors[17] have been exposed to various vapor mixtures in a low headspace volume chamber. In this configuration, the sensor material acts to separate the analyte to produce a space- and time-dependent signal response from the sensors in the array. A collection of such sensor arrays were first exposed to, and trained against, pure vapor species, each exposed at 5% of their saturated vapor pressure, $P/P^o = 0.050$, where P is the partial pressure and P^o is the saturated vapor pressure of the analyte of interest.[43] The sensor arrays were then challenged by exposures to various mixtures of these test vapors. Two linear, statistically based chemometric methods, non-negative least squares (NNLSQ)[44] and extended disjoint principal components regression (EDPCR),[31] were evaluated for their utility in analysis of the data. In each case, no *a priori* information was used regarding which vapors in the training set were present in the challenge mixtures. The performance of the ST array arrangement in speciation of mixtures was then compared to the performance of an SS array having an equal number of sensor response descriptors. The strengths and weaknesses of each data analysis approach employed for ST data were then compared.

3.3. Experimental

3.3.1. Materials

The insulating materials for fabrication of the sensor films consisted of tetracosane (99%), lauric acid (99.5%), and dioctyl phthalate (99%), purchased from Aldrich; as well as propyl gallate (98%) and quinacrine dihydrochloride dihydrate (97%), purchased from Acros Organics. Reagent grade toluene, tetrahydrofuran (THF), and chloroform, received from Aldrich, were used as solvents in the sensor suspensions. Hexane (95%), decane (99%), ethanol (95%), *n*-butanol (99.9%), ethyl acetate (99.5%) and chlorobenzene (99%), purchased from Aldrich, were used to generate vapors for delivery to the sensor arrays. Black Pearls 2000 (BP 2000), a furnace carbon black material donated by Cabot Co. (Billerica, MA), was used as the conductive phase in the sensor composites. All materials were used as received.

3.3.2. Detectors

Four suspensions, each comprised of a non-conductive sorption phase and a conductive carbon-black phase (Table 3.1), were used to fabricate the sensors evaluated in this work. First, the non-conductive (non carbon-black) sensor material(s) were placed in ≈ 60 mL of solvent and the suspension was sonicated for > 10 min. Carbon black was added to the suspension, and the resulting mixture was sonicated for > 30 min to produce a well-dispersed suspension. Dioctyl phthalate was used as a component of some of the sensor films to serve as a plasticizer and to add chemical diversity to the films.

Detector array substrates were fabricated by evaporating 30 nm of Cr and then 70 nm of Au onto glass microscope slides. A custom-made mask was used to produce the electrode pattern shown in Figure 3.1. The slide was masked with Teflon tape and sprayed with a single sensor solution using an airbrush (Iwata, Inc.). Several pairs of electrodes were monitored with an ohmmeter, and spraying was continued until the resistance across the 0.4 mm sensor electrode gaps was 500 - 1500 Ω. This created an overall sensor film of 75 x 5 mm in length and width, having a film thickness of ≈ 1-3 μm as measured with a Dektak 3030 profilometer (Sloan Technology Corp., Santa Barbara, CA). Four such detector substrates were made. Each glass slide had 15 identical sensors made of a thin film of one of the four suspensions listed in Table 3.1.

Four arrays were then placed into the custom-made aluminum sensor chamber depicted in Figure 3.2. The chamber was 110 mm long and 25 mm wide. In this study, only one side of the glass slide was coated with sensor material, so a total of 60 sensors were available for monitoring. A symmetric Teflon gasket (Figure 3.2a) was used to divide the incoming flow evenly among each of the eight vapor flow pathways. Additionally, for each of the vapor flow pathways, the inside of the aluminum chamber was covered with a film of Teflon tape (Figure 3.2c). Two weeks passed between the spraying of the sensor films and the initiation of the train/test phase.

3.3.3. Vapor Generation and Delivery

An automated flow system, controlled by LabVIEW 5.0 software, was used to deliver pulses of diluted streams of solvent vapor to the detectors.[41] The carrier gas (background stream) was oil-free air obtained from the house compressed air source (1.10 ± 0.15 parts per thousand of water vapor) controlled with a mass-flow controller (MFC) (UNIT) that could vary the flow from 6 to 625 mL min^{-1}. For generation of analyte vapors, a foreground stream of carrier gas was passed through a 220 mL bubbler filled with the desired solvent, controlled by a MFC that could vary the flow from 1 to 60 mL min^{-1}. The height of solvent in the bubblers was the same before and after each set of exposures.

Pure analyte vapors were presented to the sensor arrays at P/P^o = 0.010, 0.030, and 0.050 (1, 3, and 5 parts of foreground saturated vapor flow combined with 99, 97, and 95 parts of background air, respectively). Eight exposures were performed at each analyte concentration, for a total of 24 exposures per vapor. These pure analyte exposures were randomly delivered over all analytes and concentrations, and served as the "training" exposures. Exposures for the training period occurred over a 16-hour period.

Analyte vapor mixtures were generated by mixing equal volumes of each component. Background air bubbled through these mixtures was presented to the sensor arrays at P_{mix}/P^o_{mix} = 0.050, where P_{mix} is the sum of the partial pressure of the analytes and P^o_{mix} is the vapor pressure of the mixture. Twenty exposures of each of the mixtures were presented to the sensor arrays. These mixture exposures served as the "testing" exposures. Three exposure periods occurred over 13 h each, randomly exposing 2- and 3-component mixtures, 3- and 4- component mixtures, and 4- and 5- component mixtures. All training and testing data were collected during a five day period.

A total flow rate of 150 ± 5 mL min^{-1} (19 mL min^{-1} per chamber vapor flow pathway) was provided to the sensor chamber during the flow of either background or analyte vapor (Figure 3.2). Higher flow rates would provide more rapid sensor responses and can decrease the required exposure times.[45] However, hardware limitations of the sensor multiplexing speed and the resulting 3 s cycling frequency per sensor required longer exposure times to ensure the extraction of the desired transient response. To achieve flows with minimal variance in the rates of both the background and foreground streams, the mixtures were first generated at flow rates of 400 mL min^{-1}. A small Teflon-lined sampling pump (Science Pump Corporation) was used to withdraw vapor from the 400 mL min^{-1} stream and to present it to the sensor chamber at 150 ± 5 mL min^{-1}. Flow meters (Gilmont) were used to monitor the flow rates of the background and undiluted vapor streams, as well as to monitor the flow rate immediately prior to the entrance to the sensor chamber.

Gas chromatography-mass spectrometry (GC-MS) (Hewlett Packard 6890 GC system; Hewlett Packard 5973 Mass Selective Detector) was used to independently validate the compositions of the vapor mixtures. For each of the pure test analytes, the vapor stream delivered at a setting of P/P^o = 0.050 to the sensor chamber was sampled and manually injected into the GC. The GC-MS spectral peaks were then integrated to provide a calibration for that analyte at the specified fractional vapor pressure. Streams of the various mixtures were then also sampled and injected into the GC, and the GC-MS spectral peaks of each individually eluted analyte were integrated. The fractional vapor pressure of each species *i* in the mixture was calculated with eq (1):

$$\left(\frac{P}{P^o}\right)_i = \frac{A_i}{A_i^{cal}} \times 0.050\,, \tag{1}$$

where A_i was the integrated area of species *i* in the mixture and A_i^{cal} was the integrated area of species *i* in the calibration performed at $P/P^o = 0.050$. For all mixtures, a standard error propagation was performed on eq (1).[46] For the calibration of pure analyte vapors, at least six measurements were taken, while for mixtures, at least three measurements were taken. Mixtures consisting of 2, 3, 4, and 5 components were generated from the six test analyte vapors.

3.3.4. Measurements and Data Pre-Processing

Sensor film resistances were measured using a Keithley 2002 multimeter and a Keithley 7001 multiplexer. Each sensor substrate was connected to the multiplexer through shielded, twisted pair cables and a rotary ZIF connector (Tyco Electronics). To increase the overall sampling frequency, two Keithley 2002/7001 combinations were used to record the sensor response data. Each Keithley 2002/7001 combination recorded the responses from two of the four arrays, or 30 of the 60 total sensors. Sensor films were intentionally sprayed to produce film resistances within the same resistance range, 1000±500 Ω, to increase the multiplexing speed. Each sensor was sampled approximately every 3 s. Train and test exposures consisted of 70 s of pure background flow over the sensor arrays to establish a baseline resistance, followed by 150 s of analyte vapor flow at the desired fractional vapor pressure ($P/P^o = 0.010$, 0.030, or 0.050), followed by a stream of pure background flow for 230 s to restore the sensors to their initial states. Due to the 3 s cycling frequency, any exposure could experience up to a 3 s delay for vapor exposure initiation. For example, the pure background flow window prior to analyte exposure could range from 70 to 73 s. Prior to data collection, the sensors were subjected to 24 h of randomized exposures to all of the test vapors.[47] Pure vapors were first used to train the sensor arrays, followed by testing of the array with exposures to mixtures. Each set of exposures was randomized.

The response of each vapor detector for each analyte exposure was expressed as $S(t) = \Delta R(t)/R_b(t)$, where $R_b(t)$ is the baseline corrected resistance of the detector in the absence of analyte and $\Delta R(t)$ is the time-varying, baseline-corrected, resistance change upon exposure to the analyte. The baseline resistance, $R_b(t)$, was obtained by fitting a straight line to the data obtained during the pre-exposure period. The slope of this line was then used to extrapolate the resistance recorded immediately prior to exposure to any time *t*. $\Delta R(t)$ was calculated by subtracting $R_b(t)$ from the measured sensor resistance at time *t*. The actual times at which the sensor resistances were recorded varied with each exposure, and were different for each sensor in the array. The responses at the times used in the data analysis were calculated by interpolating between the measured data points.

In a control study, only the first, middle, and last sensor in each array (i.e., detectors 1, 8 and 15 in Figure 3.1) were sampled. This data set also captured ST sensor response information, albeit with fewer data points. These data were then compared to traditional SS data obtained from the responses of the first three sensors in each array. SS data was acquired when all sensors were in equilibrium with the initial vapor stream, near the end of the 150 s vapor exposure period. To facilitate comparisons between the two methods, the same total number of *S(t)* values were extracted in both cases. This procedure produced an equal number of total response descriptors from the three-sensor subarrays used to compare the ST and traditional SS sensing approaches.

3.3.5. Vapor Classification

The responses of these chemiresistive composite vapor sensors have been shown to be linear with the concentration of analyte, over the range of concentrations of interest in this work.[15,17] Statistical, linear-based pattern recognition techniques were therefore used to determine the identity and relative amounts of each analyte present in the vapor mixtures. Non-linear, neural network-based pattern recognition implementations may potentially provide enhanced performance in such tasks, but linear-based algorithms provide a more objective measure of performance. Hence, non-negative least squares (NNLSQ), as well as extended disjoint principal components regression (EDPCR) methods, were used to analyze the ST array responses of analyte mixtures.

3.3.5.1. Nonnegative Least Squares (NNLSQ)

For NNLSQ, training data collected at P/P^o = 0.050 were used to generate a vapor response library. Averaged baseline-corrected responses to the six vapors, *S*(*t*), extracted at four times (t = 80, 90, 130 and 150 s) for each of the fifteen sensors along each of the four arrays, were used to create a 240×6 library, **A**, of responses to the six pure analyte vapors. NNLSQ finds the linear combination of each of the pure response vectors that best matches, in a least squares sense, the mixture response vector. NNLSQ minimizes $||\mathbf{Ax}-\mathbf{b}||$, where **b** is the 240×1 measured sensor response vector to the mixture, and **x** is a 6×1 vector of concentrations of the analytes that minimizes the objective function, subject to $x_i \geq 0$.[44] NNLSQ was performed in MATLAB using a pre-programmed function.

3.3.5.2. Extended Disjoint Principal Components Regression (EDPCR)

EDPCR has been applied to determine the individual components of *simulated* vapor mixtures based on the linear addition of pure vapor exposures to an array of polymer-coated SAW sensors, with superimposed Gaussian error.[31] Briefly, EDPCR involves modeling, via Principal Components (PCs), the responses to each of the pure vapors, with a different set of PCs generated for each pure analyte. In principal, if there were no noise and if all of the sensors were perfectly linear, only a single PC would be required for each pure vapor. The direction of this PC in sensor response space would then be different for each pure vapor. All training data (P/P^o = 0.010, 0.030, 0.050) were used to generate EDPCR models for each of the pure vapors. These models describe the sensor response during exposure to a pure vapor, as a function of the vapor concentration. A leave-one-out cross-validation scheme, based on the residual squared variance between the actual and modeled array response as a function of the number of PCs employed, was used to determine the optimal number of PCs required to accurately describe the array response toward each pure vapor. The responses to analyte mixtures were then fit to all combinations of pure vapors for the 6 analytes of interest, ranging from single pure component models to the 6-component mixture model, thus encompassing 63 total possible vapor combinations.

Fitting the unknown array response to each of the 63 models consisted of projecting the response spectrum onto the PC-space of each model. A reconstructed response spectrum was obtained by projecting the PC-space data back on the transformation vectors for the model. If the model was incorrect, the reconstructed response differed significantly from the original response. A residual squared variance between the original response spectrum and the reconstructed response spectrum gave a measure of goodness of fit. The model with the lowest residual squared variance was then selected as the proper analyte combination. The analyte concentration was then determined based on the position of each analyte along the first PC for each of the vapors.

This algorithm is essentially equivalent to the more general soft independent modeling of class analogy (SIMCA) method, except that in SIMCA, each model that well-describes the data is accepted as a possibility, and more than one possibility may exist.[48-50] Hence, it is not necessary in SIMCA to obtain a single, unique solution. In contrast, in EDPCR, only the single model that provides the lowest residual squared error between an actual response vector and a modeled response vector is chosen for use. EDPCR was performed in Matlab, using custom-written code.

3.4. Results

3.4.1. Sensor Response

Figure 3.3 shows the baseline-corrected response of a sensor array made from propyl gallate and carbon black (sensor material 3, Table 3.1) to a) pure hexane and b) pure decane, each presented at P/P^o = 0.050, as a function of time. Responses were observed on two times scales: an immediate rapid response, and a slower drifting response. The latter was due to diffusive broadening of the vapor front as it progressed, through tubing, to the inlet of the sensor chamber. For visualization, the responses have been normalized by the response of each of the sensors at t = 220 s. The experimental setup produced a delay of ≈ 7 s between the initiation of analyte delivery and the response of the first sensor in the array. On exposure to hexane (P^o = 130 mm Hg), the vapor concentration rapidly became uniform over all sensors in the array, as evidenced by the similar response profiles vs. time for all 15 sensors. In contrast, on exposure to decane (P^o = 1 mm Hg), the response across the array varied significantly. The first sensor exhibited a fairly rapid response. However, decane has a partition coefficient into the sensor film that is roughly 100 times that of hexane,[41] hence, sorption of decane into the first sensor depleted the vapor sampling stream of the analyte. The sorption of analyte vapor by the earlier sensors therefore produced an altered, and delayed, arrival of the vapor front to the subsequent sensors along the array. In essence, the sensor material acted as a GC stationary phase, taking up and establishing equilibrium with the components of the vapor phase flow stream.

Figure 3.4 depicts the baseline-corrected response of the first and ninth sensors in an array of lauric acid/carbon black chemiresistive sensors (sensor type 2, Table 3.1) upon exposure to pure ethyl acetate, pure decane, or a mixture of ethyl acetate and decane, all delivered at P/P^o (or P_{mix}/P^o_{mix}) = 0.050. The first sensor in the array was exposed to a vapor stream that contained either ethyl acetate (P^o = 80 mm Hg), decane, or a mixture of both, at their original concentrations. For each vapor stream, this sensor exhibited a response that rapidly became nearly independent of time. In contrast, the ninth sensor exhibited a rapid response upon exposure to pure ethyl acetate, due to ethyl acetate having a relatively high vapor pressure and low partition coefficient into the sensor film in the array. However, consistent with Figure 3.3, this sensor exhibited a delayed response upon exposure to pure decane. Exposure to the mixture of ethyl acetate and decane showed a two-step sensor response, with one step occurring when the ethyl acetate arrived at the sensor, at ≈ 80 s, and with the other step occurring when the decane arrived at the sensor, at ≈ 130 s. The slow rise-time of the ninth sensor on exposure to decane and the mixture of decane and ethyl acetate is indicative of the slow increase in concentration of decane in the gas phase at the ninth sensor.

Figure 3.3 and Figure 3.4 each show differences in response profiles along the sensor array on exposure to decane. This is likely due to pulse broadening along the vapor flow pathway.[45] Pulse broadening could be minimized by operation at higher flow rates, which would provide sharper response profiles and greater analyte discrimination (assuming the hardware in place is able to capture these faster transient responses). However, pulse broadening is well-defined for a given geometry-flowrate-sensor-vapor combination, and therefore introduces response differences that are repeatable across repeated exposures.

3.4.2. Analysis of Mixtures

A series of fourteen mixtures was analyzed using the ST response data produced on the sensor array. For all exposures, the observed *S(t)* values were interpolated to fixed times having 10 s intervals, to produce 15 array responses in the time interval of t = 80 to 220 s. Subsets of these data, comprised of *S(t)* responses extracted at various times, ranging from single time-response descriptions to multiple time-response descriptions (including up to 7 different times), were then subjected to analysis using either NNLSQ or EDPCR.

The sum of the squared residual variance, S^2, was calculated between the mean mixture composition calculated by either NNLSQ or EDPCR, and the mixture composition indicated by GC-MS measurements (Table 3.2). The cross validation procedure for the EDPCR method yielded 3-7 PCs to describe each pure vapor response model. For all time combinations, NNLSQ was the better-performing algorithm. The optimal time combination was chosen as the combination of times that provided the lowest S^2 (best fit) between the deduced sensor array mixture composition and the mixture composition indicated by GC-MS. The optimal time combination was four times, at t = 80, 90, 130 and 150 s. The *S(t)* values extracted at t = 80 and 90 s (i.e., immediately after vapor delivery) provided information about the movement of higher-vapor pressure (lower-partitioning) analytes such as hexane, ethyl acetate and ethanol (P^o = 130, 80 and 50 mm Hg, respectively) down each array. The *S(t)* values at t = 130 and 150 s provided information on the progress of lower-vapor pressure (higher-partitioning) analytes, such as chlorobenzene, butanol and decane (P^o = 10, 5 and 1 mm Hg, respectively) down each array. Although this combination of analysis times provided the lowest S^2 between the sensor array and the GC-MS results, many combinations using 2-7 times (instead of 4), spanning approximately the same range of overall analysis times, provided comparable overall performance to that of the optimal 4-time data set.

For exposure of the sensor arrays to 2-, 3-, 4-, and 5-component mixtures, Figures 3.5-8, respectively, show the identified analytes and their estimated concentrations (●) in the analyte mixtures, as obtained from analysis of the response data using various sensor/pattern recognition combinations. These figures also display the concentrations of analytes as revealed by GC-MS analysis (□). Figures 3.5a-8a display these analyses using ST responses from all 15 sensors in each array, with responses at t = 80, 90, 130 and 150 s, and NNLSQ to identify the vapor constituents and reveal their respective concentrations.

The cross-validation for EDPCR yielded 3-7 PCs as optimal for each of the pure vapors. No clear-cut choice was available to determine the "best" fit, in that the model offering the lowest residual variance was chosen, but only because the residual error obtained with the selected model was marginally lower than that produced by the second-best fit. Figures 3.5b-8b display analyses that used all 15 sensors per array and EDPCR, using the optimal number of PCs to identify and quantify the constituents of each analyte mixture. In each case, analysis using NNLSQ more closely matched the results obtained with GC-MS. The use of NNLSQ also yielded smaller variances than were obtained through the use of EDPCR.

For exposures to all fourteen mixtures investigated in this work, Tables 3.3 and 3.4 list the estimated concentrations and standard deviations for each vapor, as well as the error (S^2), produced by using the various data analysis methods. Analysis of the data using NNLSQ generally produced smaller residual errors than were obtained from analysis using EDPCR. This measure of the residual variance was calculated using the mean concentration estimations for each of the mixtures, and therefore contains no information regarding the variance for individual components. The standard deviations however, showed a significantly larger variance in mixture analyses for the EDPCR method (Table 3.4) than for the NNLSQ method (Table 3.3).

Analyses were also performed using the responses produced by a limited number of sensors along each array. In this approach, the responses from the first, middle and last sensor along each array (detector positions 1, 8 and 15 in Figure 3.1) were sampled at t = 80, 90, 130 and 150 s, thus providing a data set that contained ST information from only three sensors per array. The twelve sensors that were not sampled provided a GC stationary phase equivalent that acted to separate vapors as they progressed along the length of each array. To compare ST and traditional data, SS data was measured using the first three sensors in each array, sampled at t = 205, 210, 215 and 220 s. Of the vapors present in the sampled mixtures, decane possessed the lowest vapor pressure. Vapor partitioning into sorption-based sensor films is inversely correlated with the vapor pressure of the analyte,[51] thus decane progressed most slowly along each sensor array. Figure 3.3b shows the response of such a sensor array upon exposure to decane, indicating that the first three sensors showed a fairly rapid response, and by t = 205 s essentially came to a steady-state equilibrium response. For the ST and SS method, responses were calculated at four separate times, to provide each method with an equal number of total response descriptors.

NNLSQ was used to identify and quantify the partial pressure of each vapor in the analyte, for ST and SS analysis using three-sensor arrays. Figures 3.3c-6c display the results obtained using ST detection with only three sensors per array on exposure to 2-, 3-, 4-, and 5- component mixtures, respectively. Comparing these results with those presented in Figures 3.3a-6a indicates that approximately the same results were obtained using this limited ST data set as were obtained using the ST response information from the full 15-sensor data set. In contrast, Figures 3.3d-6d display the results obtained using the traditional SS approach produced using the data from only the first three sensors in each array. SS analysis produced a marked decrease in the ability of the sensor array to correctly identify and quantify the presence of vapors in the tested mixtures.

Tables 3.5 and 3.6 list the estimated concentrations for each vapor obtained from exposures to each of the fourteen mixtures investigated in this work, using three sensors per array, with ST or SS data, respectively. These tables also present the sum of the squared residual error (x1000) between the estimated analyte concentrations obtained with each of the pattern recognition approaches, and the actual analyte concentration values obtained using GC-MS. For every situation investigated, the use of the ST aspects of the array response produced significantly better performance in the identification and quantification of the components of the vapor mixtures than the results obtained using traditional SS response information.

Figure 3.9 displays the sum of the squared residual error for each of the fourteen mixtures analyzed in this work. The results are presented for data analyzed with (1) NNLSQ using ST detection with the full 15 sensors per array, (2) EDPCR using ST detection with the full 15 sensors per array, (3) NNLSQ using ST detection with only 3 sensors per array, and (4) NNLSQ using SS detection (first three sensors). Vapor detection using NNLSQ on the ST response data, whether employed with the full 15 sensors per array or the limited 3 sensors per array, yielded approximately equal errors for each of the mixtures, as well as approximately the same variance in vapor estimation (Tables 3.3 and 3.5). ST vapor detection analyzed with EDPCR, using all 15 sensors per array, generally yielded a higher error as well as a much larger variance in vapor estimation (Table 3.4). The largest error in each of the mixture analyses was obtained using only the SS response information provided by the first three sensors in each array.

3.5. Discussion

Sorption-based sensors such as these have been shown to be linear with respect to pure analyte vapors over relatively low concentration ranges, as well as linear with respect to multiple vapors.[15,17,24] The development of the ST approach described here is predicated on this response linearity. Sensor response is essentially a dual-step process. First, chemical thermodynamics cause a given vapor to partition into the sensor film which causes it to volumetrically expand. This expansion causes an immediate change in the conductive properties of the film, which is linear when operating beyond the percolation regime of the composite film and over small volumetric changes.[16,52] The simplest model for the sorption behavior is given by the Langmuir adsorption isotherm, which predicts linear sorption at low concentrations over all species adsorbed. [53,54] This model assumes no interaction between any sorbed species and for all sorption sites to be energetically uniform.[53,54] At high concentrations, or in situations where interactions exist between adsorbed species (which alters the energetics of sorption), this linearity breaks down. Figures 3.3-6 show an unprecedented ability to identify mixtures, however deviations are evident between actual mixtures presented and sensor array perception. For example, Figure 3.6a shows an over-estimation of ethanol. This over-perception is likely due to a change in the adsorption energy of ethanol caused by the adsorption of ethyl acetate and chlorobenzene in the sensor film. These energy differences will vary to different degrees depending on species adsorbed, occasionally causing the sensor to wrongly perceive any given mixture.

The use of ST response data for vapor detection allowed for the extraction of important chemical information not available using traditional SS data. In this work, unnormalized baseline-corrected response data were used, to preserve information regarding the concentrations of the vapors in the test analytes. However, often sensor data is normalized:

$$\overline{S_i(t)} = \frac{S_i(t)}{\sum_{i=1}^{ns}\sum_{t=1}^{nt} S_i(t)}, \qquad (2)$$

where $S_i(t)$ is the response extracted from sensor *i* at time *t*, *ns* is the number of sensors used, and *nt* is the number of times used to extract a response $S_i(t)$. This normalization procedure creates a unit feature vector response for each exposure, largely independent of concentration, that can be used as a fingerprint for an individual analyte.[55,56] Performing principal components analysis (PCA) on normalized response data often reveals that the vast majority of the array response variance is contained in only a few principal component (PC) vectors.[17,24,57]

Performing PCA on normalized ST pure vapor training data consisting of 15 sensors per array indicated that the first five PCs contained 85% of the total response variance (35, 22, 15, 9 and 4%, respectively). Performing the same PCA using the normalized 3 sensors per array ST vapor training data similarly revealed that 86% of the total response variance was contained in the first 5 PCs (37, 21, 14, 8 and 6 %, respectively). The similarities in the PC eigenvalues for the two cases suggests that no significant additional information was extracted by using the data from the full 15 sensors in the array.

These findings are reflected in Figures 3.5-9, which show approximately equal performance for mixture identification for the two approaches. When PCA was performed on pure vapor training exposures for the normalized SS data set using the first three sensors in each array, the first five PCs contained 98% of the total response variance (65, 20, 9, 2 and 2%, respectively). While 85% of the variance is contained in only the first 2 PCs, 5 PCs were required for the ST data (note that both data sets have an equal number of response descriptors). The differences in mixture analysis performance reflect the limited amount of information obtained using the traditional SS sensing approach (Figures 3.5-9). Spreading out the variance over more PCs translates into more unique information, and an increased ability to analyze vapor mixtures, in accord with the results and conclusions reported herein.

NNLSQ yielded better mixture analysis than that obtained using EDPCR. EDPCR models the response of a sensor array toward pure vapors by loading the total response variance into a few chosen PCs, and truncates the data set by eliminating what is hopefully mostly noise. The EDPCR algorithm requires a response that is linear with concentration, and additive with respect to multiple vapors, to correctly identify analyte mixtures. Given the results obtained with NNLSQ, the sensors appear to be largely linear in their response. For responses perfectly linear with concentration, in the EDPCR method, only a single PC would be required to accurately model each pure vapor. This has been observed for polymer-coated QCM responses on exposure to pure vapors at varying concentrations.[31] The cross-validation used here to determine the optimal number of PCs yielded 3-7 PCs for each pure vapor, indicating the non-linearity in the data. The non-linear sensor responses cause EDPCR to incorrectly determine the components of analyte mixtures because of the lack of a perfect fit.

Whereas EDPCR condenses the dominant modes of the data variance into a few PCs, NNLSQ averages out relevant chemical information, and noise, over each of the individual response descriptors, and does not require an exact fit. Given the relatively large number of response descriptors in the data sets (48 for 3-sensor ST, 240 15-sensor ST), this averaging effect appears to aid the identification of the constituents of the vapor mixtures. Thus, EDPCR may be better-suited to simply describe possible fits, as in the SIMCA algorithm, rather than forcing the algorithm to decide on the single model that offers the lowest residual variance.

PCA is well-documented to be useful for displaying differences in the equilibrium response properties of an array exposed to various pure vapors.[15,17,57] However, when modeling the sensor responses of pure vapors collected using the ST method, some of the information important for mixture analysis may only be present in the lower PCs. Thus, a cross-validation procedure was used to determine the optimal number of PCs needed to model the pure vapors that yielded relatively poor results (Figures 3.5-9). For a more thorough analysis of the ability of EDPCR to identify and quantify mixture components, pure analytes were each modeled using different number of PCs ranging from 1 to 12. However, no combination of PCs offered a significant increase in array performance. The poor performance across all combinations of PCs suggests that EDPCR performed poorly due to non-perfectly linear sensor responses.

A single flow rate was employed in this study. The six pure analyte vapors chosen represented various chemistries and additionally spanned two orders of magnitude in vapor pressure. Hence, the given sensor arrangement and implementation proved sufficient to identify and quantify mixtures of the chosen vapors. However, as shown in Figure 3.3, for some analytes the latter sensors may never reach equilibrium with the concentration in the initial vapor stream. Thus, additional response information, including finer resolution of the progression of the analyte vapor down each array, could be extracted by using longer length arrays or through the use of multiple vapor flow rates. The limit of implementation of this approach would involve the use of either a flow rate slow enough that all sensors are in equilibrium with the vapor stream they are in contact with, limited by the rate of diffusion of vapors down the array; or the use of an infinitely long sensor array, which would be the equivalent of a GC column with sensors located along its length. An improved geometric implementation maximizing the information extracted per sensor employed would involve an exponentially spaced array of sensors, to better capture the ST response differences for analyte vapors having a large range of partition coefficients with the sensor films. Because approximately the same results were obtained when the ST responses were described by 3 or 15 sensors per array, only a limited number of sensor response descriptors were required for the vapors evaluated, to capture the necessary information to perform mixture decompositions down each array. Hence, when the chosen vapors possessed sufficient differences in chemistries and vapor pressures, only a fraction of the total sensors were required. Alternatively, had the vapors been more similar chemically, greater detail may have been required for mixture identification, and differences between the analysis of the ST response set containing the full 15 sensors per array and the limited set containing only 3 sensors per array would be observed.

In this work, the sensor material was a sorption-based composite comprised of small organic molecules and carbon black. Such films not only performed the vapor sensing function, but also served as the stationary phase into which the vapors partitioned and separated. The chemiresistive sensor films can be deposited and remain functional in most any form factor, making them especially attractive for use in the ST array response implementation. Many sorption-based sensors, however, are restricted in terms of the forms in which they can be fabricated. Coated quartz crystal microbalances, for instance, are restricted to the shape of the underlying quartz substrate. In situations such as these, in which one can not assemble the sensors in the type of array used here, ST mixture analysis could be applied by maintaining a low vapor headspace volume, and coating the walls of the chamber with various stationary phases. In this case, the only restriction would be that the stationary phases and sensors must exhibit mass uptake and sensor response, respectively, that is linear with concentration and additive in their response with respect to multiple vapors. If these criteria are met, the ST approach could be used with a wide variety of sensor types.

The ST data reported herein were obtained in a controlled laboratory setting. The flowrate to the sensor chamber was set at 150 mL min^{-1}, and ranged from 145 to 155 mL min^{-1} with random fluctuations throughout the training and testing periods. These fluctuations were accounted for during the training phase of the sensors. Had higher flowrate precision been achieved, the sensors would exhibit less variance in their responses, providing enhanced ability to correctly identify mixtures. Additionally, if the flowrate exhibited a systematic drift, the fingerprint response of each of the vapors could change significantly, causing degradation in the ability to correctly identify mixtures. To better understand how well this ST approach would perform in the real world, further work should be done to investigate how sensitive and/or robust this approach is to changes in exposure flowrate, as well as fluctuations/changes in temperature, humidity, and sensor response (drift). Furthermore, previous work has shown the ST method can readily detect low concentrations (ppb) of low vapor pressure analytes in the presence of higher concentrations (ppm) of high vapor pressure analytes.[41] Additional studies should be performed to better understand to what extent the ST method is able to identify a vapor present at low concentration in the presence of vapor(s) present at higher concentrations.

3.6. Conclusions

Use of ST data has been shown to provide enhanced performance in analysis of vapor mixtures relative to the traditional SS response of an array of broadly cross-responsive vapor sensors. In a low-volume headspace chamber that allows each sensor to be exposed to a well-defined, time-varying vapor stream, the sensor material acted as a chromatographic stationary phase, causing vapors to be retarded in progression along the array. The retardation was proportional to the sensor/vapor partition coefficient. The resulting sensor responses at long times and/or at positions close to the inlet captured the traditional SS sensor response differences to an unchanging vapor stream, but at shorter times and positions further from the inlet, also measured the progress of each vapor down the sensor array. Under such conditions, significantly more information was obtained on analyte mixtures relative to the information obtained using traditional sensor responses alone. The observed behavior was in accord with the physics of the time-varying convection-diffusion equation, with the flux dictated by the concentration of vapor in each of the phases and the vapor/sensor partition coefficient.[58] Modeling of the ST method to better understand its limitations, as well as explore its potential in microfluidic application, is currently underway.[58]

Previous reports using cross-responsive sensor arrays have addressed pure vapor identification, or the identification of complex mixtures as a whole, but the identification and analysis of mixtures containing more than three components has not been previously achieved. The ability of an ST sensor array consisting of only four sensor types to correctly identify and quantify 2-, 3-, 4- and 5-component mixtures, using a library consisting of responses to six pure analyte vapors, demonstrates the importance of the use of ST information. The implementation of the ST method thus clearly increases the possible application space of such sensor arrays for analysis of analyte mixtures, as well as for analysis of pure analytes.

3.7. References

(1) Rose-Pehrsson, S. L.; Grate, J. W.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801-2811.
(2) Patrash, S. J.; Zellers, E. T. *Anal. Chem.* **1993**, *65*, 2055-2066.
(3) Grate, J. W.; Patrash, S. J.; Kaganove, S. N.; Abraham, M. H.; Wise, B. M.; Gallagher, N. B. *Anal. Chem.* **2001**, *73*, 5247-5259.
(4) Getino, J.; Horrillo, M. C.; Gitierrez, J.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Sens. Act. B* **1997**, *43*, 200-205.
(5) Srivastava, R.; Dwivedi, R.; Srivastava, S. K. *Sens. Actuators, B* **1998**, *50*, 175-180.
(6) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1995**, *92*, 2652-2656.
(7) Harris, P. D.; Arnold, W. M.; Andrews, M. K.; Partridge, A. C. *Sens. Actuators, B* **1997**, *42*, 177-184.
(8) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, A* **1995**, *51*, 57-66.
(9) Fu, Y.; Finklea, H. O. *Anal. Chem.* **2003**, *75*, 5387-5393.
(10) Grate, J. W.; Nelson, D. A.; Skaggs, R. *Anal. Chem.* **2003**, *75*, 1868-1879.
(11) Baller, M. K.; Lang, H. P.; Fritz, J.; Gerber, C.; Gimzewski, J. K.; Drechsler, U.; Rothuizen, H.; Despont, M.; Vettiger, P.; Battiston, F. M.; Ramseyer, J. P.; Fornaro, P.; Meyer, E.; Guntherodt, H.-J. *Ultramicroscopy* **2000**, *82*, 1-9.
(12) Battiston, F. M.; Ramseyer, J. P.; Lang, H. P.; Baller, M. K.; Gerber, C.; Gimzewski, J. K.; Meyer, E.; Guntherodt, H.-J. *Sens. Actuators, B* **2001**, *77*, 122-131.
(13) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 2501-2508.
(14) Doleman, B. J.; Sanner, R. D.; Severin, E. J.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 2560-2564.
(15) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658-668.
(16) Sisk, B. C.; Lewis, N. S. *Langmuir* **2006**, *22*, 7928-7935.
(17) Gao, T.; Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Chem. Mat.* **2006**, *18*, 5193-5202.
(18) Sasaki, I.; Tsuchiya, H.; Nishioka, M.; Sadakata, M.; Okubo, T. *Sens. Actuators, B* **2002**, *86*, 26-33.
(19) Gardner, J. W.; Hines, E. L.; Tang, H. C. *Sens. Actuators, B* **1992**, *9*, 9-15.
(20) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177-4190.
(21) Lewis, N. S. *Acc. Chem. Res.* **2004**, *37*, 663-672.
(22) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, B* **1994**, *18*, 211-220.
(23) Koscho, M. E.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **2002**, *74*, 1307-1315.

(24) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mat.* **1996**, *8*, 2298-2312.

(25) Zhang, C.; Bailey, D. P.; Suslick, K. S. *J. Agr. Food. Chem.* **2006**, *54*, 4925-4931.

(26) Lamagna, A.; Reich, S.; Rodriguez, D.; Scoccola, N. N. *Sens. Actuators, B* **2004**, *102*, 278-283.

(27) Lozano, J.; Fernandez, M. J.; Fontecha, J. L.; Aleixanre, M.; Santos, J. P.; Sayago, I.; Arroyo, T.; Cabellos, J. M.; Futierrez, F. J.; Horrillo, M. C. *Sens. Actuators, B* **2006**, *120*, 166-171.

(28) Zhang, Q.; Zhang, S.; Xie, C.; Zeng, D.; Fan, C.; Li, D.; Bai, Z. *Sens. Actuators, B* **2006**, *119*, 538-546.

(29) Pardo, M.; Sberveglieri, G. *IEEE Trans. Inst. Meas.* **2002**, *51*, 1334-1339.

(30) Dutta, R.; Hines, E. L.; Gardner, J. W.; Kashwan, K. R.; Bhuyan, M. *Sens. Actuators, B* **2003**, *94*, 228-237.

(31) Zellers, E. T.; Pan, T.-S.; Patrash, S. J.; Han, M.; Batterman, S. A. *Sens. Actuators, B* **1993**, *12*, 123-133.

(32) Sotzing, G. A.; Phend, J. N.; Grubbs, R. H.; Lewis, N. S. *Chem. Mat.* **2000**, *12*, 593-595.

(33) Falcitelli, M.; Benassi, A.; Di Francesco, F.; Domenici, C.; Marano, L.; Pioggia, G. *Sens. Actuators, B* **2002**, *85*, 166-174.

(34) Scott, S. M.; James, D.; Ali, Z.; O'Hare, W. T. *J. Therm. Anal. Cal.* **2004**, *76*, 693-708.

(35) Sobel, N.; Khan, R. M.; Saltman, A.; Sullivan, E. V.; Gabrieli, J. D. E. *Nature* **1999**, *402*, 35-35.

(36) Stitzel, S. E.; Stein, D. R.; Walt, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 3684-3685.

(37) Covington, J. A.; Gardner, J. W.; Hamilton, A.; Pearch, T. C.; Tan, S. L. *IET Nanobiotechnology* **2007**, *1*, 15-21.

(38) Eklov, T.; Martensson, P.; Lundstrom, I. *Anal. Chim. Acta* **1999**, *381*, 221-232.

(39) Eklov, T.; Lunstrom, I. *Sens. Actuators, B* **1999**, *57*, 274-282.

(40) Eklov, T.; Lundstrom, I. *Anal. Chem.* **1999**, *71*, 3544-3550.

(41) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *82*, 54-74.

(42) Lewis, N. S.; Freund, M. S.; Briglin, S. M. U.S. Patent No. *6962675*; USPTO, Ed.: U.S.A., 2005.

(43) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 5442-5447.

(44) Lawson, C. L.; Hanson, R. J. *Solving Least Square Problems*; Prentice Hall: Englewood Cliffs NJ, 1974.

(45) Tan, S. L.; Covington, J. A.; Gardner, J. W. *IEE Proc.-Sci. Meas. Technol.* **2006**, *153*, 94-100.

(46) Aikens, D. A.; Bailey, R. A.; Moore, J. A.; Giachino, G. G.; Tomkins, R. P. T. *Principles and Techniques for an Integrated Chemistry Laboratory*; Waveland Press Inc.: Prospect Heights, IL, 1984.

(47) Burl, M. C.; Sisk, B. C.; Vaid, T. P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *87*, 130-149.

(48) Brereton, R. G. *Chemometrics - Data Anlysis for the Laboratory and Chemical Plant*; John Wiley & Sons Ltd: West Sussex, England, 2003.

(49) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*; Wiley-VCH: New York, 1999.

(50) Wold, S. *Patt. Rec.* **1976**, *8*, 127-139.

(51) Littlewood, A. B. *Gas Chromatography: Principles, Techniques, and Applications*, 2nd ed.; Academic Press: New York, 1970.

(52) Lundberg, B.; Sundqvist, B. *J. Appl. Phys.* **1986**, *60*, 1074-1079.

(53) Houston, P. L. *Chemical Kinetics and Reaction Dynamics*; McGraw-Hill: New York, 2001.
(54) Schmidt, L. D. *The Engineering of Chemical Reactions*, 2 ed.; Oxford University Press: New York, 2005.
(55) Sisk, B. C.; Lewis, N. S. *Sens. Actuators, B* **2005**, *104*, 249-268.
(56) Gardner, J. W.; Bartlett, P. N. *Electronic Noses: Principles and Applications*; Oxford University Press: New York, NY, 1999.
(57) Hopkins, A. R.; Lewis, N. S. *Anal. Chem.* **2001**, *73*, 884-892.
(58) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *manuscript in preparation*.

Table 3.1: Sensor suspensions used to spray sensor films. For each suspension, non-carbon black components were combined with ≈ 60 mL of the listed solvent and the suspension was sonicated for > 10 min. Carbon black (CB) was then added to the solution, which was then sonicated for > 30 min to obtain a well-dispersed suspension. Each suspension was 75% carbon black by weight. DP = dioctyl phthalate.

suspension	sensor materials	solvent
1	35 mg tetracosane, 15 mg DP, 150 mg CB	toluene
2	35 mg lauric acid, 15 mg DP, 150 mg CB	tetrahydrofuran
3	50 mg propyl gallate, 150 mg CB	tetrahydrofuran
4	50 mg quinacrine dihydrochloride, 150 mg CB	chloroform

Table 3.2: Fractional vapor pressures (x1000) of analyte vapors present in each of the mixtures, as determined by GC-MS sampling (calculated by eq (1)) immediately prior to entry into the sensor chamber.

mixture	hexane	decane	ethyl acetate	chloro-benzene	ethanol	butanol
1	0	23±11	27±15	0	0	0
2	0	0	13±2	24±1	0	0
3	0	0	0	23±1	0	22±1
4	0	20±1	0	0	0	31±2
5	0	11±3	8±7	20±4	0	0
6	23±2	0	0	0	29±8	44±16
7	27±8	14±8	0	0	0	44±13
8	0	0	4±1	12±1	23±6	0
9	22±7	0	0	28±11	45±9	0
10	18±4	12±4	0	0	27±4	11±3
11	9±8	14±2	0	22±6	0	69±27
12	13±8	11±4	8±2	0	31±6	20±5
13	12±0	0	4±0	8±1	23±2	13±2
14	8±0	5±0	4±0	8±0	27±2	0

Table 3.3: Fractional vapor pressures (×1000) of analyte vapors present in each of the mixtures, as determined by NNLSQ, using all 15 sensors per sensor array with *S(t)* extracted at *t* = 80, 90, 130 and 150 s. Mean and standard deviations are given, calculated over 20 exposures to each mixture. Squared residual error (S^2) for each mixture, calculated between the listed means and those obtained with GC-MS (Table 3.2), are given in the final column.

mixture	hexane	decane	ethyl acetate	chloro-benzene	ethanol	butanol	S^2
1	4±2	18±8	34±3	2±2	7±10	2±2	0.15
2	0	4±3	17±4	26±2	8±6	6±5	0.14
3	0	3±4	0	29±8	3±4	31±6	0.14
4	0	31±3	0	1±1	1±2	35±4	0.14
5	1±2	11±3	18±3	22±2	8±5	3±2	0.18
6	32±3	1±1	3±3	0±1	29±14	11±9	1.18
7	30±5	3±3	1±2	1±2	37±10	8±8	2.80
8	1±1	2±3	10±2	17±4	44±9	3±6	0.52
9	30±6	1±2	1±2	16±4	36±6	2±3	0.30
10	26±2	8±3	2±2	0	24±9	15±3	0.11
11	19±3	4±2	2±3	17±3	15±4	22±4	2.66
12	22±2	5±3	13±2	0±1	22±8	14±6	0.26
13	26±3	0	10±3	7±3	24±8	13±4	0.23
14	16±3	3±2	5±2	10±3	31±10	10±5	0.19

Table 3.4: Fractional vapor pressures (×1000) of analyte vapors present in each of the mixtures, as determined by EDPCR, using all 15 sensors per sensor array with *S(t)* extracted at *t* = 80, 90, 130 and 150 s. Mean and standard deviations are given, calculated over 20 exposures to each mixture. Squared residual error (S^2) for each mixture, calculated between the listed means and those obtained with GC-MS (Table 3.2), are given in the final column.

mixture	hexane	decane	ethyl acetate	chloro-benzene	ethanol	butanol	S^2
1	0±2	7±10	44±7	1±3	1±2	0±2	0.55
2	7±9	1±4	19±7	14±11	2±5	0	0.19
3	5±8	5±15	0	30±23	0	15±22	0.15
4	0	21±14	0	2±7	0	36±14	0.03
5	8±17	5±10	21±13	10±10	3±12	0	0.38
6	27±8	0	4±9	0	13±13	11±16	1.38
7	29±12	0	4±7	1±4	0±23	28±17	0.47
8	8±6	2±5	30±10	0	16±23	0	0.94
9	35±10	1±3	20±7	0	5±12	0	2.95
10	29±18	3±5	5±9	0	19±22	0	0.41
11	28±7	0	3±7	5±10	1±4	21±20	3.16
12	17±16	0	23±11	0	13±22	0	1.09
13	23±8	0	24±8	1±5	8±20	0	0.96
14	20±8	1±4	18±11	1±5	9±19	0	0.73

Table 3.5: Fractional vapor pressures (×1000) of analyte vapors present in each of the mixtures, as determined by NNLSQ, using the first, middle and last sensor in each array with *S(t)* extracted at t = 80, 90, 130 and 150 s. Mean and standard deviations are given, calculated over 20 exposures to each mixture. Squared residual error (S^2) for each mixture, calculated between the listed means and those obtained with GC-MS (Table 3.2), are given in the final column.

mixture	hexane	decane	ethyl acetate	chloro-benzene	ethanol	butanol	S^2
1	3±2	20±8	36±2	1±2	7±10	3±3	0.16
2	0	5±5	15±4	25±3	11±7	7±5	0.20
3	0	5±5	0	29±8	3±3	32±6	0.17
4	0	30±3	0	0±	1±2	38±3	0.15
5	0±1	13±4	17±2	21±2	9±5	5±3	0.19
6	29±3	0±1	4±4	1±1	33±13	10±7	1.23
7	28±4	2±4	1±2	3±4	37±9	10±8	2.68
8	0±1	2±3	6±4	18±4	51±10	3±5	0.84
9	27±5	1±2	1±2	17±4	42±6	1±2	0.16
10	24±3	5±4	2±2	0	30±9	13±5	0.10
11	20±5	3±2	1±2	15±4	15±5	24±2	2.54
12	18±3	4±3	15±2	0	28±8	11±5	0.21
13	27±5	0	10±2	4±4	28±9	14±3	0.30
14	16±4	3±3	5±2	8±4	35±10	10±5	0.23

Table 3.6: Fractional vapor pressures (×1000) of analyte vapors present in each of the mixtures, as determined by NNLSQ, using the first three sensors in each array with *S(t)* extracted at *t* = 205, 210, 215 and 220 s. Mean and standard deviations are given, calculated over 20 exposures to each mixture. Squared residual error (S^2) for each mixture, calculated between the listed means and those obtained with GC-MS (Table 3.2), are given in the final column.

mixture	hexane	decane	ethyl acetate	chloro-benzene	ethanol	butanol	S^2
1	35±21	5±9	4±7	5±6	37±9	0	3.47
2	6±7	1±3	8±13	25±4	25±15	1±2	0.69
3	11±10	0	0±1	20±6	33±12	9±12	1.39
4	42±17	3±8	2±3	6±6	25±5	0	3.68
5	21±11	2±5	3±8	24±7	22±8	0	1.05
6	15±12	3±5	2±3	5±8	47±15	3±3	2.11
7	15±15	10±10	1±2	3±6	52±11	1±1	4.72
8	2±4	3±4	4±6	7±7	72±10	0	2.44
9	7±11	11±11	2±3	15±13	48±14	0±1	0.53
10	40±8	0	0	7±9	27±8	0±1	0.80
11	37±9	0	0	11±9	35±4	0	7.09
12	31±7	0	0	9±7	39±5	0	1.05
13	31±8	0	0	5±7	49±9	0±1	1.23
14	24±11	2±5	0	12±10	40±11	0	0.47

Figure 3.1: Sensor substrate layout for a single constant-composition sensor array. The sensor solution was sprayed using an airbrush to generate a 5 mm × 75 mm thin film array 1-3 μm in thickness. 15 pairs of underlying Au electrodes allowed for the monitoring of 15 sensors along the array.

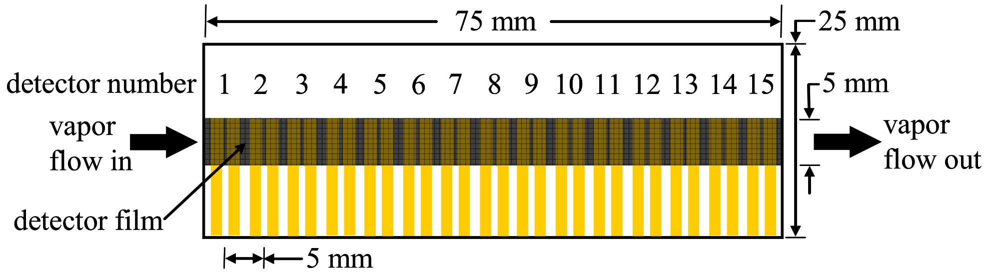


Figure 3.2: Chamber design used to accommodate multiple vapor flow pathways and sensor multiplexing: (a) three-dimensional chamber geometry with dimensions shown; b) head-on view of (a), showing all four sensor arrays as well as the vapor flow pathways, with the dashed section shown in detail in (c), which shows individual vapor flow pathways and dimensions, as well as a sensor film (not to scale).

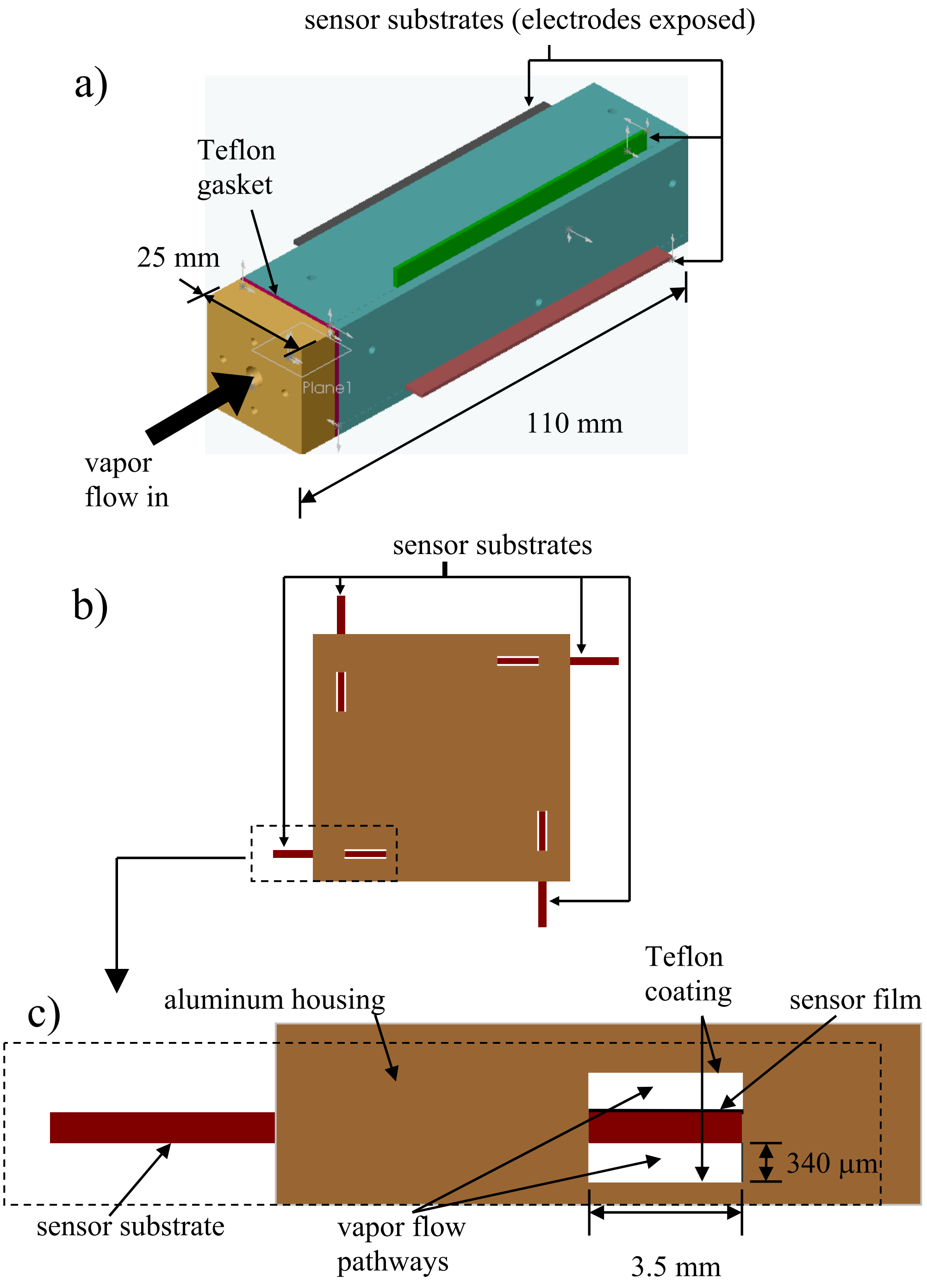


Figure 3.3: Response of a 15-sensor array of propyl gallate/carbon black (sensor composition 3, Table 3.1) to a) pure hexane and b) pure decane, each delivered at $P/P^o = 0.050$. S_1 and S_{15} denote the 1st and 15th sensor in the array, respectively. To put all sensors on the same scale, for visualization, each sensor response was scaled by its response at $t = 220$ s.

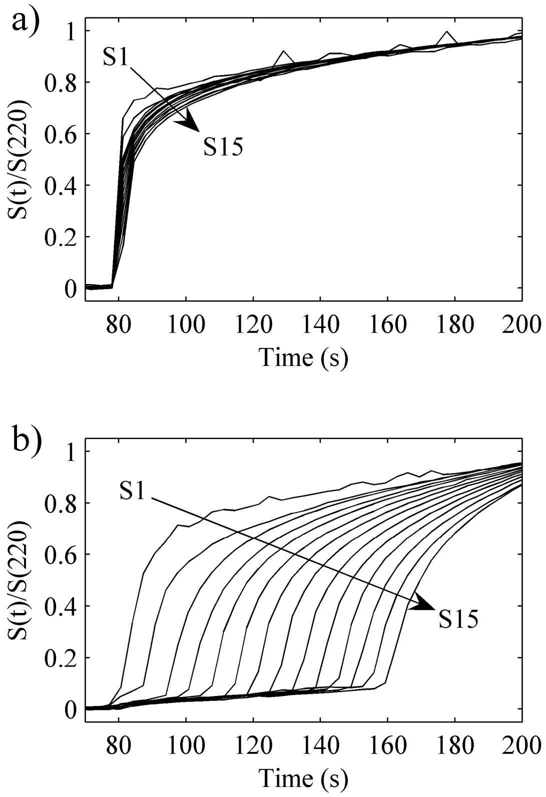


Figure 3.4: Response of the first and ninth sensors (detectors 1 and 9, Figure 3.1) along a lauric acid/dioctyl phthalate/carbon black sensor array (sensor type 2, Table 3.1) to pure ethyl acetate delivered at $P/P^o = 0.050$, pure decane delivered at $P/P^o = 0.050$ and a mixture of ethyl acetate and decane exposed at $P_{mix}/P^o_{mix} = 0.050$.

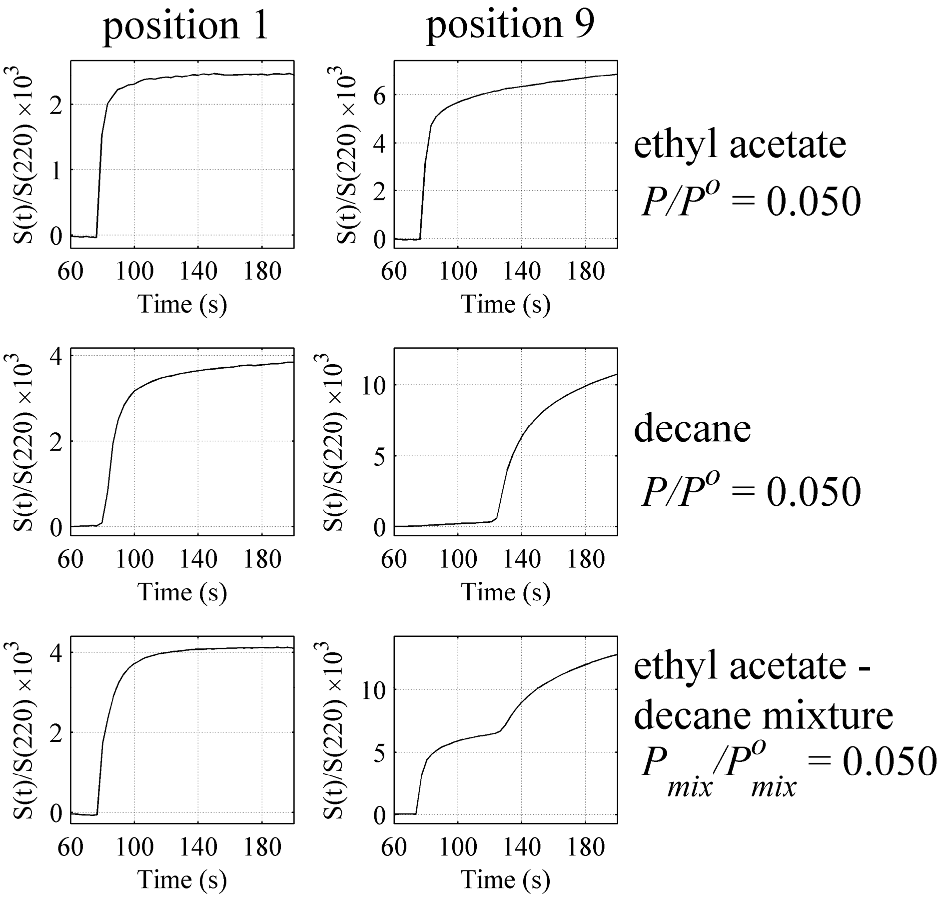


Figure 3.5: Comparison between various sensor/pattern recognition configurations in their analysis of 2-component mixture 3 (Tables 3.2-6): a) 15 sensors along each array, responses sampled at t = 80, 90, 130 and 150 s, analyzed with NNLSQ; b) 15 sensors along each array, responses sampled at t = 80, 90, 130 and 150 s, analyzed with EDPCR; c) first middle and last sensor along each array (detector numbers 1, 8 and 15, Figure 3.1), responses sampled at t =80, 90, 130 and 150 s, analyzed with NNLSQ; and d) first three sensors along each array (detector numbers 1-3, Figure 3.1), responses sampled at t =205, 210, 215 and 220 s, analyzed with NNLSQ. Results obtained with each sensor/pattern recognition configuration (●) and GC-MS (□) are given, with error bars. A1 – A6 denote hexane, decane, ethyl acetate, chlorobenzene, ethanol, and butanol, respectively.

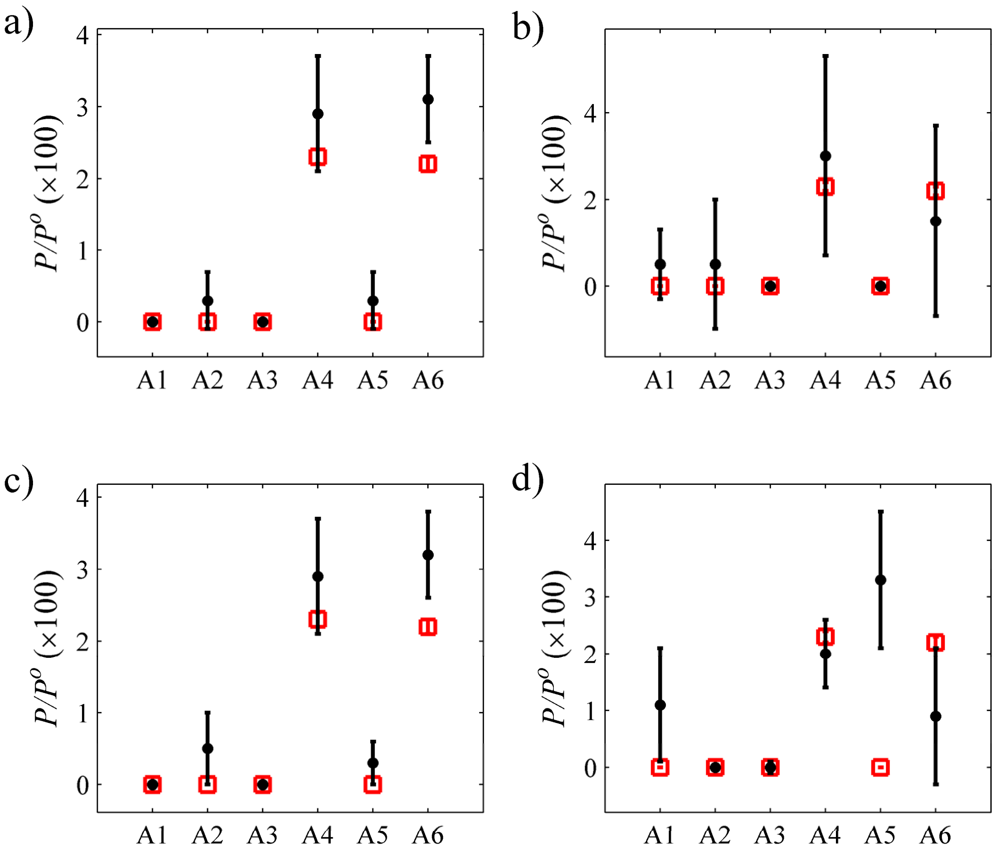


Figure 3.6: Comparison between various sensor/pattern recognition configurations in their analysis of 3-component mixture 8 (Tables 3.2-6). See Figure 3.5 caption for descriptions.

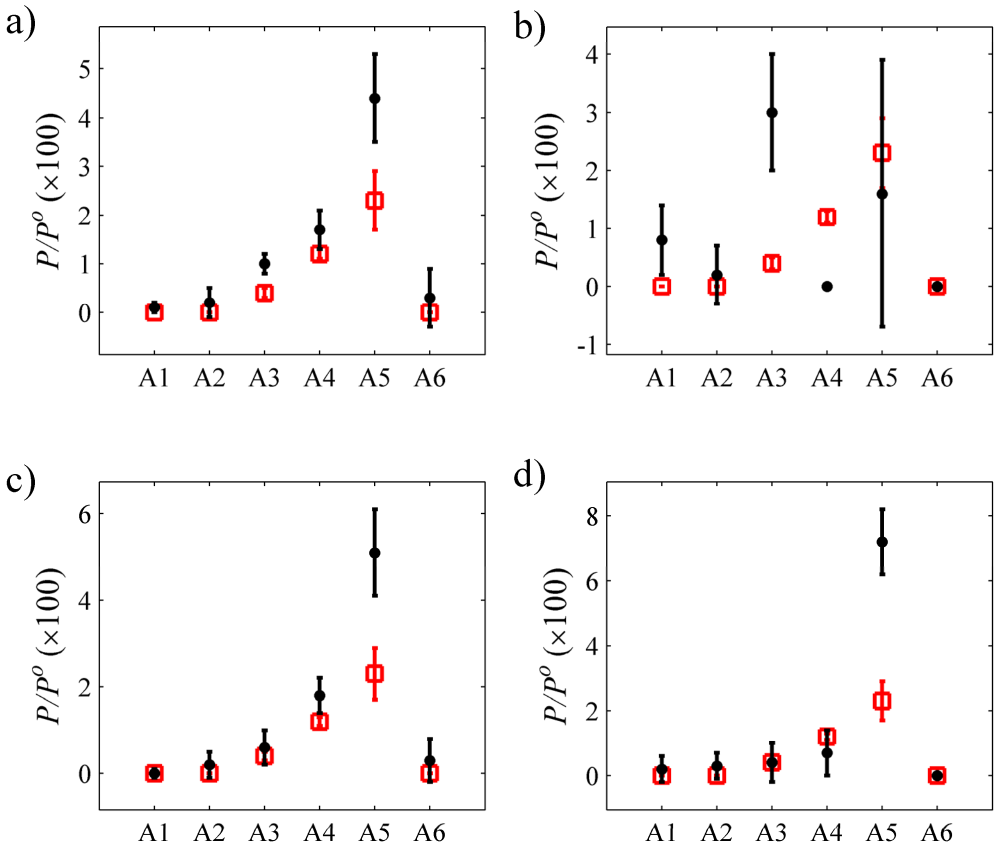


Figure 3.7: Comparison between various sensor/pattern recognition configurations in their analysis of 4-component mixture 10 (Tables 3.2-6). See Figure 3.5 caption for descriptions.

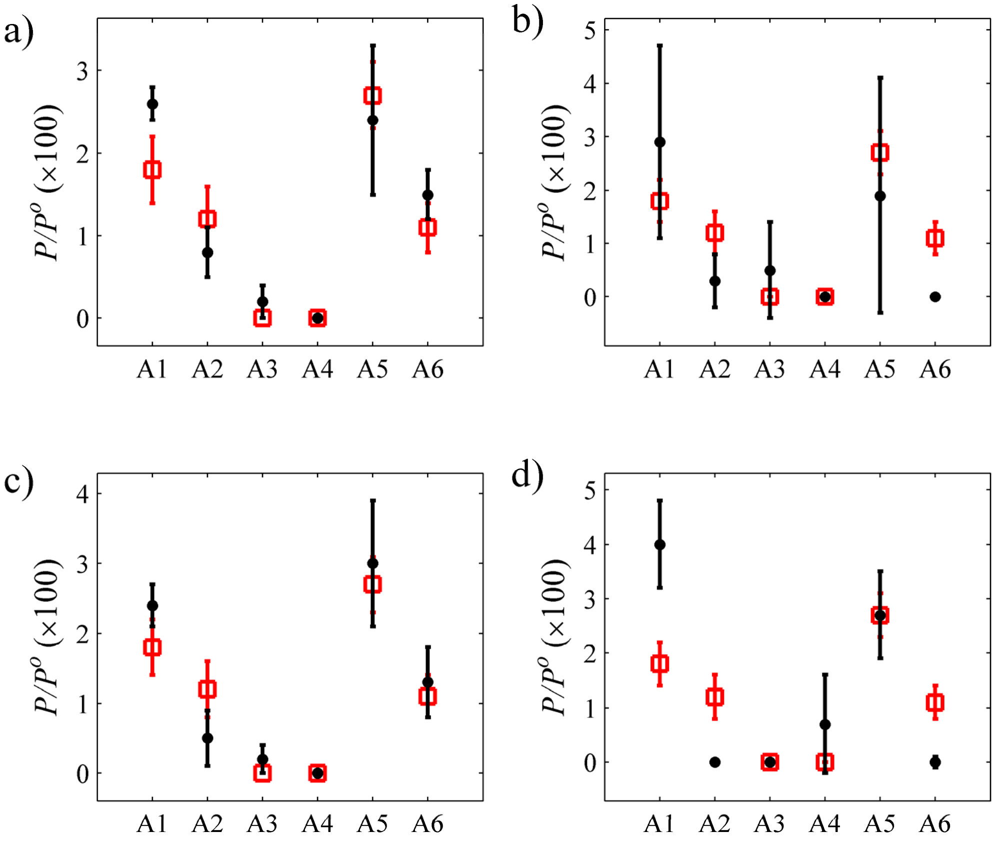


Figure 3.8: Comparison between various sensor/pattern recognition configurations in their analysis of 5-component mixture 12 (Tables 3.2-6). See Figure 3.5 caption for descriptions.

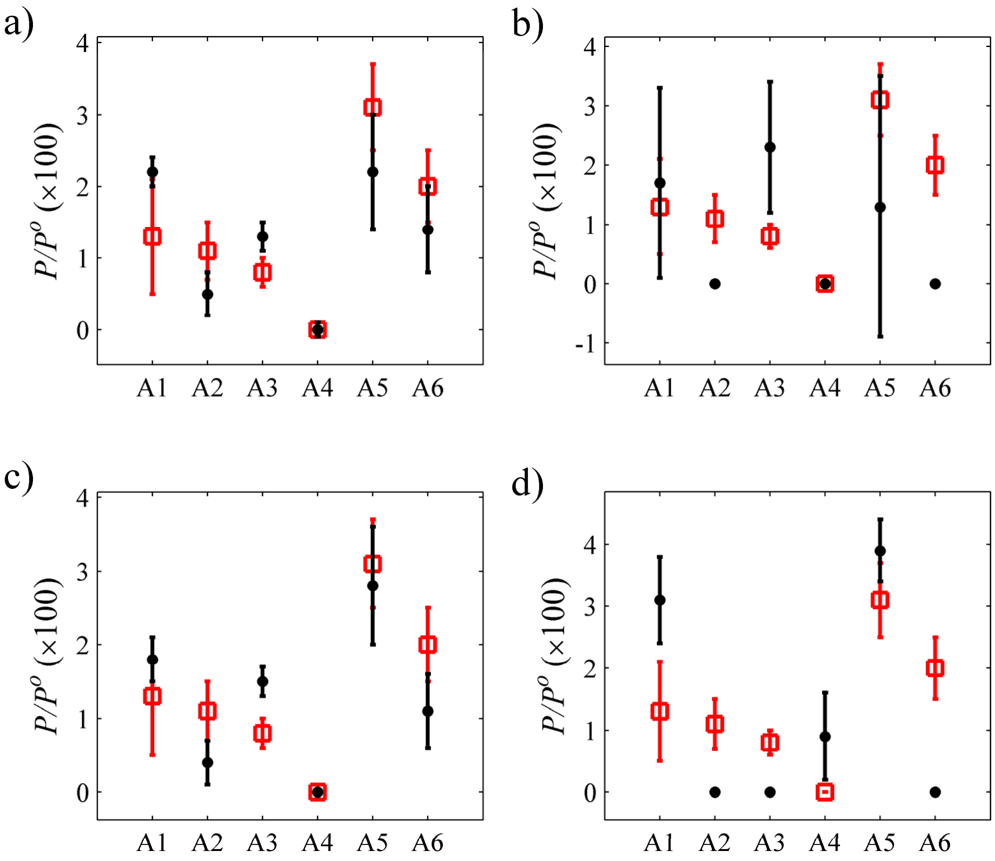
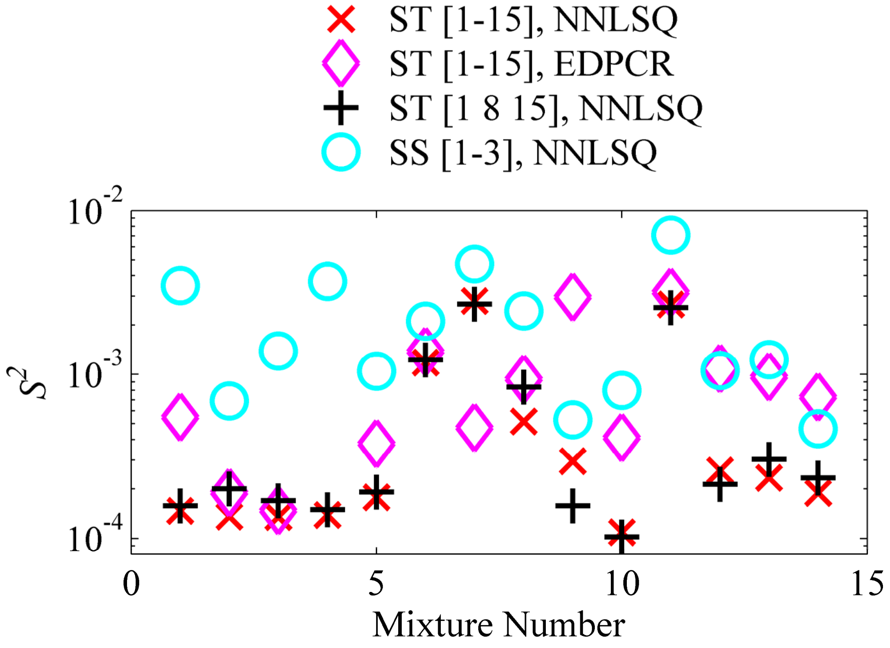


Figure 3.9: Residual squared error (S^2) observed between mean analyte estimation using various sensor/pattern recognition configurations and mean analyte estimation using GC-MS for each of the fourteen mixtures attempted in this work (last columns of Tables 3.3-6).



Chapter 4

Modeling of Spatiotemporal Response and the Definition of an Optimal Operational Regime for Mixture Analysis

4.1. Abstract

A model of the mass flow in a small volume chamber with small molecule/carbon black composite (SMCBC) chemiresistor sensors has been developed. The model was used to predict the spatiotemporal (ST) response of the sensors. The model used the convection-diffusion equation in the vapor space above each sensor film, the diffusion equation in the sensor film, and boundary conditions that forced analyte flow across the gas/sensor interface. Four SMCBC films were fabricated using 75% (wt) CB, consisting of lauric acid, propyl gallate, quinacrine dihydrochloride dihydrate, tetracosane, and dioctyl phthalate. Partition coefficients and sensor response slopes were measured for these films in response to hexane, benzene, octane, decane, methanol, ethanol, propanol, butanol, chloroform, ethyl acetate, and chlorobenzene, and the modeled responses were in good agreement with experimental responses. Pure vapors and vapor mixtures were modeled for a wide range of chamber geometries and vapor delivery flow rates. Pure vapor responses were then used to train the sensors. A set of 16 vapor mixtures that consisted of 2, 3, 4, and 5 components were then modeled assuming response additivity, and used to challenge the sensors. Using nonnegative least squares, the sensor arrays were tasked with the identification of each of the modeled mixtures. The sum of the squared residual between the modeled and calculated analyte components for all mixtures, S^2, was calculated and used to characterize the mixture analysis performance of the array. Two dimensionless Peclet number analogs, Pe_{yz} and Pe_{zz}^{-1}, were used to characterize the sensor response. Pe_{yz} was the ratio of the diffusion time from the top to bottom of the vapor channel (perpendicular to vapor flow) vs. the convection time in the direction of vapor flow. When the Pe_{yz} was large, vapor in the space overhead of the sensors was not well mixed and desired analyte depletion effects were not

observed. Pe_{zz}^{-1} was the ratio of the convection time vs. the diffusion time in the direction of vapor flow. When Pe_{zz}^{-1} was large, significant mass transport in the vapor space was due to diffusion rather than convection, and the sharp progression of the vapor front along the arrays was diminished. Chambers with values of the Peclet parameters below certain critical values were predicted to be optimum for ST sensor operation, and can be used to design a sensor cavity optimized for the analysis of mixtures.

4.2. Introduction

The widespread use of broadly responsive sensor arrays to identify and quantify vapors in the real world has been frustrated by their inability to correctly identify vapor mixtures. Sensor types include tin oxide sensors,[1-3] coated surface acoustic wave devices,[4-7] conducting organic polymers,[8-10] coated micromachined cantilevers,[11,12] dye-impregnated polymers coated onto optical fibers or beads,[13,14] polymer/carbon black composite chemiresistors,[15-17] and low-volatility small molecule/carbon black composite (SMCBC) chemiresistors.[18,19] For each of these sensor architectures, different materials, which each uniquely interact with the vapor phase, are employed to create a sensor array. Rather than using a traditional "lock-and-key" approach wherein a single sensor is required for each vapor, such array-based sensors are broadly responsive to many vapors. Each vapor creates a unique response pattern across the sensor array. Various forms of pattern recognition are then used to match the array response from an unknown to the responses from known vapors. These arrangements have been termed "artificial" or "electronic" noses due to their similarities to mammalian olfactory processes.[20,21]

Typical sensor studies to date have focused on obtaining rapid, steady-state responses between each sensor and the vapor stream being sampled,[2,13,16] through the use of relatively large-volume sensor chambers and exposure of the vapors to the sensors at high flow rates. Large-volume sensor chambers ensure that the partitioning of the vapor into the sensor array does not significantly change the concentration of analyte in the sampling stream. The high flow rate ensures that all sensors along the array are exposed to the sampling stream at essentially the same time. Studies utilizing computational fluid dynamics have sought to determine sensor/chamber designs and operating conditions which best provide this evenly distributed rapid response.[22,23]

The ability of a broadly responsive vapor sensor array to identify and/or discriminate between pure vapor species is critical. In this mode of operation, the array is first trained and then tested against the pure species of interest.[6,11,16,18,24] An alternative mode of sensor array implementation seeks to discriminate between different, complex mixtures. In this mode of operation, the array is trained for and then tested against different mixtures, however, decomposition of the mixture into constituent components is not performed. For example, sensor arrays are able to detect subtle differences in odors generated from different types of coffees,[25,26] beers,[10,27] hops,[28] wines,[29,30] vinegars,[31,32] and teas.[2,33] Under this traditional implementation, whether the array is tasked with pure specie or complex mixture identification, the array must first be trained against what it will encounter during testing. Many types of cross-responsive sensor types are highly capable under this mode of operation.

However, when the number of analytes or mixtures that need to be classified becomes large, extensive prior training is required. Whereas predictable and reproducible vapor presentations are easily obtainable in the laboratory, field operation presents the possibility of encountering a large number of mixtures, with the number growing exponentially as the pure specie sampling pool increases. The number of potential mixture types grows according to $n_m = 2^{n_p} - 1$, where n_p is the number of pure species present in the sampling library, and n_m is the number of mixture types comprised of unique analyte combinations, irrespective of analyte concentration. The number of mixture types one could encounter in the field based on a library containing only 20 pure vapor species is 10^6! These numbers are relevant only for the determination of whether a given component is present in a mixture. When a variable concentration is allowed for each vapor and its concentration must be determined, the number of vapor mixtures one may encounter effectively becomes infinite and beyond the means of any reasonable training protocol.

Broadly responsive sensor arrays often exhibit a response that is linear with vapor concentration, and additive with respect to multiple vapors.[4,18,34] This behavior should allow for training on only n_p pure components, and thereafter the sensor array should be able to identify combinations of the pure components at various concentrations. However, utilizing traditional rapid steady state responses, broadly responsive sensor arrays have failed to correctly identify multi-component mixtures.[35,36] On exposure to vapor mixtures, the arrays are quickly overwhelmed and unable to correctly assign analyte identity. Modeling of polymer-coated surface acoustic wave (SAW) sensor responses that were assumed to be linear with concentration and additive with respect to multiple vapors has exhibited difficulty in the correct identification of 3-component mixtures.[35]

The use of a flow stream that generates a time-dependent response profile may provide an additional dimensionality to each vapor's sensory response space. This approach could thus provide a method to improve the ability of broadly responsive sensor arrays to identify complex vapor mixtures. In mammals, the flow dynamics of sniffing, combined with differences in odor receptor binding affinities, are important for odor perception.[37,38] In humans, for example, each individual has a low- and high-flow rate nostril. This difference is caused by blood flow-induced nostril occlusions, which vary periodically with time. These varying flow patterns have been shown to affect odor perception.[38] A sensor chamber modeled after a canine nasal cavity, with sensors placed throughout the cavity, has been shown to provide enhanced discrimination in various classification tasks relative to sensors located only at the cavity inlet.[39]

To measure the composition of vapor mixtures, distributed chemical sensing has been employed with metal-oxide-semiconductor field-effect transistor (MOSFET) sensors. In this design, a metal catalyst was deposited along the interior of a sensor chamber, in the proximity of several MOSFET sensors. As select vapors progressed through the chamber, the vapors reacted to form other analytes that were not easily detected by the sensors. While the first sensor along the array responded to the original vapor stream, the last sensor along the array responded to an altered vapor stream. Significant improvements in the identification of binary mixtures were observed, and moderate levels of 3- and 4-component mixture identification were achieved for a library consisting of 4 pure vapors.[40-42] However, this approach was limited to vapors that selectively reacted at a catalyst to form vapors that were not readily detectable at the downstream sensors.

A more widely applicable approach has been employed with carbon black-based composite chemiresistive sensors. Using a low volume sensor chamber with linearly arranged sensor arrays, the sensor material has been shown to act similarly to a chromatographic stationary phase.[17,19] Vapors partitioned into the sensor material as dictated by the partition coefficient, K_{eq}, between the sensor and vapor phases. Strongly partitioning analyte vapors were significantly sorbed by the sensor phase, slowing the progression of analyte along each sensor array. While initial sensors along the array responded to the concentrations of the incoming vapor stream, latter sensors exhibited delayed responses due to significant analyte uptake.[17,19]

The progress of each vapor along the sensor array was dictated to first order by the flow rate, chamber geometry, and mass uptake capacity of the sensor film. This behavior created a space and time dependent, or spatiotemporal (ST), sensor response. This ST response increased the dimensionality of response space. Recent work has shown that the training of such sensor arrays toward pure vapors allowed robust identification of mixtures containing up to 5 components.[19]

This work defines an optimized operational regime for capturing an ST response using SMCBC sensors. A model for the ST sensor response was generated and was implemented for sensor geometries and vapor flow rates spanning several orders of magnitude. Simulating the exposure and identification of several vapor mixtures for the various modeled conditions, two dimensionless numbers were used to correlate the mixture analysis ability to relationships that involved the sensor chamber geometry, vapor delivery flow rate, and appropriate physical properties. These dimensionless numbers allow for the definition and fabrication of sensor arrays optimized for ST analysis and vapor mixture identification.

4.3. Experimental

4.3.1. Materials

The insulating materials for the fabrication of the sensor films consisted of tetracosane (99%), lauric acid (99.5%), and dioctyl phthalate (DP), purchased from Aldrich; as well as propyl gallate (98%) and quinacrine dihydrochloride dihydrate (97%), purchased from Acros Organics. Reagent grade toluene, tetrahydrofuran, and chloroform, received from Aldrich, were used as solvents in the sensor suspensions. Hexane (Hx, 95%), benzene (Bz, 99%), octane (Oc, 98%), decane (Dc, 99%), methanol (MeOH, 99.8%), ethanol (EtOH, 95%), 1-propanol (PrOH, 99.5%), 1-butanol (BuOH, 99.9%), chloroform (Cf, 99.8%), ethyl acetate (EA, 99.5%), chlorobenzene (Cb, 99%), and acetone (99.9%) were purchased from Aldrich and used to generate vapors for delivery to the sensor arrays. Black Pearls 2000, a furnace carbon black (CB) material donated by Cabot Co. (Billerica, MA), was used as the conductive phase in the sensor composites. All materials were used as received.

4.3.2. Detector Fabrication: Chemiresistive and Quartz-Crystal Microbalance Vapor Detectors

Table 4.1 lists the four suspensions, each comprised of a non-conductive sorption phase and a conductive CB phase, used to fabricate the different linear sensor arrays (LSA) used in this work. First, the non-conductive (non-CB) sensor material(s) were placed in ~ 60 mL of solvent and sonicated for > 10 min. CB was added to this solution and sonicated for > 30 min to produce a well-dispersed suspension.[18,19] Dioctyl phthalate was used as a component in two of the sensor films, to serve as a plasticizer (Table 4.1). These four sensor materials, when employed under ST conditions, have demonstrated the ability to correctly identify complex vapor mixtures containing up to 5 components with training on only the pure vapors.[19]

Standard 7.5 cm × 2.5 cm microscope slides were cleaned with acetone and methanol, and 7.5 cm of 0.1 cm wide drafting tape was placed along the center of the length of each slide. Detector substrates were fabricated by evaporating 30 nm of Cr followed by 70 nm of Au onto the microscope slides. After evaporation, the tape was removed and the glass slides were cut into 1.0 cm × 2.5 cm substrates. An airbrush (Iwata, Inc.) was used to spray[16,24] one of the sensor suspensions across the 0.1 cm substrate gap until the resistance across the electrode measured 1-3 kΩ. This process was performed for each of the four sensor suspensions listed in Table 4.1 to generate the four chemiresistive vapor sensors used.

Quartz Crystal Microbalance (QCM) crystals (10 MHz, blank diameter 13.6 mm) were obtained from International Crystal Manufacturing (ICM, Oklahoma City, OK). For each sensor type, the resonant frequency of the uncoated substrate, $f_{uncoated}$, was recorded prior to sensor deposition. The resonant frequency of the coated substrate, f_{coated}, was recorded after the sensor film was sprayed using an airbrush onto one side of the QCM substrate. The frequency shift due to sensor deposition, which gave a measure of film mass deposited, was calculated as $\Delta f_{sens} = f_{coated} - f_{uncoated}$.[17,43]

These sensors often experience their most significant sensor drift immediately after fabrication.[44] Thus, before the calculation of sensor partition coefficients K_{eq} and sensor response slopes M_R, both the chemiresistive and QCM vapor detectors were continuously and randomly exposed, over a 24-hour period, to the test analytes listed in Table 4.2. After this period, f_{coated} was recorded and Δf_{sens} calculated. Two weeks elapsed between the initial spraying of the chemiresistive and QCM detectors and their exposures to the test analytes for the calculation of partition coefficients and sensor response slopes.

4.3.3. Vapor Generation and Delivery

An automated flow system,[16,24] controlled with LabVIEW software, was used to deliver pulses of diluted streams of solvent vapor to the chemiresistive and QCM vapor detectors. The carrier gas (background stream) was obtained from the house compressed air source controlled with a mass flow controller (MFC) (UNIT) that could vary the flow from 0.3 to 28 L min^{-1}. For generation of saturated analyte vapors, a foreground stream of carrier gas was bubbled via a porous glass frit through a 220 mL bubbler filled with the desired analyte, controlled by a MFC that could vary the flow from 3 to 280 mL min^{-1}. The height of the analyte in the bubblers before and after each set of exposures did not change (bubbler solvent height ~ 35 cm). Teflon tubing and stainless steel Swagelok fittings were used throughout the system.

Analyte exposures consisted of a 70 s pre-exposure of background air, followed by 80 s of exposure to a test analyte, followed by a 70 s exposure to background air. Exposures were generated by combining saturated analyte vapor with background air at the volumetric flow rates needed to provide a total flow of 4 L min^{-1} and the desired fractional vapor pressure. Volumetric flow rates of background, saturated, and mixed vapor streams were all monitored with flow meters (Gilmont) to ensure accuracy in the delivered concentrations and flow rates.

The vapor concentrations for these types of sensors are typically reported in terms of the fractional vapor pressure, P/P^o, where P is the partial pressure and P^o is the saturated vapor pressure of the analyte of interest.[34,45] Analytes were exposed to the chemiresistor/QCM sensor pair at $P/P^o = 0.010$, 0.030, 0.050 and 0.080. These values corresponded to 1, 3, 5 and 8 parts of saturated analyte vapor combined with 99, 97, 95 and 92 parts of background air, respectively. Each vapor was exposed a total of 20 times, consisting of 5 exposures at each of the four chosen concentrations.

A chemiresistive sensor and QCM substrate coated with one of the four sensor materials (Table 4.1) were placed in an aluminum chamber with internal dimensions of 8.0 x 5.0 x 4.0 cm and exposed to one set of analyte vapors: Hx, Oc, Dc and EtOH (Table 4.2). This process was repeated for each of the four sensor materials listed in Table 4.1. Each of the sensors were then similarly exposed to a second and third set of analyte vapors: Bz, MeOH, PrOH and BuOH; followed by Cf, EA and Cb, respectively (Table 4.2). Five repetitions per sensor were performed for each vapor exposed at a given concentration. Each set of analyte exposures lasted roughly 6 h, with two or three exposure sets occurring each day. Exposures were presented in random order with respect to analyte identity and the delivered analyte concentration. All response data were collected within a 2 week period.

4.3.4. Measurements and Data Processing

The direct current resistances of the chemiresistors were measured using a Keithley 2002 multimeter and a Keithley 7001 multiplexer. The sensor substrate was connected to the multiplexer through a shielded, twister pair cable. QCM resonant frequencies were obtained using a Hewlett Packard 53181A frequency counter. While recording data, measurements alternated between resistance measurements of the chemiresistor and frequency measurements of the coated QCM. 5 s passed between each resistance/frequency measurement cycle.

Raw response data for both the chemiresistor and QCM detectors were first baseline-corrected to account for sensor drift during an exposure. A line of slope m was fit to > 6 points, spanning > 30 s, immediately prior to each vapor delivery and detector response. If the confidence intervals of the calculated slope m did not include zero at the 95% confidence level, a line of slope m was extrapolated from the point immediately prior to vapor arrival and subtracted from the measured resistance readings. Roughly 10% of the measured responses required this baseline drift correction.

The shift in the resonant frequency of a QCM on exposure to a vapor was required for the calculation of partition coefficients.[17,43] This shift was calculated as $\Delta f_{vapor} = f_{vapor} - f_{coated}$, where f_{vapor} was the steady state resonant frequency of the QCM on exposure to a given vapor. The value of f_{vapor} was obtained as the average of ≥ 3 data points starting 30 s after the vapor exposure was initiated. The value of f_{coated} was obtained as the average of the 6 points recorded immediately prior to vapor delivery.

The typical response descriptor for these types of chemiresistive vapor sensors is $\Delta R_{ss}/R_b$, where R_b is the pre-exposure baseline resistance as described above and ΔR_{ss} is the difference between the steady state response resistance and the baseline resistance, $R_{ss} - R_b$. For the calculation of the chemiresistive sensor responses to each of the vapors, R_{ss} was taken to be the resistance averaged over ≥ 3 data points starting 30 s after the vapor exposure was initiated.

4.3.5. Partition Coefficient and Sensor Response Slope Calculation

For each set of QCM/vapor exposures, the changes in frequency during analyte exposure, Δf_{vapor}, were plotted against the concentration (in parts per thousand, ppth) of the presented vapor. For example, methanol presented at $P/P^o = 0.080$, with $P^o = 108$ mmHg at room temperature (T = 22°C), corresponded to 1000 ppth x (0.080 x 108 mmHg) / 760 mmHg, or 11.4 ppth methanol. The QCM response slope m_{QCM} (Hz/ppth) was calculated from the 20 Δf_{vapor} data points using a linear least-squares fit with a zero intercept. This slope was converted into a partition coefficient using the Sauerbrey equation:[17,43]

$$K_{eq} \equiv \frac{C_s}{C_v} = \frac{10^6 \rho R T m_{QCM}}{M_w \Delta f_{sens} P_{atm}}, \qquad (1)$$

where C_s and C_v were the concentrations of analyte in the sensor film and vapor phases, respectively, ρ was the weight-averaged density (g ml^{-1}) of the sensor, R was the ideal gas constant (L atm mol^{-1} K^{-1}), T was the temperature (K), M_w was the molecular weight of the analyte (g mol^{-1}), Δf_{sens} was the frequency shift due to the sensor film deposition, P_{atm} was the atmospheric pressure (1 atm), and 10^6 was a conversion factor (ppth cm^3 L^{-1}).

The chemiresistive sensor response slopes M_R were calculated by plotting the sensor response, $\Delta R_{ss}/R_b$, vs vapor concentrations (ppth) for each set of chemiresitor/vapor exposures. The slope was calculated using a linear least-squares fit with a zero intercept.

4.3.6. Optical Determination of the Vapor Delivery Profile

To determine the profile of the delivered vapor front during experimental ST analysis, the vapor stream was monitored by observing its optical absorbance in the UV. The vapor delivery configuration used in previous experimental ST work was employed, to the extent possible, to determine the vapor front profile.[19] The flow-splitting gasket to the sensor array chambers[19] was removed and the flow was allowed to pass through a rectangular quartz cuvette (model 3-3.45-Q-3, Starna, Atascadero, CA) that had the bottom removed, shown in Figure 4.1.[46] The vapor delivery end cap was connected to the quartz cuvette with Teflon tubing, and sealed with Teflon tape. The cuvette wall thickness was 0.125 cm, with internal dimensions of 0.3 × 0.3 × 2.5 cm (W × H × L) and an internal cross-sectional area of 9 mm^2. The cuvette was housed in an aluminum body, which contained a 1.0 cm × 0.3 cm hole cut perpendicular to the cuvette, allowing for the optical path to be monitored in a direction perpendicular to the vapor flow.[46] The cross-sectional area of delivery tubing was 2 mm^2. The cross-sectional area in the sensor array delivery chamber, prior to splitting a single vapor stream into two, was 6 mm^2.[19] The cross-sectional area of the vapor pathway over each sensor array was 1 mm^2.

A low-pressure Hg lamp (model UVG-4, UVP Inc., Upland, CA) with a low-pass filter ($\lambda <$ 253.7 nm) was used. To minimize the lamp output drift, the lamp was powered by 5 lantern batteries wired in parallel. The lamp was mounted against one side of the hole in the aluminum housing. On the opposite side, a UV-sensitive silicon photodiode (UV50, UDT Inc.) was pressed against the exit hole. The unbiased photodiode was connected with a BNC cable to a Hewlett Packard data acquisition unit, model HP 34970A ($1 \times 10^6\ \Omega$ input impedance), to monitor the voltage. The photodiode voltage was recorded every 0.7 s.

Analyte vapor was generated at $P/P^o = 0.050$ with a flow rate of 150 ml min^{-1} and delivered to the optical monitoring system.[19] The analytes listed in Table 4.2 were delivered under these conditions, however their optical absorbance yielded a photodiode response with a low signal-to-noise ratio. Because of its strong UV absorbance, acetone was thus used to generate the analyte vapor. The diffusivity of acetone in air ($D_v = 0.083$ cm^2 s^{-1}) was within the range of the analytes used in this work, such as hexane ($D_v = 0.068$ cm^2 s^{-1}) and methanol ($D_v = 0.115$ cm^2 s^{-1}).[47]

The voltage response of the photodiode was baseline-corrected to subtract drift. The adjusted voltage was converted to an absorbance using $A = -\log(V/V_o)$, where V_o and V are the voltages produced in the absence of analyte vapor, and at time t, respectively. The absorbance was proportional to the concentration of analyte present in the quartz cuvette. The vapor response section of the absorbance was fit to:

$$A = a_1\left(1 - \frac{1}{1 + k_t(t + a_2)}\right), \tag{2}$$

where a_1 was the saturation absorbance of the vapor delivery front, a_2 allowed for a time offset due to the vapor delivery delay, and k_t was the vapor delivery rate constant (s^{-1}).

4.4. Results

4.4.1. Partition Coefficient and Sensor Response Slope Calculation

Figure 4.2 displays representative steady-state chemiresistive responses and QCM resonant frequency shifts for thin films of LSA$_{TC/DP}$ (Table 4.1). The analytes represent distinct alkane, alcohol, and ester chemical classes, thus encompassing a variety of sensor/analyte interactions. Sensor responses for these types of sensors are typically linear at low concentrations, with deviations from linearity at higher concentrations.[15,34] The responses to each of the vapors were well-grouped and generally well-fit by a straight line. The partition coefficients K_{eq} and the chemiresistive response slopes M_R, as well as the R^2 values for each of the respective slope calculations, are given in Table 4.3 for all of the sensors and analytes studied. The standard deviation in the $\Delta R_{ss}/R_b$ responses was calculated to be $\sigma_{resp} = 2 \times 10^{-5}$.

4.4.2. Optical Determination of the Vapor Delivery Profile

Figure 4.3 displays the data obtained in the various steps employed to determine the vapor delivery profile. Figure 4.3a shows the raw photodiode voltage response recorded while acetone progressed through the modified sensor chamber, with the highlighted points used to perform the baseline correction. Figures 4.3b-c show the baseline-corrected voltage response and the absorbance response, respectively. The arrival profile of acetone to the modified chamber was well fit to eq (2) with values of $a_1 = 0.00041$, $a_2 = -0.076$ s, and $k_t = 0.32$ s^{-1} (Figure 4.3c).

4.5. Modeling Formulation

4.5.1. Chamber and Sensor Mass Uptake Modeling

4.5.1.1. Model Description

Figure 4.4 depicts the sensor chamber and the main equations used for chamber modeling. For all modeled cases, $W >> H_v$, so the chamber was modeled in 2 dimensions, y and z, with edge effects along the width of the vapor channel (x-direction, Figure 4.4) neglected. Appendix 4.9.1 describes the detailed development of the mass uptake model that was used.

Briefly, in the bulk vapor phase, a transient two-dimensional convection-diffusion equation was employed:[47-49]

$$\frac{\partial C_v}{\partial t} + V_z(y)\frac{\partial C_v}{\partial z} = D_v\left(\frac{\partial^2 C_v}{\partial y^2} + \frac{\partial^2 C_v}{\partial z^2}\right), \tag{3}$$

where $C_v(y,z,t)$ was the vapor-phase concentration of the analyte of interest, $V_z(y)$ was the laminar velocity profile as a function of chamber height (Appendix 4.9.1), and D_v was the diffusivity of the analyte vapor of interest in air. For all modeled geometry/flow rate combinations, the Reynolds number, Re, was < 100, well below the Re value of 2000 that marks the transition from laminar to turbulent flow (Appendix 4.9.1).[47,48]

In the bulk sensor phase, a transient two-dimensional diffusion equation was used:[47-49]

$$\frac{\partial C_s}{\partial t} = D_s\left(\frac{\partial^2 C_s}{\partial y^2} + \frac{\partial^2 C_s}{\partial z^2}\right), \tag{4}$$

where $C_s(y,z,t)$ was the concentration of analyte in the vapor and D_s the diffusivity of analyte in the sensor material. For each of the sensor/vapor combinations, the value of D_s used was 10^{-7} cm^2 s^{-1}. This value fell between experimental diffusivities of gases diffusing through liquid and solid states, and was of the same order of magnitude of gases diffusing through polymers.[48] Increases or decreases in D_s by an order of magnitude did not change the observed mass uptake profiles. Thus, the model was mass-transfer limited by transport through the vapor phase, rather than by the rate of diffusion through the sensor film.

Appropriate boundary conditions were implemented at all phase boundaries, and are discussed in Appendix 4.9.1. Additionally, the model was scaled to place all important lengths (H_v, H_s, and L) on the same scale (Appendix 4.9.2). The concentrations of test analyte in the vapor and sensor material, $C_v(y,z,t)$ and $C_s(y,z,t)$, respectively, were calculated using COMSOL Multiphysics, a commercial finite-element analysis software package, coupled with MATLAB.

4.5.1.2. Model Validation Implementation

Mass uptake profiles were generated for previously reported experimental conditions: long, thin vapor flow pathways of dimensions 75 × 3.5 × 0.34 mm ($L \times W \times H_v$), and sensor films having H_s = 1-3 μm, and vapor delivered at 19 ml min^{-1}.[19] The optically determined absorbance delivery front (Figure 4.3) was used to generate an analyte concentration profile at the inlet of the chamber:

$$C_{v,inlet}(t) = C_{v,o}\left(1 - \frac{1}{1 + k_t t}\right), \tag{5}$$

where $C_{v,o}$ is the vapor concentration for the given exposure being modeled, k_t is the vapor delivery rate constant described in eq (2), and $C_{v,inlet}(t)$ is the time-dependent concentration delivery profile for $t \geq 0$.

4.5.1.3. Optimization Implementation

For chamber optimization, mass uptake profiles were generated with a wide range of modeled conditions. Tables 4.4 and 4.5 list L-W-H_v geometry combinations and flow rates modeled for various macro- and micro-dimensioned ST chamber conditions, respectively. Each of the L-W-H_v geometries was combined with each of the within-table flow rates. Thus, a total of 54 (9 geometries × 6 flow rates) and 78 (13 geometries × 6 flow rates) modeled conditions were possible from Table 4.4 and 4.5, respectively. Of these possibilities, only those that generated negligible pressure drops along the array were modeled (Appendix 4.9.1). For optimization modeling, a step function was used for the inlet boundary condition such that $C_{v,inlet}(t) = 0$ for $t < 0$ and $C_{v,inlet}(t) = C_{v,o}$ for $t \geq 0$.

ST analysis required the use of the array response to a vapor phase changing with time. To ensure that ST analysis consistently focused in on this transient regime for all modeled conditions, a balance was performed on the number of moles of analyte required to bring the sensor film to equilibrium with the sampled vapor phase (Appendix 4.9.1). This mole balance provided $t_{sat,median}$, the median time required to saturate the sensor arrays with the vapors for each modeled condition. The model was solved for 200 evenly spaced time steps between $t = 0$ and $t = 2.4 \times t_{sat,median}$.

4.5.2. Sensor Response Modeling

4.5.2.1. Sensor Response Generation

Sensor responses were generated by converting modeled sensor concentrations to sensor signals, $S(t)$. It has been shown that when probing the response of polymer-carbon black composite sensors on the time scale of diffusion through the sensor film, the resistance can be approximated by the sum of the resistances, wired in parallel, across the thickness of the film.[46] For all cases modeled herein, the time scale of the array response was several orders of magnitude slower than the time scale for diffusion through the sensor film (Appendix 4.9.1). This caused the analyte concentration in the sensor film, at any point along the z-direction, to be consistent across all sensor depths in the y-direction (Figure 4.4), which in turn caused the sum of the parallel resistances to be equal to each of the individual resistances. Thus, the analyte concentrations at the bottom layer of the sensor film, i.e. the bottom node of the chamber $C_s(t, y = 0, z)$ (Figure 4.4), were converted to response signals $S(t,z)$ according to $S(t,z) = C_s(t, y = 0, z) \times M_R$, where M_R was the sensor response slope for the analyte of interest given in Table 4.3. For each sensor array, $S(t)$ responses were collected from 15 nodes that were evenly distributed along the 200 node length.

To generate representative pure-vapor responses, Gaussian noise with an amplitude proportional to the standard deviation in the $\Delta R_{ss}/R_b$ responses, σ_{resp}, was superimposed on the modeled pure-vapor responses. To generate responses to vapor mixtures, responses were generated for each of the pure vapor constituents at their representative concentration and were summed, with the addition of noise proportional to σ_{resp}.

4.5.2.2. Response Extraction

To fully capture the transient response at various times along the array, $S(t,z)$ was extracted at four different times for each sensor.[19] These times were chosen as $t_{xt} = (0.3, 0.6, 1.8 \text{ and } 2.4) \times t_{sat,median}$. To describe the total sensor array response to a single analyte, responses from all four sensor arrays were used (Table 4.1). Each modeled exposure was described by 4 arrays × 15 sensors/array × 4 times/sensor, or 240 response descriptors.

4.5.3. Mixture Analysis

4.5.3.1. Mixture Generation

Table 4.6 lists the simulated mixtures, in terms of the fractional vapor pressure of each analyte. 16 different mixtures were used, ranging from 2 to 5 vapors per mixture. For each mixture,

P_{mix} / P^o_{mix} was less than 0.060, where P_{mix} was the total pressure of the mixture presented and P^o_{mix} was the saturated vapor pressure exerted by the mixture. This value was within the range where response linearity and additivity are typically observed[15,34] Mixture identification using the ST approach has been demonstrated under these circumstances.[19]

The training data were generated by simulating ten responses to each of the pure vapors. The testing data were generated by simulating twenty responses to each of the 16 modeled mixtures. For each pure vapor, the 10 training responses were averaged to create an analyte signature library. Each testing response was then fit to the training library using the nonnegative least-squares algorithm.[50] This approach produced estimates for the analyte concentration (fractional vapor pressure) of each of the test analytes.

4.5.3.2. Chamber Performance

For each single exposure, the difference between the estimated and actual analyte concentrations was squared and summed across all analytes to provide a measure of the residual error, S^2_{single}. S^2_{single} was summed across all twenty repetitions per mixture, and across all 16 mixtures, to give a single measure of mixture performance, S^2, for each of the modeled flow and geometric conditions. For each geometry/flow rate combination, response modeling and the calculation of S^2 was repeated 8 times to test the model reproducibility.

4.6. Modeling Results

4.6.1. Model Validation

Figure 4.5 displays the experimental[19] and modeled responses of all of the sensors along a 15-sensor array LSA_{PG} (Table 4.1) during exposure to pure hexane at $P/P^o = 0.050$. Experimental responses were obtained with sensor arrays ranging from 1-3 μm in thickness.[19] For validation, H_s was varied within this range: modeled results were obtained for various $H_s = 1 - 3$ μm and compared against experimental results, and reasonable agreement was found at $H_s = 2.6$ μm in thickness. For visualization, the responses were normalized to each sensor's response at $t = 140$ s. In both cases, hexane was not depleted, and all 15 sensors rapidly responded to the vapor stream.

Figure 4.6 displays the experimental[19] and modeled ($H_s = 2.6$ μm) responses of the same 15-sensor array made from LSA_{PG} (Table 4.1) during exposure to decane at $P/P^o = 0.050$. Decane, with a vapor pressure two orders of magnitude lower than hexane, possessed a partition coefficient approximately two orders of magnitude greater then hexane (Table 4.3). For both analyses, the response of the first sensor along the 15-sensor array rose rapidly upon exposure to the delivered vapor stream. In contrast, subsequent sensors along the array showed a significantly delayed response to decane, as it slowly progressed along the array.

Differences between experimental and modeled responses of LSA_{PG} to decane were evident for the 3rd sensor along the array, and became more significant for sensors positioned further downstream on the vapor flow path. Experimentally, all 15 of the sensors showed a sharp, near-vertical response to the decane vapor front. In contrast, the modeled sensor responses rose progressively more gradually toward the end of the array. This behavior was indicative of the modeled laminar flow. Experimentally, roughness along the sensor films and channel walls would produce localized eddies that would enhance the diffusion of analyte vapor. This would provide better mixing in the y-direction (Figure 4.4), leading to more complete vapor stream depletion, and correspondingly sharper response curves.

Figure 4.7 displays experimental[19] and modeled responses for exposure of $LSA_{TC/DP}$ to pure ethyl acetate, pure decane, and a mixture of the two, respectively. Only the responses of sensor 1 and 9 are displayed. The black points depict the experimental response data, and the dark solid lines depict the modeled responses for a film of thickness H_s = 1.5 μm and for the optically observed vapor delivery profile (k_t = 0.32 s^{-1}). Using the optically observed vapor delivery profile, the modeled results for the 1st sensor exhibited a more rapid response than that observed experimentally. This behavior suggested that the modifications made to the vapor delivery path for optical detection affected the vapor delivery properties of the chamber.

To better model the response of the first sensor to the analyte delivery profile, the value of k_t was modified to 0.18 s^{-1}. This procedure produced a more gradual vapor delivery front in the modeled response (Figure 4.3c). Figures 4.7a-c show the modeled responses for this altered delivery profile. Figures 4.5c and 4.6c display the modeled responses of LSA_{PG} (Table 4.1) to hexane and decane, respectively, using this modified vapor delivery rate constant. In all cases, better agreement was obtained with the altered vapor delivery profile. Thus, vapor delivery to the sensor chamber exhibited a slower rise time to the final concentration than the chamber modified for optical detection. In Figures 4.7b-c, sensor 9 again displayed a near vertical initial response shape, indicative of a well-mixed turbulent flow stream that allowed for more analyte depletion as the front progressed along the array.

4.6.2. Chamber Characterization

Parameters that characterized the channel in terms of dimensions and flow rate were needed to summarize the modeling. The Peclet number, *Pe*, is a ratio of the rate of transport due to convection relative to the transport due to diffusion.[48] The standard Peclet number captures the competition between these processes along the same direction of flow. To fully describe transport within the channel, an additional, alternative Peclet number was needed to address diffusion in the direction perpendicular to the direction of flow. Hence, two Peclet numbers were defined in the modeling.

4.6.2.1. *Pe_{yz}*: Latitudinal Diffusion vs. Longitudinal Convection

Pe_{yz} was the ratio of the time for material to diffuse vertically (y-direction, top to bottom, Figure 4.4) to the time to convect horizontally (z-direction, inlet to outlet, Figure 4.4) along the vapor channel:

$$Pe_{yz} = \frac{t_{diffusion}^{top \to bottom}}{t_{convection}^{inlet \to outlet}} = \frac{H_v^2 / D_v}{L / U} = \frac{UH_v^2}{D_v L} = \frac{H_v Q}{D_v LW} \,. \tag{6}$$

A large value of Pe_{yz} indicated that the vapor stream did not significantly mix vertically and the concentration at the top of the stream became different from that at the bottom, as the front progressed along the channel. As a result, the analyte concentration at the bottom of the vapor stream was determined primarily by the slow vertical diffusion, which was approximately constant. The vapor stream would not get significantly depleted, and only small ST differences should be observed between the various sensors along the array.

A small value of Pe_{yz} indicated that analyte diffused quickly from the top to bottom of the chamber. This situation provided a well-mixed vapor front with negligible concentration differences in the y-direction, and allowed for significant depletion of the vapor stream and hence should produce a large ST response.

4.6.2.2. *Pe_{zz}^{-1}*: Longitudinal Convection vs. Longitudinal Diffusion

Pe_{zz} denotes the traditional Peclet number, and is the ratio of the diffusion vs. convection time in the direction of vapor flow. Its inverse is given by:

$$Pe_{zz}^{-1} = \frac{t_{convection}^{inlet \to outlet}}{t_{diffusion}^{inlet \to outlet}} = \frac{L / U}{L^2 / D_v} = \frac{D_v}{UL} = \frac{D_v W H_v}{QL} \,. \tag{7}$$

When this ratio was unity, material progressed from the inlet to the outlet due to equal contributions from convection and diffusion. This situation still allowed for mass depletion along the sensor array, assuming Pe_{yz} was sufficiently small. However, the sharp, well-separated responses of individual sensors would give way to more gradual sensor responses with significant overlap. A small Pe_{zz}^{-1} value indicated that convection dominated vapor transport. Hence, a sharp vapor front was maintained as analyte progressed along the chamber, and well-separated responses should be observed along the array.

4.6.3. Chamber Optimization

Figure 4.8 displays log-log plots for simulated film thicknesses of 0.1 μm. The residual error, S^2, was plotted against Pe_{yz}, with various thresholds chosen to display Pe_{zz}^{-1} effects: Pe_{zz}^{-1} values above the chosen threshold were displayed as "×." The "macro" and "micro" labeled points designate the modeled results from Tables 4.4 and 4.5, respectively. Figure 4.8a displays the residual error vs. Pe_{yz} with a Pe_{zz}^{-1} threshold of 0.1. Only a single modeled condition fell outside of the Pe_{zz}^{-1} threshold. Figures 4.8b and 4.8c display results for Pe_{zz}^{-1} thresholds of 10^{-2} and 5×10^{-3}, respectively. As the threshold value decreased, more modeled conditions fell outside of the threshold and were therefore disqualified.

From these figures, a region can be described that provides an optimal ST mixture analysis. For sensor film thicknesses of 0.1 μm, optimal performance was achieved when $Pe_{yz} < 10^{-3}$ and $Pe_{zz}^{-1} < 5\times10^{-3}$. At the critical value of $Pe_{yz} = 10^{-3}$, sufficient mass uptake was achieved for optimal mixture analysis. Beyond this value, for $Pe_{yz} < 10^{-3}$, additional mass uptake did not serve to improve the ST analysis of mixtures.

The above criteria were obtained for sensor film thicknesses of 0.1 μm. Pe_{yz}, Pe_{zz}^{-1}, and their associated criteria do not, however, directly take into account varying film thickness. Thicker films would provide more mass uptake and accordingly longer response times. These longer times may cause enhanced vapor front broadening effects, which could inhibit the ability of the sensor arrays to correctly identify mixtures. The same analysis was therefore performed for a modeled film thickness of 1 μm. As shown in Figures 4.9a-c, the region of optimal ST performance was not significantly different from that defined for 0.1 μm films, $Pe_{yz} < 10^{-3}$ and $Pe_{zz}^{-1} < 5\times10^{-3}$.

4.7. Discussion

4.7.1. General Features of the Modeling

This work has identified an optimized ST operational regime for small molecule/carbon black composite vapor sensors of various thicknesses. This defined region allows for the design of sensor cavities that are optimized for ST detection. Two dimensionless numbers were used to define this regime: the ratio of the latitudinal diffusion time to the longitudinal convection time, Pe_{yz}; as well as the ratio of the longitudinal convection time to the diffusion time, Pe_{zz}^{-1}. The Pe_{yz} and Pe_{zz}^{-1} values required for optimal ST performance and the optimal mixture analysis appear to be consistent for the two sensor film thicknesses modeled, 0.1 and 1 μm. For both cases, optimal mixture analysis was achieved when $Pe_{yz} < 10^{-3}$ and $Pe_{zz}^{-1} < 5\times10^{-3}$. This behavior suggests that the given criteria also define a region of optimal ST mixture analysis for similar polymer/carbon black composite sensors (20-40% CB), or other sorption-based cross-responsive sensor modality, as long as their sorptive capacities are not too different (i.e., several orders of magnitude) from those observed here. The only requirement is that the partition coefficients of the sensor and vapors are known, so that the time frame required to capture the ST response can correctly be calculated (Appendix 4.9.1). Additionally, the diffusion time through the sensor material must be much less than the transient time frame needed to capture the ST response.

4.7.2. Film Thickness

Films of two thicknesses, encompassing the range of film thickness typically employed for these sensor types, were modeled. Thinner films (0.1 μm) are useful due to their more rapid response times. Alternatively, thicker films (1 μm or thicker) produce slower response times, that can be more easily captured with slower data acquisition systems. Previous ST experimental work has employed films in the 1-3 μm range to ensure that the transient responses were sufficiently captured with the available data acquisition hardware.[19] It is important that the data acquisition instrumentation be able to respond significantly faster than the transient sensor response time. For sensors of a given film thickness, the vapor delivery flow rate could be modulated to generate faster or slower responses, as desired. The only requirement would be that Pe_{yz} and Pe_{zz}^{-1} should have values for the system of interest below their respective optimization thresholds.

4.7.3. Macro- vs. Micro- Scaled Chambers

Two sets of sensor chambers were modeled: one with macroscopic dimensions, and one with microscopic dimensions (except in length). In most cases, micro-scaled chambers fit within the desired Pe_{yz} region, and were therefore favorable (Figures 4.7-9). This was due to the smaller height, relative to the sensor array length of the vapor channel. For these chambers, lower flow rates were required due to the higher pressure drops generated during flow along the channel (Appendix 4.9.1). Because of the lower flow rates used, the Pe_{zz}^{-1} criterion became more critical (eq (7)).

For micro-scaled chamber fabrication, the process must not include elastomeric-based materials (e.g., polymers) typically used in microfabrication processes.[51] Such materials readily sorb vapors and thus could significantly deplete the vapor stream of analyte. Instead, hard, non-sorbing materials are required. One such approach would include an etched-silicon flow pathway, bound to a glass surface via anodic bonding.[52]

4.7.4. Major Assumptions

The ST optimization criteria described herein are entirely dependent on the validity of the model. Major assumptions in the model include: 1) sensor responses linear with analyte concentration, and additive with respect to multiple vapors; 2) smoothly sprayed sensor films of uniform thickness; and 3) laminar flow in the vapor phase above the sensor films. Additionally, sensor response drift due to aging, temperature and humidity fluctuations, etc. was not taken into account. Further testing of the effects of aging on the ST-based performance of mixture analysis is needed for a thorough analysis of the ST method.

4.7.4.1. Response Linearity

Previous work has demonstrated response linearity and additivity for these sensor types during exposure to pure and 2-component mixtures.[18,34] Additionally, previous work has demonstrated that response additivity is a reasonable approximation, for most complex mixtures.[19] For cases in which the sensor responses are highly nonlinear and non-additive, mixture analysis from sensor training on pure vapors would be problematic. Additional training would be needed to train the sensors for this nonlinearity. However, the advantage of ST analysis, relative to traditional steady state analysis, would still be evident.

4.7.4.2. Smooth Sensor Films

Sensor fabrication typically does not produce smooth sensor films.[53,54] When looking at the sensor films with the naked eye, agglomerations were apparent. Sensor thickness profiles have similarly been observed to be uneven.[19] Black Pearls 2000 CB typically exists in particles ranging from 8 to 20 nm in diameter, but also forms aggregates ranging from 50 to 200 nm in diameter.[55] When spray depositing sensor films, larger aggregates tend to temporarily clog the airbrush outlet. The various particle/aggregate sizes, coupled with occasional airbrush outlet clogging, create the rough sensor films typically observed. For previous ST work using these sensor types, films were observed to range between 1-3 μm in thickness.[19] Film thicknesses were therefore varied within these bounds for model validation (Figures 4.5-7). While actual films are not of uniform thickness, the modeled responses should be representative of an average sensor film thickness, which captures the average sorption capacity of the experimentally sprayed films.

4.7.4.3. Laminar Flow

The assumption of laminar flow is reasonable for all geometry/flow rate combinations modeled in these works (Tables 4.4-5). For all cases, the Reynolds number, *Re*, was < 100, well below the transitional regime from laminar to turbulent flow (Appendix 4.9.1). Model validation as shown in Figures 4.5-7 displayed excellent agreement between experimental and modeled responses for the first sensors along each array. All 15 sensors along the array exhibited identical response curves for a high vapor pressure, low-partitioning analyte such as hexane (Figure 4.5). In this case, analyte progressed so rapidly down the sensor array that differences in response due to turbulent rather than laminar flow were not observable on the timescale used. The progression along an array of a low vapor pressure, high-partitioning analyte, like decane (Figures 4.6 and

4.7), showed differences between the experimental and modeled responses. These differences can be attributed to non-perfectly laminar flow in the vapor stream.

The transition from laminar to turbulent flow occurs in the neighborhood of Re ~ 2000.[47,48] This transition is for flow through smooth, flat surfaces. The sensor film, with thickness ranging from 1-3 µm, has aggregates on the order of microns due to carbon black aggregation and sensor deposition. All other exposed walls were coated with Teflon tape, which possessed small lumps produced by the formation of air pockets between the Teflon tape and the aluminum chamber housing. Additionally, the experimental design involved constantly changing the cross-sectional area of flow as vapor was delivered to the sensor arrays, going from 2 mm^2 in the delivery tubing, to 6 mm^2 prior to delivery to the arrays, to 1 mm^2 in the area overhead the arrays.[19] The unsmooth chamber surfaces, combined with the experimental design, appeared to create, and maintain, some level of turbulent flow in the vapor channels.

4.7.5. Effect of Turbulence

The goal of ST analysis was to time-separate the responses of sensors along an array. Figures 4.6-7 show that a perfect laminar flow assumption lessens the degree of separation between sensors in an array by making individual response fronts more gradual relative to those observed experimentally. Thus, for these experimental conditions, some turbulence results in better ST based implementation, by increasing the effective diffusivity of the vapor. For these experimental conditions,[19] Pe_{yz} and Pe_{zz}^{-1} were calculated to be 5×10^{-2} and 4×10^{-4}, respectively. If the turbulence caused the effective vapor diffusivity to increase by a factor of 10, the effective Pe_{yz} and Pe_{zz}^{-1} would become 5×10^{-3} and 4×10^{-3}, respectively (eq (6) and eq (7)). Thus, under these conditions, such turbulence was favorable for ST sensor response. Under certain conditions, however, turbulence could act to lessen the desired ST response, by causing longitudinal diffusive transport to become significant and increasing Pe_{zz}^{-1} beyond the critical value. One could attempt to model individual flow rate/geometry cases and take into account various levels of turbulence. However, the fact that the laminar-turbulent transition slowly occurs over a wide range of Reynolds numbers, coupled with the varying degrees of roughness present on the walls of the sensor chamber, would make this a difficult task.

4.8. Conclusions

A model of ST sensor response has been developed and implemented to define an optimized ST regime which provides sorption-based broadly responsive sensor arrays with the ability to analyze vapor mixtures to the fullest extent possible. Two dimensionless parameters were employed. The first, Pe_{yz}, was a ratio of the time required for analyte to diffuse from the top to bottom of the vapor channel, relative to the time required for analyte to convect from the inlet to outlet of the vapor channel. If this ratio was too large, the vapor front was not well-mixed as it progressed along each array, and analyte depletion effects were not observed. The second, Pe_{zz}^{-1}, was a ratio of the time required for an analyte to convect, relative to the time required for analyte to diffuse, from the inlet to outlet of the vapor channel. If this ratio was too large, significant mass transport of material along the array was due to diffusion, and vapor front progression along the arrays was broadened, causing sensor responses to overlap. As turbulence is introduced, enhanced vapor diffusivities are generated, which act to relax (increase) the Pe_{yz} criterion, while further restricting (decreasing) the Pe_{zz}^{-1} criterion.

For modeled sensor film thicknesses of 0.1 μm and 1 μm, a regime was defined based on Pe_{yz} and Pe_{zz}^{-1} which provided the sensor arrays with the ability to analyze vapor mixtures to the fullest extent possible, given training on only pure vapor species. The critical values of Pe_{yz} and Pe_{zz}^{-1} denote maximum values which allow for optimum ST analysis, assuming laminar flow. The critical values of Pe_{yz} and Pe_{zz}^{-1} were the same for the two film thicknesses, $< 10^{-3}$ and $< 5\times10^{-3}$, respectively. This suggests that these same critical values can be used to define regions of optimal ST response for other sensor types of similar (i.e. within a few orders of magnitude) sorption capacities, as long as the sorptive capacities of the films towards the vapors of interest, given by the partition coefficients K_{eq}, are known.

4.9. Appendix

4.9.1. Model Details

In the bulk vapor phase, a transient two-dimensional convection-diffusion equation was employed,[47-49]

$$\frac{\partial C_v}{\partial t} + V_z(y)\frac{\partial C_v}{\partial z} = D_v\left(\frac{\partial^2 C_v}{\partial y^2} + \frac{\partial^2 C_v}{\partial z^2}\right), \tag{A1}$$

where $C_v(y,z,t)$ was the vapor-phase concentration of the analyte of interest, $V_z(y)$ was the velocity profile as a function of chamber height, and D_v was the diffusivity of the analyte vapor of interest in air. For each of the analytes listed in Table 4.2, vapor-air diffusivities were estimated using the Chapman-Enskog equation for diffusivity.[47,48] From all sensor-analyte combinations, the average value of D_v used was 0.08 cm^2 s^{-1}, and all vapor diffusivities fell within a factor of 2 of this value.

The Reynolds number is a dimensionless number representing a ratio of the inertial to viscous forces acting on a fluid.[47-49] This was calculated as $Re = (H_v U\rho)/\mu$, where $U = Q/(WH_v)$ was the mean linear flow velocity (cm s^{-1}), Q was the volumetric flow rate of vapor delivery (cm^3 s^{-1}), and ρ and μ are the density (g cm^{-3}) and viscosity (g cm^{-1} s^{-1}) of air, respectively. The value of Re

is important in determining the transition from laminar to turbulent flow, which occurs when inertial forces significantly overpower viscous forces. This transitional regime occurs at $Re \sim 2000$.[47,48] Below this value, laminar flow is typical. For all modeled geometry/flow rate combinations, $Re < 100$. Thus, laminar flow was assumed and the Navier-Stokes equation was solved for the problem of plane Poiseuille flow between parallel plates.[49] For a coordinate system that begins at the sensor/vapor boundary, the equation of this parabolic velocity profile is:[49]

$$V_z(y) = -6U\left((y/H_v)^2 - (y/H_v)\right). \tag{A2}$$

Flow through the sensor chamber is pressure-driven. For certain geometry/flow rate combinations, the pressure drop generated along a sensor array could become significant. A significant pressure drop along a sensor chamber makes it difficult and/or impractical to generate these flow rates in practice due to the high pressure required at the chamber inlet. A significant pressure drop also creates a density and concentration gradient along a chamber, according to the ideal gas relation $C_v(z) = P(z)/RT$, where $P(z)$ was the pressure of the analyte of interest at position z. Tables 4.4 and 4.5 allow for a total of 54 (9 geometries × 6 flow rates) and 78 (13 geometries × 6 flow rates) geometry/flow rate combinations for the macro- and micro-dimensioned models, respectively. The relationship between the linear flow velocity and the pressure drop was $\Delta P = -12\mu UL/H_v^2$.[49] Only geometry/flow rate combinations that generated < 10% change in pressure along an array, i.e., $|\Delta P| < 0.1$ atm, were considered.

In the bulk sensor phase, a transient two-dimensional diffusion equation was used,[47-49]

$$\frac{\partial C_s}{\partial t} = D_s\left(\frac{\partial^2 C_s}{\partial y^2} + \frac{\partial^2 C_s}{\partial z^2}\right), \tag{A3}$$

where $C_s(y,z,t)$ was the concentration of analyte in the vapor and D_s the diffusivity of analyte in the sensor material. The initial condition for the sensor and vapor phase was $C_s(y,z,t = 0) = 0$. For each of the sensor/vapor combinations, the value of D_s used was 10^{-7} cm^2 s^{-1}. This value falls between the experimental diffusivities of gases diffusing through liquid and solid states, and was of the same order of magnitude for gases diffusing through polymers.[48] To determine the sensitivity of the model to the value chosen for D_s, the value of D_s was varied for many sensor/analyte arrangements. The same mass uptake responses were observed for $D_s = 10^{-6} - 10^{-8}$ cm^2 s^{-1}. For a diffusivity of 10^{-7} cm^2 s^{-1}, the time scales for diffusion through a 0.1 and 1.0 micron film are 10^{-3} and 10^{-1} s, respectively. Responses of similar polymer/carbon black composite sensor films of these thicknesses have been shown to occur on these time scales.[46] Comparatively, for the various geometries and flow conditions modeled, most (> 90%) exposure times ranged from 10 to 1000 s. Thus, for the film thicknesses used, the mass uptake and the diffusion through the sensor film was not limited by diffusivity, but rather was limited by the rate of mass transfer from the vapor phase to the sensor surface.

At the top vapor/wall boundary of the vapor flow pathway, an impermeable boundary condition (BC) of the form of $\partial C_v / \partial y = 0$ was employed. At the bottom sensor/wall boundary of the sensor film, an impermeable BC of the form $\partial C_s / \partial y = 0$ was employed (Figure 4.4). Additionally, at the sensor/inlet and sensor/outlet boundary, an impermeable BC of the form $\partial C_s / \partial z = 0$ was used. This BC implies that analyte can only access the sensor film by depleting the vapor phase.

For model validation, a time-dependent concentration profile $C_{v,inlet}(t)$ was used at the vapor inlet, to simulate the vapor delivery profile encountered experimentally. For all other modeled situations, a step function was used for vapor introduction, such that for $t > 0$, $C_v = C_{v,o}$, where $C_{v,o}$ was the concentration of vapor exposure. $C_{v,o}$ was converted from a fractional vapor pressure to concentration units of mol m^{-3} using the ideal gas law. At the vapor outlet, a convective flux BC was used, which is useful in convection-dominated situations in which the outlet concentrations are unknown. This sets the diffusive component of flux across the boundary to zero, $D_v\left(\partial^2 C_v / \partial z^2\right) = 0$, forcing material to pass through the boundary only via convection.

The flux across the vapor/sensor boundary was driven by the difference between the conditions of the sensor and vapor phase; that is, the difference between the concentration of the vapor phase and the concentration in the sensor that would be at equilibrium with the concentration in the vapor. This flux was dictated by the instantaneous values of C_v and C_s, as well as the partition coefficient K_{eq} that defined equilibrium between the two phases. On the vapor side of the boundary, $D_v\left(\partial C_v / \partial y\right) = K_m(C_s - K_{eq}C_v)$, while on the sensor side of the boundary, $D_s\left(\partial C_s / \partial y\right) = -K_m(C_s - K_{eq}C_v)$, where K_m is the mass-transfer coefficient at the surface.[49] These BCs maintained a continuity of flux across the surface, in addition to forcing the sensor/vapor equilibrium condition.

The mass-transfer coefficient K_m was estimated for each geometry/flow rate combination using a Sherwood (*Sh*) number correlation for laminar, forced flow along a flat plate.[47] This correlation was given by $K_m = Sh \times D_v / H_v$, where the Sherwood number $Sh = 0.323 Re^{1/2} Sc^{1/3}$, and $Sc = \mu / \rho D_v$ was the Schmidt number, the ratio of momentum diffusivity to mass diffusivity.[47] This relationship took into account various competing transport processes for each of the modeled conditions.[47] To determine the sensitivity to the value of K_m chosen, the value of K_m was allowed to vary by several orders of magnitude larger and smaller from that estimated for many geometric arrangements. In all cases tested, the same mass uptake responses were obtained from the modeling results. This again suggested that mass uptake by the sensor film was limited by the rate of mass transfer to the surface via the bulk vapor phase.

Modeling of the transient concentration profiles was required for modeling of the ST sensor responses. Given the wide range of geometries and flow rates, the time frame for capturing this transient response varied widely. For example, in situations where all sensors rapidly come to equilibrium with the sampled vapor stream, due to a high flow rate, the appropriate time frame would be relatively short. In situations where a long time is required for all sensors to come to equilibrium with the sampled vapor stream, due to a low flow rate, the time frame would be orders of magnitude longer. To assure that all of the modeled geometry/flow rate combinations captured the transient response, the time required to saturate each sensor array with an analyte vapor of interest was calculated. For this calculation, the molar flow rate, Q_m, of analyte through the vapor entrance was calculated as $Q_m = QC_{v,o}$, where Q is the volumetric flow rate defined as $Q = UWH_v$. The number of moles required to bring the entire length of the linear sensor array to equilibrium with the sampled stream was calculated as $n_{sat} = C_{v,o}WL(K_{eq}H_s+H_v)$, which accounts for saturation of the sensor and vapor phase (first and second part within the brackets, respectively), given the experimental conditions. For a sensor array responding to a test analyte, the time required for the sensor array to establish equilibrium with the vapor phase was calculated as:

$$t_{sat} = \frac{n_{sat}}{Q_m} = \frac{WL\left(K_{eq}H_s + H_v\right)}{Q}. \qquad \text{(A4)}$$

This approach assumed that analyte would not reach the outlet of the vapor channel until the entire length of the sensor film was in equilibrium with the sampled concentration of analyte. If this assumption was invalid due to chamber geometry, for example if H_s were of the same order as L, analyte vapor would not be depleted and would rapidly distribute along the length of the sensor array. Under such cases, the observed response time scale would not allow for saturated sensor responses; however, the desired ST responses would be lost due to Pe_{yz} being large. For a given geometry/flow rate combination, the maximum time the model solved for, t_{max}, was calculated as $t_{\max} = 2.4 \times t_{sat,median}$, where $t_{sat,median}$ is the median time, out of 44 total times (4 sensors x 11 analytes), required to saturate the sensor arrays towards the analytes of interest.

For all modeled geometries, the length (L) and height (H_v and H_s) dimensions were different by several orders of magnitude. Thus, each problem required three length scales of interest: L, H_v, and H_s. To obtain accurate modeled responses in a timely manner, scaled geometries were introduced for the governing equations and appropriate boundary conditions. Appendix 4.9.2 offers a detailed description of the scaling approach employed. Additionally, all concentrations were made dimensionless with respect to the inlet concentration $C_{v,o}$ and were scaled to have values between 0 and 1.[48,49]

The concentrations of test analytes in the sensor material, $C_s(y,z,t)_s$, were solved for using COMSOL Multiphysics, a commercial finite element analysis software package, coupled with MATLAB. Because the equations were made dimensionless with respect to $C_{v,o}$, for a given geometry/flow rate/sensor/analyte combination, the response profile was calculated once for an inlet concentration of 1. Thereafter, the responses for different inlet concentrations were obtained by scaling this response profile accordingly with respect to the inlet vapor concentration. A rectangular mesh, with 10 nodes along the sensor thickness, 20 nodes along the vapor channel (both in the y-direction, Figure 4.4), and 200 nodes along the vapor flow pathway (z-direction, Figure 4.4), was used. Concentrations were solved for 200 times that were evenly spaced in time between $t = 0$ and $t = t_{max}$.

4.9.2. Model Scaling

For the modeled geometries, three length scales of interest were present: the chamber length, L; the vapor channel thickness, H_v; and the sensor film thickness, H_s. Re-scaling will be demonstrated for a single hypothetical case, however the approach and implementation was identical for all modeled cases. Assume one wants to model a sensor chamber of length $L = 1000$ au, vapor channel height $H_v = 1$ au, and sensor thickness $H_s = 0.01$ au, where au is some arbitrary length unit (Figure 4.A1).

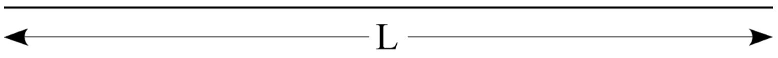


Figure 4.A1: To-scale representation of hypothetical chamber of $L = 1000$ au, $H_v = 1$ au, and $H_s = 0.01$ au.

From Figure 4.A1, it is clear that the length L is the dominant length scale for the problem. The use of the finite element method requires the generation of a mesh with individual nodes present throughout the boundaries of interest. To capture the resolution required to successfully solve the problem in the y-direction (Figure 4.4), which includes the vapor channel thickness and sensor film thickness, an excessive number of nodes would be required in the z-direction. This excessive number of nodes would require an excessive time to solve the problem. A scaled geometry was therefore introduced.

The problem was scaled so that the three length scales (L, H_v, and H_s) were approximately on the same scale. To do this, the vapor channel height, H_v, was used as a reference point for the scaling of the other two dimensions. The goal was to introduce new scaled geometries, $\overline{L}$ $\overline{H_v}$ and $\overline{H_s}$, so that excessive nodes and solution times were not required for each of the modeled conditions. The equations were scaled so that in all cases, $\overline{L} = 10\overline{H_v}$ and $\overline{H_s} = 0.5\overline{H_v}$, while

$\overline{H_v} = H_v$. Under this scaling, the geometry shown in Figure 4.A2 (to scale) was modeled each time. In this scaled geometry, nodes were evenly spaced along each of the dimensions.

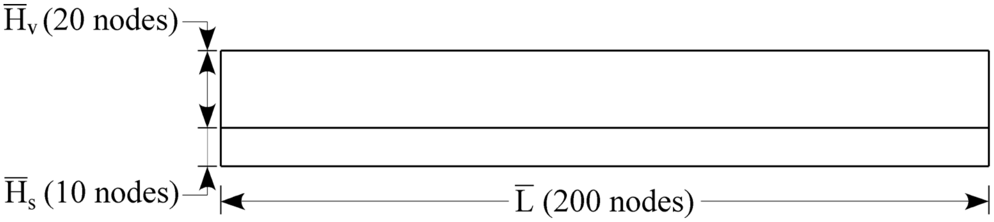


Figure 4.A2: To-scale representation of scaled chamber geometries, where $\overline{H_v} = H_v$, $\overline{L} = 10\overline{H_v}$, and $\overline{H_s} = 0.5\overline{H_v}$.

The original equations governing mass transfer in the vapor and sensor phases are given by eq (A1) and eq (A3), respectively. Under these conditions, with original geometries in place, z scales with L, and y scales with H_v and H_s in the vapor and sensor phase, respectively. To transfer the geometry to the desired relationships shown in Figure 4.A2, new scaling factors were introduced: L_{scale} and H_{scale} for the length and height scales, respectively. Desiring that $\overline{L}/\overline{H_v} = 10$, L_{scale} was defined as $L_{scale} = (L/H_v)/10$, and was calculated to be L_{scale} = 100. Similarly, desiring that $\overline{H_s}/\overline{H_v} = 0.5$, H_{scale} was defined as $H_{scale} = 2H_s/H_v$, and was calculated to be H_{scale} = 0.02. Using these scaling factors, the new scaled dimensions are given in eq (A5)-(A7):

$$\overline{L} = \frac{L}{L_{scale}} = \frac{1000}{100} = 10 \tag{A5}$$

$$\overline{H_s} = \frac{H_s}{H_{scale}} = \frac{0.01}{0.02} = 0.5 \tag{A6}$$

$$\overline{H_v} = H_v = 1 . \tag{A7}$$

Using the scaled dimensions, eq (A1) and (A3) were adjusted accordingly. In eq (A1) (vapor phase), z scales with the new length scale L_{scale}. The y dimension scales with 1 and thus remained unchanged, because H_v was the reference dimension for all other dimensions. Introducing these scalings, eq (A1) became eq (A8):

$$\frac{\partial C_v}{\partial t} + \overline{V_z}(\overline{y})\frac{\partial C_v}{\partial \overline{z}} = \overline{D_v} \bullet \left(\frac{\partial^2 C_v}{\partial \overline{y}^2} + \frac{\partial^2 C_v}{\partial \overline{z}^2} \right) . \tag{A8}$$

Here, $\overline{V_z}(\overline{y}) = V_z(y)/L_{scale}$ was the scaled velocity, and

$$\overline{D_v} = \begin{bmatrix} \overline{D_{v,y}} \\ \overline{D_{v,z}} \end{bmatrix} \tag{A9}$$

was a scaled diffusivity vector. Because y was not scaled,

$$\overline{D_{v,y}} = D_v \tag{A10}$$

and the relationship

$$\overline{D_{v,z}} = D_v \big/ L_{scale}^2 \tag{A11}$$

accounted for the new z-coordinate scaling. Note that the dot product was used for the distribution of the scaled vapor diffusivities.

In eq (A3) (sensor phase), z scaled with the new length scale L_{scale}, while y scaled with the height scale H_{scale}. Introducing these scalings, eq (A3) transformed to eq (A12):

$$\frac{\partial C_s}{\partial t} = \overline{D_s} \bullet \left(\frac{\partial^2 C_s}{\partial \bar{y}^2} + \frac{\partial^2 C_s}{\partial \bar{z}^2} \right). \tag{A12}$$

Here,

$$\overline{D_s} = \begin{bmatrix} \overline{D_{s,y}} \\ \overline{D_{s,z}} \end{bmatrix} \tag{A13}$$

is a scaled diffusivity vector, where

$$\overline{D_{s,y}} = D_v \big/ H_{scale}^2 \tag{A14}$$

and

$$\overline{D_{s,z}} = D_v \big/ L_{scale}^2\ . \tag{A15}$$

Again note the use of the dot product for the distribution of the scaled sensor diffusivities.

In addition to scaling the main governing equations (eq (A1) and eq (A3)), the vapor/sensor boundary condition was scaled. Effectively, scaling H_s to $\overline{H_s}$ increases the volume in the sensor phase by $1/H_{scale}$. Thus, the partition coefficient between the two phases was adjusted accordingly:

$$\overline{K_{eq}} = H_{scale} K_{eq}\ . \tag{A16}$$

For the current geometry being evaluated, $H_{scale} = 0.02$, and the sensor thickness (and volume) increased by $1/0.02 = 50$ to place it on the same scale as the vapor channel thickness. To account for this, the partition coefficient was scaled, decreased by multiplication by 0.02, to keep the problem consistent.

4.10. References

(1) Getino, J.; Horrillo, M. C.; Gitierrez, J.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Sens. Act. B* **1997**, *43*, 200-205.

(2) Dutta, R.; Hines, E. L.; Gardner, J. W.; Kashwan, K. R.; Bhuyan, M. *Sens. Actuators, B* **2003**, *94*, 228-237.

(3) Harrison, P. G.; Willett, M. J. *Nature* **1988**, *332*, 337-339.

(4) Rose-Pehrsson, S. L.; Grate, J. W.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801-2811.

(5) Patrash, S. J.; Zellers, E. T. *Anal. Chem.* **1993**, *65*, 2055-2066.

(6) Grate, J. W.; Patrash, S. J.; Kaganove, S. N.; Abraham, M. H.; Wise, B. M.; Gallagher, N. B. *Anal. Chem.* **2001**, *73*, 5247-5259.

(7) Wohltjen, H. *Sens. Act.* **1984**, *5*, 307-325.

(8) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1995**, *92*, 2652-2656.

(9) Harris, P. D.; Arnold, W. M.; Andrews, M. K.; Partridge, A. C. *Sens. Actuators, B* **1997**, *42*, 177-184.

(10) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371-377.

(11) Baller, M. K.; Lang, H. P.; Fritz, J.; Gerber, C.; Gimzewski, J. K.; Drechsler, U.; Rothuizen, H.; Despont, M.; Vettiger, P.; Battiston, F. M.; Ramseyer, J. P.; Fornaro, P.; Meyer, E.; Guntherodt, H.-J. *Ultramicroscopy* **2000**, *82*, 1-9.

(12) Jensenius, H.; Thaysen, J.; Rasmussen, A. A.; Veje, L. H.; Hansen, O.; Boisen, A. *Appl. Phys. Lett.* **2000**, *76*, 2615-2617.

(13) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 2501-2508.

(14) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*.

(15) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mat.* **1996**, *8*, 2298-2312.

(16) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177-4190.

(17) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *82*, 54-74.

(18) Gao, T.; Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Chem. Mat.* **2006**, *18*, 5193-5202.

(19) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Langmuir* **2007**, *23*, 13232-13241.

(20) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, B* **1994**, *18*, 211-220.

(21) Lewis, N. S. *Acc. Chem. Res.* **2004**, *37*, 663-672.

(22) Falcitelli, M.; Benassi, A.; Di Francesco, F.; Domenici, C.; Marano, L.; Pioggia, G. *Sens. Actuators, B* **2002**, *85*, 166-174.

(23) Scott, S. M.; James, D.; Ali, Z.; O'Hare, W. T. *J. Therm. Anal. Cal.* **2004**, *76*, 693-708.

(24) Sisk, B. C.; Lewis, N. S. *Sens. Actuators, B* **2005**, *104*, 249-268.

(25) Pardo, M.; Sberveglieri, G. *IEEE Trans. Inst. Meas.* **2002**, *51*, 1334-1339.

(26) Gardner, J. W.; Shurmer, H. V.; Tan, T. T. *Sens. Actuators, B* **1992**, *6*, 71-75.

(27) Zhang, C.; Bailey, D. P.; Suslick, K. S. *J. Agr. Food. Chem.* **2006**, *54*, 4925-4931.

(28) Lamagna, A.; Reich, S.; Rodriguez, D.; Scoccola, N. N. *Sens. Actuators, B* **2004**, *102*, 278-283.

(29) Lozano, J.; Fernandez, M. J.; Fontecha, J. L.; Aleixanre, M.; Santos, J. P.; Sayago, I.; Arroyo, T.; Cabellos, J. M.; Futierrez, F. J.; Horrillo, M. C. *Sens. Actuators, B* **2006**, *120*, 166-171.

(30) Di Natale, C.; Davide, F. A. M.; D'Amico, A.; Nelli, P.; Groppelli, S.; Sberveglieri, G. *Sens. Actuators, B* **1996**, *33*, 83-88.

(31) Zhang, Q.; Zhang, S.; Xie, C.; Zeng, D.; Fan, C.; Li, D.; Bai, Z. *Sens. Actuators, B* **2006**, *119*, 538-546.

(32) Xiaobo, Z.; Jiewen, Z.; Shouyi, W.; Xingyi, H. *Sensors* **2003**, *3*, 101-109.

(33) Yu, H.; Wang, J. *Sens. Actuators, B* **2007**, *122*, 134-140.

(34) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658-668.

(35) Zellers, E. T.; Pan, T.-S.; Patrash, S. J.; Han, M.; Batterman, S. A. *Sens. Actuators, B* **1993**, *12*, 123-133.

(36) Park, J.; Groves, W. A.; Zellers, E. T. *Anal. Chem.* **1999**, *70*, 4191-4201.

(37) Sobel, N.; Prabhakaran, V.; Desmond, J. E.; Glover, G. H.; Goode, R. L.; Sullivan, E. V.; Gabrieli, J. D. E. *Nature* **1999**, *392*, 282-286.

(38) Sobel, N.; Khan, R. M.; Saltman, A.; Sullivan, E. V.; Gabrieli, J. D. E. *Nature* **1999**, *402*, 35-35.

(39) Stitzel, S. E.; Stein, D. R.; Walt, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 3684-3685.

(40) Eklov, T.; Martensson, P.; Lundstrom, I. *Anal. Chim. Acta* **1999**, *381*, 221-232.

(41) Eklov, T.; Lunstrom, I. *Sens. Actuators, B* **1999**, *57*, 274-282.

(42) Eklov, T.; Lundstrom, I. *Anal. Chem.* **1999**, *71*, 3544-3550.

(43) Grate, J. W.; Snow, A.; Ballantine, D. S.; Wohltjen, H.; Abraham, M. H.; McGill, R. A.; Sasson, P. *Anal. Chem.* **1988**, *60*, 869-875.

(44) Burl, M. C.; Sisk, B. C.; Vaid, T. P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *87*, 130-149.

(45) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 5442-5447.

(46) Briglin, S. M.; Lewis, N. S. *J. Phys. Chem., B* **2003**, *107*, 11031-11042.

(47) *Perry's Chemical Engineer's' Handbook*, Seventh ed.; McGraw-Hill: New York, NY, 1997.

(48) Bird, R. B.; Stewart, W. E.; Lightfoot, E. N. *Transport Phenomena*, 2nd ed.; John Wiley & Sons, Inc.: New York, 2002.

(49) Deen, W. M. *Analysis of Transport Phenomena*; Oxford University Press: New York, 1998.

(50) Lawson, C. L.; Hanson, R. J. *Solving Least Square Problems*; Prentice Hall: Englewood Cliffs NJ, 1974.

(51) Quake, S. R.; Scherer, A. *Science* **2000**, *290*, 1536-1540.

(52) Schmidt, M. A. *Proc. IEEE* **1998**, *86*, 1575-1585.

(53) Loffredo, F.; Quercia, L.; Massera, E.; Di Francia, G. *Macromol. Symp.* **2005**, *228*, 139-146.

(54) Quercia, L.; Loffredo, F.; Di Francia, G. *Sens. Actuators, B* **2005**, *109*, 153-158.

(55) Zhu, W.; Miser, D. E.; Chan, W. G.; Hajaligol, M. R. *Carbon* **2004**, *42*, 1841-1845.

Table 4.1: Organic molecule/carbon black composite linear sensor arrays used in this study. CB = carbon black; DP = dioctyl phthalate; THF = tetrahydrofuran.

suspension	sensor materials	solvent
$LSA_{LA/DP}$	35 mg lauric acid, 15 mg DP, 150 mg CB	THF
LSA_{PG}	50 mg propyl gallate, 150 mg CB	THF
LSA_{QDD}	50 mg quinacrine dihydrochloride dihydrate, 150 mg CB	chloroform
$LSA_{TC/DP}$	35 mg tetracosane, 15 mg DP, 150 mg CB	toluene

Table 4.2: Analytes and abbreviations used in this work.

Analyte	
Hx	hexane
Bz	benzene
Oc	octane
Dc	decane
MeOH	methanol
EtOH	ethanol
PrOH	1-propanol
BuOH	1-butanol
Cf	chloroform
EA	ethyl acetate
Cb	chlorobenzene

Table 4.3: Partition coefficients K_{eq} ($\times10^{-3}$) and sensor response slopes M_R ($\times10^{4}$) with respect to part per thousand (ppth) of each of the analyte vapors and linear sensor arrays. Below K_{eq} and M_R the R^2 values are displayed for each respective slope fit.

sensor		Hx	Bz	Oc	Dc	MeOH	EtOH	PrOH	BuOH	Cf	EA	Cb
$LSA_{LA/DP}$	K_{eq} ($\times10^{-3}$)	1.1	2.0	12.1	126.1	0.5	1.4	4.2	15.9	1.1	0.9	26.5
	R^2	0.88	0.91	0.88	0.88	0.96	0.75	0.98	0.97	0.93	0.83	0.91
	M_R ($\times10^{4}$)	5	8	73	890	5	8	34	146	2	9	90
	R^2	0.98	0.97	0.96	0.95	0.99	0.98	0.99	0.99	0.98	0.99	0.98
LSA_{PG}	K_{eq} ($\times10^{-3}$)	0.4	0.6	4.5	46.1	0.4	0.9	2.3	7.3	0.4	0.5	23.8
	R^2	0.93	0.93	0.95	0.95	0.98	0.96	0.95	0.96	0.94	0.98	0.92
	M_R ($\times10^{4}$)	5	11	61	640	11	27	67	202	7	22	229
	R^2	0.94	0.97	0.95	0.95	0.98	0.92	0.99	0.99	0.98	0.96	0.97
LSA_{QDD}	K_{eq} ($\times10^{-3}$)	2.1	3.5	18.3	193.9	1.7	7.3	10.0	38.6	2.2	2.8	28.2
	R^2	0.97	0.95	0.88	0.87	0.89	0.98	0.98	0.95	0.88	0.97	0.91
	M_R ($\times10^{4}$)	5	13	44	506	9	20	92	403	4	15	144
	R^2	0.91	0.92	0.87	0.96	0.92	0.96	0.98	0.98	0.98	0.96	0.96
$LSA_{TC/DP}$	K_{eq} ($\times10^{-3}$)	1.0	1.6	11.3	119.2	0.3	1.1	2.8	10.2	0.8	0.6	47.2
	R^2	0.96	0.94	0.95	0.92	0.98	0.86	0.97	0.98	0.91	0.97	0.91
	M_R ($\times10^{4}$)	5	5	59	700	3	5	20	79	1	7	107
	R^2	0.95	0.96	0.94	0.91	0.98	0.91	0.99	0.99	0.98	0.97	0.95

Table 4.4: Geometry combinations and flow rates used for macro-dimensioned modeling.

L (mm)	W (mm)	H_v (mm)	Q (mL min^{-1})
75	3.5	0.34	1
75	8	0.34	2
75	35	0.34	4
75	3.5	0.1	10
75	3.5	0.8	19
30	3.5	0.034	30
750	3.5	0.034	
7.5	3.5	0.034	
2	3.5	0.034	

Table 4.5: Geometry combinations and flow rates used for micro-dimensioned modeling.

L (mm)	W (mm)	H_v (mm)	Q (μL min^{-1})
50	0.2	0.01	1
25	0.2	0.01	5
10	0.2	0.01	10
5	0.2	0.01	25
100	0.2	0.01	50
50	0.1	0.01	100
50	0.05	0.005	
50	0.4	0.01	
50	0.8	0.01	
50	0.2	0.006	
50	0.2	0.003	
50	0.2	0.02	
50	0.2	0.04	

Table 4.6: Modeled mixtures ($P/P^o \times 100$) used for the determination of an optimal ST operational regime.

mix #	Hx	Bz	Oc	Dc	MeOH	EtOH	PrOH	BuOH	Cf	EA	Cb
1	3	0	0	0	0	0	2	0	0	0	0
2	0	0	0	4	0	0	0	1	0	0	0
3	0	0	0	0	2	0	0	0	0	0	1
4	0	0	3	0	0	0	0	0	2	0	0
5	0	1	0	0	0	1	0	0	0	4	0
6	1	0	0	0	2	0	0	0	0	2	0
7	0	0	1	0	0	0	1	0	0	0	3
8	0	1	0	0	0	1	0	3	0	0	0
9	1	0	0	1	0	0	2	0	2	0	0
10	0	1	0	3	0	0	0	1	0	0	1
11	0	0	0	0	2	1	0	0	1	2	0
12	0	1	0	0	1	0	1	0	0	0	2
13	1	0	1	2	1	0	1	0	0	0	0
14	0	1	0	1	0	2	0	1	0	1	0
15	1	1	1	0	0	0	0	2	1	0	0
16	0	0	1	0	0	1	0	0	1	2	1

Figure 4.1: Arrangement used to optically monitor the delivery of analyte vapor to the sensor chamber. The vapor delivery arrangement used to deliver vapor to the sensor chamber was modified to deliver vapor to the quartz cuvette, as shown.

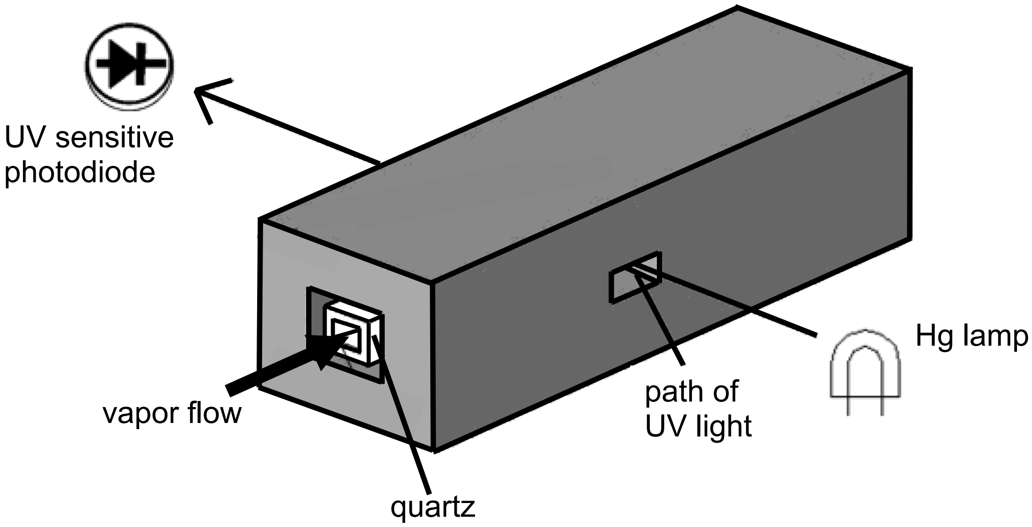


Figure 4.2: a) Sensor response and b) QCM frequency shifts as a function of concentration of various analyte vapors presented to $LSA_{TC/DP}$ (Table 4.1). See Table 4.2 for analyte identity.

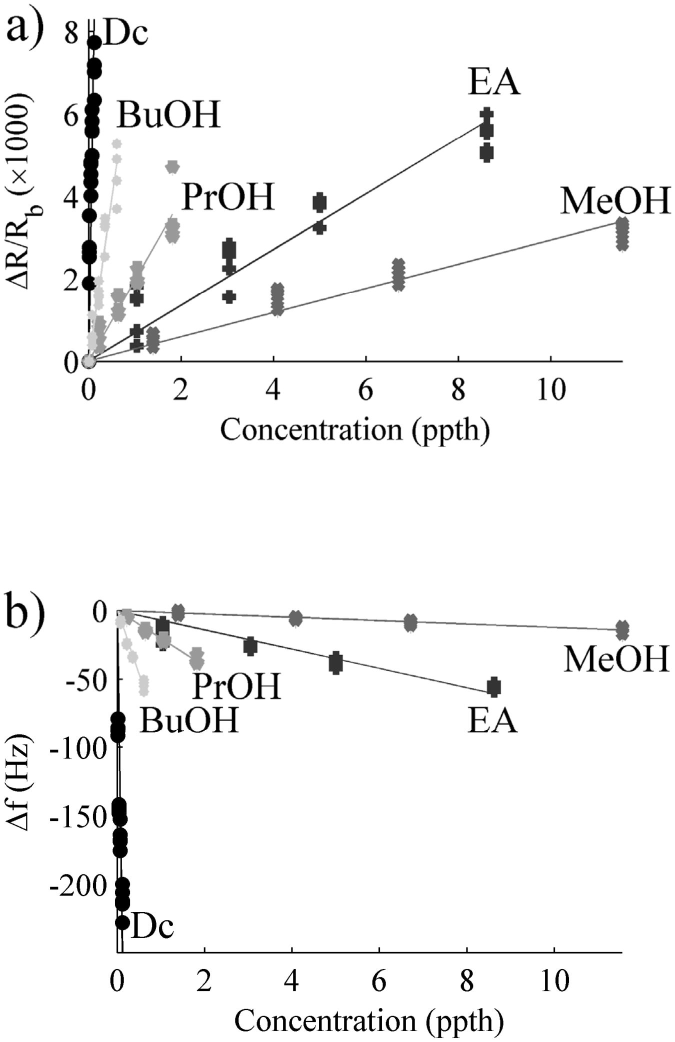


Figure 4.3: a) Raw photodiode voltage response during delivery of acetone vapor. b) Baseline-corrected photodiode voltage response during delivery of acetone vapor. c) Absorbance response profile of the acetone front arrival. Raw data are shown as points. The fitting in c) yielded a_1 = 0.00041, a_2 = -0.076 s, and k_t = 0.32 s^{-1}.

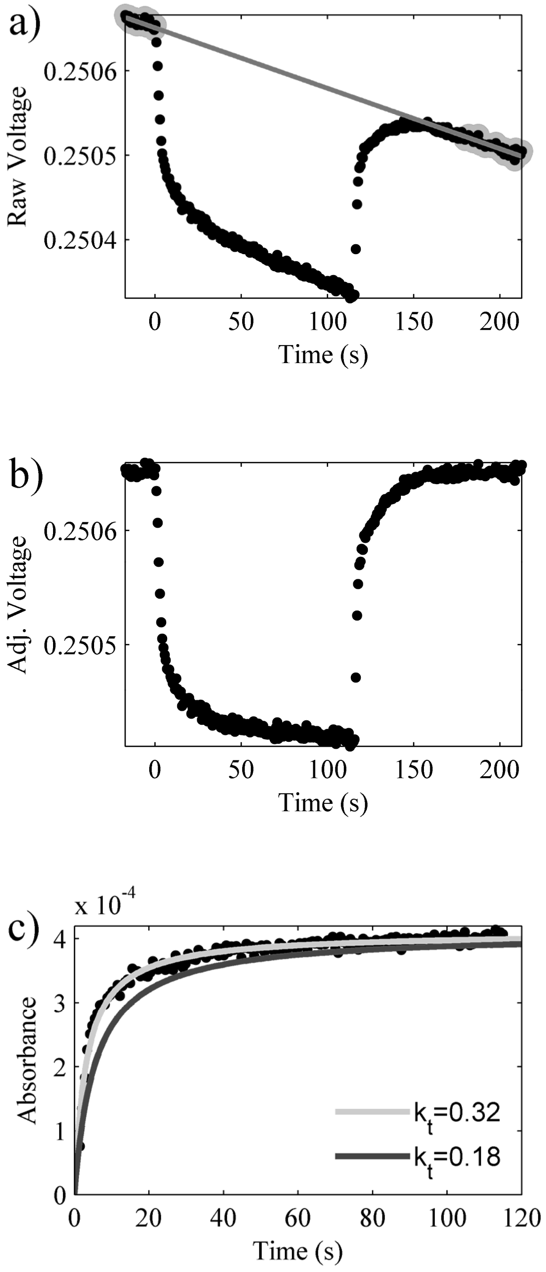


Figure 4.4: Modeling scheme, not to scale. Main equations employed are shown. See Appendix 4.9.1 for details on boundary conditions not shown in the figure.

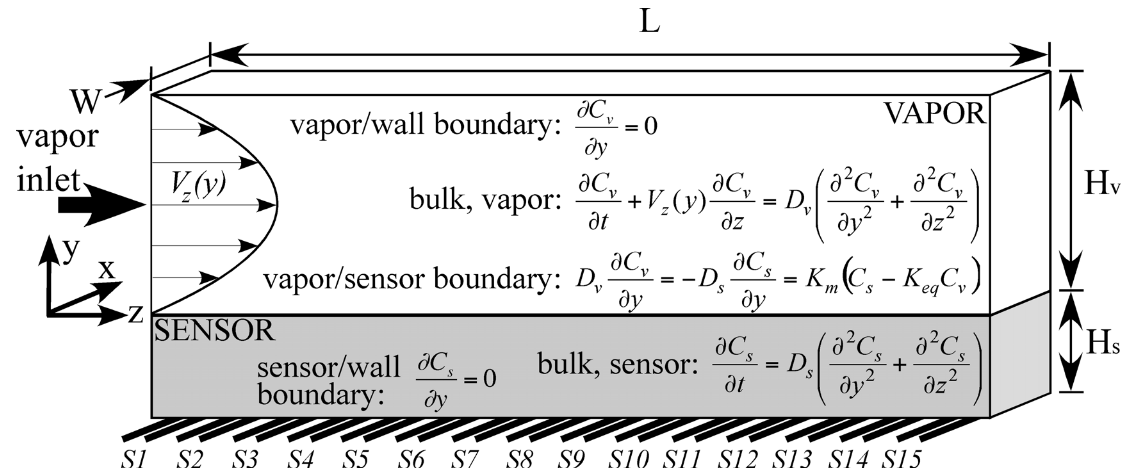


Figure 4.5: a) Experimental, b) modeled (H_s = 2.6 μm, k_t = 0.32 s^{-1}), and c) modeled (H_s = 2.6 μm, k_t = 0.18 s^{-1}) responses of all 15 of the sensors in a LSA_{PG} array to hexane. All responses were scaled to the response value obtained at t = 140 s.

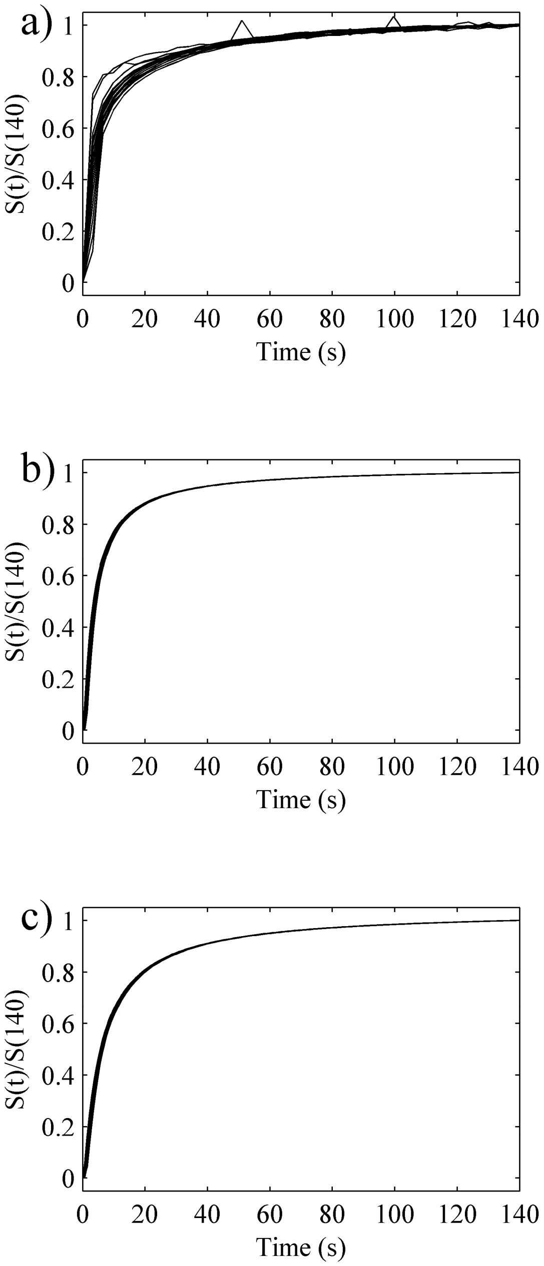


Figure 4.6: a) Experimental, b) modeled (H_s = 2.6 μm, k_t = 0.32 s^{-1}), and c) modeled (H_s = 2.6 μm, k_t = 0.18 s^{-1}) responses of LSA_{PG} to decane. All responses were scaled to the response value obtained at t = 140 s.

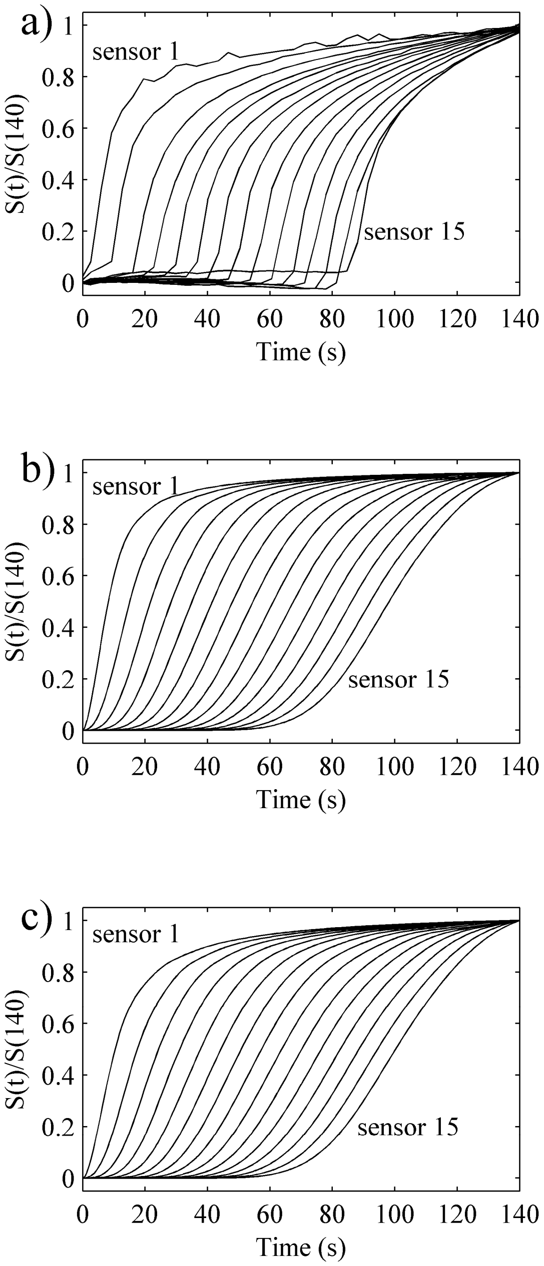


Figure 4.7: Exposure to a) pure ethyl acetate, b) pure decane, and c) ethyl acetate / decane mixture. Experimental (points) and modeled (lines) responses of first and ninth sensor along a 15-sensor $LSA_{TC/DP}$ array. Modeled sensor film thickness were 1.5 μm. All responses were scaled by the individual sensor response at t = 140 s.

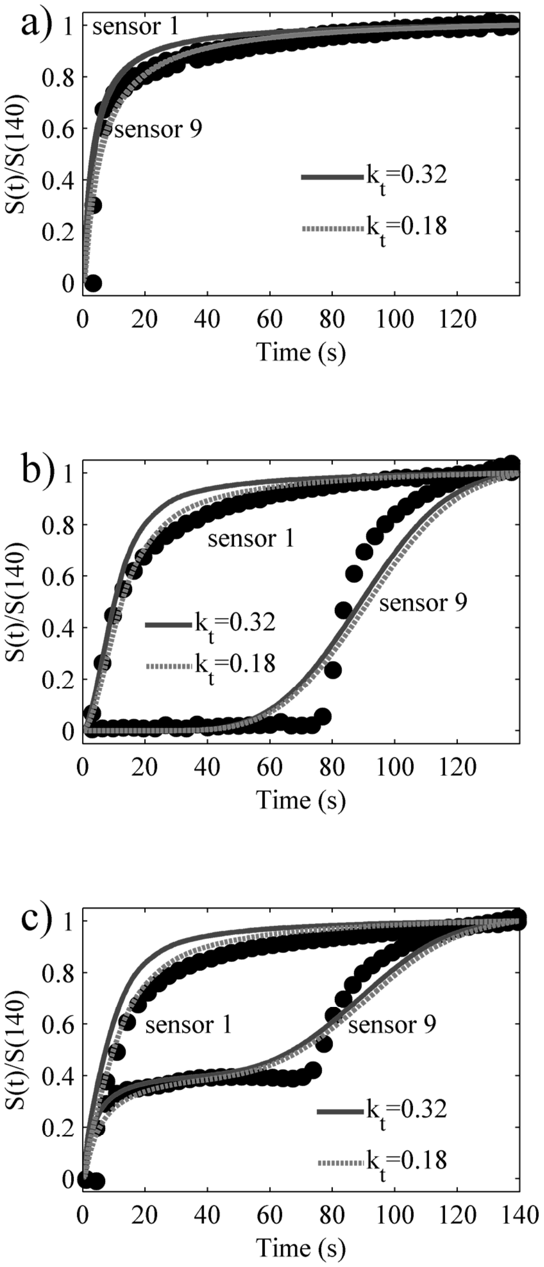


Figure 4.8: Residual error of mixture analyses for the modeling of 0.1 μm films, with various Pe_{zz}^{-1} cutoff points.

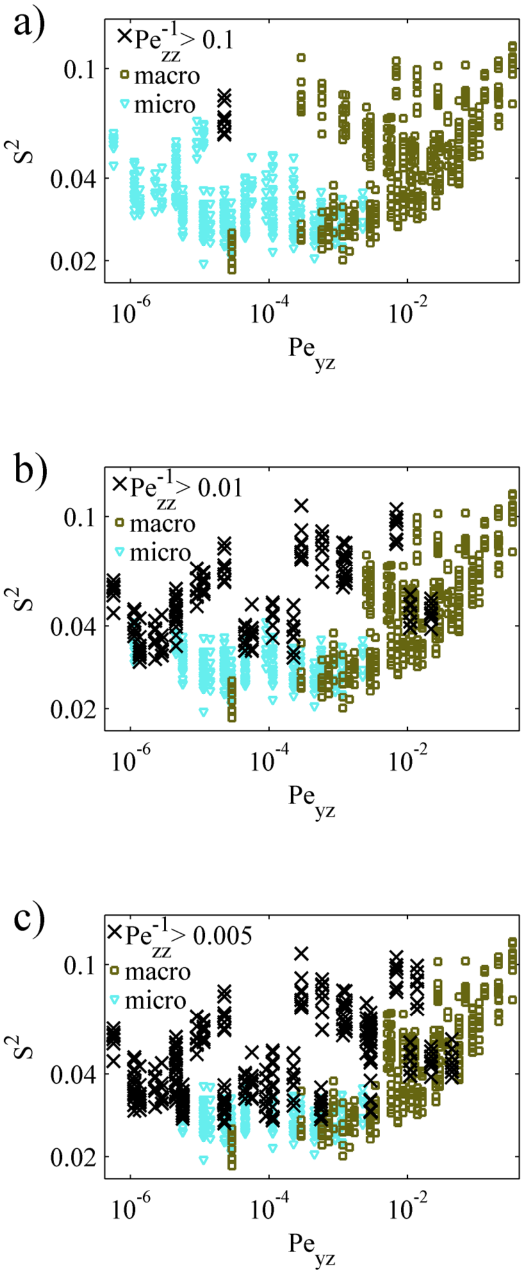
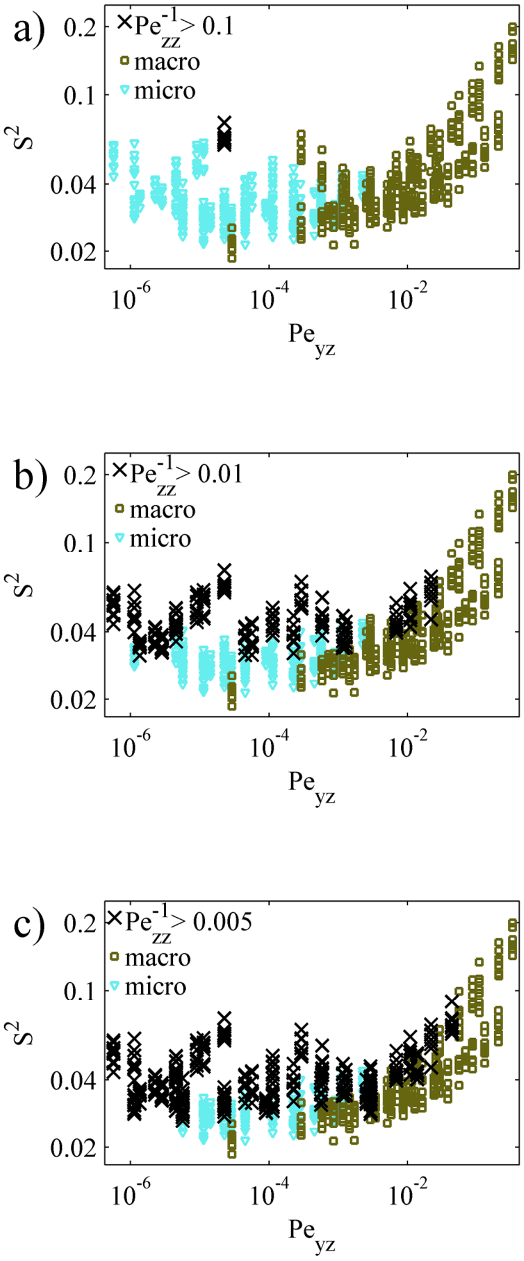


Figure 4.9: Residual error of mixture analyses for the modeling of 1 μm films, with various Pe_{zz}^{-1} cutoff points.



Chapter 5

Evaluation of Pattern Recognition Techniques for Analysis of Vapor Mixtures Using Spatiotemporal Response

5.1. Abstract

Modeled responses of pure vapors to small molecule/carbon black composite vapor sensors were characterized to determine the most useful features for analysis of vapor mixtures. Characterization techniques included the number of principal components (n_{PC}) in response space and the minimum resolution factor (*min rf*) among all vapor combinations in the pure vapor training library. Across libraries that contained different numbers of analytes and responses with varying levels of noise, calculation of n_{PC} did not provide a robust correlation with the ability to analyze mixtures. In contrast, the *min rf* value did produce consistent correlations with mixture analysis performance. For typical levels of noise, excellent mixture analysis performance was obtained when the *min rf* was ≥ 15. The responses were analyzed using nonnegative least squares (NNLSQ) and extended disjoint principal component regression (EDPCR). In all cases, NNLSQ was superior for mixture analysis, with a drawback of EDPCR resulting from the increased number of vapor response models required as the number of analytes in the library increased. An alternative method of EDPCR implementation, in which only a single fit to all vapors in the analyte library (PCR-SF), was also evaluated. PCR-SF provided significant improvements relative to EDPCR, and significantly decreased the computational time for the classification algorithm. NNLSQ still, however, provided optimal mixture classification results.

5.2. Introduction

Array-based sensing employing broadly responsive vapor sensor arrays has received significant attention in the recent literature. Sensing modalities include tin oxide sensors,[1,2] metal-oxide-semiconductor field-effect transistors,[3,4] coated optical fibers[5] and optical beads,[6] coated acoustic wave devices,[7,8] coated microcantilevers,[9,10] intrinsically conductive polymer chemiresistors,[11,12] and polymer/carbon black and small molecule/carbon black composite chemiresistors.[13,14] In this approach, in contrast to the traditional "lock and key" approach in which a single sensor responds to a single vapor, each sensor responds to a broad array of vapors and the response pattern across the array determines the identity of the vapor. Such arrays have been termed "electronic noses" due to similarities to mammalian olfaction,[15,16] in which the olfactory receptor proteins are broadly responsive and the collective response pattern dictates the identity of the vapor.[17]

Broadly responsive arrays have two typical modes of operation: discriminating between pure species or discriminating between complex mixtures. In each case, a training period is required in which multiple responses are generated toward each target vapor class. For identification of pure vapors, target classes include vapors generated from alkanes and alcohols,[5,7,11,13] thiols,[18] explosives,[19] and nerve agent simulants.[7,20] For analysis of complex mixtures, target classes include vapors from different coffees,[21] teas,[1] vinegars,[2] beers,[6,12] hops,[22] and wines;[23] as well as freshness monitoring of meats,[24] fish,[25] and produce.[26,27] In either case, during the training phase, the array is exposed to each vapor (or mixture) it will encounter during later testing phases. During testing, the response for an unknown is compared against the training vapor response library, using various pattern recognition algorithms. The vapor identity is then determined by the best fit from the training library.

A benefit of this approach is that a broadly responsive array can generally generate a unique response pattern for any vapor (pure or mixture) that may be encountered,[6,8,12,13,28] assuming the vapor is present at a reasonable concentration. This property obviates the need to develop a new sensor specifically for the target of interest. A downside to this approach, compared to lock-and-key sensors, is that mixtures are difficult to analyze in terms of their pure components. Thus, for any anticipated mixture of species, the array requires prior training on the anticipated vapors.

For real-world applications, multiple vapors will generally be present in a sampled vapor stream. Training for all possibilities of mixtures is an unreasonable burden, due to the extensive time required. The ideal situation is to train an array on pure vapors only, and to use these pure response signatures to identify mixtures of the pure species. The response of such sensors is often linear with vapor concentration,[28,29] and additive with respect to multiple vapors.[30] While this should allow for facile mixture analysis with training on only pure vapors, sensor arrays in their typical implementation can only analyze mixtures containing less than four species.[31]

Most reports dealing with sensor arrays extract a time-independent response for each sensor, to the vapor stream, along the array.[1,6,8,12,13] This time-independent response describes the partitioning of the analyte into each sensor along the array, in equilibrium with an unchanged vapor stream. The use of time-dependent response descriptors, including analysis of the kinetic response information from slow responding sensors,[2] or incorporation of a means for vapor stream modification to introduce a time-dependence to the analyte concentration in the vapor stream,[3,4,14] has been demonstrated to enhance the ability of sensor arrays to analyze vapor mixtures.[3,4,14] Presumably, the time dependence increases the sensor response space, and the amount of unique or meaningful information for analysis.

We have recently generated a mathematical model for the response to various vapors of small molecule/carbon black composite chemiresistor arrays operated in the spatiotemporal (ST), or space- and time-dependent, regime.[32] Under ST operation, the sensors were arranged into linear arrays, with low headspace volume and low vapor flow rates.[14,32,33] As vapor flowed over a sensor array, the analyte partitioned into the sensor film and a time-dependent response profile for all sensors along the array was observed.[33] When multiple vapors flowed over an array, the responses along the array were separated for each vapor, based on the differences in analyte partitioning into the film.[14,32,33] This additional dimension of sensor response data was shown to provide an enhanced ability to correctly analyze mixtures in terms of their pure components.[14,32]

ST responses were generated for a large number of geometric and vapor delivery flow rate combinations.[32] In the sensors, vapor concentration vs. time profiles were generated for pure vapors for particular flow rate and cavity dimensions. Sensor responses vs. time signatures were generated from the concentration vs. time profiles using measured sensor sensitivities to a particular analyte. Responses for pure vapors and vapor mixtures were then generated by adding Gaussian noise to the response vs. time signatures, and extracting the response at chosen times. The pure vapor response signatures were then used to analyze vapor mixture signatures. For each modeled geometry and flow rate, a single measure of the sum of the squared residual error between the modeled and estimated vapor concentrations, S^2, was calculated, to provide a metric for ranking mixture analysis performance. Using all modeled geometries and flow rates, an optimal region for ST mixture analysis was defined, based on two dimensionless Peclet number analogs that incorporated the chamber geometry and vapor delivery flow rate of the system of interest.[32]

In this work, we have evaluated the relative performance of various data characterization methods, on ST sensor response data to pure vapors, with respect to the ability to robustly predict the utility of pure vapor training for the determination of the components in vapor mixtures. Pure vapor response data was generated for a variety of modeled conditions. Principal components analysis (PCA) was first used with four different methods to determine the number of principal components (n_{PC}) in the pure vapor training library.[34] Fisher's linear discriminant analysis (FLDA) was used to calculate the resolution factor (rf) along an optimally separated dimension in response space between each of the pure vapors in the training library.[35] These tools were used to characterize the response space of pure training responses from each of the modeled geometry/flow rate combinations, which was compared to the ability of the modeled conditions to correctly identify the vapor mixtures. By simulating vapor libraries with different numbers of analytes and with different levels of superimposed noise, a method best suited to be a predictor for mixture analysis ability was introduced. Additionally, the performance of two different pattern recognition approaches, non-negative least squares (NNLSQ) and extended disjoint principal components regression (EDPCR), were evaluated for use in mixture analysis.

5.3. Experimental

5.3.1. Sensor Mass Uptake Model Development and Implementation

The model for the ST response of linear arrays of small molecule/carbon black composite vapor sensors used a time-dependent convection-diffusion equation with a laminar velocity profile to model the flow of vapor through the chamber and over the linear sensor arrays, with a time-dependent diffusion equation to model analyte transport in the sensor film.[32] Appropriate boundary conditions were implemented, and the equations were solved using COMSOL Multiphysics, a commercial finite element solver package, for the time-dependent mass uptake profile in the sensor films.

Tables 5.1 and 5.2 list all sensor materials and analytes used, respectively, for the modeling of responses in this work. Sensor fabrication has been described elsewhere.[14,28,32] Partition coefficients and response sensitivities for all sensor/analyte combinations have previously been reported.[32]

ST responses for 0.1 μm sensor films were generated for > 100 chamber geometries/vapor flow rate combinations, as previously reported.[32] These modeled responses constituted all generated data sets used in this study, and included optimized and non-optimized ST operational conditions. Additionally, for comparison against traditional, time-independent, steady state (SS) vapor sensing, this same model was used, with the responses sampled at times after each sensor had reached equilibrium with the unmodified vapor stream.

The initial modeling study used 4 linear arrays with 15 sensor nodes per array and with each sensor sampled at 4 different times. NNLSQ was used to match modeled “unknown” exposures to the known analyte library. Thus, a total of 240 (4 sensor arrays × 15 sensors/array × 4 times/sensor) response descriptors, n_{desc}, were available. To ensure that the data was not over-trained or over-fit, only a subset of the response descriptors were used. Responses from only three sensors along each array and at only three times were considered, to give 36 response descriptors (4 arrays × 3 sensors/array × 3 times/sensor).

For ST analysis, responses from the first, middle, and last sensor along each array were used and extracted at three response times, $t = (0.3,\ 1.0,\ \text{and}\ 2.4) \times t_{sat,median}$, where $t_{sat,median}$ was the median time required to saturate each of the vapor/sensor combinations.[32] Previous work has demonstrated that use of this limited number of sensors captured the same ST response as that of the full array.[14]

For SS analysis, the same model was employed with a chamber geometry of $0.075 \times 3.5\times10^{-3} \times 3.4\times10^{-4}$ m (L × W × H).[14] The responses from the first three sensors were used and sampled at three times each: $t = (9.8,\ 9.9,\ \text{and}\ 10) \times t_{sat,max}$, where $t_{sat,max}$ was the largest of the times required to establish vapor-sensor equilibrium for all vapor/sensor combinations. Sampling at these later times, combined with sampling only the first three sensors, ensured that only responses in equilibrium with the undepleted vapor stream were used. For ST and SS analysis, $n_{desc} = 36$, where n_{desc} is the number of response descriptors, which allowed a direct comparison of the two approaches.

5.3.2. Response Generation

5.3.2.1. Pure vapors

Pure vapor responses for 11 analytes were generated as described previously.[32] The pure vapor responses were assumed to be linear with analyte concentration.[28,30] Vapor concentrations were denoted as the fractional vapor pressure of an analyte, P/P^o, where P and P^o were the partial pressure and the vapor pressure of the analyte of interest, respectively.[15] Sensor mass uptake profiles were converted into sensor responses using previously reported sensitivities, and the modeled responses were representative of experimentally obtained $S(t) = \left(R(t) - R_b\right)R_b$ responses.[14,32] Unless otherwise noted, Gaussian noise representative of the noise typically observed in $\Delta R/R_b$ responses from small molecule/carbon black composites ST responses, $\sigma_{resp} = 2\times10^{-4}$, was superimposed onto each simulated analyte exposure.[32]

For each modeled geometry and flow rate, a subset of vapors from Table 5.2 was used to generate a pure vapor data set. For pure vapor response characterization, for each pure vapor in the training library, 200 responses were modeled at $P/P^o = 0.050$. These same 200 modeled responses per vapor were separately averaged and used as pure vapor response signatures for mixture analysis using NNLSQ. For EDPCR mixture analysis, concentration-dependent training responses were required.[31] Responses for each vapor were generated at $P/P^o = 0.0050$, 0.030 and 0.050, with twenty repetitions generated at each concentration.

5.3.2.2. Vapor Mixtures

Responses for vapor mixtures were generated as described previously.[32] Response additivity was assumed for multiple vapors,[30] and vapor mixture responses were generated by the addition of pure vapor responses. Data sets employing 5 to 10 analytes in the training library were generated for analysis. For each geometry/flow rate, 10 unique *n*-analyte libraries were chosen. For example, for 5 analyte libraries, Hx/Oc/MeOH/PrOH/Cf was used in one analysis, and Bz/Dc/EtOH/BuOH/EA was used in another analysis (Table 5.2). Eight additional analyte combinations were used to complete the 5 analyte library study. For each of the different 5-analyte libraries, 12 different mixtures were simulated: mixture 1 consisted of [0 2 0 3 0] parts of components 1-5, respectively, regardless of their identity. For example, if the analytes were Bz/Dc/EtOH/BuOH/EA, a mixture of 2 parts Dc and 3 parts BuOH was modeled; if the analytes were Hx/Oc/MeOH/PrOH/Cf, a mixture of 2 parts Oc and 3 parts PrOH was modeled. Eleven additional analyte mixtures were modeled, or 12 total analyte mixtures. This protocol allowed analysis of vapor libraries with different degrees of similarity present amongst the vapors. Additionally, because the same concentrations were used for each set of mixtures, regardless of analyte identity, the generated S^2 error, indicative of how well a given data set performed in terms of mixture analysis, was on the same scale across all modeled n_{an}-component analyte libraries.

For analyses using 5 analytes in the training library, 12 different vapor mixtures were simulated and analyzed: 4 binary, 4 ternary, and 4 quaternary mixtures. For analyses using > 5 analytes in the training library, 4 additional 5-component mixtures were simulated to make a total of 16 different vapor mixtures. For each simulated mixture, 50 exposure repetitions were generated. All response generations were performed with COMSOL multiphysics in conjunction with MATLAB.

5.3.3. Pure Vapor Response Data Set Characterization

A complete set of mixture analyses were performed for each of the over 100 modeled geometry/flow rate combinations. Some modeled conditions offered enhanced mixture analysis due to optimized ST conditions, defined by the relationships between the chamber geometry and the flow rate of vapor delivery.[32] Some situations offered enhanced mixture analysis due to the presence of dissimilar analytes in the training library. This work aimed to determine what was contained in pure vapor response data that allowed for higher levels of mixture analyses. PCA and FLDA were used to characterize the pure vapor response data sets generated for each of the modeled geometry/flow rate combinations. For a given set of modeled conditions, the 200 modeled pure exposures delivered at $P/P^o = 0.05$ were used for each pure vapor in the analyte library to characterize the properties of the n_{an}-analyte library. These same responses were used as training data for analysis using NNLSQ.

For analysis using PCA and FLDA, all responses were first concentration normalized by the sum of all response descriptors according to eq (1):

$$\overline{S_i}(t) = \frac{S_i(t)}{\sum_{i=1}^{n_{desc}} S_i(t)}, \tag{1}$$

where $S_i(t)$ was the original modeled response, i was the response descriptor number, and $\overline{S_i}(t)$ the normalized response. Normalization ensured that differences in responses generated by the different vapors were due to differences in the response fingerprints, rather than differences in the response magnitudes.[36] These characterization techniques were evaluated as a function of the number of vapors in the training library, and for various levels of superimposed noise. All analyses, including PCA and FLDA, were performed in MATLAB.

5.3.3.1. PCA and Determination of the Number of PCs

PCA allows for compression of n_{desc}-dimensional data by reorganization into n_{PC}-dimensional data, where $n_{PC} << n_{desc}$ and the n_{PC} dimensional data set contains most of the information in the original data set. PCA rotates the natural array response space from one where each dimension is the response of a single sensor, onto a space in which the first dimension, or PC, lies along the direction that contains the largest variance in the data set.[34,35,37] The second PC is along the direction that captures the second largest variance in the data set, and is orthogonal to the first PC.[34,35,37] This process is repeated for all n_{desc} dimensions in the original data set. From this rotated space, the first n_{PC} PCs ideally contain all (or most) of the useful information in the original data set, while the remaining dimensions, $n_{PC} + 1$ to n_{desc}, primarily contain noise.[34,35,37] The dimensionality of any data set is expected to be indicative of the amount of unique information present in the data and is defined by the number of PCs in the data set, n_{PC}.[34] The challenge is to determine the proper number of PCs to use. To calculate n_{PC}, the eigenvalue criterion and explained variance schemes were used.[34,37,38] Other schemes, such as the Scree-test, were avoided due to their higher degrees of subjectivity.[34,37,38]

A 2D data set of size $n_{exp} \times n_{desc}$ was the starting point, where n_{exp} was the total number of modeled exposures present in the data set. The n_{exp} was given by $n_{an} \times n_{rep}$ where n_{an} was the number of analytes in the library and n_{rep} was the number of training repetitions per analyte. Depending on the size of the training analyte library, $n_{an} = 5$, 7, 9 or 10, while $n_{desc} = 36$.

To perform PCA on this $n_{exp} \times n_{desc}$ data set, the data was mean-centered and autoscaled,[35] from which the correlation matrix was calculated. The eigenvectors and eigenvalues of this correlation matrix represented the new dimensions of maximized variance and the amounts of variance captured in each of these dimensions, respectively.[34,35,37] The autoscaled data were then projected onto the eigenvectors to create the PC-transformed data set. The eigenvalues describe the amount of variance captured along each PC.

5.3.3.1.1. Eigenvalue Criterion

For autoscaled data, the average eigenvalue is 1, and the sum of the eigenvalues is equal to n_{desc}.[37] For an array of specific lock-and-key sensors, where each sensor responded only to a single analyte, all the eigenvalues would equal 1. For cross-responsive sensor arrays, the eigenvalues instead decrease monotonically. Thus the initial eigenvalues were significantly larger than the last few.

The number of PCs to retain can be set using a cutoff value for the eigenvalues. The classical cutoff value is 1, known as "Kaiser's rule."[34,35,37,39] According to Kaiser's rule, PCs are retained only if they contain more information than the average component. Kaiser's rule will underestimate n_{PC} in cases in which the descriptors are mainly independent of one another.[34,40] Conversely, when descriptors are highly collinear, the value of n_{PC} will be overestimated.[34] To adjust for the former and latter cases, eigenvalues cutoffs of 0.7 and 2.0, respectively, have been used.[38] For this work, n_{PC} was defined as the number of components possessing eigenvalues ≥ 1 and ≥ 2. The two methods will be denoted as eigenvalue (≥ 1) and eigenvalue (≥ 2), respectively.

5.3.3.1.1. Explained Variance

Alternatively, the value of n_{PC} can be defined as the number of PCs required to explain a given fraction of the total variance.[34,35,37] The fraction of the total variance captured in a component is equal to that component's eigenvalue divided by n_{desc}. Starting with the 1st PC, this method used as many components as were required to explain a set fraction of the total variance in the data.[34,35,37]

Most implementations use sufficient PCs to describe 70% to 90% of the total variance.[34,38] These values can be adjusted for particular data sets. For example, when only one or two PCs explain most of the variation, the percentage could be increased to 95% to insure that less dominant, though significant, components are captured.[34] For this work, the value of n_{PC} was defined as the number of PCs required to explain 60% and 80% of the cumulative response variance. The two methods will be denoted as variance (60%) and variance (80%), respectively.

5.3.3.2. Fisher's LDA and Calculation of the rf Between Analytes

Fisher's linear discriminant analysis (FLDA) takes two clusters of data in *n*-dimensional space, and projects them onto a single dimension that maximally separates the clusters.[35,41] The optimal separation vector can be identified by finding the direction that maximizes the resolution factor, *rf*:

$$rf = \frac{\left|\overline{y_1} - \overline{y_2}\right|}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \qquad (2)$$

where $\overline{y_1}$ and $\overline{y_2}$ are the means, and σ_1 and σ_2 the standard deviations, respectively, of clusters 1 and 2 along this optimized dimension.[41,42] Assuming Gaussian distributions of cluster projections along the optimized dimension, resolution factors of 1, 2, and 3 correspond to correct classifications between the two clusters of 72, 92, and 98%, respectively.

The *rf* value was calculated between all binary combinations of analytes present in the training library. For a lock-and-key sensor array in which each sensor responded only to a single analyte, and when responses are concentration-normalized according to eq (1), all vapor combinations would report infinite *rf*'s.[43] For a broadly responsive sensor array, the *rf* will be some finite value, this providing a measure of how distinct the response patterns of various vapors are in sensor response space.

For FLDA, a train-test scheme was employed to insure that the reported *rf*'s were not artificially large due to overfitting of the response data. For each binary vapor combination, 150 of the 200 generated responses for each pure vapor were used to generate a FLDA model. The remaining 50 exposures were then projected onto the optimized dimension that maximized the *rf* among training data, and the resolution factor was calculated from these projections according to eq (2). All reported *rf*'s were from testing data analyses. For each of the modeled conditions, the minimum testing *rf* of all binary vapor combinations, *min rf*, was used to describe the situation for which confusion of the analytes would most likely occur.

5.3.4. Mixture Analysis

Responses were generated based on the assumption of linearity with concentration and additivity with respect to multiple vapors. Two linear, statistically based methods, NNLSQ[44] and EDPCR,[31] were therefore chosen for mixture identification. NNLSQ has been applied successfully to experimental mixture analyses using ST response information,[14] while EDPCR has been applied successfully to modeled low-order mixture analyses using SS response information.[31,45,46] Alternate, nonlinear forms of pattern recognition, such as neural network-based implementations, could provide enhanced performance in mixture analysis. However, such methods often require long training times and depend on the method of user implementation.[47] Statistically based methods provide a more objective measure of performance and are thus favored for an objective analysis of response data. For both NNLSQ and EDPCR, raw non-normalized response vectors were used to retain concentration information.

NNLSQ and EDPCR were used to determine the vapor concentration in each of the unknown mixture data. In cases using 5 analyte libraries, 12 vapor mixtures with 50 repetitions per mixture, or 600 mixture repetitions, were attempted. In cases using more than 5 analyte libraries, 16 vapor mixtures with 50 repetitions per mixture, or 800 mixture repetitions, were attempted. To generate a single metric that described how well each modeled geometry/flow rate/analyte combination performed in mixture analysis, the sum of the squared residual error, S^2, was calculated. This was the sum of the squares between the deduced and known vapor concentrations for individual mixture identifications, summed over all 600 or 800 mixture repetitions:

$$S^2 = \sum_{n_{mix}} \left(\sum_{n_{rep}} \left(\sum_{n_{an}} \left(C_{est} - C_{input} \right)^2 \right) \right), \tag{3}$$

where n_{mix} was 12 or 16 for libraries with 5 or > 5 analytes, respectively; $n_{rep} = 50$ the number of repetitions per mixture; n_{an} the number of analytes in the vapor library; and C_{est} and C_{input} the estimated and input vapor concentrations, respectively. Lower values of S^2 denoted lower estimation errors and thus indicated better performing mixture analyses.

5.3.4.1. Non-Negative Least Squares

For each pure vapor in the training library, the 200 modeled pure-vapor responses, used for PCA and FLDA, were averaged to generate a response signature for exposure at $P/P^o = 0.050$. This process created a pure vapor signature library **A** of size $n_{desc} \times n_{an}$. For an array response to an unknown mixture **b** of size $n_{desc} \times 1$, NNLSQ found the linear combination of the pure vapor response vectors that best matched the mixture response vector, such that all contributions from each pure vapor were ≥ 0.[44] NNLSQ was performed using a MATLAB function.

5.3.4.1. Extended Disjoint Principal Components Regression

The method of implementation for EDPCR has been well-described in the literature.[31] The approach is similar to the more general soft independent modeling of class analogy (SIMCA) method, except that in SIMCA each model that well-describes the data is accepted as a possibility, and more than one possibility may exist.[35,37,48] Briefly, EDPCR involves individually modeling, via PCs, the responses to each of the pure vapors. For each pure vapor, PCA was performed on mean-centered pure vapor response data: twenty repetitions per vapor at each of three concentrations, $P/P^o = 0.0050$, 0.030 and 0.050. The first m dimensions in PC space, where m was the number of relevant PCs, would traditionally serve as the model for the given pure vapor. Because responses were generated with an assumption of linearity with respect to concentration and Gaussian noise, a single PC model, or $m = 1$, was chosen for each pure vapor.[31] For training data, the position along this single PC was correlated with the training concentration, and used to calculate vapor concentrations of testing data.

Mixture response signatures were fit to all pure vapor models, and all higher-order models consisting of up to five different pure vapor models, which represented vapor mixtures. For multi-component models, the degree of overlap between the different analyte models was taken into account by the dot product between their respective PC dimensions.[31] For each model, the PC-transformed data was used to re-construct the original mixture response vector. The correct model, which established mixture component analyte identities, was chosen as the model that was best able to reconstruct the original mixture response vector, or the model offering the lowest squared residual between the original and reconstructed mixture response vector.[31] After analyte identities were established, analyte concentrations were calculated based on the projected distance along each of the pure-vapor PC models.[31]

EDPCR works because PCA ideally conserves all relevant response information. For a response to any pure vapor, if the response were projected onto the correct vapor PC model, no relevant response information would be lost, and the original response vector could be perfectly re-constructed from the projected data. If the vapor were, however, projected onto the wrong vapor PC model, information would be lost on re-construction.

2- 3- 4- and 5-component mixtures were simulated for data sets containing 5, 7, 9 and 10 vapors in the library. For mixture analysis of each of the unknown response vectors, this required fitting to all possible 1-, 2-, 3-, 4-, and 5-component combinations of the vapors in the library. For 5, 7, 9 and 10 vapors in the library, this process corresponded to a total of 31, 119, 381 and 637 mixture fittings, respectively. EDPCR was performed in MATLAB.

5.4. Results

5.4.1. Initial Screening of Methods

Figures 5.1a-e display S^2 values obtained from mixture analyses for simulated 10 analyte libraries using NNLSQ on ST and SS response data, and EDPCR on ST response data, as a function of the data set classification method. The same mixture sets were classified using ST and SS data in each case. ST results analyzed with NNLSQ demonstrated lower S^2 values that varied widely depending on chamber dimensions and flow rate. SS results analyzed with NNLSQ showed higher and more consistent error levels, due to a consistent means of extracting the steady-state sensor response information. For all modeled ST conditions, NNLSQ provided lower error identification rates than EDPCR.

Figures 5.1a-d display plots of S^2 vs. n_{PC} based on the variance (60%) and (80%) methods, and the eigenvalue (≥ 1) and (≥ 2) methods, respectively. All figures demonstrate that n_{PC} was larger for the ST data than the SS data, regardless of how the n_{PC} were calculated. In Figures

5.1a-c, aside from the initial benefit obtained from using ST rather than SS data, no clear trend was evident between n_{PC} and S^2. For example, Figure 5.1b shows that the error in mixture analysis was the same despite n_{PC} ranging from 3-8. In contrast, Figure 5.1d shows that as n_{PC} increased, S^2 in general decreased. Figure 5.1e displays S^2 vs. the *min rf* from all binary vapor combinations in the modeled libraries, showing that as the *min rf* of all binary vapor combinations increased, the modeled sensor array was better able to analyze vapor mixtures.

With 10 analytes in the vapor library, a maximum of 10 PCs can be present in any data set. However, fewer PCs might have relevant data in a 10-element broadly responsive array. The eigenvalue ($\geq$ 1) method resulted in n_{PC} spanning a range from 5 to 10 for ST data (Figure 5.1c), likely an overestimation due to the high collinearity of the sensor response. Using the variance (60%) rule likely underestimated the value of n_{PC} with 1-3 reported for the various cases (Figure 5.1a).

5.4.2. Different Sized Analyte Libraries

We now analyze how the characterization techniques changed for libraries that contained different numbers of analytes, and for response data that was generated with different levels of noise. Calculation of n_{PC} was performed with two methods: variance (80%) and eigenvalue ($\geq$ 2). Additionally, the *min rf* method was used. NNLSQ rather than EDPCR was used for mixture analysis due to its low S^2 values.

Figures 5.2a-c display S^2 vs. n_{PC} for 5, 7 and 9 analyte training libraries, respectively, where n_{PC} was defined using the variance (80%) method. Both ST and SS analyses are displayed. In all cases, the n_{PC} were larger and S^2 lower for ST data. Analysis of 5 analyte libraries (Figure 5.2a) yielded n_{PC} generally ranging from 2 to 10, despite the fact that no more than 5 PCs were possible. Additionally, for the 5, 7 and 9 analyte library cases, no trend was observed between increased n_{PC} and S^2 for ST analyzed data.

Figures 5.3a-c display S^2 for NNLSQ analysis vs. n_{PC} for the same 5, 7 and 9 analyte mixture analyses, using the eigenvalue ($\geq$ 2) method. In all cases, n_{PC} was less than the number of analytes in the library. As more analytes were introduced to the library, n_{PC} slowly increased. Thus this method correctly responded to the increase in response information possible due to larger analyte libraries. Additionally, as n_{PC} increased, mixture analyses improved. However, higher n_{PC} values were not required to optimally analyze mixtures. For example, Figure 5.3c shows that mixture analysis error was consistently the lowest when 5 PCs were in the data set, but the same level of mixture error was achievable with only 3 PCs in many data sets.

Figures 5.4a-c display S^2 for NNLSQ anlaysis vs. the *min rf* in the analyte library for the same 5, 7 and 9 analyte mixture analyses. In all cases, S^2 decreased as the *min rf* increased. At *rf*

~ 15, the lowest errors and optimal mixture identifications were achievable, but were not guaranteed. A region of optimal mixture analysis became better defined for *rf* > 15. The *min rf*'s reported for SS analyses were representative of the minimums reported for similar pure-vapor analyses using a 7 sensor array of small molecule/carbon black composite sensors.[28]

5.4.3. Different Levels of Noise

Figure 5.5 displays S^2 for NNLSQ analysis of simulated 10 analyte libraries vs. the characterization techniques discussed in Figures 5.2-4, with superimposed noise of σ_{resp} = (1, 4, and 16) × 10^{-4}. Figure 5.5a displays S^2 vs. n_{PC} using the variance (80%) method. As the response noise increased, n_{PC} and S^2 increased. When the noise was set to 16×10^{-4}, the n_{PC} were routinely overestimated, ranging from 1 to 20. As the noise increased, more and more of the response information (variance) was due to noise; thus the explained variance method ultimately resulted in the inclusion of noise-dominated dimensions.

Figure 5.5b displays S^2 vs. n_{PC} based on the eigenvalue criterion (≥ 2). With increased noise, the n_{PC} remained approximately constant while S^2 increased. When the noise was 16×10^{-4}, n_{PC} exhibited a range of 1 to 6. While n_{PC} fell within the required range of ≤ 10 for the simulated 10 analyte libraries, this definition of n_{PC} offered no clear correlation with the ability to analyze vapor mixtures.

Figure 5.5c displays S^2 vs. the *min rf* in the analyte library. For ST data, as the noise was increased, the *min rf* decreased and S^2 increased. Thus, for SS and ST broadly responsive array response data, the *min rf* calculated from all pure vapors in the analyte library can be used as an indicator for the ability of the array to correctly analyze mixtures.

5.5. Discussion

5.5.1. Broadly Responsive Array Mixture Analysis Prediction

An attempt was made to characterize the pure vapor response data of a broadly responsive array in terms of the array's ability to analyze mixtures. Two approaches were used: calculation of n_{PC}, and calculation of the *min rf* amongst all vapors in a training library. For the calculation of n_{PC}, two methods were used: how many PCs were required to explain a certain amount of the variance, and the number of PCs with eigenvalues above a chosen value.

The methods for calculating n_{PC} sought to determine the number of dimensions needed to capture class-specific non-noise response information. The explained variance method in many cases produced more dimensions than analytes (Figure 5.2c; Figure 5.5a). The eigenvalue method performed better in terms of providing realistic results for n_{PC} (Figures 5.3a-c; Figure 5.5b), however no trend was consistently observed between n_{PC} and the ability of an array to analyze vapor mixtures (Figure 5.5b). These findings suggest that as noise increased, the n_{PC} of the response space remained unchanged. However, signatures of the pure vapors overlapped to higher degrees. This increased overlap caused higher levels of analysis error.

The *min rf* gave a relative measure of the level of separation between pure vapors in response space, which accounted for response noise with σ_1 and σ_2 (eq (2)). As the *rf* between two vapors increased, the response signature for the two vapors became increasingly different, and the likelihood that the two vapors would be confused decreased. Across various library sizes and levels of noise, the *min rf* value was consistently correlated with the ability to successfully analyze mixtures. While higher n_{PC} values may allow for a lower *min rf*, this is not the rule. Rather, it is the number of unique pure vapor response clusters that can fit within those n_{PC} that is important for mixture analysis.

5.5.2. NNLSQ vs. EDPCR

In previous experimental work, NNLSQ consistently outperformed EDPCR for mixture analysis, offering lower variability and greater accuracy in mixture analysis.[14] It was concluded that EDPCR broke down due to nonlinearities exhibited in the observed response to multiple vapors.[14] Figure 5.1 gives a comparison of EDPCR and NNLSQ applied to identical data for the analysis of mixtures from their pure vapor signatures. In all cases, NNLSQ outperformed EDPCR. These data were generated with an assumption of linear and additive responses and Gaussian noise. Thus, even with linear data, NNLSQ outperformed EDPCR.

In this work, EDPCR was implemented by fitting the response of each unknown mixture to the full complement of all 5-component and lower order mixtures possible from the n_{an}-component training library. For the modeled cases using n_{an} = 5, 7, 9 and 10 vapors in the library, this corresponded to 31, 119, 381 and 637 mixture fittings, respectively. The usual method of EDPCR implementation has limited the number of allowed vapor combinations based on user information about the unknown exposure. For an *m*-component mixture chosen from an n_{an} component library, two implementations have been used: full fitting to all mixtures of *m* components or fewer possible from the n_{an} component analyte library,[31] and a limited fitting to the *m*-component mixture and all lower-order mixtures composed of only the *m* components (EDPCR-limited).[45,46,49] For example, under the former method, a binary mixture chosen from a 5 component library would be fit to all binary mixtures and all pure vapors possible from the analyte library, for a total of 15 different combinations.[31] Under the EDPCR-limited method, the binary mixture would be fit to the single 2-component mixture, and only the two pure vapors, for a total of 3 different combinations.[45,46,49] In this study, and the previous experimental study where

NNLSQ proved superior to EDPCR,[14] each unknown mixture response was fully fit to all mixtures that consisted of ≤ 5 components, comprised of all analytes in the training library.

Figure 5.6 offers a comparison between the performance of NNLSQ and EDPCR for a modeled 3-component vapor mixture with ST response information, using a 6 component vapor library. Responses for the vapor mixture were generated twenty times, and mean and standard deviations of analysis using various methods are shown. Responses were generated and analyzed for three levels of noise: the noise typically observed in these sensor's ST responses (2×10^{-4}) as well as noise one and two orders of magnitude less than that observed.

Figure 5.6a shows analysis using NNLSQ, and displays the modeled concentrations of all 6 vapors, as well as analysis results for the three levels of noise. In this case, no *a priori* information about the vapor mixture was known – the algorithm freely fit to all vapors in the library. For all three levels of noise simulated, NNLSQ correctly identified mixture components at their correct concentrations. Figure 5.6b shows analysis using EDPCR-limited, the most typically implemented method: allowing the modeled array to fit to the one 3-component mixture, the three 2-component mixtures, and the three pure vapors. In this case, the algorithm knew beforehand that the mixture consisted of Oc, Cf and PrOH (Table 5.2). For the level of noise representative of sensor responses, the algorithm was able to roughly pick out the correct concentrations of the components, with higher levels of error and more variance exhibited than NNLSQ analysis. Responses generated with one and two orders of magnitude less noise performed much better, consistently picking the correct concentrations with little analysis variance. Based on the user input of which three analytes were present in the mixture, this method required fitting to 7 different vapor models.

Figure 5.6c shows the analysis using the full EDPCR fitting method (EDPCR-full), with fits allowed to all 1-, 2-, 3-, 4-, 5- and 6-component models, or a total of 63 models. In this implementation, only the modeled responses with two orders of magnitude lower noise were able to pick out the correct vapors and concentrations. The two higher levels of noise were not able to correctly identify or quantify the components in the mixtures. This method is indicative of the implementation of EDPCR in this work and in a previous experimental ST analysis.[14]

NNLSQ allowed for a single simultaneous fit to all vapors in the library. EDPCR required a fit to all combinations of vapors in the library, and chose vapor concentrations based on the modeled vapor(s) that provided the lowest residual. Differences in analyses observed in Figures 5.6b-c suggest that EDPCR performed poorly due to the increased number of models fit to the unknown exposure. Implementation of EDPCR in a manner similar to NNLSQ, offering a single fit to a model composed of all vapors in the library (PCR-SF), might provide enhanced mixture analysis. For PCR-SF, the unknown mixture response vector was projected onto a single model comprised of all individual pure vapor models from the analyte library, and analyte concentrations were calculated based on the distance along each of the pure-vapor PC models. This is in contrast to EDPCR-limited or EDPCR-full, where analyte identity was first calculated by the vapor model offering the best ability to reconstruct the original mixture vector, and concentrations were then calculated based on the distance along the PCs of the model constituents.[31] Such an analysis was therefore performed, fitting each unknown exposure to only the collective 6 analyte model. Figure 5.6d shows that using PCR-SF, mixture analysis was greatly enhanced relative to EDPCR.

Figure 5.7 shows the same analyses performed in Figure 5.6, with a simulated 4-component mixture chosen from a 6-component library. NNLSQ again estimated vapor concentrations correctly (Figure 5.7a). EDPCR-limited now had to fit to one 4-component, four 3-component, and six 2-component mixtures, and four pure vapors, or 15 different vapor models, compared to 7 different vapor models for the 3-component mixture. Only the lowest level of noise was able to consistently identify the correct vapors and concentrations (Figure 5.7b). Extra model fittings required for higher-order mixtures appear to limit the performance of mixture analysis using EDPCR, rather than the true ability of an array to analyze a mixture. EDPCR-full (Figure 5.7c) yielded approximately the same results as EDPCR-limited. PCR-SF again performed significantly better (Figure 5.7d) than the other implementations of EDPCR.

The breakdown in the ability of EDPCR to analyze mixtures when going from 3- to 4-component mixtures (Figure 5.6b and 5.7b, respectively) was consistent with the breakdown point of EDPCR observed previously, where it was concluded that using EDPCR, lower-order mixtures (up to 2 or 3 components) could be successfully identified, but higher order mixtures (containing 4 or 5 components) could not.[45,46,49] Figures 5.6-7 suggest that these conclusions are algorithm-dependent, and are not true for alternative algorithms such as NNLSQ (Figures 5.6a-7a) or PCR-SF (Figures 5.6d-7d). Additionally, PCR-SF was much more computationally efficient than EDPCR, which would require over 1000 and 30,000 model fittings for a 10- and 15-analyte library, respectively, compared to a single fit for PCR-SF.

Figure 5.8 shows the ST response data from Figure 5.1, analyzed with EDPCR, PCR-SF, and NNLSQ. Using PCR-SF, significant improvements were observed relative to the traditional EDPCR approach. NNLSQ, however, still reported the lowest mixture analysis error. The differences in mixture analysis ability arose from differences in how the algorithms handled response noise: whereas NNLSQ retained all dimensions, PCR-SF found the single dimension that described the array response toward each pure vapor. For Gaussian noise, all n_{desc} dimensions in NNLSQ, and the single dimension defining each of the PCR-SF pure-vapor models, contain the same level of noise. Whereas NNLSQ was able to average out the effects of noise over n_{desc} dimensions and minimize its effect on mixture analysis, PCR-SF retained only a single dimension per pure vapor and mixture analysis was therefore more easily troubled by response noise.

The results shown in Figures 5.6-8 suggest that the traditional implementation of EDPCR is well-suited for pure vapors only. Operation in this manner has been proven to work,[31,45,46,49,50] and the computational time required for an *n* analyte library grows linearly as *n* and is therefore managable. However, for mixture analysis, EDPCR performs poorly, due to the large number of fits required, and becomes computationally inefficient because the computational time grows as 2^n. For mixture analysis using EDPCR, the alternative single fit method introduced here, PCR-SF, is advantageous.

5.6 Conclusions

Modeled sensor responses were used to determine the most useful features for the analysis of vapor mixtures. Characterization techniques were used to find a predictor, which worked across different sized analyte libraries and for different levels of response noise, for the ability of broadly responsive array pure-vapor training to analyze vapor mixtures. Characterization techniques included calculation of *min rf* among all vapor combinations in the pure vapor training library, and different means to calculate the n_{PC} in the pure-vapor response space. Calculation of n_{PC} did not provide a robust correlation with the ability to analyze vapor mixtures, whereas the *min rf* provided a correlation consistent across different sized analyte libraries, and responses generated with different levels of noise. For response noise typical of small molecule – carbon black composite chemiresistors, analysis of mixtures with pure vapor training became feasible when *min rf* > 15. Mixtures were analyzed using NNLSQ and EDPCR. A comparison between these two algorithms was presented for analysis of a 3- and 4-component mixture generated from a vapor library consisting of 6 components. In cases where EDPCR required fitting to a limited number of vapor models, correct mixture analysis was obtained. As the number of vapor models increased, the ability of EDPCR to correctly identify mixtures decreased. In all cases, NNLSQ was able to correctly identify the mixtures. An alternative single fit method for EDPCR implementation was tested, PCR-SF. This method offered significant improvements in mixture analysis ability, and a significant decrease in the required computational time for vapor classification.

5.7. References

(1) Dutta, R.; Hines, E. L.; Gardner, J. W.; Kashwan, K. R.; Bhuyan, M. *Sens. Actuators, B* **2003**, *94*, 228-237.

(2) Xiaobo, Z.; Jiewen, Z.; Shouyi, W.; Xingyi, H. *Sensors* **2003**, *3*, 101-109.

(3) Eklov, T.; Lundstrom, I. *Anal. Chem.* **1999**, *71*, 3544-3550.

(4) Eklov, T.; Lunstrom, I. *Sens. Actuators, B* **1999**, *57*, 274-282.

(5) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*.

(6) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 2501-2508.

(7) Carey, W. P.; Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 149-153.

(8) Rose-Pehrsson, S. L.; Grate, J. W.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801-2811.

(9) Baller, M. K.; Lang, H. P.; Fritz, J.; Gerber, C.; Gimzewski, J. K.; Drechsler, U.; Rothuizen, H.; Despont, M.; Vettiger, P.; Battiston, F. M.; Ramseyer, J. P.; Fornaro, P.; Meyer, E.; Guntherodt, H.-J. *Ultramicroscopy* **2000**, *82*, 1-9.

(10) Adams, J. D.; Parrott, G.; Bauer, C.; Sant, T.; Manning, L.; Jones, M.; Rogers, B.; McCorkle, D.; Ferrell, T. L. *Appl. Phys. Lett.* **2003**, *83*, 3428-3430.

(11) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1995**, *92*, 2652-2656.

(12) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371-377.

(13) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mat.* **1996**, *8*, 2298-2312.

(14) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Langmuir* **2007**, *23*, 13232-13241.

(15) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 5442-5447.

(16) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, B* **1994**, *18*, 211-220.

(17) Axel, R. *Sci. Amer.* **1995**, *273*, 154-159.

(18) Briglin, S. M.; Gao, T.; Lewis, N. S. *Langmuir* **2004**, *20*, 299-305.

(19) Toal, S. J.; Trogler, W. C. *J. Mater. Chem.* **2006**, *16*, 2871-2883.

(20) Hopkins, A. R.; Lewis, N. S. *Anal. Chem.* **2001**, *73*, 884-892.

(21) Gardner, J. W.; Shurmer, H. V.; Tan, T. T. *Sens. Actuators, B* **1992**, *6*, 71-75.

(22) Lamagna, A.; Reich, S.; Rodriguez, D.; Scoccola, N. N. *Sens. Actuators, B* **2004**, *102*, 278-283.

(23) Lozano, J.; Fernandez, M. J.; Fontecha, J. L.; Aleixanre, M.; Santos, J. P.; Sayago, I.; Arroyo, T.; Cabellos, J. M.; Futierrez, F. J.; Horrillo, M. C. *Sens. Actuators, B* **2006**, *120*, 166-171.

(24) Winquist, F.; Hornsten, E. G.; Sundgren, H.; Lundstrom, I. *Meas. Sci. Technol.* **1993**, *4*, 1493-1500.

(25) Di Natale, C.; Brunink, J. A. J.; Bungaro, F.; Davide, F.; D'Amico, A.; Paolesse, R.; Boschi, T.; Faccio, M.; Ferri, G. *Meas. Sci. Technol.* **1996**, *7*, 1103-1114.

(26) Berna, A. Z.; Lammertyn, J.; Saevels, S.; Di Natale, C.; Nicolai, B. M. *Sens. Actuators, B* **2004**, *97*, 324-333.

(27) Hines, E. L.; Llobet, E.; Gardner, J. W. *Electron. Lett.* **1999**, *35*, 821-823.

(28) Gao, T.; Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Chem. Mat.* **2006**, *18*, 5193-5202.

(29) Severin, E. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 2008-2015.

(30) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658-668.

(31) Zellers, E. T.; Pan, T.-S.; Patrash, S. J.; Han, M.; Batterman, S. A. *Sens. Actuators, B* **1993**, *12*, 123-133.

(32) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Manuscript in preparation*.

(33) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *82*, 54-74.

(34) Jolliffe, I. T. *Principal Component Analysis*, Second ed.; Springer: New York, 2002.

(35) Brereton, R. G. *Chemometrics - Data Anlysis for the Laboratory and Chemical Plant*; John Wiley & Sons Ltd: West Sussex, England, 2003.

(36) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177-4190.

(37) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*; Wiley-VCH: New York, 1999.

(38) Ferre, L. *Comput. Stat. Data Anal.* **1995**, *19*, 669-682.

(39) Kaiser, H. F. *Educ. Psychol. Meas.* **1960**, *20*, 141-151.

(40) Jolliffe, I. T. *Appl. Statist.* **1972**, *21*, 160-173.

(41) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*, 5th ed.; Prentice Hall: Upper Saddle River, NJ, 2002.

(42) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.

(43) The normalized response from any sensor in the array would be zero or one for the unresponding and responding sensors, respectively, with only a single responding sensor per analyte. Therefore, for comparison between any two vapor classes, σ_1 and σ_2 from eq (2) would be zero.

(44) Lawson, C. L.; Hanson, R. J. *Solving Least Square Problems*; Prentice Hall: Englewood Cliffs NJ, 1974.

(45) Park, J.; Groves, W. A.; Zellers, E. T. *Anal. Chem.* **1999**, *70*, 4191-4201.

(46) Jin, C.; Kurzawski, P.; Hierlemann, A.; Zellers, E. T. *Anal. Chem.* **2008**, *80*, 227-236.

(47) Shaffer, R. E.; Rose-Pehrsson, S. L.; McGill, R. A. *Anal. Chim. Acta* **1999**, *384*, 305-317.

(48) Wold, S. *Patt. Rec.* **1976**, *8*, 127-139.

(49) Hsieh, M.; Zellers, E. T. *Anal. Chem.* **2004**, *76*, 1885-1895.

(50) Zellers, E. T.; Batterman, S. A.; Han, M.; Patrash, S. J. *Anal. Chem.* **1995**, *67*, 1092-1106.

Table 5.1: Organic molecule/carbon black composite sensors used in this study. CB = carbon black; DP = dioctyl phthalate; THF = tetrahydrofuran.

suspension	sensor materials	solvent
1	35 mg lauric acid, 15 mg DP, 150 mg CB	THF
2	50 mg propyl gallate, 150 mg CB	THF
3	50 mg quinacrine dihydrochloride dihydrate, 150 mg CB	chloroform
4	35 mg tetracosane, 15 mg DP, 150 mg CB	toluene

Table 5.2: Analytes and their abbreviations used in this work.

Analyte	
Hx	hexane
Bz	benzene
Oc	octane
Dc	decane
MeOH	methanol
EtOH	ethanol
PrOH	1-propanol
BuOH	1-butanol
Cf	chloroform
EA	ethyl acetate
Cb	chlorobenzene

Figure 5.1: a-d: S^2 vs. n_{PC} for various 10 analyte training libraries based on the a) variance (60%), b) variance (80%), c) eigenvalue (≥ 1), and d) eigenvalue (≥ 2) methods. e) S^2 vs. the minimum *rf* in the data set. SS analysis using NNLSQ (SS, LSQ), ST analysis using NNLSQ (ST, LSQ) and ST analysis using EDPCR (ST, PCR), are indicated by a black ×, red *, and blue ∇, respectively.

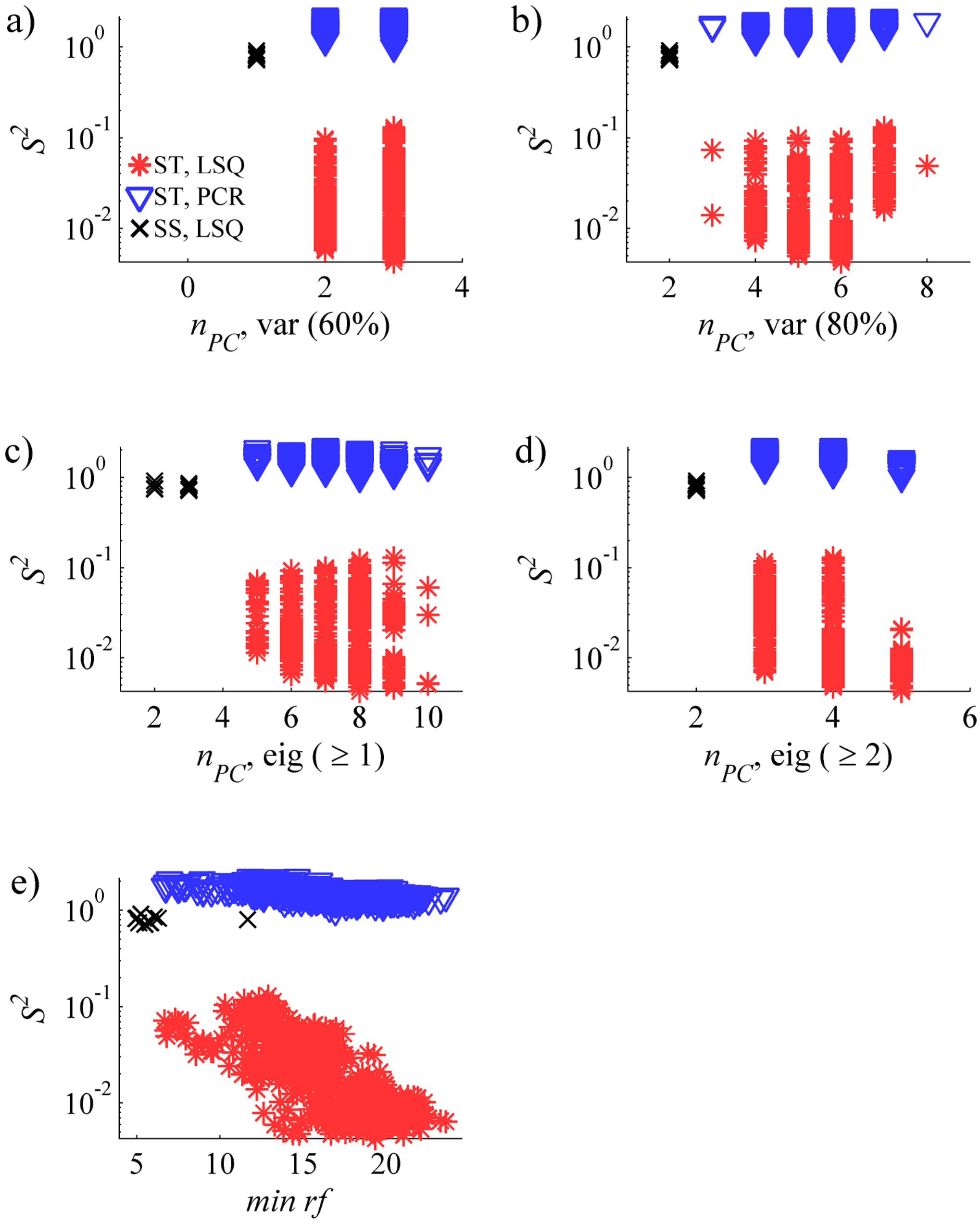


Figure 5.2: NNLSQ S^2 values vs. n_{PC} using the variance (80%) method for 5, 7, and 9-analyte training libraries for a-c, respectively. SS and ST data are indicated by a black × and red *, respectively.

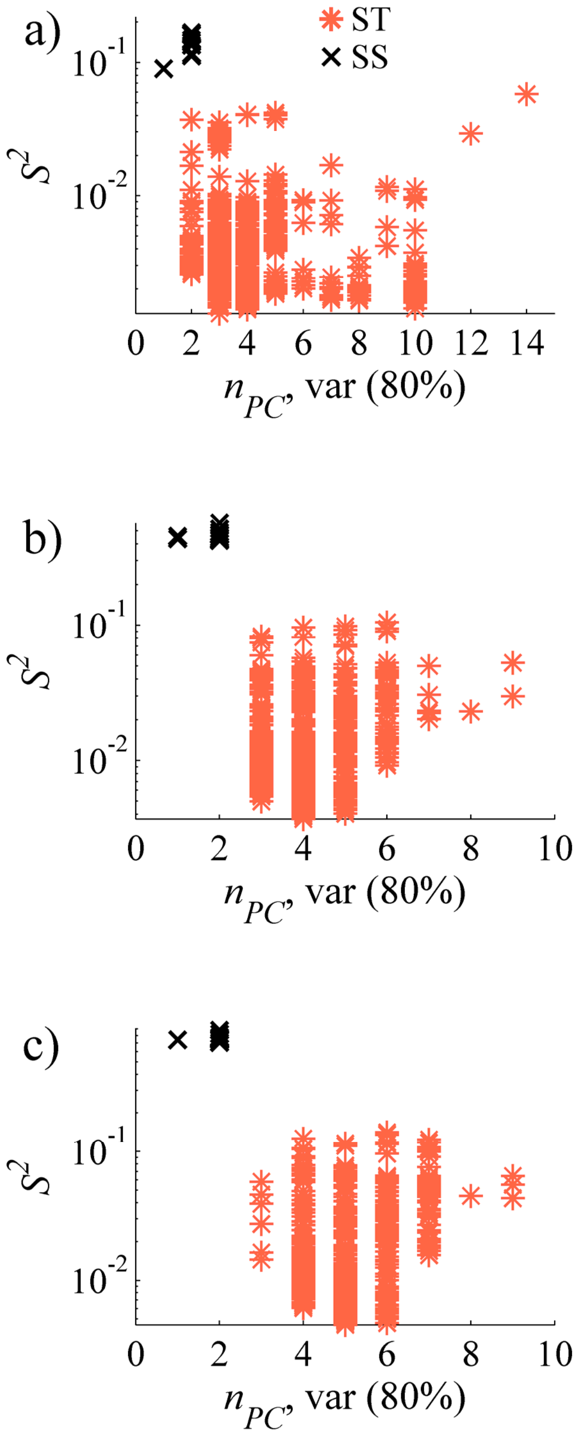


Figure 5.3: NNLSQ S^2 values vs. n_{PC} using the eigenvalue (≥ 2) method for 5, 7 and 9-analyte training libraries for a-c, respectively. SS and ST data are indicated by a black × and red *, respectively.

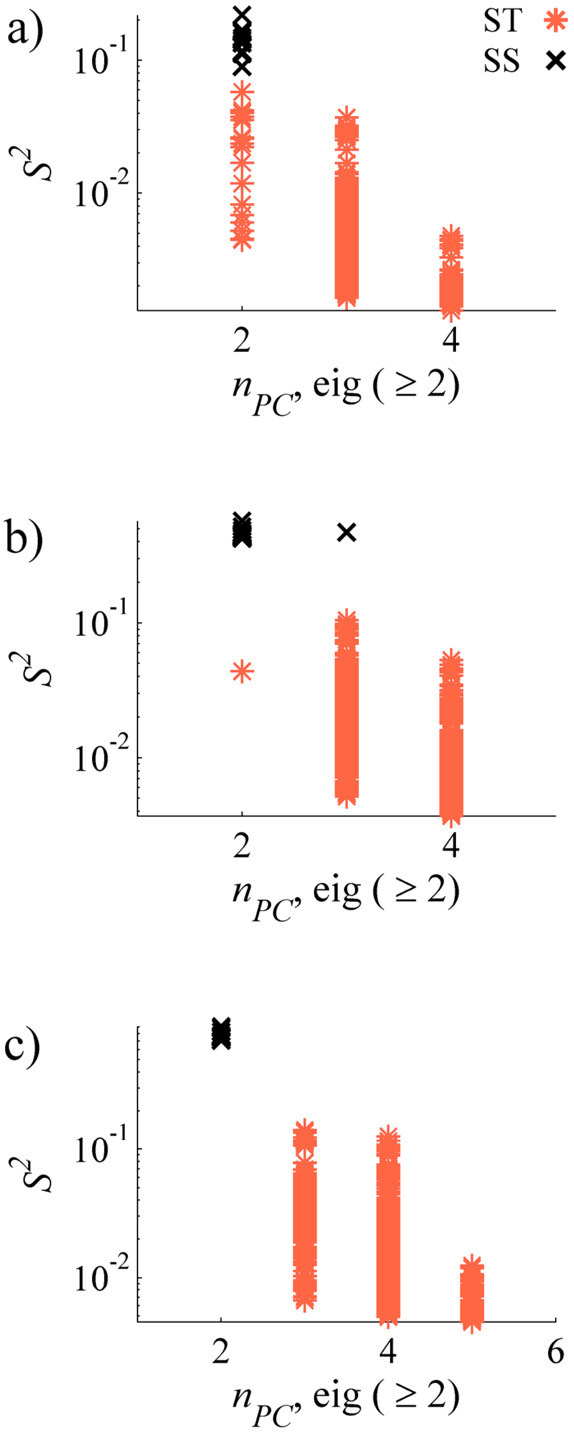


Figure 5.4: NNLSQ S^2 values vs. the minimum *rf* amongst pure analyte vapors in 5, 7 and 9-analyte training libraries for a-c, respectively. SS and ST data are indicated by a black × and red *, respectively.

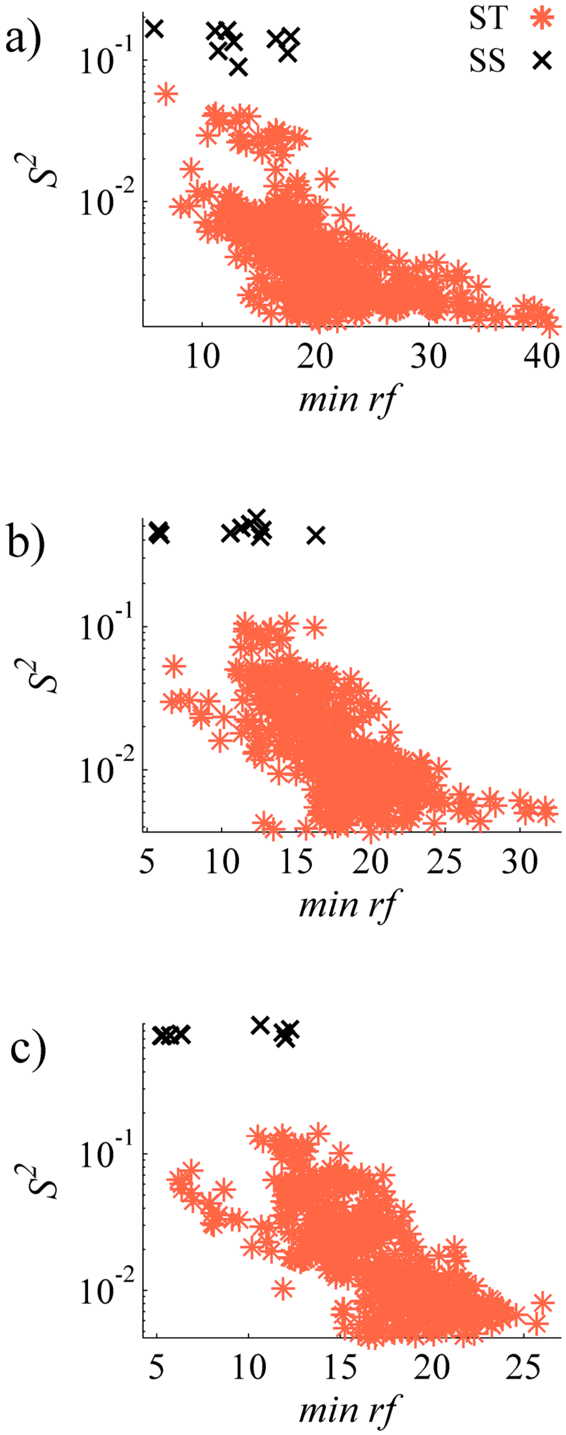


Figure 5.5: NNLSQ S^2 values vs. a) n_{PC} using the variance (80%) method; b) n_{PC} using the eigenvalue (≥ 2) method; and c) the minimum *rf* in the vapor library. Noise levels added to data of 1×10^{-4}, 4×10^{-4}, and 16×10^{-4} are indicated by ○, ∇, and □, respectively. Vapor libraries consisted of 10 analytes, with three different levels of noised superimposed on sensor responses.3

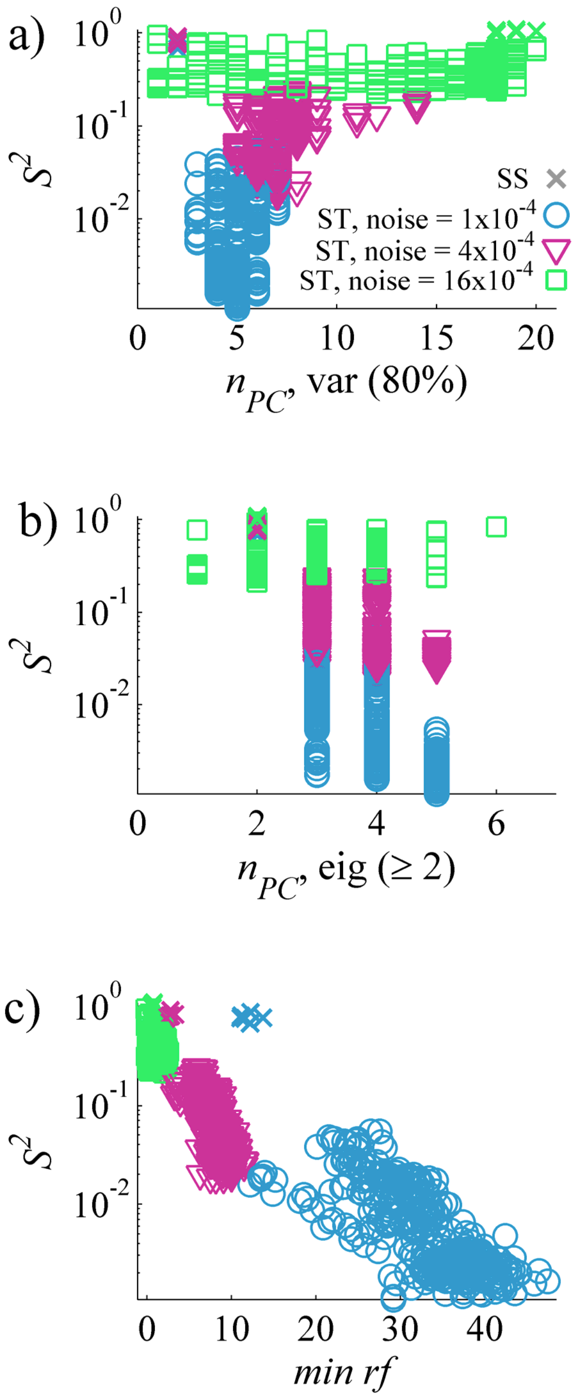


Figure 5.6: Analysis of a 3-component mixture simulated from a 6-component analyte library, with three levels of Gaussian noise superimposed to generate training and testing responses, analyzed using a) NNLSQ; b) EDPCR analysis allowing for a fit to all combinations of the modeled mixture; c) EDPCR analysis allowing a fit to all combinations of the 6 analyte library; d) PCR-SF. Means and standard deviations were calculated from 20 mixture analysis repetitions. Noise levels of 2×10^{-6}, 2×10^{-5}, and 2×10^{-4} were used.

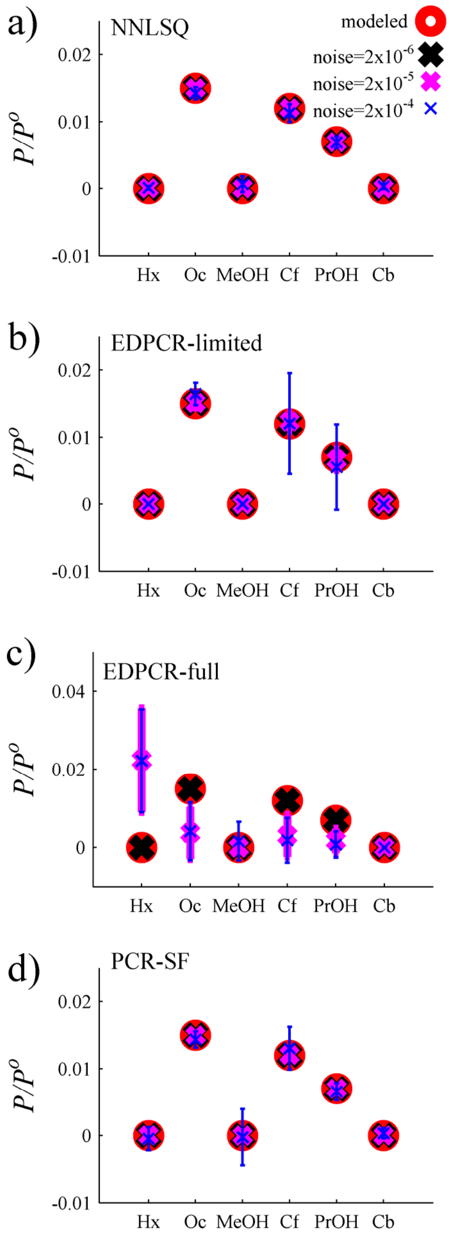


Figure 5.7: Analysis of a 4-component mixture simulated from a 6-component analyte library. See Figure 5.6 for descriptions.

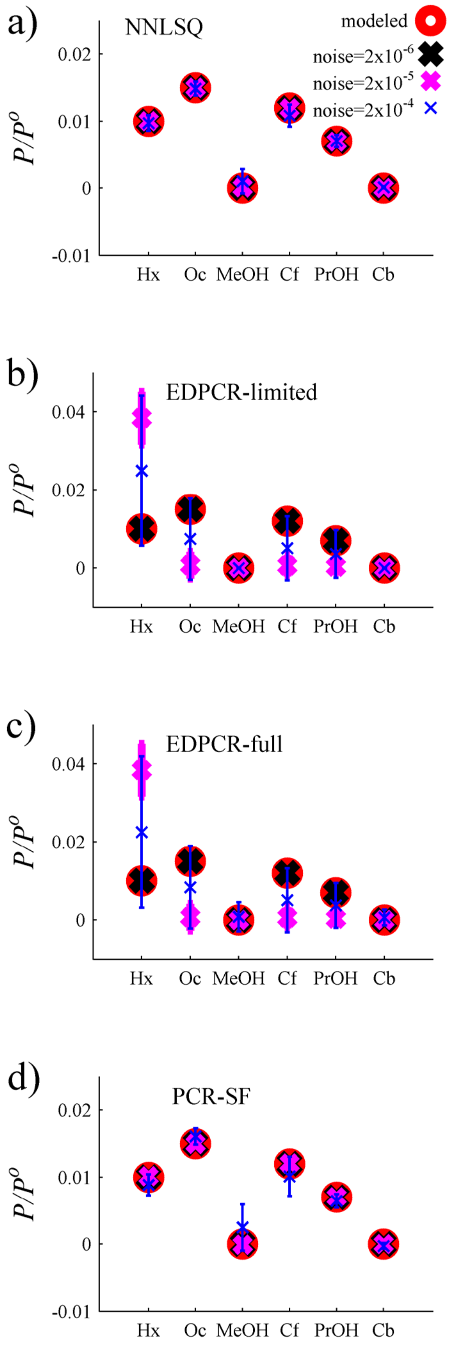
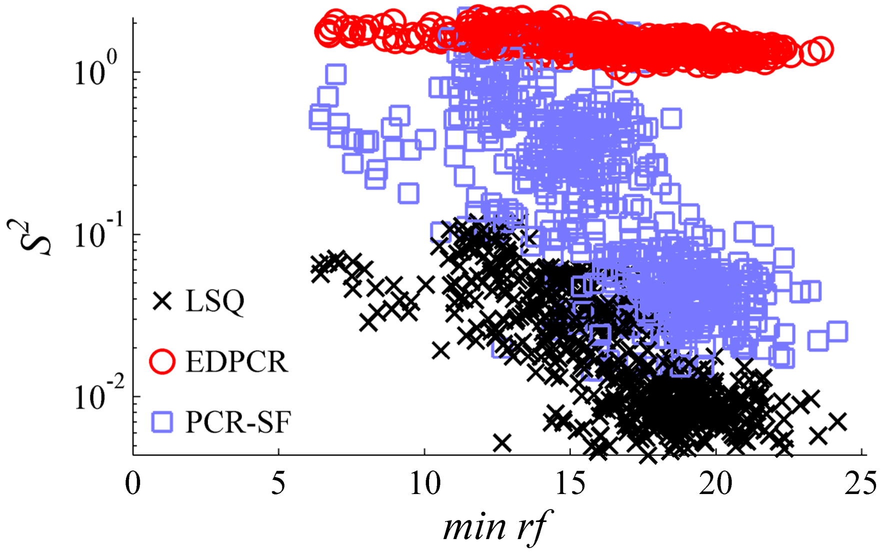


Figure 5.8: S^2 analyzed using NNLSQ (LSQ), EDPCR, and PCR-SF, vs. the minimum *rf* amongst analyte vapors in various 10-analyte training libraries. Analysis using the NNLSQ, EDPCR-full, and PCR-SF methods are indicated by the black ×, red ○, and blue □, respectively.



Chapter 6

Enhancement of Pure Vapor Classification Rates Using Spatiotemporal Response

6.1. Abstract

Spatially linear sensor arrays made from small molecule/carbon black composite thin film chemiresistors were used to discriminate between and classify among the vapor classes of alcohols (methanol, ethanol, 1-propanol, 1-butanol, and 1-hexanol), alkanes (hexane, heptane, octane, nonane, and decane), chloromethanes (dichloromethane, trichloromethane, and tetrachloromethane), and xylenes (m-xylene, p-xylene, and o-xylene). Traditional steady-state (SS) sensor responses, and space- and time- dependent spatiotemporal (ST) sensor responses, were extracted from the same sensor arrays, and their ability to discriminate and classify among the vapors was compared. Using Fisher's linear discriminant (FLD) and ST response information, for 109 of the 120 binary pure vapor combinations, an increased ability to discriminate between the vapors, denoted by the testing resolution factor (rf_{test}), was observed. Three classification techniques were used: *k* nearest neighbors (*k*NN), a feed-forward back-propagation neural network (FFBNN), and a bioinspired hierarchical classification (HC) scheme based on FLD, FLD-HC. FLD-HC exhibited equal or better classification rates than *k*NN for 14 out of 16 and 15 out of 16 of the vapors when SS and ST response information was used, respectively. FLD-HC yielded zero misclassifications of a target vapor as belonging to a different chemical class, for example classifying hexane as an alcohol, while *k*NN and NN analyses yielded 4 out of 800 and 9 out of 800 wrong-class misclassifications, respectively. Non-Gaussian response distributions for ST response data caused the two linear methods, *k*NN and FLD-HC, to exhibit equal or better correct classification rates for 15 out of 16 and 13 out of 16 of the vapors, respectively, when ST response information was used. In contrast, for all 16 vapors, increased classification rates were observed when ST response information was used with the non-linear FFBNN. Classification rates were well-correlated with the FLD rf_{test} value. When the

minimum rf_{test} between a target vapor and all other vapors in the library was > 8.5, > 95% correct classification rates were achieved for the target vapor. Alternatively, when the rf_{test} between the target vapor and an alternate vapor was < 6, the likelihood of the alternate vapor being wrongly classified as the target vapor increased significantly.

6.2. Introduction

Broadly responsive array-based vapor sensing has received significant attention due to its ability to detect and discriminate between various analyte vapors. Array-based sensing has been explored using chemoresponsive dyes,[1,2] tin oxides,[3,4] gold nanoparticles,[5,6] metal-oxide-semiconductor field-effect transistors,[7,8] optical fibers or beads coated with thin polymer films,[9,10] polymer-coated surface acoustic wave devices,[11,12] polymer-coated micro-machined cantilevers,[13,14] intrinsically conductive polymers,[15,16] and conductive polymer composite chemiresistors.[17,18] In all cases, each sensor responds to a large number of vapors, and the individual sensors possess essentially no classification ability. The array response captures the relative response magnitudes of all sensors in the array, and when array training is combined with appropriate pattern recognition techniques, the sensor arrays can provide vapor classification and identification. Due to similarities to mammalian olfactory processes, such architectures have been termed "electronic" or "artificial" noses.[19,20]

One attractive signal transduction mode incorporates the use of chemically sensitive resistors, or chemiresistors. Such sensors are inherently low power,[17] compact,[20] compatible with very large-scale integration (VLSI) processes,[4] can be deposited onto a variety of substrates including glass,[21] ceramics,[17] or interdigitated electrodes,[5] and can be fabricated in a variety of form factors for the optimization of the sensor response.[22,23]

One chemiresistive sensing approach employs carbon black composites (CBC), where carbon black (CB) serves as the conductive phase, and either polymers or small, non-polymeric, organic materials serve as the insulating phase.[17,18,20-22,24] In these CBC sensors, the insulating material in each sensor determines the sensor/vapor selectivity and sensitivity, while the CB offers an inexpensive method for transducing the chemical response information into an electrical resistive response signal. When several insulating materials are used across an array of sensors, the array is able to discriminate between different vapors or vapor mixtures.[17,18,20-22,24]

Recent work has shown that by employing CBC vapor sensors in a linear spatial arrangement, with small volumes overhead of the sensor arrays and with low vapor flow rates, the sensor films act similar to packed columns in gas chromatography, serving to time-separate responses at different sensors along the array.[22] This arrangement thus creates a space- and time- dependent, or spatiotemporal (ST), response pattern across each sensor array. The use of multiple sensor arrays in ST mode has produced an increased ability to analyze vapor mixtures, with training on only the pure vapor components.[24] ST responses have also been modeled, and optimized operational conditions for mixture analysis have been defined based on two dimensionless numbers that relate the chamber geometries, vapor delivery flow rate, and physical properties of the vapors.[23]

While the benefits of the ST approach to vapor sensing have been clearly demonstrated for the analysis of vapor mixtures, they have not yet been demonstrated for classification of pure vapors. In this work we have recorded the ST responses from linear sensor arrays, to evaluate whether enhanced recognition rates of pure vapors are produced using this method relative to the same sensors exposed simultaneously to the vapor stream using traditional, steady-state (SS) response information. Arrays were exposed to vapors in four vapor classes: alcohols (methanol, ethanol, 1-propanol, 1-butanol, and 1-hexanol), alkanes (hexane, heptane, octane, nonane, and decane), chloromethanes (dichloromethane, trichloromethane, and tetrachloromethane), and xylenes (o-xylene, m-xylene, and p-xylene). Within these sets, the vapors possess the same chemical functionalities. Discriminating vapors from within such classes is among the most difficult pure vapor classification tasks,[21,25] and the utility of ST response information for pure vapor classification was evaluated on such a set of tasks. A biologically inspired hierarchical classification scheme based on Fisher's Linear Discriminant (FLD), FLD-HC, is introduced, and its advantages and disadvantages are discussed, relative to the alternative *k*-nearest neighbors (*k*NN) and feed-forward back-propagation neural network (FFBNN) classification algorithms. Additionally, this work demonstrates the utility of the FLD resolution factor (*rf*) for predicting where analyte misclassification errors will be made.

6.3. Experimental

6.3.1. Materials

The insulating materials for the fabrication of the sensor films consisted of tetracosane (99%), tetracosanoic acid (99%) and dioctyl phthalate (DP, 99%), purchased from Aldrich; and propyl gallate (98%) purchased from Acros Organics. Reagent grade toluene and tetrahydrofuran, received from Aldrich, were used as solvents in the sensor suspensions. n-hexane (Hx, 95%), n-heptane (Hp, 99%), n-octane (Oc, 98%), n-nonane (Nn, 99%), n-decane (Dc, 99%), methanol (MeOH, 99.8%), ethanol (EtOH, 95%), 1-propanol (PrOH, 99.5%), 1-butanol (BuOH, 99.9%), 1-hexanol (HxOH, 99%), and tetrachloromethane (CCl_4, 99.9%) were purchased from Aldrich, and dichloromethane (CH_2Cl_2, 99%), trichloromethane ($CHCl_3$, 99.5%), m-xylene (m-xyl, 99%), p-xylene (p-xyl, 99%), and o-xylene (o-xyl, 99%) were purchased from Fluka. These pure solvents were used to generate the vapors for delivery to the sensor arrays. Black Pearls 2000, a furnace carbon black material donated by Cabot Co. (Billerica, MA), was used as the conductive phase in the sensor composites. All materials were used as received

6.3.2. Detector Fabrication

Four suspensions, each composed of a nonconductive sorption phase and a conductive CB phase (Table 6.1), were used to fabricate the sensors used in this work. Each sensor suspension consisted of 50 mg of sensor material supported in 20 ml of solvent. 12.5 mg of the nonconductive sorption phase was added to 20 ml of either tetrahydrofuran or toluene and sonicated for > 10 min. Next, 37.5 mg of CB was added to each solution and the suspension was sonicated for > 30 min.

Four sensor array substrates were fabricated and arranged as described previously.[24] Briefly, 30 nm of chromium and then 70 nm of gold were evaporated onto glass microscope slides (75 × 25 mm) using a custom-made mask to create an array of 15 pairs of sensor electrodes along the length of the slide.[24] Each sensor pair was evenly spaced every 5 mm, with a gap of 0.4 mm present between each pair of electrodes forming the basis for each individual sensor.

Each sensor array substrate was masked with Teflon tape and sprayed along the entire 15 electrode pair length with a single sensor suspension, until the resistance of all 15 sensors was between 1 and 3 kΩ. This process created a sensor film area of 75 mm × 5 mm. The procedure was repeated for each of the four sensor materials listed in Table 6.1, providing a total of four different linear sensor arrays. These arrays were placed into a custom-made aluminum sensor chamber that has been described previously,[24] with the vapor flow channels coated with Teflon tape. After sensor spraying and chamber assembly, the arrays were randomly exposed to the alkanes for a 24 h period.[26] One week passed between the initial sensor spraying and the collection of train/test sensor response data.

6.3.3. Vapor Generation and Delivery

An automated flow system controlled by LabVIEW 5.0 software was used to deliver pulses of diluted streams of solvent vapor to the detectors.[18,24] The carrier gas was obtained from the house compressed air source and controlled with a mass flow controller (MFC) (UNIT) that could vary the flow from 6 to 625 ml min^{-1}. To generate analyte vapors, a foreground stream of carrier gas was passed through a porous glass frit and bubbled through a 220 ml bubbler filled with the desired solvent. The foreground stream was controlled by another MFC (UNIT) that could vary the flow from 6 to 625 ml min^{-1}. The height of solvent was essentially the same (42 cm) before and after each set of exposures.

The analytes used to generate vapors, and abbreviations used in the remainder of this work, are listed in Table 6.2. Four sets of analyte exposures were performed: alkanes (Hx, Hp, Oc, Nn, Dc), alcohols (MeOH, EtOH, PrOH, BuOH, HxOH), xylenes (m-xyl, p-xyl, o-xyl), and chloromethanes (CH_2Cl_2, $CHCl_3$, CCl_4). Vapor concentrations were denoted as P/P^o, where P is the partial pressure of the analyte and P^o is the vapor pressure of the analyte.[27] Vapors were delivered to the sensor arrays at $P/P^o = 0.050$, or 5 parts of saturated vapor flow combined with 95 parts of background air. During each set of analyte exposures, each vapor was delivered to the sensor chamber 150 times; this corresponded to 750 total exposures for the alkanes and alcohol exposures, and 450 total exposures for the chloromethane and xylene exposures. Within each set of exposures, the vapors were delivered in random order. Each set of exposures occurred within a 72 h period. All sets of analyte exposures occurred within a 5 week period.

A total flow rate of 85 ± 3 ml min^{-1} was delivered to the collective sensor chamber, which was divided into 8 symmetric flow channels.[27] This produced a flow rate of 11 ml min^{-1} for each linear sensor array channel (note: only four of the eight individual flow channels contained a sensor film). To more reproducibly deliver vapors, vapors were initially generated at a flow rate of 260 ml min^{-1}, and a small Teflon-lined sampling pump (Science Pump Corporation) was used to withdraw 175 ml min^{-1} from this flow stream. Flow meters (Gilmont) were used to calibrate the vapor delivery system.

Each vapor exposure consisted of four steps: 1) 20 s of background air flow at 85 ml min^{-1}, with no data collection, to re-establish flow dynamics throughout the vapor channels; 2) 50 s of background air flow at 85 ml min^{-1}, with data collection, to define the pre-exposure sensor resistances; 3) 120 s of vapor delivery at $P/P^o = 0.050$ and a flow rate of 85 ml min^{-1}, with data collection, to capture the sensor response; and 4) 150 s of background air flow at 350 ml min^{-1}, with no data collection, to ensure the removal of all analyte from the sensor films.

6.3.4. Measurements and Data Processing

The data collection arrangement has been described previously.[27] Sensor film resistances were measured using a Keithley 2002 multimeter and a Keithley 7001 multiplexer. Each sensor substrate was connected to the multiplexer through shielded, twisted pair cables and a rotary ZIF connector (Tyco Electronics). Two Keithley 7001/2002 combinations, controlled with LabVIEW, were used to increase the sensor sampling frequency. Each 7001/2002 combination recorded sensor resistances from two of the four arrays. Each sensor was sampled approximately every 3 s.

The sensor response was expressed as $S(t) = \Delta R(t)/R_b(t)$, where $\Delta R(t)$ is the time-varying, baseline-corrected, resistance change upon exposure to the analyte, and $R_b(t)$ is the baseline-corrected resistance of the sensor in the absence of analyte.[24] To perform the baseline correction, a line was fit to resistance values obtained during the 50 s prior to analyte delivery. The slope of the line was calculated and extrapolated from the point immediately prior to vapor delivery, to calculate the baseline drift-adjusted resistance at any time t, $\Delta R_b(t)$. $\Delta R(t)$ was calculated by subtracting $R_b(t)$ from the measured sensor resistance at time t.[24]

The sensor responses using ST and SS information were then used to compare the two approaches for their respective ability to correctly classify pure vapors. For ST response information, the responses from the first, middle, and last sensors (positions 1, 8 and 15) along each 15 sensor array were used and sampled at 10 s and 105 s into the vapor delivery.[24] For SS response information, the responses from the first, second, and third sensors (positions 1, 2 and 3) along each 15 sensor array were used and were sampled 100 s and 105 s into the vapor delivery.[24] This approach provided each of the two methods with 24 total response descriptors (4 arrays × 3 sensors/array × 2 times/sensor), allowing for a direct comparison between the two methods. The actual times that the resistances were recorded varied between all exposures and along the sensors in the array. The responses at the times used in the data analysis were therefore calculated by linearly interpolating between the measured data points.

To ensure that the sensor arrays discriminated between vapors because of differences in response patterns, rather than perceived concentration, a normalization procedure was used for each vapor exposure.[21,24,26] This procedure divided each response descriptor by the sum of the array response descriptors for that exposure:

$$S(t)'_j = \frac{S(t)_j}{\sum_{j=1}^{n} S(t)_j}, \tag{1}$$

where $S(t)_j$ is response descriptor j, and $n = 24$ is the total number of response descriptors.[21,24,26] This normalization procedure takes advantage of the sensor's linear response,[21] creating a concentration-independent feature vector for each pure vapor. Array responses for all exposures were normalized according to eq (1). All response manipulation was performed in MATLAB.

6.3.5. Cluster Analysis: Fisher's Linear Discriminant

Fisher's linear discriminant (FLD) was used on response data normalized by eq (1) to find the dimension of optimal separation between the clusters of all of the binary combinations of vapors. The dimension of optimal separation is defined as the direction that maximizes the resolution factor, *rf*, between the two response clusters:

$$rf = \frac{\overline{y_1} - \overline{y_2}}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{2}$$

where $\overline{y_1}$ and $\overline{y_2}$ are the means, and σ_1 and σ_2 the standard deviations, respectively, of the projected values of vapor clusters 1 and 2 onto the optimal dimension.[28,29] Identification of this optimal dimension produces a projection vector that transforms the original 24-dimensional data onto a single dimension. A train/test scheme was employed for all binary combinations, to ensure that results were not subject to overfitting of the data. The first 100 randomized exposures for each analyte served as training data. For each binary vapor combination, training data were used to calculate the optimized projection vectors based on eq (2). These projection vectors were then used to project the testing data (the last 50 randomized exposures to each analyte) for the same two vapors onto the dimension of optimal separation. The population statistics of projected testing clusters were then calculated for the two classes of data: the projection means $\overline{y_1}$ and $\overline{y_2}$, and the projection standard deviations σ_1 and σ_2, for vapors 1 and 2, respectively. The population statistics of the test data were then used to calculate the testing *rf* metric, rf_{test}. This procedure was repeated for all combinations of pure vapors. All of the reported *rf*'s were calculated from testing data. FLD was performed in MATLAB using a custom-written code.

6.3.6. Vapor Classification

To gauge the ability of the sensor arrays to correctly identify vapors, three classification techniques were used. Two linear methods, *k*-nearest neighbors (*k*NN) and a FLD hierarchical classification (FLD-HC) scheme, as well as a nonlinear feed-forward back-propagation neural network (FFBNN), were used. In all cases, the first 100 exposures to each vapor served as training data to generate a model, while the last 50 exposures to each vapor served as testing data for the model validation. All classification schemes were performed in MATLAB.

6.3.6.1. *k* Nearest Neighbors (*k*NN)

The *k*NN algorithm calculates the distance between the response of a single test analyte and the responses from all of the training data from all of the vapors. The algorithm classifies unknown vapors based on which vapor class makes up the majority of the *k* closest training points, or the *k* nearest neighbors.[30,31] The Euclidean distance was used, which in two or three dimensions is the linear distance between two points. When choosing from more than two classes, ties are possible for any integer value of $k > 1$. Additionally, regardless of sample distribution type, no integer value of $k > 1$ performs better than the single neighbor situation of $k = 1$.[31] For these reasons, $k = 1$ was used. Training responses were normalized according to eq (1), and provided a 24 dimensional map in response space. Each normalized testing response was placed within this map, and was classified as belonging to the closest vapor.

6.3.6.2. FLD Hierarchical Classification (FLD-HC)

The signals generated from the olfactory bulb evolve with time.[32] Specifically, response patterns generated from the olfactory receptor proteins are continuous throughout vapor exposures and responses, while their expression to the brain is modulated in a vapor-specific way by the mitral cells, with the initial mitral cell expressions providing information optimal for initial class groupings, and the evolved mitral cell expressions providing finer information optimal for identification within classes.[32,33] The FLD-HC approach applies these principals to the hierarchical classification of vapors.

A series of binary classification steps was employed until an unknown vapor was classified. Each step used a different FLD model that weighed the sensor responses according to their ability to optimally separate between the classes. Each of these FLD models allowed for an optimal expression of the sensor array response for each step of vapor classification. The different sensor weightings were captured in the different Fisher projection vectors at each binary step, which projected the 24 dimensional array response onto a single optimized dimension for each cluster separation. The evolution of the projection vector used at each binary step, which continually altered and optimized the expression of different sensors along the array, is analogous to the evolution of olfactory receptor protein expression performed via mitral cells for subsequent steps of vapor classification.

The first binary classification step was to break down the four-class classification problem into 2 two-class classification problems. A model was generated using normalized training data, in which class 1 consisted of all alcohols and chloromethanes, and class 2 consisted of all alkanes and xylenes (800 total training exposures each class). Class 2 included all vapors that consisted of only carbons and hydrogens, while class 1 included vapors that possessed additional functionality. From the training data, a decision boundary was generated:

$$D = \overline{y_1} - \frac{rf \times \sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{3}$$

where *rf* and all other variables were calculated from the training projections and are defined in eq (2). For clusters with Gaussian distributions, this decision boundary was the point along the one-dimensional optimized separation direction at which a response had an equal likelihood of belonging to either class 1 or class 2. For classification, normalized test exposures were then projected onto this optimized dimension. Exposures were classified according to which side of the D-boundary they were projected onto.

Once classified as either an alcohol/chloromethane or an alkane/xylene, a series of tests were then performed to further classify exposures within the group designation. For the former case, classification proceeded with class 1 defined as alcohols and class 2 defined as chloromethanes; for the latter case, class 1 was defined as alkanes and class 2 as xylenes. For both cases, a model was generated in which class 1 and class 2 used 500 and 300 training exposures, respectively. The unknown exposure was projected onto the optimal dimension for separation, and was classified as belonging to one of the four vapor classes.

When the vapor class of the unknown was determined, the final classification scheme grouped the lightest member (lowest molecular weight) within each class against all other heavier within-class members. For alkane classification, a model was generated in which class 1 initially consisted of 100 training exposures to hexane, and class 2 initially consisted of 100 training exposures to each of heptane, octane, nonane and decane (400 total exposures). A new projection vector was calculated by finding the dimension that maximized eq (2) for these training exposures. The unknown alkane was then projected onto this new separation dimension. If the projection fell on the hexane side (class 1) of the newly calculated D-boundary (eq (3)), classification was complete. If the projection fell on the class 2 side of the D-boundary, classification continued, and a new model was constructed in which all heptane training exposures (100 exposures, the new class 1) were compared against all octane/nonane/decane training exposures (300 exposures, the new class 2). If necessary, classifications were continued in this manner until a single binary classification task of nonane vs. decane was achieved, and a decision was made. Alcohols, chloromethanes, and xylenes were classified in an analogous manner.

6.3.6.3. Feed-Forward Back-Propagation Neural Networks (FFBNN)

Artificial neural networks, modeled after their biological counterparts, are constructed by connecting and training artificial neurons, or nodes.[34,35] The user chooses the number and nature of the hidden nodal transfer functions, for example linear, sigmoidal, etc. Training data are then used to adjust the transfer functions, and force the desired outputs for different classes of interest. Once the model is developed, it can be applied to unknown testing data that is classified according to the neural network output.

A feed-forward back-propagation neural network (FFBNN) was used.[34,35] One input was used for each sensor, for a total of 24 input neurons. Two hidden layers were used. The first hidden layer consisted of 24 nodes, with each node utilizing a tangent-sigmoidal transfer function. The second hidden layer consisted of a single node that employed a linear transfer function. Normalized training exposures for each of the analytes, and their respective class identities, were input to the neural network for training. Transfer functions were optimized over 500 training iterations, decreasing the classification error rate for the training data in each iteration. Testing exposures were then fed to the FFBNN, and the data were classified according to the network output.

6.4. Results

6.4.1. Sensor Response

Figure 6.1 shows responses from sensors 1, 2, 3, 8 and 15 along a 15 sensor array of tetracosane/CB composite (sensor 3, Table 6.1) during exposure to a series of alkanes at P/P^o = 0.050. The responses of the individual sensors were scaled by their respective responses at t = 100 s to aid in visual comparison. The time scale was adjusted so that the first sensor along the array exhibited its first response at t ~ 0 s. Different response patterns were clearly observed for each of the five alkanes.

As alkane chain length increased, the analyte vapor pressure decreased, and the sensor/analyte partition coefficients increased.[22,36] The increased partition coefficient caused an increase in the analyte uptake by the sensor array. This in turn caused an analyte-dependent time delay in the response along the sensors in the array, generating the desired ST response effects. For example, on exposure to hexane, having the highest vapor pressure and lowest partition coefficient of the alkanes studied, no response separation was observed between the first three sensors along the array. In contrast, as the alkane chain length and partition coefficient increased, the responses from the first three sensors along the array become more separated.

The ST response analysis utilized sensors 1, 8 and 15, and their responses were calculated at t = 10 s and 105 s to describe the array response. The SS response analysis utilized sensors 1, 2 and 3, and calculated their responses at t = 100 s and 105 s. The SS analysis clearly sampled the steady-state response from the sensors. In contrast, the ST analysis sampled a transient response in addition to the steady-state response. From the figures, without any information on sensor response magnitude, one is clearly able to discriminate between the different alkanes to various degrees. Similar, time-dependent, ST response profiles were observed for the alcohols, chloromethanes, and xylenes.

6.4.2. Analyte Discrimination: FLD

Table 6.3 lists the rf_{test} values for all binary vapor combinations. The within-class combinations for the four vapor classes are highlighted along the diagonal. In all cases, the bottom number in parenthesis lists the rf_{test} value obtained using SS response information, whereas the top number lists the rf_{test} value obtained using ST response information. For Gaussian projection clusters, rf_{test}'s of 1, 2 and 3 correlate with the ability to correctly classify class A from class B 72, 92 and 98% of the time, respectively, along the optimized dimension of separation.[18]

In 109 of the 120 binary combinations, the rf_{test} value increased when ST rather than SS response information was used. The average percentage decrease in rf_{test} values for the remaining 11 cases was 8%, ranging from a decrease of 0.8% to 25%. In contrast, the average percentage increase for the 109 cases was 60%, ranging from an increase of 5% to 232%. The decreased rf_{test} values can be attributed to non-Gaussian projections along the optimized dimension for the ST response data. The SS analyses sampled the steady-state response of each sensor, which exhibited a Gaussian distribution. In contrast, the ST analyses sampled a transient response, which exhibited a non-linear rise profile at different stages of the response (Figure 6.1), thus leading to a non-Gaussian distribution of sensor responses. These non-Gaussian distributions caused σ_1 and σ_2 in eq (2) to increase, leading to a reduction in the reported rf_{test} values for some of the cases. Importantly, the increases in rf_{test} values that resulted from the use of the ST response information greatly outweighed the decreases in rf_{test} values that resulted from difficulties encountered in extraction of the ST response data.

When ST rather than SS data was used, marked increases were observed for discrimination between EtOH and PrOH (rf_{test} increased 232% from 2.2 to 7.3), Hp and Oc (rf_{test} increased 110% from 2.8 to 5.9), and Oc and Nn (rf_{test} increased 130% from 1.4 to 3.2). These three cases represented discrimination between extremely similar analytes, for which the only difference was the presence of an extra carbon along the alkane's length. Extremely similar chemistries were present, and use of the SS response information produces rf_{test} values < 3, indicating a rapid degradation in the ability to classify between the two response clusters. In contrast, the use of ST response information produced rf_{test} values > 3 for all cases, yielding $> 98\%$ classification ability along the optimized dimension for these Gaussian distributed clusters. These increases in rf_{test} offer the potential for significant improvements in classification rates for many of the difficult tasks shown in Table 6.3.

6.4.3. Analyte Classification: *k*NN, FLD-HC, and FFBNN

Tables 6.4-6 show the confusion matrices for all of the classification analyses. For each confusion matrix, the row heading designates the actual vapor exposed to the array, whereas the number in each column designates the fraction of testing exposures (out of 50) that were predicted to belong to the respective vapor. Hence, perfect classification would produce an identity matrix in this representation of the data. The sum of each row equals 1, accounting for all 50 testing exposure classifications. The bottom (parenthesis) and top number denote the analyses using SS and ST response information, respectively. The within-class confusion matrices are highlighted, shown along the diagonal. The instances in which vapors were misclassified as belonging to the wrong vapor class are shown in bold.

6.4.3.1. *k*NN

Table 6.4 shows the confusion matrix for vapor classification using the *k*NN method with $k = 1$. For 15 of the 16 vapors, increased classification rates were observed when ST rather than SS response data was used. The exception was hexane, which demonstrated one extra misclassification for the ST analysis, or a decrease of 0.02 in classification rate. Significant increases in vapor identification rates were observed using ST rather than SS analyses for ethanol (1.00 vs. 0.82), propanol (0.96 vs. 0.70), octane (0.82 vs. 0.52), and m-xylene (0.84 vs. 0.64). In most cases, classification errors were due to misclassifications within the same vapor class, the few exceptions being chloromethane classifications that were wrongly classified as alcohols, alkanes, or xylenes.

6.4.3.2. FLD-HC

Table 6.5 shows the confusion matrix for vapor classification using the FLD-HC method. For 13 of the 16 vapors, increased classification rates were observed using ST rather than SS analysis. Exceptions were hexane, tetrachloromethane, and p-xylene, each of which exhibited one extra misclassification when ST data was used, or a decrease of 0.02 in classification rate. These decreases in classification performance were insignificant compared to some of the increases. For example, comparing SS and ST analyses, correct classification rates increased from 0.72 to 1.00 for propanol, from 0.86 to 1.00 for ethanol, and from 0.70 to 0.94 for nonane.

In contrast to *k*NN, in no cases were vapors misclassified by FLD-HC as belonging to the incorrect vapor class. Figures 6.2a-c show the projection of all testing exposures onto the initial FLD models used to determine class identity. This approach clearly correctly classified all testing vapors as belonging to their proper vapor class. Figure 6.2a shows the initial classification of the testing vapor responses, setting alkanes/xylenes against alcohols/chloromethanes. The 50 testing exposures to each vapor are shown and labeled, as well as the Fisher decision boundary based on training data (eq (3)). Figures 6.2b-c depict the classification of all testing alcohols vs. chloromethanes, and alkanes vs. xylenes, respectively. In all cases, vapors were projected onto the correct side of the decision boundary, and were always classified as belonging to their correct class. Figure 6.2d details the first classification step of testing alkanes, once they were properly classified as alkanes. Four misclassifications of hexane as heptane, and one misclassification of heptane as hexane, were observed. The ST rf_{test} values for the classification steps detailed in Figures 6.2a-d were 7.4, 9.5, 25, and 3.1, respectively. SS analyses similarly yielded no class misclassifications, and rf_{test}'s for the four situations detailed in Figures 6.2a-d were 5.2, 6.3, 14.9, and 2.9, respectively.

6.4.3.3. FFBNN

Table 6.6 shows the confusion matrix for vapor classification using a FFBNN approach. In all 16 cases, increased classification rates were observed using ST rather than SS data. Use of ST rather than SS data, respectively, increased the correct classification rates from 0.80 to 1.00 for heptane, 0.80 to 0.96 for heptane, 0.86 to 1.00 for propanol, and 0.80 to 0.98 for m-xylene. ST-based FFBNN alcohol classifications were performed at the same high rates achieved by the ST-based *k*NN and FLD-HC methods, while the SS-based FFBNN alcohol classification rates were lower than those obtained using the SS-based FLD-HC method. In several cases, chloromethanes were misclassified as alcohols, alkanes, or xylenes, and in one case butanol was misclassified as p-xylene.

6.5. Discussion

6.5.1. SS vs. ST Data Extraction

The decreased rf_{test} values observed when ST rather than SS response data was used were due to the presence of non-Gaussian sensor response distributions. The measurements of the ST response information required linear interpolation during the transient non-linear section of each sensor's response. Additionally, due to the time required to complete each cycle of resistance measurements, there was a ± 1.5 second window during which the vapor delivery was actually initiated. These complications produced non-Gaussian response distributions and FLD projections, increasing σ_1 and σ_2 from eq (2), and decreasing the reported rf_{test}'s using FLD. These non-Gaussian projections can be seen to some extent in Figure 6.2 for CCl_4 (Figures 6.2a-b), and octane, nonane and decane (Figure 6.2d).

*k*NN and FLD-HC, two linear classification algorithms, depend on linear relationships and Gaussian distributions amongst the response data. Accordingly, the *k*NN and FFBNN method yielded decreased classification rates for 1 out of 16 and 3 out of 16 vapors, respectively, when ST rather than SS data was used. In contrast, neural networks are much better suited for dealing with non-linear data.[37] The ability of the FFBNN to model non-Gaussian responses yielded improved classification rates for all cases when the ST response information was used. In an overwhelming majority of the cases, the benefit of using the ST response information far outweighed the difficulties encountered due to non-Gaussian response distributions.

6.5.2. FLD-HC

The use of the biologically inspired FLD-HC decision tree classification scheme is interesting because typical decision trees form seemingly random branches,[38] but the FLD-HC approach instead requires the user to implement logical branches. The first branch separated the vapors that contained only carbon and hydrogen from the vapors that had additional functionalities, while subsequent branches better defined the vapor class and ultimately defined the unknown vapor. In 14 out of 16 and 15 out of 16 of the SS and ST classification tasks evaluated, the FLD-HC algorithm worked as well or better than the *k*NN method (Tables 6.4-5). In 12 out of 16 and 11 out of 16 of the SS and ST classification tasks, respectively, the FLD-HC algorithm worked as well or better than the FFBNN approach (Tables 6.5-6). Additionally, in contrast to analysis using the *k*NN or FFBNN approaches (Tables 6.4, 6.6), the FLD-HC method yielded no misclassifications of a vapor as belonging to the incorrect vapor class (Table 6.5). This is important for field applications, decreasing the number of false positives for target vapor classes (such as explosives), and has the potential to allow for an unknown vapor to be correctly classified within its respective vapor class.

At each FLD-HC step, FLD was used to find the optimal single dimension, among the 24 dimensional array response space, for cluster separation. If a sensor were extremely useful for classification at a certain step, its dimension in response space would be heavily weighted in the calculation of the optimal FLD dimension. Alternatively, if a sensor were not useful for classification, its dimension in response space would be ignored in the calculation of the optimal FLD dimension.

An issue with the FLD-HC algorithm was the formation of non-Gaussian clusters for the multi-vapor classes when discriminating between the different vapor classes. For example, if an unknown was determined to belong to the alkane group, the exposure was then projected onto a dimension optimized by setting 100 hexane training exposures against 100 training exposures each of heptane, octane, nonane and decane. The individual vapor projections were mostly Gaussian. However, when heptane, octane, nonane and decane were grouped into one class, the heptane projection was distinct from the octane/nonane/decane cluster, creating further non-Gaussian clustering amongst the group (Figure 6.2d). For within-class classification, an alternative approach was thus attempted to address this highly non-Gaussian class clustering. Specificallly, heavy alkanes (or alcohols, xylenes, or chloromethanes) were set against light alkanes (or alcohols, xylenes, or chloromethanes). In this approach, for alkane classification, class 1 was hexane/heptane, and class 2 was octane/nonane/decane. Further classifications were performed accordingly. Using this alternative approach, non-Gaussian projections were still obtained in some cases, and approximately the same classification performances were obtained as in Table 6.5. An alternative decision boundary to eq (3) that accounts for the extent of non-Gaussian distribution experienced in each of the clusters would address these difficulties with the production of non-Gaussian data clusters, and increase the correct classification rates of the pure vapors.

There has been some debate in the literature on whether a large or small number of sensors in an array are required for optimal vapor discrimination and classification.[39-41] Differences can clearly arise from the different algorithms that have been used to assess the classification performance in each case. For example, using extended disjoint principal components regression,[41,42] a noisy sensor may contribute largely to the dimension that describes the concentration-dependent array response to different vapors. With the noisy sensor heavily weighted, poorer classification performance may occur due to the inclusion of such a sensor response in the array. Using FLD to quantitate the discriminating ability of an array, the noisy sensor would be naturally excluded from the projection vector, because its incorporation would generate a non-optimal dimension with a lower rf.[40] Due to the optimal dimension selection process, FLD-HC classification can only benefit, or at worst maintain the same level of classification, when more sensors are included in an array. This conclusion of course assumes that heavily weighted sensors do not exhibit more significant response drift than do the other sensors along the array.

6.5.3. *rf* Value As a Predictor of Classification Performance

Significant increases in vapor classification rates were observed for data obtained using ST rather than SS approaches. These increases, and the analytes that were confused in each case, correlated well with the rf_{test} values (Table 6.3). For example, using *k*NN, the ethanol classification rate increased from 0.82 to 1.00 when ST response information was used (Table 6.4). The 0.18 (or 9 out of 50) SS misclassifications of ethanol were as propanol. The lowest SS rf_{test} value exhibited by ethanol was towards propanol, which increased from 2.2 to 7.3 when ST response information was used. Generally, for all vapors, when analytes were misclassified, they tended to be misclassified as the analytes that had low rf_{test} values with respect to the analyte of interest.

Figures 6.3a-b summarize these rf_{test}-classification rate relationships for all data from Tables 6.3-6. Figure 6.3a uses all identity matrix positions from Tables 6.4-6, and displays the fraction of each target vapor that was correctly classified against the minimum rf_{test} value that was exhibited by that vapor toward its other within-class analytes. For example, for hexane classification, the minimum rf_{test} value was 3.2 and 5.5 for SS and ST analysis, respectively (Table 6.3). For SS and ST analysis using *k*NN, hexane was correctly classified 96% and 94% of the time, respectively (Table 6.4). These points were plotted as (3.2, 0.96) and (5.5, 0.94) on Figure 6.3a. The plot includes all vapor classifications, analyzed with SS and ST response information using the *k*NN, FLD-HC, or FFBNN methods.

Alternatively, Figure 6.3b displays all non-identity matrix positions from Tables 6.4-6, and details the cases in which misclassifications could be made. For example, hexane could be misclassified as the 15 alternative vapors shown in Table 6.2. The fractions that were misclassified as each of these were plotted against their respective rf_{test} values toward hexane. For hexane analyzed with ST response information using *k*NN, alkane misclassifications were plotted as (5.5, 0.06), (10.3, 0), (8.8, 0) and (12.5, 0) for heptane, octane, nonane and decane, respectively (Tables 6.3-4). All vapors are shown in Figure 6.3b, analyzed with ST and SS response information, using the three classification methods. Because of the wide range of rf_{test} values, only the region from $rf_{test} = 0$ to 14 is shown.

Figure 6.3a shows that the use of the SS response information generally yielded lower minimum rf_{test} values, and lower classification rates, than when ST response information was used. Additionally, the *k*NN method generally exhibited the poorest classification rates. Beyond a minimum rf_{test} value of ~ 7, excellent correct classification rates were observed whether ST or SS response information was used, regardless of the classification method. Thus, as long as the responses of the different vapors in a library were unique enough, it did not matter what method was used for classification. The exception was tetrachloromethane (CCl_4) classification, which yielded poor classification rates with a minimum rf_{test} value of 8.4 towards m-xylene (Table 6.3). CCl_4 exhibited the most widely-varying, non-Gaussian response clusters (Figures 6.2a-b). Figure 6.3a demonstrates that the minimum rf_{test} between a target vapor and all other vapors in the analyte library can be used to predict the likelihood of vapor misclassifications occuring.

Figure 6.3b shows that use of only the SS response information produced many more low rf_{test} instances than when ST data was used. These instances of low rf_{test} values were much more likely to be incorrectly classified as the correct target vapor. As the rf_{test} of an analyte increased toward the target vapor, the analyte was significantly less likely to be misclassified as the target vapor, whether ST or SS responses were used. Again, beyond an rf_{test} of approximately 7, analytes were rarely misclassified as the correct target vapors, the exception again being CCl_4. Figure 6.3b demonstrates that the rf_{test} between a target vapor and an alternative analyte in the signature library is an excellent predictor for the likelihood that the target will be misclassified as the alternative analyte.

6.6. Conclusions

Four linear sensor arrays made from small molecule/carbon black composite chemiresistors with low headspace volumes were tasked with discriminating and classifying among four vapor classes. ST and SS response information was used, and direct comparisons were made for pure vapor classification. Three linear algorithms (FLD, *k*NN, FLD-HC) and one nonlinear algorithm (FFBNN) were used to assess array performance: FLD was used to describe the relative distance between response clusters in response space, while the *k*NN, FLD-HC and FFBNN algorithms were used to assess the ability of the arrays to classify vapors. In 109/120, 15/16, and 13/16 of the vapor classification cases, the linear methods (FLD, *k*NN, and FLD-HC, respectively) exhibited equal or increased performances when ST rather than SS response information was used. The non-linear FFBNN exhibited increased classification rates in all cases when ST response information was used. The difficulties exhibited by the linear methods when ST rather than SS data was used were due to non-Gaussian ST response clusters.

The bioinspired classification algorithm FLD-HC compared favorably to the *k*NN and FFBNN methods. In 14 out of 16 and 15 out of 16 of the SS and ST pure vapor classification tasks evaluated, respectively, the FLD-HC algorithm worked as well or better than the *k*NN method. In 12 out of 16 and 11 out of 16 of the SS and ST classification tasks, respectively, the FLD-HC algorithm worked as well or better than the FFBNN method. When vapor misclassifications were made using the FLD-HC method, they were always made for vapors belonging to the same respective chemical class, while the *k*NN and FFBNN methods occasionally misclassified vapors as belonging to the incorrect chemical class.

6.7. References

(1) Janzen, M. C.; Ponder, J. B.; Bailey, D. P.; Ingison, C. K.; Suslick, K. S. *Anal. Chem.* **2006**, *78*, 3591-3600.
(2) Rakow, N. A.; Suslick, K. S. *Nature* **2000**, *406*, 710-713.
(3) Manorama, S.; Sarala Devi, G.; Rao, V. J. *Appl. Phys. Lett.* **1994**, *64*, 3163-3165.
(4) Bednarczyk, D.; DeWeerth, S. P. *Sens. Actuators, B* **1995**, *27*, 271-274.
(5) Wohltjen, H.; Snow, A. *Anal. Chem.* **1998**, *70*, 2856-2859.
(6) Han, L.; Daniel, D. R.; Maye, M. M.; Zhong, C. J. *Anal. Chem.* **2001**, *73*, 4441-4449.
(7) Sundgren, H.; Lundstrom, I.; Winquist, F.; Lukkari, I.; Carlsson, R.; Wold, S. *Sens. Actuators, B* **1990**, *2*, 115-123.
(8) Lundstrom, I.; Shivaraman, C.; Svensson, C.; Lundkvist, L. *Appl. Phys. Lett.* **1975**, *26*, 55-57.
(9) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2001**, *73*, 2501-2508.
(10) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*.
(11) Rose-Pehrsson, S. L.; Grate, J. W.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801-2811.
(12) Wohltjen, H. *Sens. Act.* **1984**, *5*, 307-325.
(13) Baller, M. K.; Lang, H. P.; Fritz, J.; Gerber, C.; Gimzewski, J. K.; Drechsler, U.; Rothuizen, H.; Despont, M.; Vettiger, P.; Battiston, F. M.; Ramseyer, J. P.; Fornaro, P.; Meyer, E.; Guntherodt, H.-J. *Ultramicroscopy* **2000**, *82*, 1-9.
(14) Jensenius, H.; Thaysen, J.; Rasmussen, A. A.; Veje, L. H.; Hansen, O.; Boisen, A. *Appl. Phys. Lett.* **2000**, *76*, 2615-2617.
(15) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1995**, *92*, 2652-2656.
(16) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, A* **1995**, *51*, 57-66.
(17) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mat.* **1996**, *8*, 2298-2312.

(18) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177-4190.
(19) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, B* **1994**, *18*, 211-220.
(20) Lewis, N. S. *Acc. Chem. Res.* **2004**, *37*, 663-672.
(21) Gao, T.; Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Chem. Mat.* **2006**, *18*, 5193-5202.
(22) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *82*, 54-74.
(23) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *in preparation*.
(24) Woodka, M. D.; Brunschwig, B. S.; Lewis, N. S. *Langmuir* **2007**, *23*, 13232-13241.
(25) Sisk, B. C.; Lewis, N. S. *Sens. Actuators, B* **2005**, *104*, 249-268.
(26) Vaid, T. P.; Burl, M. C.; Lewis, N. S. *Anal. Chem.* **2001**, *73*, 321-331.
(27) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 5442-5447.
(28) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.
(29) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*, 5th ed.; Prentice Hall: Upper Saddle River, NJ, 2002.
(30) Brereton, R. G. *Chemometrics - Data Anlysis for the Laboratory and Chemical Plant*; John Wiley & Sons Ltd: West Sussex, England, 2003.
(31) Cover, T. M.; Hart, P. E. *IEEE Trans. Inf. Theory* **1967**, *13*, 21-27.
(32) Friedrich, R. W.; Laurent, G. *Science* **2001**, *291*, 889-894.
(33) Laurent, G.; Stopfer, M.; Friedrich, R. W.; Rabinovich, M. I.; Volkovskii, A.; Abarbanel, H. D. I. *Ann. Rev. Neurosci.* **2001**, *24*, 263-297.
(34) Ripley, B. D. *J. R. Statist. Soc. B.* **1994**, *56*, 409-437.
(35) Ripley, B. D. In *Networks and Chaos - Statistical and Probabilistic Aspects*; Barndorff-Nielsen, O. E., Jensen, J. L., Kendall, W. S., Eds.; Chapman and Hall: London, 1993, pp 40-123.
(36) Littlewood, A. B. *Gas Chromatography: Principles, Techniques, and Applications*, 2nd ed.; Academic Press: New York, 1970.
(37) Livingstone, D. J.; Manallack, D. T.; Tetko, I. V. *J. Computer-Aided Mol. Design* **1997**, *11*, 135-142.
(38) Ho, T. K. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832-844.
(39) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R. *Chem. Rev.* **2000**, *100*, 2595-2626.
(40) Burl, M. C.; Sisk, B. C.; Vaid, T. P.; Lewis, N. S. *Sens. Actuators, B* **2002**, *87*, 130-149.
(41) Park, J.; Groves, W. A.; Zellers, E. T. *Anal. Chem.* **1999**, *70*, 4191-4201.
(42) Zellers, E. T.; Pan, T.-S.; Patrash, S. J.; Han, M.; Batterman, S. A. *Sens. Actuators, B* **1993**, *12*, 123-133.

Table 6.1: Sensor films used in this work. Materials were sonicated in 20 ml of solvent to obtain a well-dispersed suspension. CB = carbon black; DP = dioctyl phthalate.

sensor #	materials	solvent
1	25% tetracosanoic acid, 75% CB	tetrahydrofuran
2	25% propyl gallate, 75% CB	tetrahydrofuran
3	25% tetracosane, 75% CB	toluene
4	17.5% tetracosane, 7.5% DP, 75% CB	toluene

Table 6.2: Analyte abbreviations.

analyte	
Hx	hexane
Hp	heptane
Oc	octane
Nn	nonane
Dc	decane
MeOH	methanol
EtOH	ethanol
PrOH	1-propanol
BuOH	1-butanol
HxOH	1-hexanol
m-xyl	m-xylene
p-xyl	p-xylene
o-xyl	o-xylene
CH_2Cl_2	dichloromethane
$CHCl_3$	trichloromethane
CCl_4	tetrachloromethane

Table 6.3: rf_{test} values obtained for all binary vapor combinations using the FLD algorithm. The bottom number (parenthesis) and top number were obtained using SS and ST response data, respectively. Within-class combinations are highlighted.

	EtOH	PrOH	BuOH	HxOH	Hx	Hp	Oc	Nn	Dc	CH_2Cl_2	$CHCl_3$	CCl_4	o-xyl	m-xyl	p-xyl
MeOH	11.6 (7.9)	10.3 (7.4)	12.8 (9.7)	9.9 (11.9)	42.8 (22.5)	37.3 (23.5)	43.2 (26.4)	40.0 (26.3)	38.2 (29.1)	24.3 (12.4)	28.8 (12.6)	15.8 (8.6)	32.7 (26.2)	32.2 (21.2)	32.9 (22.7)
EtOH	N/A	7.3 (2.2)	9.7 (5.7)	12.1 (9.9)	46.4 (21.9)	40.8 (22.0)	48.1 (24.4)	41.5 (22.8)	41.3 (25.4)	24.8 (14.1)	28.0 (13.7)	12.4 (7.5)	37.1 (23.9)	33.0 (20.1)	31.9 (22.2)
PrOH		N/A	5.1 (3.8)	8.5 (8.8)	34.2 (16.8)	32.8 (18.0)	35.5 (19.0)	34.1 (19.0)	31.5 (21.0)	16.6 (12.2)	18.6 (12.0)	11.2 (7.6)	26.7 (19.6)	22.8 (16.6)	22.3 (17.9)
BuOH			N/A	6.6 (5.7)	31.6 (15.9)	31.3 (16.6)	30.8 (17.6)	29.7 (16.7)	26.9 (19.7)	19.1 (13.3)	19.7 (13.1)	10.7 (7.5)	21.3 (18.1)	19.4 (15.0)	20.3 (16.1)
HxOH				N/A	26.7 (15.9)	26.0 (17.2)	25.4 (18.6)	27.9 (18.3)	23.4 (21.1)	13.2 (11.4)	13.2 (12.6)	9.8 (6.8)	18.3 (14.6)	15.4 (12.3)	17.3 (13.2)
Hx					N/A	5.5 (3.2)	10.3 (6.1)	8.8 (7.2)	12.5 (9.6)	39.9 (23.1)	34.0 (24.9)	18.3 (8.2)	39.5 (20.9)	34.6 (18.4)	39.8 (22.7)
Hp						N/A	5.9 (2.8)	5.8 (4.0)	10.8 (6.1)	36.6 (29.5)	38.0 (31.2)	17.0 (10.5)	38.1 (25.6)	32.5 (22.3)	36.3 (33.2)
Oc							N/A	3.2 (1.4)	6.8 (5.1)	41.5 (32.0)	37.8 (31.3)	17.5 (10.5)	42.0 (27.3)	35.6 (21.7)	40.5 (32.7)
Nn								N/A	4.0 (2.4)	38.0 (26.6)	34.2 (27.7)	15.6 (10.7)	36.1 (25.1)	30.9 (20.1)	34.6 (28.1)
Dc									N/A	31.1 (28.8)	25.7 (34.6)	15.4 (11.4)	27.0 (28.6)	23.7 (23.9)	26.8 (31.8)
CH_2Cl_2										N/A	16.4 (5.8)	11.3 (11.5)	28.8 (19.5)	22.3 (16.6)	33.3 (17.8)
$CHCl_3$											N/A	8.7 (8.8)	25.8 (17.3)	18.4 (14.0)	24.1 (15.1)
CCl_4												N/A	12.2 (4.0)	8.4 (4.3)	12.0 (3.8)
o-xyl													N/A	5.2 (4.4)	8.3 (5.9)
m-xyl														N/A	4.2 (2.7)

Table 6.4: Testing confusion matrix for vapor classification using the *k*NN, *k* = 1 method. The bottom (parenthesis) and top numbers were obtained using SS and ST response data, respectively. Within-class matrices are highlighted. Misclassifications as the wrong vapor class are shown in bold.

	MeOH	EtOH	PrOH	BuOH	HxOH	Hx	Hp	Oc	Nn	Dc	CH_2Cl_2	$CHCl_3$	CCl_4	o-xyl	m-xyl	p-xyl
MeOH	1.00 (0.98)	0 (0)	0 (0.02)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
EtOH	0 (0)	1.00 (0.82)	0 (0.18)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
PrOH	0 (0)	0 (0.24)	0.96 (0.70)	0.04 (0.06)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
BuOH	0 (0)	0 (0.02)	0.04 (0.10)	0.92 (0.86)	0.02 (0.02)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
HxOH	0 (0)	0.02 (0.02)	0 (0.02)	0.02 (0.12)	0.96 (0.84)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hx	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.94 (0.96)	0.06 (0.04)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hp	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.14 (0.100	0.80 (0.78)	0.04 (0.12)	0.02 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Oc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.12 (0.14)	0.82 (0.52)	0.06 (0.26)	0 (0.08)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Nn	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.02)	0.02 (0.04)	0.14 (0.08)	0.68 (0.62)	0.14 (0.24)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Dc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.02)	0 (0)	0 (0.06)	0.10 (0.10)	0.90 (0.82)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
CH_2Cl_2	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.98)	0 (0.02)	0 (0)	0 (0)	0 (0)	0 (0)
$CHCl_3$	0 (0)	0 (0)	0 (0)	**0.02** (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.06)	0.98 (0.94)	0 (0)	0 (0)	0 (0)	0 (0)
CCl_4	0 (0)	0 (0)	0 (0)	0 (0)	**0.02** (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 **(0.02)**	0.02 (0)	0 (0.02)	0.96 (0.94)	0 (0)	0 (0)	0 **(0.02)**
o-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.88)	0 (0.06)	0 (0.06)
m-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.04)	0.84 (0.64)	0.14 (0.32)
p-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.06 (0.08)	0.10 (0.20)	0.84 (0.72)

Table 6.5: Testing confusion matrix for vapor classification using the FLD-HC method. The bottom (parenthesis) and top numbers were obtained using SS and ST response data, respectively. Within-class matrices are highlighted.

	MeOH	EtOH	PrOH	BuOH	HxOH	Hx	Hp	Oc	Nn	Dc	CH_2Cl_2	CHCl3	CCl_4	o-xyl	m-xyl	p-xyl
MeOH	1.00 (1.00)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
EtOH	0 (0.02)	1.00 (0.86)	0 (0.12)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
PrOH	0 (0)	0 (0.12)	1.00 (0.72)	0 (0.16)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
BuOH	0 (0)	0 (0.04)	0 (0)	1.00 (0.96)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
HxOH	0 (0.02)	0.02 (0.04)	0 (0.02)	0.02 (0)	0.96 (0.92)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hx	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.90 (0.94)	0.10 (0.06)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hp	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.04 (0.06)	0.92 (0.86)	0.04 (0.08)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Oc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.02)	0.90 (0.76)	0.10 (0.22)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Nn	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.04)	0.02 (0.08)	0.94 (0.70)	0.02 (0.18)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Dc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.02)	0 (0.02)	1.00 (0.96)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
CH_2Cl_2	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.98)	0 (0.02)	0 (0)	0 (0)	0 (0)	0 (0)
$CHCl_3$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (1.00)	0 (0)	0 (0)	0 (0)	0 (0)
CCl_4	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.02)	0.06 (0.04)	0.92 (0.94)	0 (0)	0 (0)	0 (0)
o-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (1.00)	0 (0)	0 (0)
m-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.94)	0 (0.06)
p-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.08 (0.06)	0.92 (0.94)

Table 6.6: Testing confusion matrix for vapor classification using the FFBNN method. The bottom (parenthesis) and top numbers were obtained using SS and ST response data, respectively. Within-class matrices are highlighted. Misclassifications as the wrong vapor class are shown in bold.

	MeOH	EtOH	PrOH	BuOH	HxOH	Hx	Hp	Oc	Nn	Dc	CH_2Cl_2	$CHCl_3$	CCl_4	o-xyl	m-xyl	p-xyl
MeOH	1.00 (0.98)	0 (0.02)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
EtOH	0 (0)	1.00 (0.90)	0 (0.10)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
PrOH	0 (0)	0 (0.10)	1.00 (0.86)	0 (0.04)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
BuOH	0 (0)	0 (0)	0 (0.12)	0.96 (0.88)	0.02 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	**0.02** (0)
HxOH	0 (0)	0 (0)	0.02 (0)	0.02 (0.08)	0.96 (0.92)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hx	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.94)	0 (0.06)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hp	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.04)	1.00 (0.80)	0 (0.16)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Oc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.06)	0.96 (0.80)	0.04 (0.14)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Nn	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.04)	0.06 (0.08)	0.86 (0.76)	0.06 (0.12)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Dc	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.06 (0.14)	0.94 (0.86)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
CH_2Cl_2	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (1.00)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$CHCl_3$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.04)	0.98 (0.96)	0 (0)	**0.02** (0)	0 (0)	0 (0)
CCl_4	**0.02** (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	**0.02** **(0.02)**	0.04 (0.06)	0.08 (0.14)	0.82 (0.70)	0 **(0.08)**	0 (0)	**0.02** (0)
o-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1.00 (0.96)	0 (0.04)	0 (0)
m-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.10)	0.98 (0.80)	0.02 (0.10)
p-xyl	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.04 (0.18)	0.96 (0.82)

Figure 6.1: Responses of sensors 1, 2, 3, 8 and 15 along an array of tetracosane/dioctyl phthalate/carbon black composite (sensor 4, Table 6.1) to the alkanes. For visualization, sensor responses are scaled by their response at 100 s.

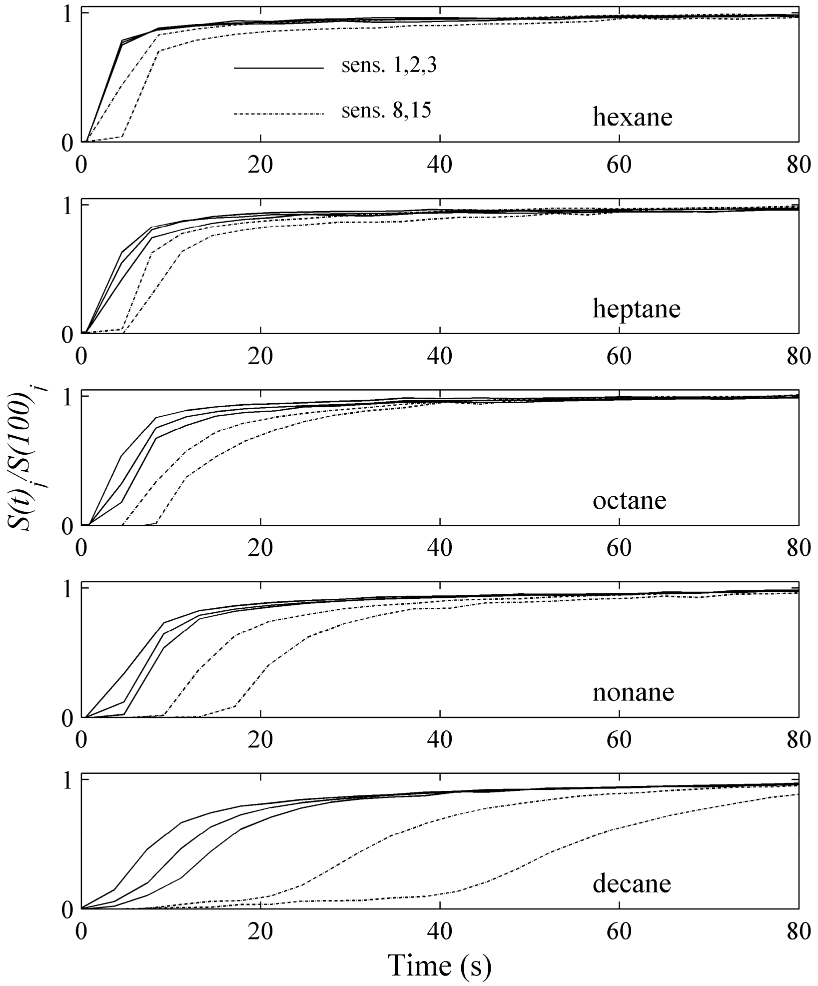
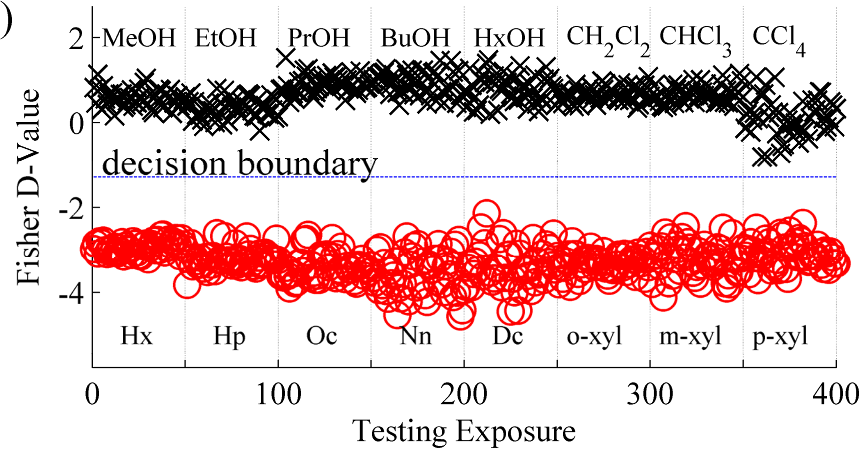


Figure 6.2: Waterfall plots for FLD-HC classification at various steps: a) alcohols / chloromethanes vs. alkanes / xylenes, b) alcohols vs. chloromethanes, c) alkanes vs. xylenes, and d) hexane vs. heptane-decane. For all cases, the FLD decision boundary (eq (3)) based on training data is shown.

a)



b)

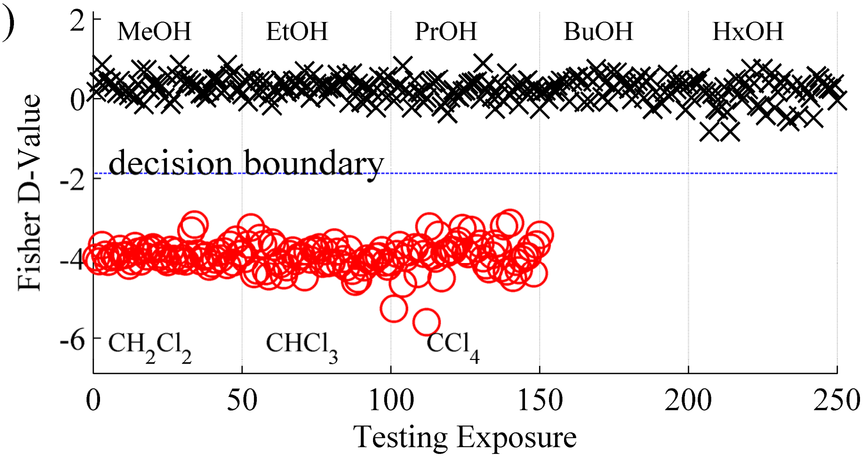


Figure 6.2:

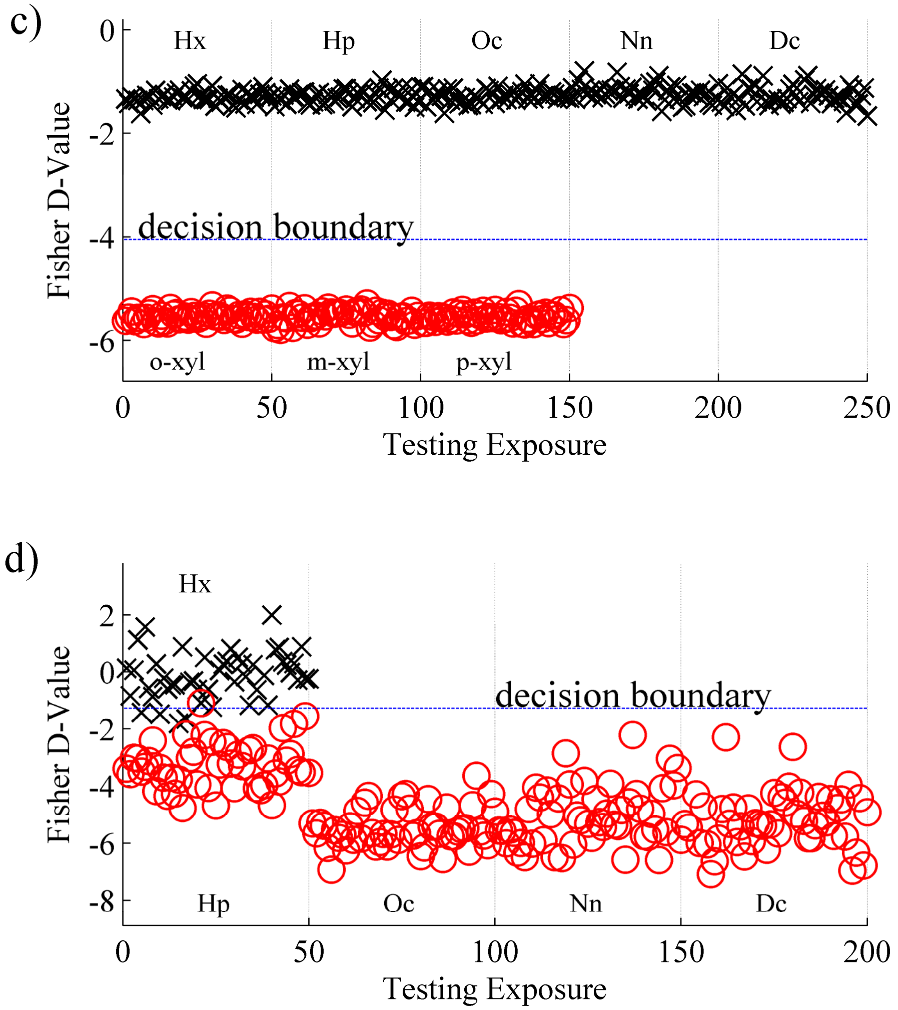
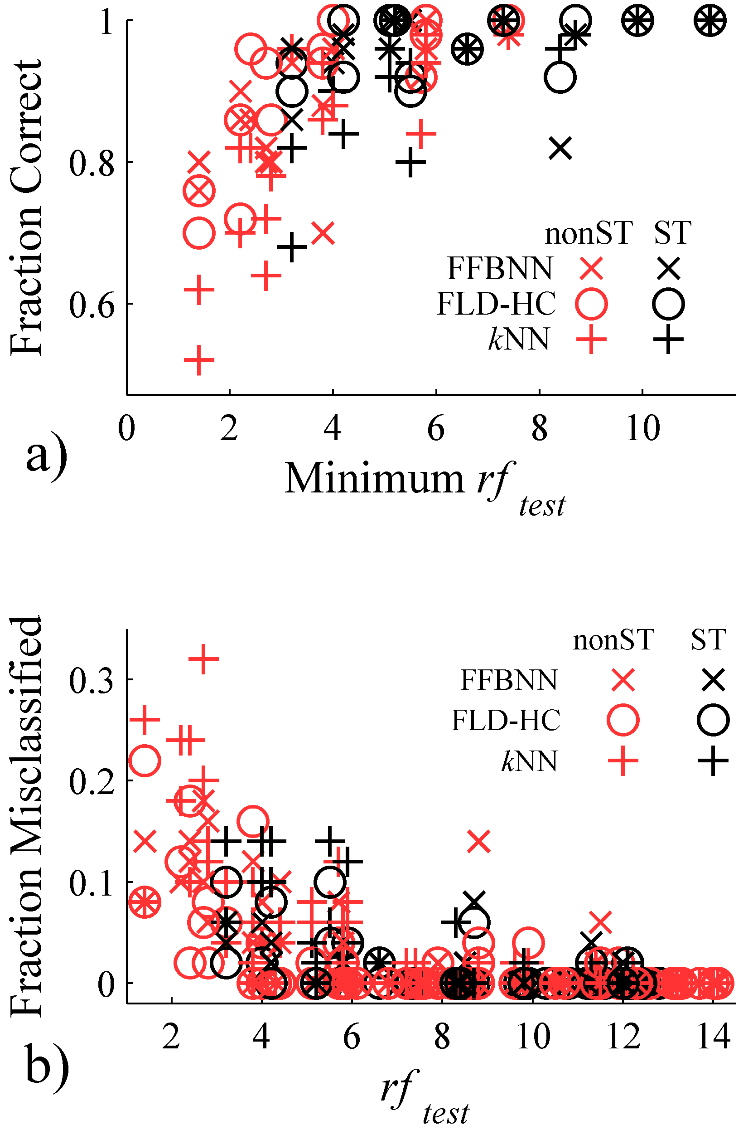

c)
Hx
Hp
Oc
Nn
Dc
0
-2
-4
-6
Fisher D-Value
decision boundary
o-xyl
m-xyl
p-xyl
0
50
100
150
200
250
Testing Exposure
d)
Hx
2
0
-2
-4
-6
-8
Fisher D-Value
decision boundary
Hp
Oc
Nn
Dc
0
50
100
150
200
Testing Exposure

Figure 6.3: Fraction of the 50 testing exposures a) correctly identified vs. the minimum rf_{test} for the target vapor, and b) misclassified vs. the target vapor – confusing vapor rf_{test}, for alkane and alcohol identification, analyzed using SS and ST response information and *k*NN, FLD-HC, or a FFBNN for classification.



Chapter 7

Thesis Summary

7.1 Summary of Thesis

Carbon black composite thin film chemiresistors implemented in an electronic nose arrangement offer a low-cost, low-power means to detect a variety of vapors in a number of environments. A major difficulty with this approach has been the inability to utilize the linear and additive response properties of the sensors, for the analysis of vapor mixtures with training on only the pure response signatures. These difficulties arise due to response signatures that are similar in nature for chemically similar analytes. When responses towards these similar analytes are observed in mixtures, the subtle differences between them become overrun by noise, and pure vapor training becomes incapable of mixture analysis. Because the number of different vapor mixtures one could encounter in field-testing grows exponentially with the number of vapors one could encounter in the field, the training protocol required for the sensors rapidly becomes infeasible when only a limited number of vapors could potentially be present. For field-practical implementation, the linear and additive response properties must be exploited for mixture analysis.

This thesis addressed these issues with two approaches. A novel class of carbon black composites, where the insulating phase consisted of small organic molecules, was demonstrated to provide enhanced pure-vapor discrimination rates due to higher function group densities, and enhanced sensor-vapor interactions. These new sensor types were then implemented in a fashion that created a strong space- and time-, or ST dependency of the sensor response. This ST data further enhanced array-vapor response differences, leading to higher pure-vapor classification rates. These enhanced array response differences, coupled with the sensors linear and additive response properties, were utilized to perform vapor mixture analyses based on only pure vapor training. With these capabilities, the utility of carbon black composite vapor sensors for field use, and electronic noses in general, is increased significantly.

In addition to demonstrating an enhanced ability to analyze vapor mixtures with pure vapor training and ST analysis, the requirements for obtaining optimized mixture analysis using pure vapor training were defined. Two dimensionless numbers, Pe_{yz} and Pe_{zz}, were used to define an optimized sensing regime based on relationships between the sensor chamber geometries, vapor delivery flow rates, and physical properties of the sensor and vapor phase. With these Pe_{yz} and Pe_{zz} definitions, a chamber optimized for ST analysis that maximally separates different vapors can be fabricated.

Pure vapor response signatures were analyzed to determine their most useful features for their use in mixture analysis. Calculation of the resolution factor between each of the pure vapor response signature provided a measure of how separated the signatures were in response space. The minimum resolution factor of all binary combinations of pure vapor response signatures provided a measure of where misclassifications were most likely to occur. Accordingly, the minimum resolution factor was an excellent predictor for where vapor misclassification would most likely occur, and for the ability to analyze vapor mixtures with only pure vapor training. These definitions defined what is needed in terms of pure-vapor response signatures for mixture analysis, and apply to all sensor modalities implementation in electronic nose fashion.

The developments presented in this thesis allow for the exploitation of the linear and additive response properties of these sensors, providing an enhanced ability to analyze pure vapors and vapor mixtures. With the ability to robustly identify pure vapors or vapor mixtures with only pure-vapor training, the required training protocol is decreased significantly. This, combined with the ability of these sensors to continually re-establish a new sensor baseline under different background conditions such as drifting background humidity, makes these sensors increasingly attractive for field use.

7.2 Future Directions

The definitions of Pe_{yz} and Pe_{zz} in Chapter 4 should be used to design and fabricate an optimized chamber for mixture analysis. The experimental ST conditions used in this thesis allow for further optimization, with respect to Pe_{yz}, by a factor of 10. An optimized chamber will likely require microfabrication to ensure that the Pe_{yz} requirement is met. The microfabricated chamber must be made from a non-sorption material to ensure that the chamber body does not act to sorb and deplete the vapor stream of analyte. Alternative nonsorbing materials would be required, such as an etched-silicon flow pathway bound to a glass surface via anodic bonding.

Due to hardware limitations, all experimental ST work performed in this thesis employed a sampling frequency for each sensor of ~ 3 s^{-1}. The ST chamber and sensors were therefore designed to assure that the transient response of the sensors occurred on time scales of up to 80 s. This assured that the 3 s^{-1} sampling frequency sufficiently captured the transient response of the vapors. Sampling times of 80 s are impractical for field application, however: ideally, vapor sampling would take place on the order of seconds. This would allow for a more continuous sampling of the vapor phase. The optimal Pe_{yz} and Pe_{zz} definitions can be used to design sensor chambers that occur on this rapid timescale. The requirement is that the sensor sampling hardware occurs on a much more rapid timescale, which is easily obtainable with alternative methods for resistance measurement.

To take advantage of this rapid sampling, a reproducible method for vapor delivery and sampling must be developed for these timescales. One approach would employ a piston-type sampling arrangement forcing vapor flow along the sensor arrays. The piston intake would initially draw outside vapor onto the sensor arrays to elicit a response, and subsequently push the vapor away from the sensor arrays to purge the sensors. In this arrangement, the sensors would not necessarily deplete the vapor stream. To ensure that the outflowing vapor stream is analyte-free, a high-capacity denuder-type sorption unit would be placed past the sensors along each array on the piston side of each array. Each intake of the piston would indicate the beginning of a measurement, and sensor response extraction could be synced accordingly to reproducibly extract the time-dependent response information. The piston could continuously cycle, which would provide continuous real-time vapor sample measurements.

Finally, an alternative sensor array implementation, where the sensors are distributed exponentially along each array, would maximize the amount of unique information extracted per sensor. This would allow for efficient analysis of analytes with vapor pressures ranging several orders of magnitude. The first few exponentially spaced sensors along each array would capture the progression of the most strongly partitioning analytes with the lowest vapor pressures, while the last sensors along each array would capture the progression of the most weakly partitioning analytes with the highest vapor pressures. An alternative implementation that would similarly maximize the amount of unique information extracted per sensor would distribute sensors linearly along each array, but employ a height above the sensor film that increases exponentially as distance along the array increases. Due to the initial low height (and small cross-sectional area) above the sensor film, the first linearly spaced sensors would provide detailed information for the progression of strongly partitioning vapors along the array, while the last linearly spaced sensors would provide detailed information of weakly partitioning vapors along the array. The two methods listed above would allow a detailed measure of the decomposition of vapors, spanning several orders of magnitude in vapor pressure, along each of the sensor arrays.

The suggestions listed above take advantage of the linear and additive response properties of the sensors, and the definition of an optimized ST regime for mixture analysis, for the design and fabrication of a device well-suited for field application. Ideally, a field-portable unit will provide rapid analysis times and the ability to analyze vapors spanning several orders of magnitude in vapor pressure. These suggestions detail various aspects of sensor response that could be exploited to fabricate such a device, and methods for their implementation, for the generation of a field-portable and field-practical device. Such a device would allow these sensors to be used in a variety of fields, such as military vapor detection tasks, space shuttle air quality monitoring, breath test-based medical disease diagnostics, and food quality/spoilage monitoring.