Pacific Northwest
National Laboratory
Operated by Battelle for the
U.S. Department of Energy

Chemometric Classification of Unknown Vapors by Conversion of Sensor Array Pattern Vectors to Vapor Descriptiors: Extension from Mass-Transducing Sensors to Volume-Transducing Sensors

J. W. Grate
B. M. Wise

June 2001

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This document was printed on recycled paper.
(8/00)

Chemometric Classification of Unknown Vapors by Conversion of Sensor Array Pattern Vectors to Vapor Descriptors: Extension from Mass-Transducing Sensors To Volume-Transducing Sensors

J. W. Grate
B. M. Wise

June 2001

Environmental Molecular Sciences Laboratory
Pacific Northwest National Laboratory
Richland, Washington 99352

Summary

Pacific Northwest National Laboratory is currently developing enabling science and technologies for high-performance chemical detection systems based on sensor arrays. This report describes progress in new pattern-recognition methods. A new chemometric method has been developed for classifying unknowns by transforming the vector containing the responses from a multivariate detector to a vector containing descriptors of the detected analyte. These descriptors characterize the detected analyte. In principle, this allows characterization and possibly identification of unknown vapors that were not tested in a training set.

Previously, this approach was derived for gravimetric polymer-coated sensors, such as polymer-coated acoustic wave devices. In the present report, we have derived mathematics showing how these new chemometric methods can be applied to other sensor types. In particular, it is shown how they can be applied to volume-transducing polymer-coated sensors. An example of this type of sensor is the chemiresistor technology with polymer/carbon black composite sensing layers.

We also discuss the application of this method for surface acoustic wave sensors where the responses of the sensors have both a gravimetric and a modulus component, the modulus component representing a volume-based transduction mechanism.

In summary, this report describes how our new chemometric pattern-recognition approach can now be applied to a variety of sensor technologies. It is not restricted solely to gravimetric polymer-coated sensors.

The work described is being published in Analytical Chemistry.

Acknowledgements

Contents

Introduction

A new chemometric method was recently described for classifying unknowns by transforming the vector containing the responses from a multivariate detector to a vector containing descriptors of the detected analyte (Grate et al. 1999). This approach was derived for sensor arrays where each sensor's signal is proportional to the amount of vapor sorbed by a polymer on the sensor surface. In this case, the response is proportional to the partition coefficient, K, and the concentration of the vapor in the gas phase, C_V, where K is the ratio of the concentration of vapor in the sorbent polymer phase, C_S, to C_V.

$$K = C_S / C_V \tag{1}$$

The descriptors of the detected analyte for sorption-based vapor sensors were taken as the solvation parameters in a linear solvation energy relationship (LSER) for vapor sorption, as given in Eq. 2 (Abraham 1993; Grate and Abraham 1991; Grate et al. 1996; Grate 2000).

$$\log K = c + r R_2 + s \, \pi_2^H + a \, \alpha_2^H + b \, \beta_2^H + l \log L^{16} \tag{2}$$

In this model, log K is modeled as a linear combination of terms, each of which is related to particular solubility interactions. The descriptors of sorbed vapors, R_2, π_2^H, α_2^H, β_2^H, and log L^{16}, are solvation parameters that characterize the solubility properties of the vapor (Abraham 1993), where R_2 is a calculated excess molar refraction parameter that provides a quantitative indication of polarizable n and p electrons; π_2^H measures the ability of a molecule to stabilize a neighboring charge or dipole; α_2^H and β_2^H measure effective hydrogen-bond acidity and basicity, respectively; and log L^{16} is the liquid/gas partition coefficient of the solute on hexadecane at 298K (determined by gas-liquid chromatography). The log L^{16} parameter is a combined measure of exoergic dispersion interactions that increase log L^{16} and the endoergic cost of creating a cavity in hexadecane leading to a decrease in log L^{16}. The values for all these descriptors except R_2 were determined from experimental measurements of hydrogen-bonding equilibria or partitioning equilibria, and they have been determined for over 2000 compounds.

The concept of transforming sensor-array patterns into these vapor descriptors is shown in Figure 1. This approach stands in contrast to the conventional approach where detected sensor-array patterns are compared only to patterns for known compounds determined in training.

The new chemometric method was described for acoustic wave sensors where the sensor's signal is proportional to the mass of vapor sorbed. In this case, the gravimetric response function can be expressed according to Eq. 3 (Grate et al. 1988; Grate et al. 1992).

$$R = \Delta f_V = \Delta f_S C_V K / \rho_s \tag{3}$$

sensor array patterns
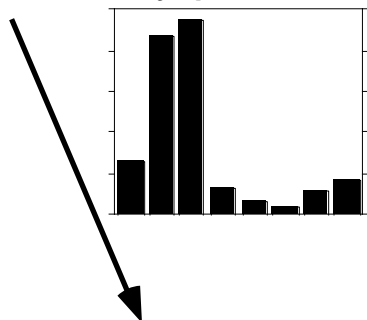
descriptors

$$R_2 \qquad \pi_2^H \qquad \Sigma\alpha_2^H \qquad \Sigma\beta_2^H \qquad \log L^{16}$$

Figure 1. Illustration of the Goal of Converting an Array Response Vector, Shown as a Bar Graph, into Descriptors of the Detected Vapor, Where the Descriptors are the Solvation Parameters from a Linear Solvation Energy Relationship for Vapor Sorption

The response, R, is a frequency shift, f_v, that is proportional to the concentration of vapor in the sorbent polymer phase $C_S = C_V K$. For these sensors, the proportionality to C_S is given by the ratio of the frequency shift due to application of the polymer, f_s, to the polymer density, $_s$.

In our previous report, two approaches were considered for determining the solvation parameters for detected vapors from sensor-array response vectors (Grate et al. 1999). In the first, which was similar to classical least squares (CLS), all the descriptors were solved for simultaneously given knowledge of the properties of the sorbent polymers on the sensors. In the second approach, inverse least squares (ILS) methods were used to develop equations for each solvation parameter individually. The ILS approach does not require knowledge of the polymer properties; it relies on training.

Although derived for *mass-transducing sensors*, the chemometric method for extracting descriptors from multivariate responses is potentially very general (Grate et al. 1999). In the present paper, we describe how solutions can be derived for *volume-transducing sensors*. Sorbent polymers loaded with conductive particles can be used as chemiresistor vapor sensors where the sensor's response, a change in resistance, is related to the fractional volume increase of the film on vapor sorption. Carbon-particle loaded polymers for this method of vapor sensing have been described (Ruschau et al. 1989; Lundberg and Sundqvist 1986; Talik et al. 1992), and they have been used for array-based sensing by Lewis and coworkers (Lonergan et al. 1996; Doleman et al. 1998; Severin et al. 2000; Albert et al. 2000). The design of a chemiresistor sensor array by varying the properties of the sorbent component of a composite film was proposed by Grate in 1990 in connection with phthalocyanine/polymer composite Langmuir-Blodgett films for organic vapor sensing (Grate et al. 1990). The phalocyanine nanoparticles served as the current carrying component, and response characteristics were correlated with vapor sorption by the polymer component.

2

Recently, Severin and Lewis described detailed studies of carbon-particle-loaded polymer chemiresistors that examined how vapor sorption, volume increases, and sensor resistance changes are related (Severin and Lewis 2000). Vapor sorption by test polymers with and without carbon-particle loading was made using the quartz crystal microbalance (QCM). These authors demonstrated that sensor resistance changes are related to the extent of volume increase regardless of the identity of the vapor producing the volume increase. Correlation of the response measurements with vapor densities as liquids supported this mechanism. Therefore, just as acoustic-wave sensors represent a very general method of sensing the mass of vapor sorbed, these carbon/polymer composite chemiresistor sensors represent a very general method of sensing the volume of vapor sorbed. As a result, the signals from an array of these sensors are directly related to vapor sorption, and they can, in principle, be processed to obtain analyte descriptors. This paper sets out the equations for doing so.

Theoretical Sections

Response Models and Solutions for Mass-Based Sensor Arrays

We have shown previously that the response model for an array of gravimetric sensors can be expressed in matrix algebra according to Eq. 4 (Grate et al. 1999).

$$\mathbf{R} = \mathbf{C}\ 10^{(\mathbf{VP} + \mathbf{1c})}\ \mathbf{D}^{-1}\ \mathbf{F} \qquad (4)$$

This equation is directly analogous to Eq. 3 for a single sensor, where matrix $\mathbf{R}$ (vapors by polymers) contains the response values, and matrix $\mathbf{C}$ (number of vapors by number of vapors) is a diagonal matrix of the concentrations of the vapors. Matrix $\mathbf{F}$ (number of polymers by number of polymers) is a diagonal matrix of the f_s values of the sensors, and matrix $\mathbf{D}$ (number of polymers by number of polymers) is a diagonal matrix of the polymer densities. The superscript of -1 denotes the inverse of the matrix. The K value in Eq. 3 is represented by the exponential term in Eq. 4.

$$\mathbf{L} = \mathbf{VP} + \mathbf{1c} \qquad (5)$$

As given in Eq. 5, log K values modeled by the LSER relationship are in matrix $\mathbf{L}$, where matrix $\mathbf{V}$ (number of vapors by five) contains the vapor solvation parameters, and matrix $\mathbf{P}$ (5 by number of polymers) contains the polymer parameters. The constants of the LSER equations are given by the vector $\mathbf{c}$ (1 by number of polymers), and $\mathbf{1}$ is a vector of ones (number of vapors by 1) .

Using the CLS approach, one can solve for $\mathbf{V}$ containing the vapor descriptors given knowledge of $\mathbf{P}$, $\mathbf{c}$, $\mathbf{D}$, $\mathbf{F}$, and $\mathbf{C}$.

$$\{\log (\ \mathbf{C}^{-1}\ \mathbf{R}\ \mathbf{D}\ \mathbf{F}^{-1}\ )\ -\ \mathbf{1c}\ \}\ \mathbf{P}^{T}(\mathbf{PP}^{T})^{-1} = \mathbf{V} \qquad (6)$$

The superscript T denotes the transpose of a matrix. The requirement to know $\mathbf{C}$ in order to solve for $\mathbf{V}$ represents a significant limitation of this solution since one is unlikely to know the concentration of an unknown vapor before identifying it. However, it is possible to rearrange the response model and the solution so that one solves for the log of the concentration simultaneously with the descriptor values. In this approach, the $\mathbf{V}$ matrix is augmented to contain the log of the vapor concentration in addition to the descriptors for each vapor, and the $\mathbf{P}$ matrix is augmented with a vector of ones. Then the response model and the solution are given as shown in Eqs. 7 and 8. Equation 8a expresses the solution for an entire matrix of responses, $\mathbf{R}$, while Eq. 8b expresses the solution where a vector of descriptors $\mathbf{v}_a$ is obtained from a single response vector, $\mathbf{r}$.

$$\mathbf{R} = 10^{(\mathbf{V}_a\,\mathbf{P}_a + \mathbf{1c})}\ \mathbf{D}^{-1}\ \mathbf{F} \qquad (7)$$

$$\{\log (\ \mathbf{R}\ \mathbf{D}\ \mathbf{F}^{-1}\ )\ -\ \mathbf{1c}\ \}\ \mathbf{P}_a^{T}(\mathbf{P}_a\mathbf{P}_a^{T})^{-1} = \mathbf{V}_a \qquad (8a)$$

$$\{\log (\ \mathbf{r}\ \mathbf{D}\ \mathbf{F}^{-1}\ )\ -\ \mathbf{c}\ \}\ \mathbf{P}_a^{T}(\mathbf{P}_a\mathbf{P}_a^{T})^{-1} = \mathbf{v}_a \qquad (8b)$$

5

Solving Eq. 8 for $\mathbf{V}_a$ gives the values of the descriptors R_2, $\,_2^{H}$, $\,_2^{H}$, $\,_2^{H}$, and $\log L^{16}$, and the log of the concentration. This method of augmenting the $\mathbf{V}$ matrix with other parameters related to the vapor, such as its concentration, provides the key to adapting the model to address volume-transducing sensors in addition to gravimetric sensors.

This approach represents a synthesis of chemometric data analysis with a detailed response model that includes the LSER model describing the interaction between vapors and polymers. It is similar to a CLS approach (Beebe et al. 1998). In the CLS formulation, the instrument response, $\mathbf{R}$, is a function of concentration, $\mathbf{C}$, and other parameters: $\mathbf{R}=\mathbf{CS}$, where $\mathbf{S}$ is the sensitivity. In this model, the effect on the response $\mathbf{R}$ of different analytes is linearly additive. This is a sensible physically interpretable model, which works in many physical systems, the most notable of these being spectroscopy where the Beer-Lambert law holds. Given knowledge of the sensitivities in $\mathbf{S}$, the concentrations, $\mathbf{C}$, can be obtained from $\mathbf{R}$: $\mathbf{C} = \mathbf{R}\,\mathbf{S}^{T}(\mathbf{SS}^{T})^{-1}$. CLS is typically used to obtain concentrations of multiple analytes in a mixture. The response model for mass-transducing sensors in Eq. 4 fits the form of a CLS model, with all parameters after the concentration $\mathbf{C}$ lumped into a single sensitivity parameter, $\mathbf{S}$. Given knowledge of the polymer properties, one can solve for the values of all the vapor descriptors simultaneously. In our approach, we essentially model a single vapor as a mixture of five "pure components" related to each solvation parameter, yielding the five parameters for the test vapor. Thus, one obtains descriptor values instead of concentrations.

It is also possible to determine equations for each descriptor individually using an appropriate ILS method such as partial least squares (PLS) regression or principal components regression (PCR). ILS models one specific parameter of interest, $\mathbf{y}$, as a function of instrumental response, $\mathbf{X}$: $\mathbf{y} = \mathbf{Xb}$, where $\mathbf{b}$ is a vector of weighs generally determined by regression (Beebe et al. 1998; Wise and Gallagher 1996). While the ILS model is not physically interpretable, it has a number of desirable properties. Perhaps the most important of these is that complete knowledge of the sources of variance in the response need not be known. Instead, it is possible to construct models from the measured response $\mathbf{R}$ and an independent measurement of the particular property of interest. In addition, there exist a variety of ILS models that tolerate, in fact take advantage of, collinearity in the data, which is a must in the system of interest given large arrays (six or more sensors). In particular, PLS and PCR methods tolerate collinearities in the data and filter out noise. In our case, the ILS approach is advantageous because prior knowledge of polymer properties in not required. We described the use of ILS models to extract vapor descriptors in our previous paper on mass-transducing sensors and conducted simulations on synthetic gravimetric sensor data showing how the precision of the vapor descriptors depends on the measurement noise (reproducibility) in the sensor responses (Grate et al. 1999). We will not elaborate on ILS approaches in this paper, but whenever it is possible to derive a CLS solution, it follows that ILS solutions could also be found.

Response Models and Solutions for Volume-Transducing Arrays

The response function for a volume-transducing sensor can be expressed according to Eq. 9.

$$R = \,_{v}\,S \tag{9}$$

The volume fraction of the vapor in the polymer/vapor solution, $_{v}$, times the sensitivity, S, gives the response, R. (It is assumed throughout this presentation that the volume increase due to vapor sorption is small relative to the initial polymer volume, and the ratio of vapor volume to polymer volume is nearly the same as the ratio of vapor volume to the volume of the vapor/polymer

6

solution.) The volume fraction is related to the amount of vapor in the polymer, $C_S = C_V K$. Therefore the volume fraction of vapor can be expressed according to Eq. 10, where v_v is the specific volume of the vapor as a liquid.

$$_v = v_v \, C_V \, K \tag{10}$$

Then the response function can be expressed so that the response is related to vapor specific volume and the concentration, as given in Eq. 11.

$$R = v_v \, C_V \, K \, S \tag{11}$$

The sensitivity, S, in Eq. 11 has a different value and different units than the sensitivity in Eq. 9 above, but the unchanged notation S is retained for simplicity.

It follows that the response function can be expressed in matrix algebra according to

$$\mathbf{R} = \mathbf{Y} \, \mathbf{C} \, 10^{(\mathbf{VP} + \mathbf{1c})} \, \mathbf{S} \tag{12}$$

The matrix $\mathbf{Y}$ is a diagonal matrix (number of vapors by number of vapors) containing the specific volumes of the vapors. Then the solution can be expressed as

$$\{\log (\mathbf{Y}^{-1} \, \mathbf{C}^{-1} \, \mathbf{R} \, \mathbf{S}^{-1}) - \mathbf{1c} \} \, \mathbf{P}^T(\mathbf{PP}^T)^{-1} = \mathbf{V} \tag{13}$$

This solution, like the initial solution for mass-based sensors described above, is of limited value because the concentration and the specific volume of the unknown vapor must be known to solve for the descriptors in $\mathbf{V}$. However, matrix $\mathbf{V}$ can be augmented so that it contains the log of the product of the vapor concentration times the vapor specific volume. This augmented matrix will be defined as $\mathbf{V_b}$. The $\mathbf{P}$ matrix must be augmented with a vector of ones as before.

$$\mathbf{R} = 10^{(\mathbf{V_b} \, \mathbf{P_a} + \mathbf{1c})} \, \mathbf{S} \tag{14}$$

$$\{\log (\mathbf{R} \, \mathbf{S}^{-1}) - \mathbf{1c} \} \, \mathbf{P_a}^T(\mathbf{P_a}\mathbf{P_a}^T)^{-1} = \mathbf{V_b} \tag{15a}$$

$$\{\log (\mathbf{r} \, \mathbf{S}^{-1}) - \mathbf{c} \} \, \mathbf{P_a}^T(\mathbf{P_a}\mathbf{P_a}^T)^{-1} = \mathbf{v_b} \tag{15b}$$

This derivation shows for the first time how an array of polymer-sorption based sensors with signals proportional to volume increases can be can be used to solve for the descriptors of sorbed vapors. One also solves for the value of the log of the product of the vapor concentration times the vapor specific volume. This is not of value for classification. If the vapor was identified from the found descriptors and the specific volume determined, then the concentration could be obtained.   In any case, the important point is that the descriptor values can, in principle, be obtained. As noted previously, the existence of this CLS type solution indicates that ILS solutions for each individual descriptor could be determined by calibration.

The product of vapor concentration in mass per volume times the vapor specific volume in volume per mass can be regarded as a vapor concentration in volume per volume units. Although this is a strange expression for concentration, regarding this product as a concentration indicates that

the solution in Eq. 15 for a volume-based sensor array is equivalent to the solution in Eq. 8 for a mass-based sensor array.

Vapor specific volume is not highly correlated with the solvation parameters used as descriptors. To verify this, the liquid densities of 43 diverse compounds were tabulated with the solvation parameters and examined for correlations. The correlation matrix is given in Table 1.

The responses of carbon particle/polymer composite chemiresistor sensors, generally taken as the change in resistance relative to the initial resistance, are proportional to the relative volume change of the polymeric insulating phase (Lonergan et al. 1996). For appropriate carbon loadings, the response is linearly proportional to the vapor concentration and the fractional volume increase of the polymer (Lonergan et al. 1996; Severin et al. 2000; Severin and Lewis 2000). Given these characteristics, this sensor technology represents a volume-transducing method that may be appropriate for this classification approach. The models above indicate a dependence on vapor specific volume, which has been demonstrated experimentally (Severin and Lewis 2000).

The experiments used to investigate vapor sorption and chemiresistor response employed QCMs modified with electrodes so that QCM and chemiresistor measurements could be made simultaneously, providing a signal related to vapor mass and a signal related to vapor volume. These studies tested a swelling-induced resistance change hypothesis. As noted by Lewis in the final sentence of his paper, this approach provides a "single-element densitometer" (Severin and Lewis 2000). This could provide a descriptive molecular property directly, provided that the measurement was calibrated. Thus, the QCM/chemiresistor combination would require training to calibrate the resistance change to frequency change ratio to the vapor densities. Then densities could be obtained directly for additional vapors not in the training.

	R_2	H_2	H_2	H_2	$\log L^{16}$	d^2
R_2		0.347	-0.264	-0.229	0.576	0.476
H_2	0	1	0.020	0.507	0.277	0.217
H_2	0	0	1	-0.030	-0.428	0.124
H_2	0	0	0	1	0.127	-0.336
$\log L^{16}$	0	0	0	0	1	0.132
d	0	0	0	0	0	1

[1] Based on a set of 43 diverse compounds.
[2] density as a liquid

Response Models and Solutions for Arrays of Sensors with Combined Mass and Volume Transduction Mechanisms

Sorption-based sensor arrays whose individual sensors respond to both the mass and the volume of the sorbed vapor can be modeled by combining Eqs. 4 and 12, obtaining:

$$\mathbf{R} = \mathbf{C} \ 10^{(\mathbf{VP} + \mathbf{1c})} \mathbf{D^{-1}} \ \mathbf{F} + \mathbf{Y} \ \mathbf{C} \ 10^{(\mathbf{VP} + \mathbf{1c})} \mathbf{S} \qquad (16)$$

Unfortunately, it does not appear possible to obtain a closed form solution for vapor parameters **V** and **Y**, and concentration **C** given the response **R** and polymer parameters **P**, **c**, **D**, **F** and **S**. If **Y** is taken as the identify matrix (specific volumes of the vapors as liquids = 1), however, the following solution is obtained:

$$\{\log (\mathbf{R} \ (\mathbf{D^{-1}} \ \mathbf{F} + \mathbf{S}) ^{-1}) - \mathbf{1c} \} \ \mathbf{P_a^T}(\mathbf{P_a}\mathbf{P_a^T})^{-1} = \mathbf{V_a} \qquad (17a)$$

$$\{\log (\mathbf{r} \ (\mathbf{D^{-1}} \ \mathbf{F} + \mathbf{S}) ^{-1}) - \mathbf{c} \} \ \mathbf{P_a^T}(\mathbf{P_a}\mathbf{P_a^T})^{-1} = \mathbf{v_a} \qquad (17b)$$

where Va is defined as before.

Using this as an initial guess, it is possible to determine Va and Y using direct fitting of the sensor responses to the model with a non-linear least squares optimization procedure. This has been verified on simulated data.

Polymer-coated acoustic wave vapor sensors (Grate 2000; Grate et al. 1993a, b) responding to both the mass of the sorbed vapor and its effect on the polymer modulus represent a combined mass and volume transducing sensor technology. The modulus effect can be modeled as a volume effect (i.e., a swelling-induced modulus change), since modulus changes are proportional to increases in polymer free volume (Grate et al. 1992; Martin and Frye 1990; Martin et al. 1994; Ferry 1980).

Models for SAW vapor sensor response expressed as the sum of mass and volume terms were reported by Grate in 1992 and 2000 (Grate et al. 1992; Grate and Zellers 2000). These assume the polymer films on the sensors are acoustically thin. The full model is given in Eq. 18.

$$\Delta f_V = (\ \Delta f_S \ C_V \ K / \ \rho_s \ ) + \ f_L (\ v_v \ C_V \ K) \ (\ \Delta f_S \ \alpha_{SAW} / \ \alpha) \qquad (18)$$

The mass term is the same as the mass term above in Eqs. 3 and 4. The volume term includes the fractional free volume of the vapor as a liquid, f_L. The product $\Delta f_S \ \alpha_{SAW} / \alpha$ gives the frequency change due to a fractional volume increase of the polymer film, where α in Eq. 4 is the coefficient of thermal expansion of the polymer and α_{SAW} represents the kHz change in frequency due to a 1°C change in temperature per kHz of coating on the device surface. Values for this variable are empirically measured by determining the effect of polymer thermal expansion on polymer-coated SAW device frequency. Assuming that the fractional free volume factor is a constant (Grate and Zellers 2000), the volume term can be reduced to the form in Eq. 11, giving

$$\Delta f_V = (\ \Delta f_S \ C_V \ K / \ \rho_s \ ) + \ v_v \ C_V \ K \ S \qquad (19)$$

Thus, this response model fits the form of the matrix model in Eq 16.[a]

As formulated above, the vapor mass and volume contribute to a single sensor signal, the resonant frequency. Accordingly, the density of the vapor cannot readily be obtained directly because separate signals for each transduction mechanism are not provided by the measurement. Using other measurement approaches on SAW sensors that provide two separate signals per sensor, Martin and coworkers obtain velocity and attenuation changes, which are related primarily to mass and volume respectively. By obtaining two signals per sensor, the vapor response behavior observed (e.g., in parametric representations) could be related to vapor density (Martin and Frye 1990; Martin et al. 1994).

Generalized Response Model and Solution

The mathematics underlying this approach are not necessarily limited to sensor arrays. In its most general form, this work represents a classification method for a multivariate detector where the response of each channel can be modeled with a linear relationship based on a set of sample descriptors. Unknown samples can then be characterized and classified in terms of those descriptors. The most general response models and solutions can be expressed as follows.

$$\mathbf{R} = \ \mathbf{S_V} \ \mathbf{C} \ 10^{(\mathbf{V'P'})} \ \mathbf{S_P} \qquad (20)$$

$$\{\log (\ \mathbf{R} \ \mathbf{S_P}^{-1}) \} \ \mathbf{P'}_a{}^T (\mathbf{P'}_a \mathbf{P'}_a{}^T)^{-1} = \mathbf{V'}_a \qquad (21)$$

The matrix $\mathbf{S_V}$ is a diagonal matrix containing analyte specific parameters that influence the response independent of the specific interactions of the analyte with each channel. Specific interactions of the analyte with detector channel properties are modeled in the general linear

---

(a) Note that the volume fraction of diluent vapor in a vapor/polymer solution is symbolized by ϕ_v in this paper, whereas v_1 was used to symbolize the same volume fraction in the recent paper on modeling SAW responses as mass and volume terms. The symbol v is used in the present paper as v_v, representing the specific volume of the vapor as a liquid, not the volume fraction of diluent vapor.

relationship $\mathbf{V'P'}$, where $\mathbf{V'}$ contains analyte descriptors (not necessarily vapor analytes) and $\mathbf{P'}$ contains parameters specific to the properties of detector channels. The matrix $\mathbf{Y}$ in Eq. 12 is an example of a specific $\mathbf{S_V}$ matrix, containing vapor specific volumes that influence sensitivity but are not part of $\mathbf{VP}$. The matrix $\mathbf{S_P}$ contains channel specific sensitivity parameters, like $\mathbf{S}$ in Eq. 12 or $\mathbf{D^{-1} F}$ in Eq. 4. Augmentation of $\mathbf{V'}$ and $\mathbf{P'}$ to capture $\mathbf{S_V C}$ leads to the solution for $\mathbf{V'_a}$ containing the analyte descriptors in $\mathbf{V'}$ as well as the log of the product of analyte specific sensitivity factor times the analyte concentration.

Discussion

The conventional approach for processing sensor array data relies on training to develop pattern recognition algorithms for specific analytes or classes of analytes, an approach used successfully in many types of multivariate analysis including electronic noses (Grate 2000; Albert et al. 2000; Jurs et al. 2000). It would be desirable, however, if sensor arrays can be designed and/or trained to recognize analyte properties, just as an infra red spectrum can indicate the presence of particular types of bonds and functionalities in an analyte. A pattern vector from a sensor array is mathematically equivalent to a spectrum, both being first order responses (Booksh and Kowalski 1994).

The chemometric approach described in this paper and its preceding paper represents a synthesis of chemometric methods with knowledge of the response mechanisms of the sensors. Given knowledge of the sensor materials and their interactions with vapors, it is possible to extract information about those vapors from the array. Thus, once an array has been suitably calibrated on known vapors, the array response to an unknown vapor can be converted to several descriptors of the detected vapor, even if the vapor was not included in the original training.

In our previous publication (Grate et al. 1999), we derived the approach for extracting analyte descriptors from multivariate instrument responses using underlying theory for LSERs and mass-transducing acoustic wave sensor responses. We claimed that the approach might be used for any multivariate instrument where the responses of each channel could be modeled with a suitable linear relationship. The mathematics in the first manuscript accounted for analyte/detector channel interactions (vapor/polymer interactions for SAW sensors) and channel specific calibration factors (such as those in D and F). No analyte specific calibration factors that were independent of analyte/detector channel interactions were included.

In this paper, we show for the first time how this new approach for obtaining chemical information from multivariate instruments can, in fact, be applied to a technology other than mass-transducing acoustic wave sensor arrays. The QCM is the most obvious example of a mass-transducing acoustic wave sensor. If the new method were restricted to only QCMs, it would be of limited application. By revising the approach to include analyte-specific calibration factors, such as the vapor specific volume, the method can be applied to sensors that respond to the amount of sorbed analyte, whether the amount is a mass or a volume. Thus, the methodology is now more general, and can include volume-transducing chemical sensors such as the polymer/carbon black composite chemiresistors.

We also show for the first time how the approach can be applied to estimate descriptors for sensors where each sensor provides one signal containing components due to the mass and volume of the sensed analyte. Optimizations can be used successfully to correctly obtain the descriptors. We will elaborate on this result using simulations in a future paper. With this capability, the method will be applicable to sensors such as the SAW device, which is a very popular platform for sensor array instruments. Thus, the method is not restricted to strictly mass-transducing acoustic wave sensors, but is more generally applicable to a variety of acoustic wave sensor array types.

The inclusion of analyte-specific calibration factors leads further to a more general formulation with pleasing simplicity and symmetry, as shown in Eq. 20 and 21. This now includes analyte specific calibration factors in SV, analyte/detector channel interactions in V'P', and channel-specific

13

calibration factors in SP. The possibility of analyte specific calibration factors was not present in our previous publication. The formulation in Eq. 20 and 21, along with derivations for additional sensor technologies, substantiates our previous claim for generality beyond mass-transducing sensor arrays.

Application of these approaches to polymer-coated sensor arrays requires that the responses are due solely to absorption in the bulk of the polymer and that the sorption process can be accurately modeled with a linear descriptor-containing relationship such as the LSER. In addition, the method requires that the polymers in the array be diverse. In our previous paper on mass-transducing sensors, we used synthetic sensor data and simulations to explore the sensitivity of the descriptor solutions to measurement noise in the array data. We found that 10 to 20% measurement noise was acceptable for obtaining descriptors to a precision that is similar to the precision with which the descriptor values are known (Grate and Zellers 2000). The root-mean-square errors of prediction (RMSEP) for each of the parameters grew approximately linearly with noise. The parameter errors for $_2^H$, $_2^H$, and $_2^H$ were approximately 0.06, 0.02, and 0.03, respectively, for 20% noise in the sensor responses. This is comparable to the error in the original parameters, which can be taken as about 0.03 units for the $_2^H$, $_2^H$, and $_2^H$ parameters. The error for the log L^{16} parameter can be taken as 0.1 units. In the simulations, the log L^{16} error at 20% sensor noise was somewhat larger at 0.3 – 0.4 log units.

In addition, it was shown that found descriptor values could be used to classify vapors by comparison with tabulated descriptor values for known compounds. In the simulations, error bounds of two times the RMSEP were constructed for each parameter. The number of test vapors falling within these error bounds for all parameters was determined as a function of noise. Averaging over all the vapors in the test set, typically two or fewer vapors were within the error bound for noise levels up to about 10%.   In addition, for the vast majority of cases where more than one match was found, the found vapors included the correct vapor plus one or more compounds within the same compound class.

If sensor array response vectors are related to solvation parameters, it follows that any property or process that can also be correlated with solvation parameters could be correlated with array responses. Solvation parameters can be used to calculate such properties as the air/water partition coefficient, sorption in biological tissues and fluids, respiratory tract irritation in mice, eye irritation and nasal pungency thresholds in man, and possibly also odor thresholds in man(Abraham 1993; Abraham and Weathersby 1994; Abraham et al. 1994; Nielsen et al. 1996; Abraham et al. 1996; Abraham 1996; Alarie et al. 1998a, b; Alarie et al. 1996; Alarie et al. 1995; Abraham et al. 1998a, b, c, d). Thus, a scientific basis exists for relating sensor array responses to parameters that are relevant to environmental fate, toxicology, environmental health, and olfaction.

References

Abraham, M. H. 1993. *Chemical Society Reviews 22*, 73–83.

Abraham, M. H. 1996. In *Indoor Air and Human Health*; R. B. Gammage and B. A. Berven, Ed.; Lewis Publishers: New York, pp 67–91.

Abraham, M. H., J. Andonian-Haftvan, J. E. Cometto-Muniz, and W. S. Cain. 1996. *Fundam. Appl. Toxicol. 31*, 71–76.

Abraham, M. H., J. Andonian-Haftvan, G. Whiting, A. Leo, and R. W. Taft. 1994. *J. Chem. Soc., Perkin. Trans. 2 **1994***, 1777–1791.

Abraham, M. H., R. Kumarsingh, J. E. Cometto-Muniz, and W. S. Cain. 1998a. *Ann. N. Y. Acad. Sci. 855*, 652–656.

Abraham, M. H., R. Kumarsingh, J. E. Cometto-Muniz, and W. S. Cain. 1998b. *Toxicology in Vitro 12*, 403–408.

Abraham, M. H., R. Kumarsingh, J. E. Cometto-Muniz, and W. S. Cain. 1998c. *Toxicology in Vitro 12*, 201–207.

Abraham, M. H., R. Kumarsingh, J. E. Cometto-Muniz, and W. S. Cain. 1998d. *Arch. Toxicol. 72*, 227–232.

Abraham, M. H., and P. K. Weathersby. 1994. *J. Pharm. Sci. 83*, 1450–1456.

Alarie, Y., J. Andonian-Haftvan, G. D. Nielsen, and M. H. Abraham. 1995. *SAR and QSAR in Environ. Res. 134*, 92–99.

Alarie, Y., G. D. Nielsen, and M. H. Abraham. 1998a. *Pharmacol. Toxicol. (Copenhagen) 83*, 270–279.

Alarie, Y., M. Schaper, G. D. Nielsen, and M. H. Abraham. 1998b. *Arch. Toxicol. 72*, 125–140.

Alarie, Y., M. Schaper, G. D. Nielsen, and M. H. Abraham. 1996. *SAR and QSAR in Environ. Res. 5*, 151–165.

Albert, K. J., N. S. Lewis, C. L. Schauer, G. A. Sotzing, S. E. Stitzel, T. P. Vaid, and D. R. Walt. 2000. *Chem. Rev. (Washington, D. C.) 100*, 2595–2626.

Beebe, K. R., R. J. Pell, and M. B. Seasholtz. 1998. *Chemometrics: A Practical Guide*; John Wiley and Sons, Inc.: New York.

Booksh, K. S., and B. R. Kowalski. 1994. *Anal. Chem. 66*, 782A–791A.

Doleman, B. J., M. C. Lonergan, E. J. Severin, T. P. Vaid, and N. S. Lewis. 1998. *Anal. Chem. 70*, 4177–4190.

Ferry, J. D. 1980. *Viscoelastic Properties of Polymers;* 3rd. Ed.; John Wiley and Sons, Inc.: New York.

Grate, J. W. 2000. *Chemical Reviews 100*, 2627–2648.

Grate, J. W., and M. H. Abraham. 1991. *Sens. Actuators B 3*, 85–111.

Grate, J. W., M. H. Abraham, and R. A. McGill. 1996. In *Handbook of Biosensors: Medicine, Food, and the Environment*; E. Kress-Rogers and S. Nicklin, Ed.; CRC Press: Boca Raton, FL, USA, pp 593–612.

Grate, J. W., M. Klusty, W. R. Barger, and A. W. Snow. 1990. *Anal. Chem. 62*, 1927–1924.

Grate, J. W., M. Klusty, R. A. McGill, M. H. Abraham, G. Whiting, and Andonian-Haftvan. 1992. *J. Anal. Chem. 64*, 610–624.

Grate, J. W., S. J. Martin, and R. M. White. 1993a. *Anal. Chem. 65*, 940A–948A.

Grate, J. W., S. J. Martin, and R. M. White. 1993b. *Anal. Chem. 65*, 987A–996A.

Grate, J. W., A. Snow, D. S. Ballantine, H. Wohltjen, M. H. Abraham, R. A. McGill, and P. Sasson. 1988. *Anal. Chem. 60*, 869–875.

Grate, J. W., B. M. Wise, and M. H. Abraham. 1999. *Anal. Chem. 71*, 4544–4553.

Grate, J. W., and E. T. Zellers. 2000. *Anal. Chem. 72*, 2861–2868.

Jurs, P. C., G. A. Bakken, and H. E. McClelland. 2000. *Chem. Rev. (Washington, D. C.) 100*, 2649–2678.

Lonergan, M. C., E. J. Severin, B. J. Doleman, S. A. Beaber, R. H. Grubbs, and N. S. Lewis. 1996. *Chem. Mater. 8*, 2298–2312.

Lundberg, B., and B. Sundqvist. 1986. *J. Appl. Phys. 60*, 1074–1079.

Martin, S. J., and G. C. Frye. 1990. *Appl. Phys. Lett. 57*, 1867–1869.

Martin, S. J., G. C. Frye, and S. D. Senturia. 1994. *Anal. Chem. 66*, 2201–2219.

Nielsen, G. D., M. H. Abraham, L. F. Hansen, M. Hammer, C. J. Cooksey, J. Andonian-Haftvan, and Y. Alarie. 1996. *Arch. Toxicol. 70*, 319–328.

Ruschau, G. R., R. E. Newnham, J. Runt, and B. E. Smith. 1989. *Sens. Actuators 20*, 269–275.

Severin, E. J., B. J. Doleman, and N. S. Lewis. 2000. *Anal. Chem. 72*, 658–668.

Severin, E. J., and N. S. Lewis. 2000. *Anal. Chem. 72*, 2008–2015.

Talik, P., M. Zabkowska-Waclawek, and W. Waclawek. 1992. *J. Mater. Sci. 27*, 6807–6810.

Wise, B. M., and N. B. Gallagher. 1996. *J. Process Control 6*, 329–348.

Distribution

No. of Copies		No. of Copies	

OFFSITE

ONSITE

2   U.S. Department of Energy
    1000 Independence Avenue SW
    Washington D.C. 20585
       Attn: M. O'Connell
            B. Hoffheins

12 Pacific Northwest National Laboratory
     G. B. Dudder (5)   K6-48
     J. W. Grate (5)   K8-93
     Technical Report Files (2)