SANDIA REPORT
SAND97-0782 • UC-606
Unlimited Release
Printed April 1997

Evaluation of an Evanescent Fiber Optic Chemical Sensor for Monitoring Aqueous Volatile Organic Compounds

Dianna S. Blair

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Approved for public release; distribution is unlimited.

Sandia National Laboratories

SF2900Q(8-81)

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.
NOTICE: 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 any of their employees, nor any of their contractors, subcontractors, or 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, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

SAND97-0782
Unlimited Release
Printed April 1997

Distribution
Category UC-606

Evaluation of an Evanescent Fiber Optic Chemical Sensor for Monitoring Aqueous Volatile Organic Compounds

Dianna S. Blair
Environmental Characterizations and
Monitoring Systems Department
Sandia National Laboratories
P. O. Box 5800
Albuquerque, NM 87185-0755

Abstract

Linear chemometric algorithms were used to model the quantitative response of an evanescent fiber optic chemical sensor in aqueous mixtures with concentrations between 20 and 300 ppm. Four data sets were examined with two different experimental arrangements. Two data sets contained trichloroethene, 1,1,2 trichloroethane, and toluene. Partial Least Squares, PLS, and Principal Component Regression, PCR, algorithms performed comparably on these calibration sets with cross-validated root mean squared errors of prediction (RMSEP) for trichloroethene, 1,1,1 trichloroethane, and toluene of approximately 26, 29 and 22 ppm, respectively. The third data set contained trichloroethene, 1,1,2 trichloroethane, toluene, and chloroform and the fourth contained these four analytes as well as tetrachloroethene. Again, both chemometric algorithms performed comparably on a given data set with RMSEP for trichloroethene, 1,1,2 trichloroethane, toluene, and chloroform of approximately 6, 6, 9, and 16 ppm from the first set, and 7, 11, 13, and 31 ppm from the second set with tetrachloroethene RMSEP of 31 ppm. The decrease in the quantitative performance of the sensor for modeling toluene and chloroform upon addition of tetrachloroethene to the sample solutions is due to increased cladding absorption features in the spectral response matrix. These features overlap with the analyte absorption features of toluene and chloroform. These results suggest one of the limitations with this type of sensing format.

Acknowledgments

The author would like to thank Drs. David Haaland and Edward Thomas, of Sandia National Laboratories (SNL), for their scientific contributions to this work and their critical reviews of this paper.

Contents

Abstract 1
Introduction 5
Theory 7
Experimental 10
Results and Discussion 13
Conclusions 30
References 31
Appendix A - Experimental Design 34

Figures

1 Diagram of original evanescent fiber optic chemical sensor holder. 11
2 Drawing of final evanescent fiber optic chemical sensor holder design. 12
3 Single beam evanescent reference spectrum of fiber. 14
4 FT-IR transmission spectra of neat analytes. 15
5 EFOCS response spectra to 50 ppm aqueous solutions. 16
6 IR transmission and eigenvector spectra for EFOCS to toluene. 18
7 IR transmission and eigenvector spectra for EFOCS to TCE. 19
8 IR transmission and eigenvector spectra for EFOCS to TCA. 20
9 TCE predicted concentrations based on PCR models. 22
10 TCA predicted concentrations based on PCR models. 23
11 Toluene predicted concentrations based on PCR models. 24
12 Reference versus predicted concentration for toluene. 25
13 Sensor response to PCE compared to inherent fiber spectrum. 26
14 EFOCS response to toluene compared to first weight loading vectors 27
15 EFOCS response to TCE compared to first weight loading vectors. 28
16 Reference versus predicted concentration for TCE. 29

Tables

1 Chemometric Performance 17
2 Independent Validation 21
3 Modeling Response Data 25

Introduction

Performing quantitative and qualitative monitoring of low level volatile hydrocarbons in aqueous solutions is both difficult and essential for environmental and public health operations. These operations include industrial waste water and process monitoring (1), site assessment and remediation (2), and disinfection byproduct control for the drinking water and beverage industries (3). However, measurements can be problematic in that the analytes of concern are present at low concentrations in a large, potentially interfering, background of water. In a typical analysis of such a sample the volatile organic contaminate (VOC) analytes are separated from the water by sample pretreatment. For example, the EPA approved method for this type of analysis, Method 8260, requires purge-and-trap sample preparation which allows analyte separation from the water phase and preconcentration prior to sample injection into a gas chromatograph. An alternate approach is to use a solvent-free extraction technique. The extraction system uses a silicone coated, silica core fiber mounted on a microsyringe in which the silicone isolates organic analytes that are then thermally injected into a gas chromatograph (4).

An alternative to these staged methods of sample pretreatment prior to a traditional analysis method is to use polymeric membranes directly integrated into a sensor platform where VOCs can preferentially partition. For platforms such as quartz crystal microbalance and surface acoustic wave devices, probing changes in the membranes, such as acoustic frequency shifts caused by chemical absorption that changes the mass and visco-elastic properties of the membrane, provides inferential information about the presence of chemical analytes (5,6). Another application that has been examined extensively in the past decade employs polymer coated optical wave guides for measurements involving evanescent wave interactions. Polymers have been used to detect organic molecules in aqueous solutions (7-10) and sol-gels have been used for pH measurements (11). Inherently weak in nature, the information obtained by evanescent interactions can be enhanced by using polymeric membranes that separate and concentrate the analytes of concern from their environment. Hydrophobic polymers are

frequently employed as cladding material for optical wave guides in both the visible and near-infrared spectral regions (12-17). These materials provide mechanical strength, environmental protection, and elimination of water interference by excluding water from reaching the optical region of the sensor.

Sensors based on this principal, known as evanescent fiber optic chemical sensors, EFOCS, or fiber optic evanescent wave sensors, FEWS, hold promise for use in environmental applications. Specifically, EFOCS can use commercially available plastic clad silica, PCS, fiber optic cables as the optical waveguide element. This results in a relatively inexpensive sensor element for alternatives to the traditional analysis methods of aqueous sample monitoring in both environmental and industrial applications. The silica core, silicone clad fibers are relatively inexpensive and have good transmission properties in the overtone/combination band region of the electromagnetic spectrum which is the region of interest for spectroscopic detection of volatile organic compounds.

This work examines the quantitative response of an EFOCS sensor to concentrations in the parts-per-million range of aqueous solutions containing volatile organic compounds (VOCs). Four different data sets will be examined. Solutions in the data sets contain various VOCs frequently present at environmental contamination sites. These VOCs include trichloroethene (TCE), 1,1,2 trichloroethane (TCA), toluene, chloroform, and tetrachloroethene (PCE). In particular, the usefulness of linear chemometric algorithms for building models to predict these analyte concentrations based on the sensor equilibrium response will be evaluated.

Theory

Evanescent Wave Spectroscopy

Light can be guided in an optical element, or wave guide, at very high efficiencies by total internal reflection. The phenomenon occurs when light, traveling in a medium of refractive index n_1, strikes the boundary of a second, relatively lower refractive index, medium, n_2, at an angle, measured with respect to normal, greater than θ_c, the critical angle, defined as (18):

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{1}$$

When light undergoes this total internal reflection, a small amount of energy penetrates into the second medium. Known as the evanescent wave it has a depth of penetration, d_p, defined as distance from the interface where intensity of the electric field is reduced to e^{-1} the intensity at the interface. This depth into the second medium is (19):

$$d_p = \frac{\lambda_1}{2\pi\left(\sin^2\theta - n_{12}^{\,2}\right)^{1/2}} \tag{2}$$

where Θ is the angle of incidence at the interface with respect to normal, λ_1 is the wavelength of light in the waveguide medium, λ/n_1, and n_{12} is the refractive index ratio of the waveguide and surrounding medium, n_1/n_2. Absorbing species at the interface will interact with the evanescent wave. Resultant changes in guided light intensity provides qualitative and quantitative information about species present in the evanescent field. Infrared spectroscopy is performed in this format when the sample to be examined is highly absorbing. Traditionally known as attenuated total reflection, ATR, the planar optical elements transport light in a narrow range of angles. Therefore, light strikes the interface between the optical element and sample at a discrete range of angles. However, when the light guiding element is an optical fiber this sampling technique is known as

evanescent wave spectroscopy, EWS. The mathematics that describe the phenomenon are more complex than traditional ATR due to the range of angles that are present at the core/cladding interface. The cone of light that an optical fiber can accept and guide from an infinite source is referred to as the numerical aperture, NA, and is defined as (20):

$$NA = n_{ext} \sin\theta_a = \left(n_1^2 - n_2^2\right)^{1/2} \tag{3}$$

where n_{ext} is the refractive index of the external medium, θ_a is acceptance angle of the fiber, and n_1 and n_2 are the refractive indices of the core and cladding, respectively.

If infrared active species are present in the evanescent field of an optical fiber, light intensity will decrease according to the following relationship (12):

$$-\log\left(\frac{I}{I_o}\right) = \alpha_e Lc + \log\left(\frac{NA_o^2}{NA^2}\right) \tag{4}$$

where I is transmitted light intensity after analyte exposure, I_o is reference intensity with no analyte present, L is fiber length, c is molar concentration, and NA and NA_o are numerical apertures of the fiber with and without analyte present, respectively. Effective molar absorptivity, α_e, is defined,

$$\alpha_e = \varepsilon\eta_p \tag{5}$$

where ε is molar absorptivity of infrared active species and η_p is the fraction of light entering the fiber that is present in the cladding,

$$\eta_p = \frac{k}{V} \tag{6}$$

In equation 6, k is a proportionality constant, determined by Gloge (21) to be 1.89 for equilibrium mode distribution in a fiber. The denominator, known as the V parameter, determines the number of optical modes a fiber can support and is defined as,

$$V = \frac{\left[2\pi r\left(n_1^2 - n_2^2\right)^{1/2}\right]}{\lambda} \tag{7}$$

where r is the fiber radius.

Methods for Modeling Spectral Data

The use of statistical and mathematical methods to extract information efficiently from chemical data is known as chemometrics. Originally based on statistical data analysis techniques borrowed from other disciplines, such as econometrics, chemometrics is a growing chemistry sub discipline. Currently, new methods and techniques are being developed to solve problems specific to the analysis of chemical data.

The method of extracting specific information about a sample from instrumental responses consists of three steps: calibration, validation, and prediction. Calibration is a procedure used to build an empirical model that correlates instrument response to known variables. Specifically, for spectroscopy the calibration process builds a model using statistical methods that correlate a physical parameter, such as analyte concentration with spectroscopic response for a set of samples. Validation examines model precision. This is performed to verify that the model is correlated to spectral features associated with the analyte, and not with spectral noise. The prediction phase then uses the model built from the calibration set to estimate the unknown concentration from a second set of spectral data.

Factor analysis decomposes a response matrix, **A**, such as the sensor spectra, of rank r into a sum of $\bar{r}$ pseudorank matrices that determine the true underlying dimensionality of the matrix and describe the number of variables that are linearly independent. The decomposition can be accomplished by a variety of methods, but the general result is that **A** is decomposed into the product of at least two matrices:

$$\mathbf{A} = \mathbf{T}\mathbf{P}^{\mathrm{T}} \qquad (8)$$

where **T** is a matrix of scores dimensioned m x r, where m is the number of samples, and $\mathbf{P}^{\mathrm{T}}$ are the loading vectors, r x n, where n is the number of spectral frequencies.

Two chemometric methods were used in this study. Principle Component Regression, PCR, and Partial least-squares (PLS) analysis. PCR is a calibration method based on either the results from a Nonlinear Iterative Partial Least Squares, NIPALS, (22)

or Singular Valued Decomposition, SVD, (23) of the response matrix. PLS analysis incorporates information from both the response and concentration matrices into the scores and loadings vectors. This results in a model that can, in general (23), model concentration variation with fewer linear terms. The response and concentration matrices are decomposed into smaller matrices in an iterative procedure that exchanges information between the two data sets. Incorporation of concentration information during the generation of the scores and loadings results in placing relevant spectral information, i.e., that associated with concentration variation, into the earlier scores and loadings.

Experimental

Over the course of this research, two experimental arrangements were used.

Experimental Arrangement 1

Spectroscopic instrumentation consisted of a Nicolet 800 FT-IR spectrometer configured for the near-infrared. A CaF_2 beamsplitter, quartz halogen near-infrared source, and a liquid nitrogen cooled InSb detector were used. The EFOCS was coupled to the spectrometer using a fiber optic interface manufactured by Nicolet. Fibers were physically attached to the interface using standard fiber optic connectors. Fine tuning of light position on the InSb detector was performed using an X, Y, and Z translational stage at the fiber insertion point. The number of scans co-added per spectra were 32 at 4 cm^{-1} resolution using Happ-Genzel apodization. Data collection, processing, and storing was performed on a Nicolet 680 computer.

The sensor was constructed by winding jacketed fiber in a 5 cm diameter loop on an aluminum holder designed to support the fiber in a rigid configuration, Figure 1. The fibers were commercially available nylon jacketed 200 μm silica core fibers clad in 230 μm

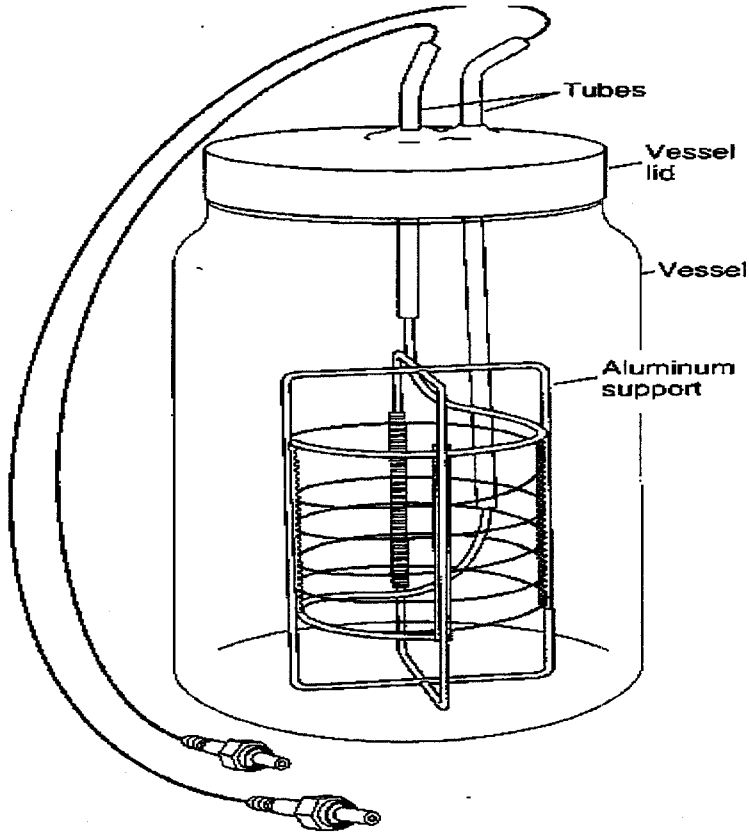

Figure 1. Diagram of evanescent fiber optic chemical sensor on an aluminum holder designed to provide mechanical support of the fiber. Included in the drawing is the glass exposure vessel used during the experiments.

silicone purchased from Fiberguide Industries. The jacket was removed from the active length of fiber by placing the wound fiber into boiling 1,2-propanediol for 5 minutes. The stripped lengths of fiber used for results reported here were 1 m. The inactive length of fiber was protected from the jacket stripping procedure by passing the ends through a rubber septum on the end of a 0.5 cm I.D. glass tube.

Experimental Arrangement 2

Spectroscopic instrumentation consisted of a Bomem MD-155 Fourier-transform infrared (FT-IR) spectrometer configured for the near-infrared. A KCl beamsplitter, quartz halogen near-infrared source, and a liquid nitrogen cooled InSb detector were used. The EFOCS was coupled to the spectrometer using a Bomem fiber optic interface

accessory manufactured by Bomem. Fibers were physically attached to the interface using standard fiber optic connectors. The number of scans coadded per spectra were 64 at 8 cm^{-1} resolution using cosine apodization. Data collection, processing, and storing was performed on a PC based computer using a Galactic software package.

The sensor was constructed by winding 8 m of jacketed fiber in a 6.4 cm diameter loop onto a slotted anodized aluminum holder designed to support the fiber in a rigid configuration, Figure 2. This holder was designed and built for the EFOCS for field deployment of the sensor. Of primary concern in the designing and building of the holder was breakage of the fiber element. During early stages of development and evaluation of the EFOCS technology it was found that the silica core of these fibers, stripped of their

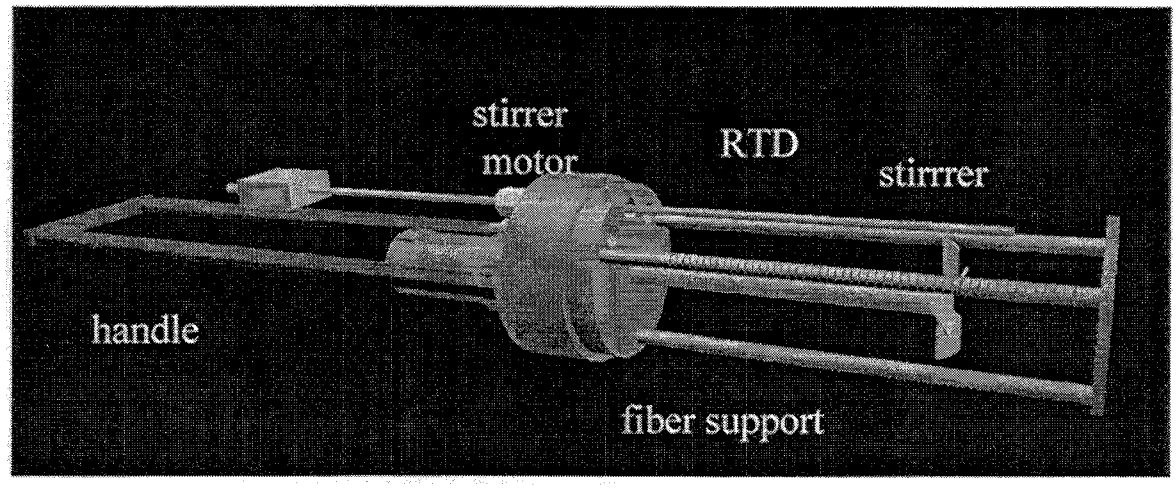

Figure 2. Evanescent fiber optic chemical sensor holder.

protective nylon jackets and wound in a relatively tight geometry, underwent stress, crack, corrosion due to attack by water. The result was frequent mechanical failure of the fiber. Therefore, this design was made to eliminate any tight bends in the fiber that would cause such failure. A stirrer was incorporated into the EFOCS holder design to reduce water phase diffusion times of the analytes, and solution temperatures were monitored using an Omega RTD.

General

For both experimental setups, 400 μm silica core/480 μm silica clad fibers were used to couple the sensing fiber to the optical benches. These glass-on-glass fibers were used to minimize changes in modal distribution of the sensing fiber caused by temperature fluctuations and fiber movements. Further, this coupling technique also allowed the sensing fiber to be used in a remote location since no inherent C-H absorption bands are present in the evanescent field of glass-on-glass fibers to attenuate transmission intensity. Coupling between fibers was accomplished by using bulk head type connectors with 1,2-propanediol as a coupling gel.

Toluene, TCE, TCA, chloroform, and PCE were all purchased from Aldrich Chemical Company. Aqueous solutions were made volumetrically. Based on earlier diffusion studies (10), it was determined 20 minutes exposure time was sufficient the analyte concentration in the 230 μm silicone cladding fiber to reach equilibrium. Therefore, during these quantitative studies the EFOCS was exposed to the mixture solution for this period of time prior to collecting the spectrum. The experimental designs that were used to generate the calibration data for prediction modeling are outlined in Appendix A.

Results and Discussion

All reference spectra were taken with the EFOCS in water. A typical single beam reference spectrum is shown in Figure 3. This spectrum illustrates the wavelength dependent light loss determined by source characteristics, absorption and scattering losses in the optical path, and detector responsivity. The decreased light transmission observed between 5500 and 6000 cm^{-1} is due to energy absorption by the overtone bands of the fundamental methyl functional groups present in the silicone cladding. It is a primary factor contributing to the sensor detection limit. Total analyte signal is a function of fiber length, equation 4. However, light transmission in the C-H overtone spectral region,

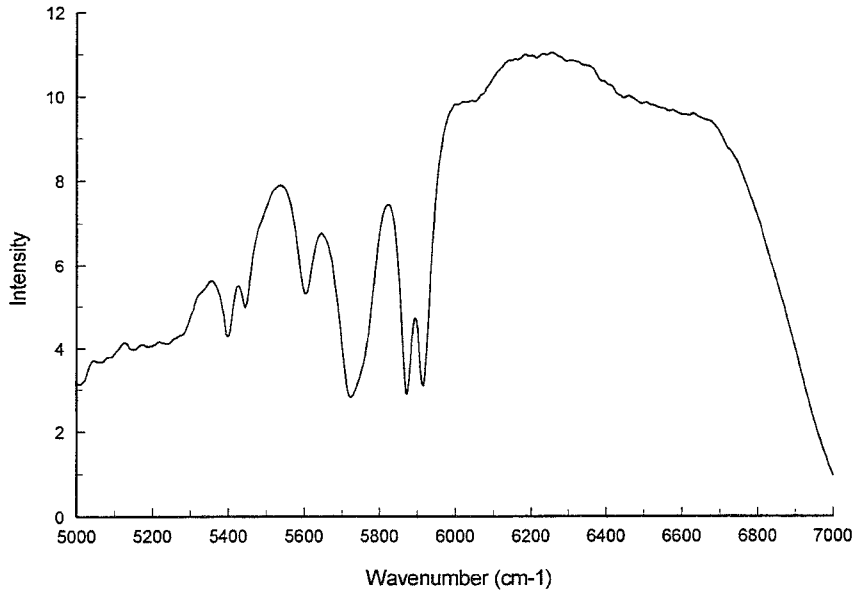

Figure 3. Single beam evanescent reference spectrum of 1 m plastic clad silica fiber in near-infrared spectral region.

where the analytes of concern absorb energy, determines the useful fiber length. For example, Figure 4 shows the near infrared transmission spectra of the five organic compounds studied. Clearly, the active spectral region for these compounds is between 5500 and 6200 cm^{-1}.

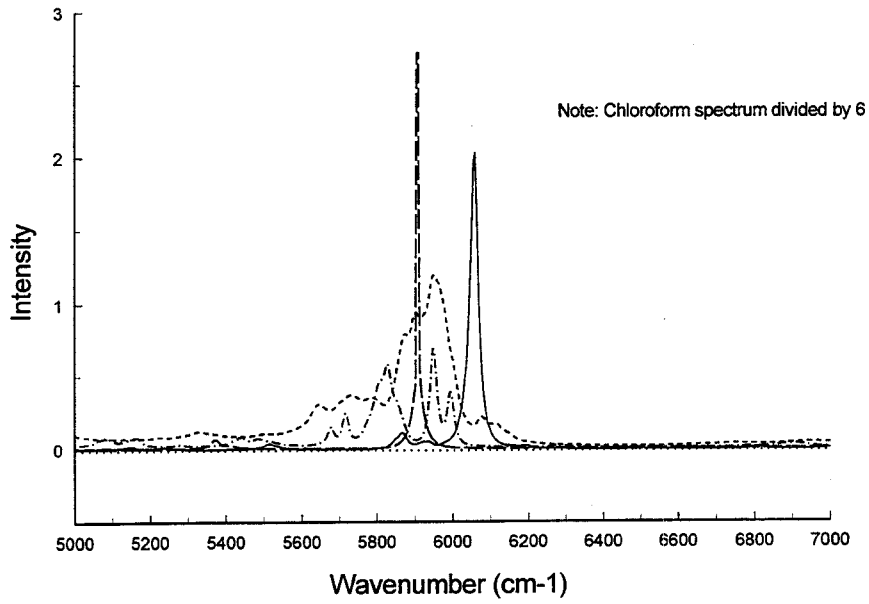

Figure 4. FT-IR transmission spectra of neat analytes in 2 mm cuvette for a) TCE, b) TCA, c) toluene, d) chloroform, and e) PCE.

The equilibrium sensor responses to 50 ppm aqueous solutions of the five analytes studied are shown in Figure 5. The sensor response to TCE, TCA, toluene, and chloroform resemble the neat analyte spectra in Figure 4. However, each spectra also has additional spectral features related to the silicone cladding. The presence of cladding absorption features in the sensor analyte response spectra is caused by increasing evanescent field penetration depth after analyte exposure due to changes in the effective cladding refractive index, see equation 2. The changing penetration depth is evidenced

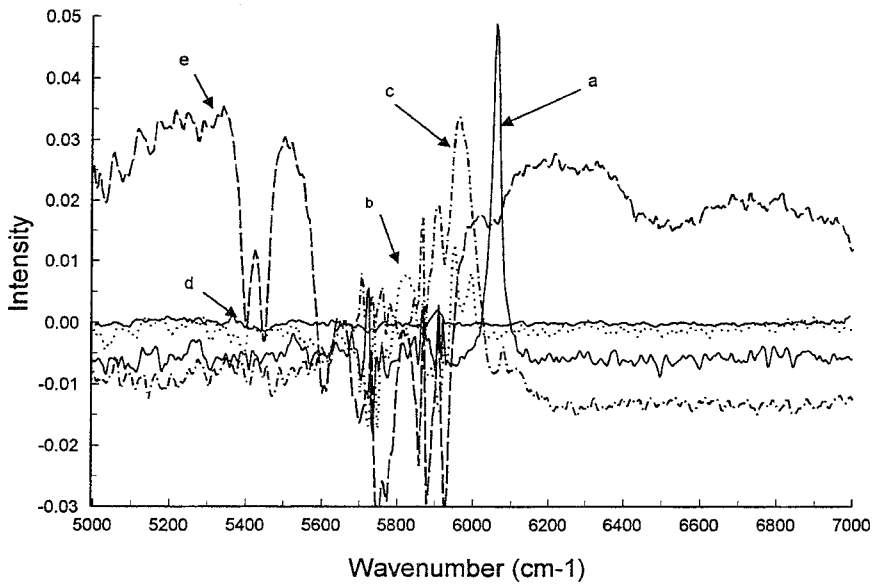

Figure 5. 8 m EFOCS response spectra to 50 ppm aqueous solutions of a) TCE , b) TCA, c) toluene, d) chloroform, and e) PCE.

most strongly in the sensor response to PCE which has no spectral features in the region of interest but has a relatively high refractive index, 1.51. The effective refractive index of the cladding after exposure to PCE is much greater than the cladding refractive index during acquisition of the reference spectrum. This perturbation results in what appears to be negative absorption features and a positive baseline shift. The relative intensity of the spectral features that result from exposing the sensor to a 50 ppm solution of PCE compared to the 50 ppm solutions of the other analytes suggests that this analyte has a higher partitioning coefficient in the silicone than the other analytes. Without the presence of absorption features in the analyte itself, the presence of this chemical in a mixture of other analytes in an aqueous sample clearly represents the most difficult

analyte to quantitate. Sensor response to the other analytes, all with refractive indices greater than the cladding, show negative baseline shifts. This is probably due to lack of temperature control in the exposure bath. The reference spectra were taken with the sensor immersed in a water bath at room temperature while the sensor responses to the analytes were taken in water actively stirred by a motor on the sensor holder. The warmed motor was observed to raise the water temperature up to 6° C. Temperature increase is known to increase the throughput of fiber optic cables due to the temperature dependent material dispersion and could explain the baseline shift (24).

Both PCR and PLS were used to build calibration models from all data sets. Models were generated on mean centered data using cross-validation. Outlier detection was performed by examining the concentration F-ratio statistic (23). The results of the calibrations from Set 1 with the three analytes TCE, TCA, and toluene, is summarized in Table 1 based on the chemometric algorithm used. No outliers were detected and both

Table 1
Performance of chemometric
models generated using Set 1.

	PCR		PLS	
Analyte	RMSEP (ppm)	factors	RMSEP (ppm)	factors
trichloroethene	26	3	25	3
trichloroethane	29	4	30	4
toluene	22	3	23	3

algorithms performed comparably on this data set. The RMSEP for all three analytes did not differ significantly between each other or between the different algorithms. The concentration residuals for this data set indicates an approximate 20 ppm error at the two lower concentration levels and about 30 ppm at the higher level with the mean value for residuals scattered around zero. The optimal number of factors determined for each analyte model did not change between algorithms. The spectral variation of the factorial

design, as determined by the first three eigenvectors obtained from PCR, is shown in Figs. 6 to 8. Included in each of the figures are one neat analyte spectrum obtained in

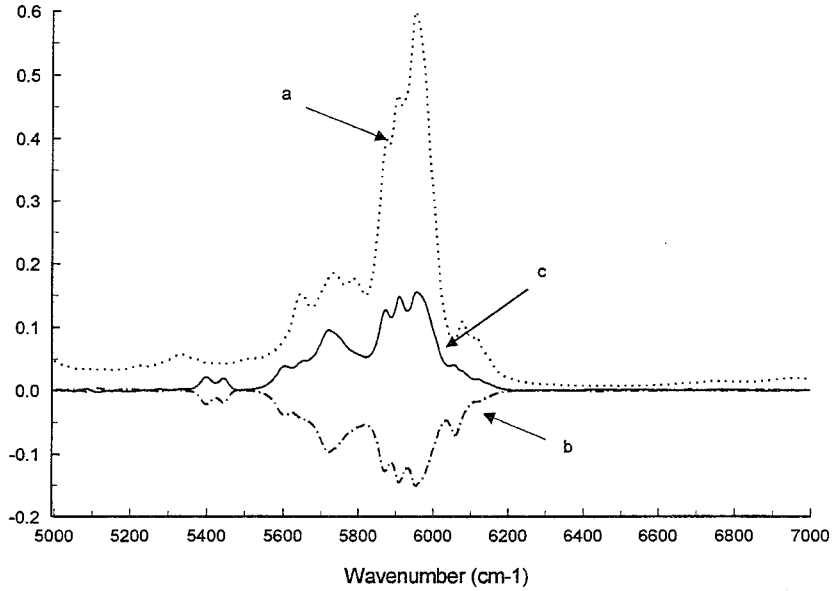

Figure 6. (a) Infrared transmission spectrum of neat toluene taken in transmission through a 1 mm cuvette, (b) first PCR eigenvector from calibration model generated using Set 1, and (c) first weight loading vector from PLS generated toluene calibration model from Set 1.

transmission mode through a 1 mm cuvette sample. As is apparent, spectral features related to toluene are present in the first eigenvector while, in general, the dominant spectral features in the second and third eigenvectors are related to TCE and TCA, respectively. An obvious exception is the presence of the trichloroethene related

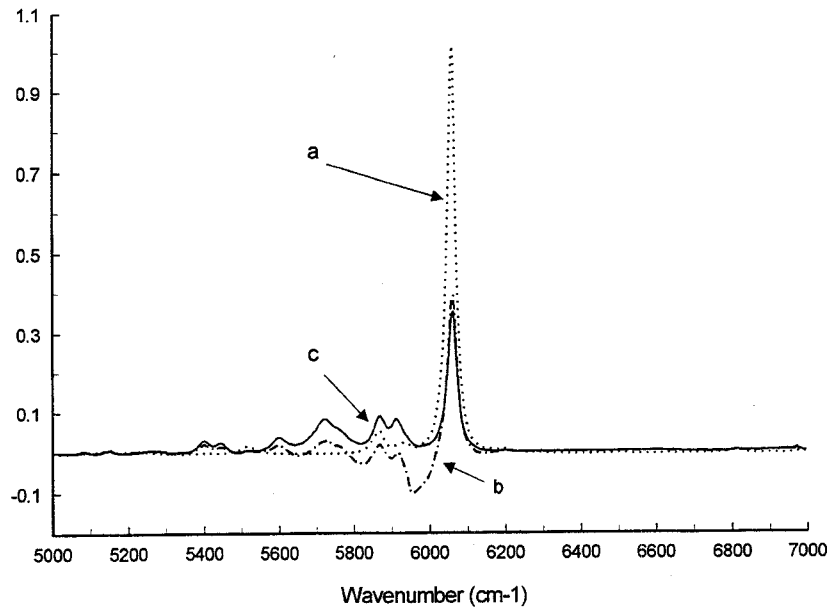

Figure 7. (a) 1 mm cuvette infrared transmission spectrum of neat trichloroethene, (b) second PCR eigenvector from calibration model generated based on Set 1, and (c) first weight loading vector from PLS generated trichloroethene calibration model from Set 1.

absorption feature at 6063 cm^{-1} present in the third eigenvector. Also included in each of the three figures is the first weight loading vector of the PLS generated calibration model for the respective analyte. The spectral similarities between these loading vectors, the transmission spectrum of their respective analyte, and the corresponding PCR eigenvector are readily apparent.

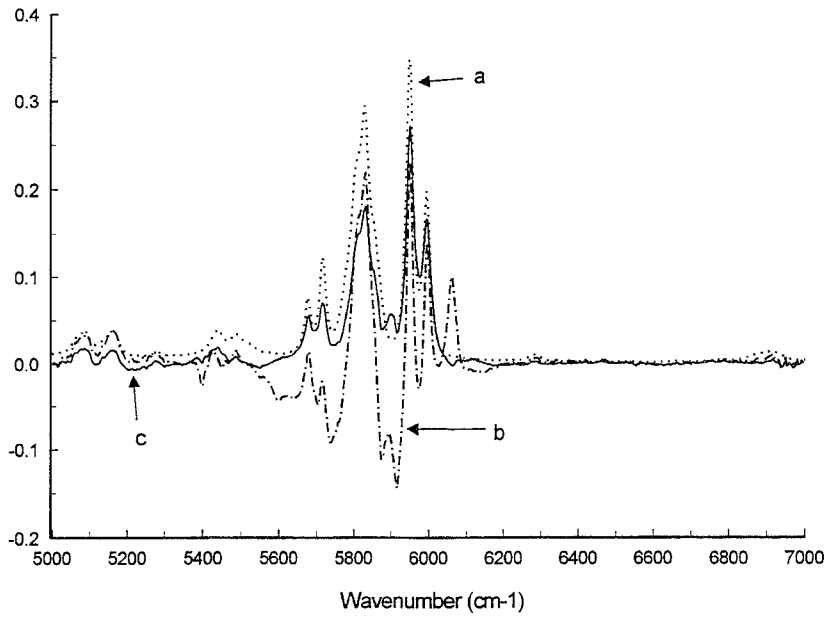

Figure 8. (a) Infrared transmission spectrum of 1 mm cuvette sample of neat 1,1,1 trichloroethane, (b) third PCR eigenvector from calibration model generated using Set 1, and (c) first weight loading vector from PLS generated trichloroethene calibration model from Set 1.

PCR and PLS calibrations on Set 2, were also performed and compared to the models generated using an independent experimental design, Set 1, but using the same three analytes of concern. One of the 15 samples in this data set was identified as an outlier and eliminated from the data set prior to calibration. The resultant loading vectors were similar to those obtained on the full factorial data set. However, to directly compare the performance of the models generated using the Set 2 data with the Set 1 data requires a correction for the different number of samples used to generate the model. The higher number of samples included in the factorial design data set results in models with less errors. Just like signal averaging, increasing the number of samples in a data set lowers the error, or noise, as a square root of the number of samples used. Therefore, to partially correct for this discrepancy between the data sets, the RMSEP for each model was multiplied by the square root of the ratio of number of samples in Set 2 to those in Set 1. The RMSEPs from the resultant PCR models were, 29, 23, and 32 and from PLS were 29, 20, and 32 for the three analytes trichloroethene, trichloroethane, and toluene, respectively.

Independent validation of both the PCR and PLS generated models was obtained by predicting aqueous concentrations in Set 2 with the model generated on Set 1 data .

Table 2
Independent validation of models generated using data from Set 1 to predict on Set 2 data.

	PCR	PLS
Analyte	RMSEP (ppm)	RMSEP (ppm)
trichloroethene	29	29
trichloroethane	27	27
toluene	40	39

The results are listed in Table 2. As can be seen, the performance of the PCR and PLS model, on the independent data set, is comparable to that on the model building data set.

Plots of the predicted concentration versus reference concentration are shown in Figures 9-11 for TCE, TCA and toluene, respectively. Included in each figure is the least squares

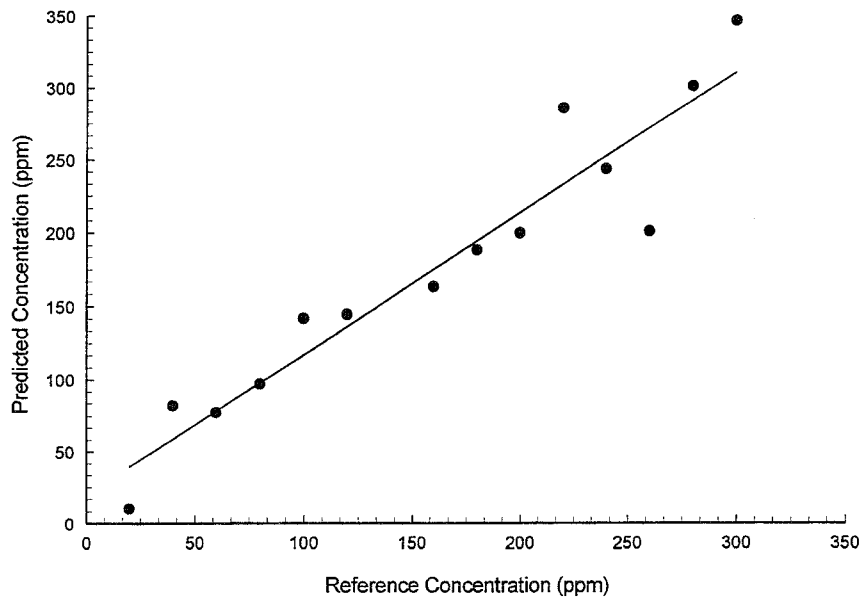

Figure 9. Trichloroethene concentration, reference vs. concentration predicted by PCR model using data from Set 1.

fit of the predicted versus reference concentrations. The models for both TCE and TCA appear to have no relative bias with respect to prediction, with calculated slopes for the least squares lines of 0.97 and 0.98, respectively. The toluene model, however, exhibited a tendency to underpredict the concentration. This was demonstrated as a relative bias and a least squares slope of 0.70. In general, the models performed well with limited failures observed. At the lowest trichloroethane concentration both models predicted a concentration close to zero. The model also underpredicted the lowest trichloroethene concentration. Whereas these data points are not tagged as outliers in the generation of

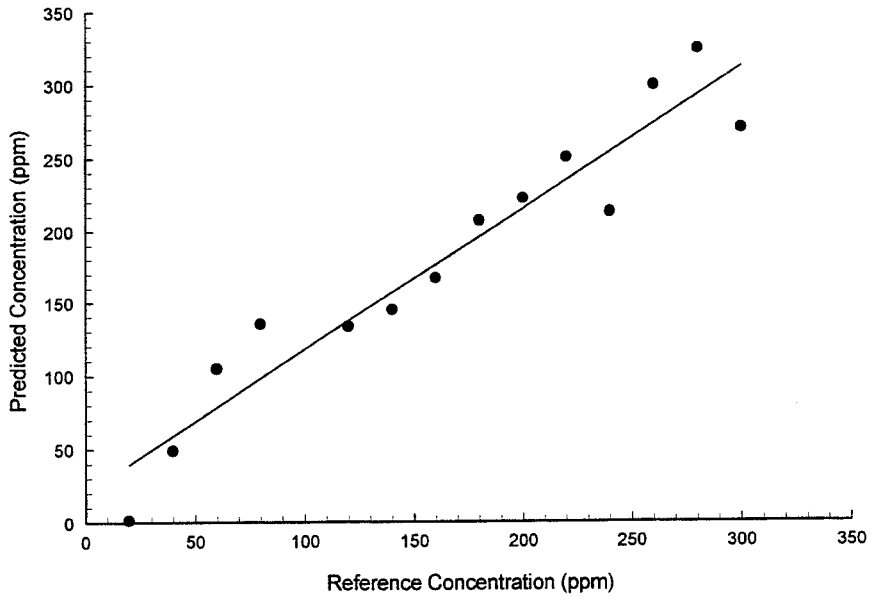

Figure 10. Reference versus predicted concentration for 1,1,1 trichloroethane. Concentration predicted by PCR model generated using Set 1.

any of the models, these concentrations are close to the sensors detection limit for these analytes and therefore could pose problems in both the calibration and prediction phases.

The same procedure, discussed above, was used to generate calibration models on Sets 3 and 4. The results of these calibrations are summarized in Table 3. No significant difference in the precision of the models generated using either algorithm for a particular data set was observed. In general, it can be observed that PCR required a higher number of factors to obtain comparable performance to PLS. This is due to the incorporation of

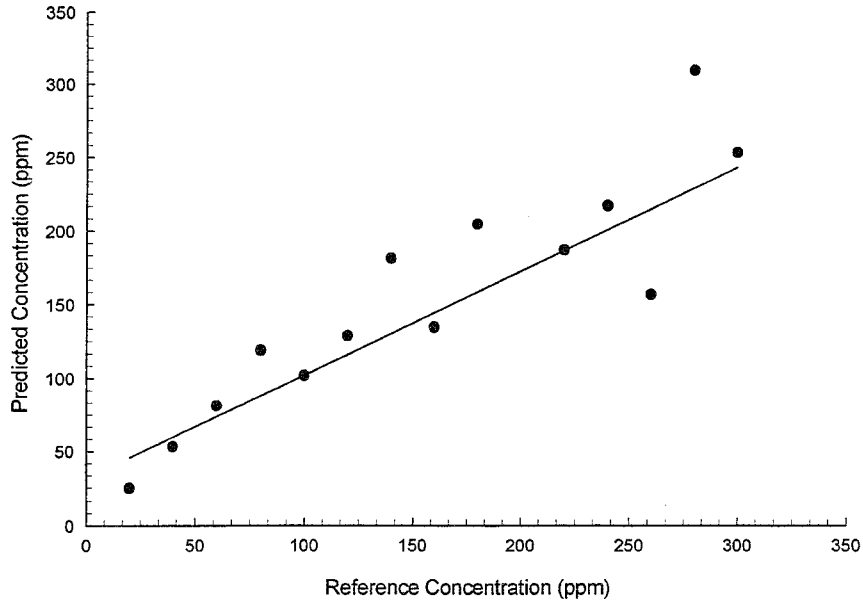

Figure 11. Toluene concentration, reference vs. concentration predicted on Set 1 by PCR.

chemical information earlier in the model building with PLS (23). However, the models generated on data Set 3 performed significantly better, as determined by F-ratio test at the 95% confidence limit, for toluene and chloroform than the models generated on data Set 4. This reduced precision is illustrated by plotting reference concentration versus the PLS model predicted concentration for both data sets, Figure 12. The greater error associated with the predicted concentrations can be observed across the entire data set with what appears to be a positive bias at the low concentration range.

Table 3
Modeling of EFOCS response data
from two data sets using different
chemometric algorithms

Analyte	PLS RMS Set 1	No. of Factors	PCR RMS Set 1	No. of Factors	PLS RMS Set 2	No. of Factors	PCR RMS Set 2	No. of Factors
toluene	6	5	6	9	13	5	12	7
TCE	6	3	7	2	7	3	7	4
TCA	9	4	9	6	11	6	13	9
chloro	16	8	19	13	31	6	28	7
PCE					31	3	28	4

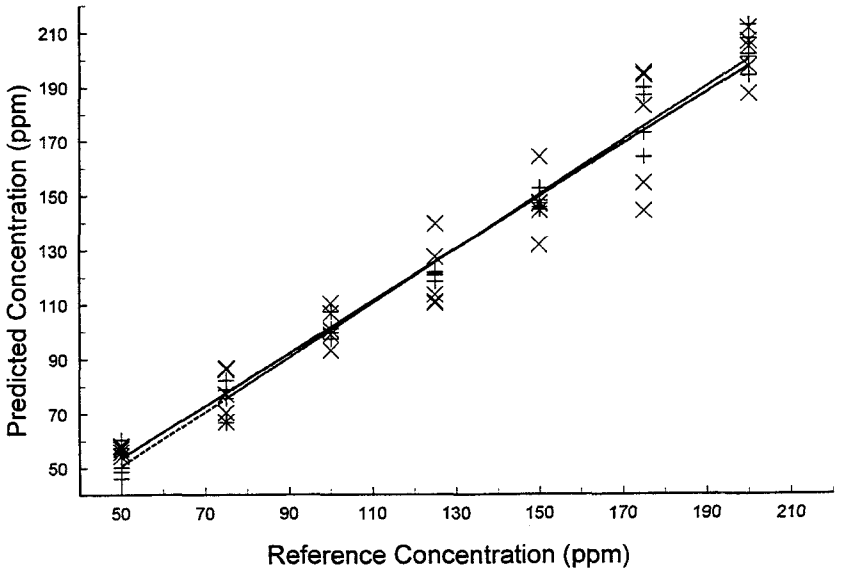

Figure 12. Reference versus predicted concentration for toluene based on PLS generated models from Set 3 (+) and Set 4 (x).

As discussed above, the presence of PCE in the solution mixtures resulted in a significant contribution of the cladding spectral features to spectral variation of the sensor response. The relationship between the cladding absorption bands and the PCE concentration is best illustrated in Figure 13 where the first PLS weight loading vector for the PCE model, Set 4, is plotted. Also included in the figure is the ratioed spectrum of two

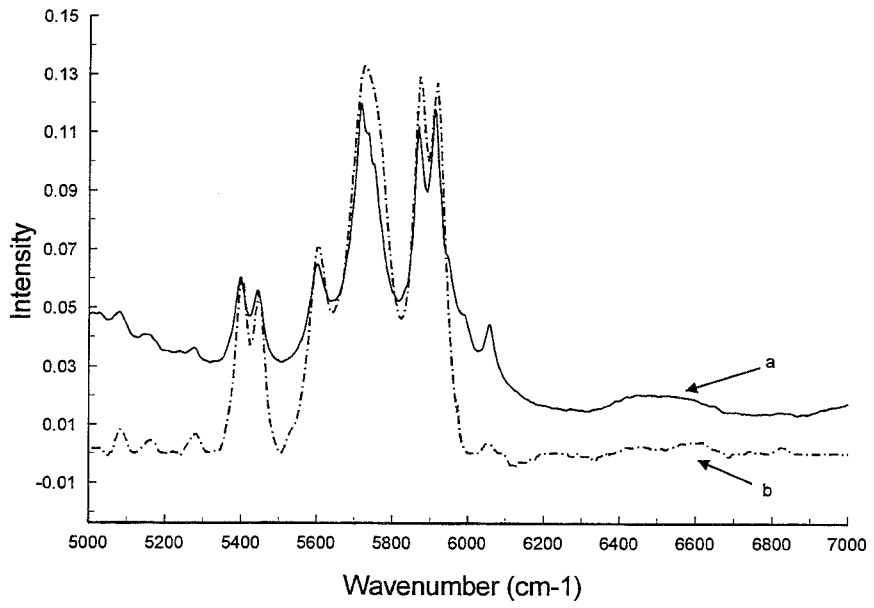

Figure 13. (a) First weight loading vector of PLS model from Set 4 and (b) ratioed EFOCS spectra with silicone clad fibers of different length.

fibers of different lengths which shows the spectral contribution due to the silicone cladding. The spectral similarities between these two spectra clearly shows that the model generated for PCE concentration is related primarily to absorption bands overlapping

with the cladding spectral features that undergo a relatively high variation in Set 4. Figure 14 shows the first weight loading vectors for the toluene models generated from both data sets. Also included in the figure is the sensor response spectrum to a 50 ppm aqueous toluene solution for comparison. As can be observed, similar spectral features are present in all vectors. However, intensity of features associated with the analyte is of lower relative intensity in Set 4. This was also observed for chloroform. Therefore, analytes with absorption features that overlap with the cladding absorption features become more difficult to model when a relatively high refractive index analyte is present due to spectral interference, especially at low toluene or chloroform concentrations where the spectral contribution from the analyte of concern is minimal compared to the spectral variance caused by the PCE.

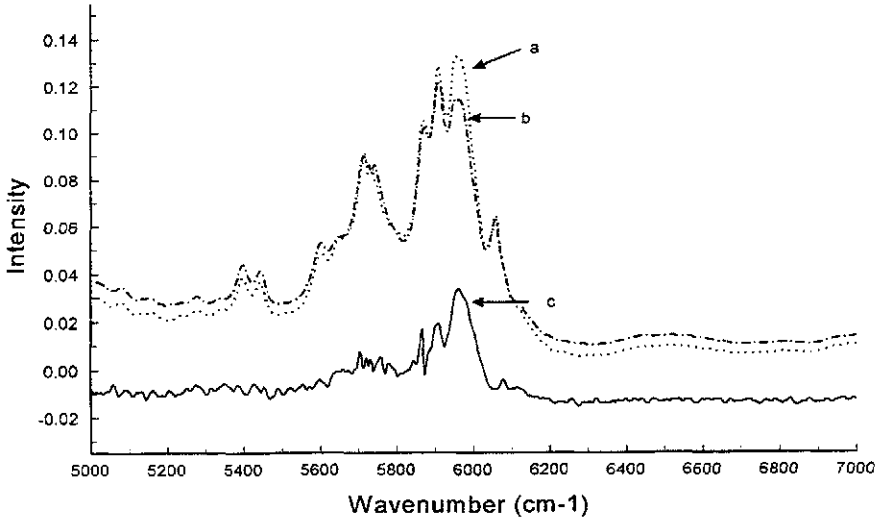

Figure 14. (a) First weight loading vector for toluene PLS model from Set 1, (b) first weight loading vector for toluene PLS model from Set 2, and (c) EFOCS response to 50 ppm aqueous toluene solution.

The ability to model TCE and TCA concentrations using PLS or PCR comparably in both data sets is due to the presence of analyte spectral features shifted from the cladding bands. This is best illustrated in Figure 15 where the first weight loading vectors for the TCE calibration models for both data sets are plotted. Also included in this plot is the sensor response spectrum to a 50 ppm aqueous solution of TCE. The spectral features associated with the analyte and silicone cladding for the loading vectors are of comparable relative intensity, suggesting that spectral variance caused by the presence of PCE in the mixtures did not interfere with the ability of PLS to model the TCE concentration. This

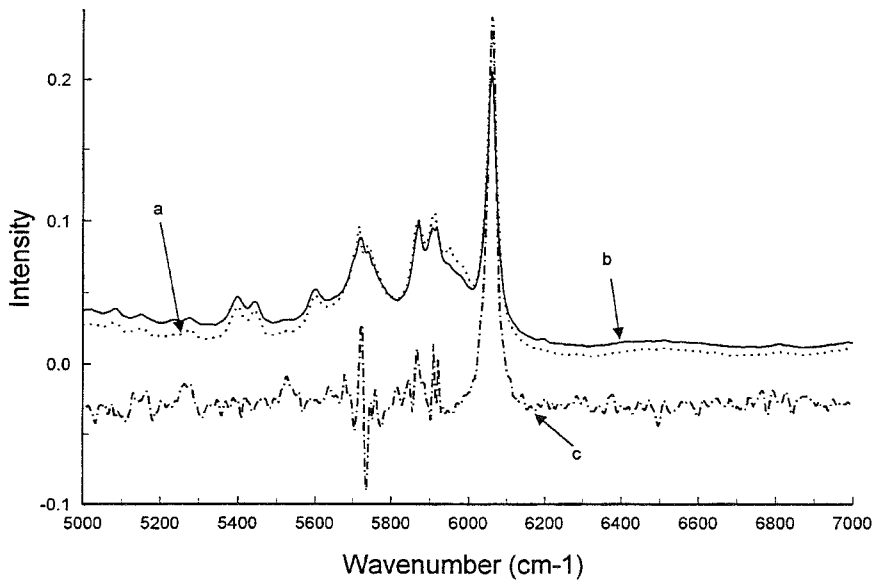

Figure 15. (a) First weight loading vector for TCE PLS model from Set 3, (b) first weight loading vector for TCE PLS model from Set 4, and (c) EFOCS response to 50 ppm aqueous TCE solution.

reduced interference is illustrated by plotting reference concentration versus the PLS model predicted concentration for both data sets, Figure 16. The data points overlay each other almost exactly. However, the reduced precision at the higher concentrations shows that the linear methods used here may not fully fit the data. This lack of fit may be due to a number of nonlinear effects that are occurring. The reference analyte concentrations used

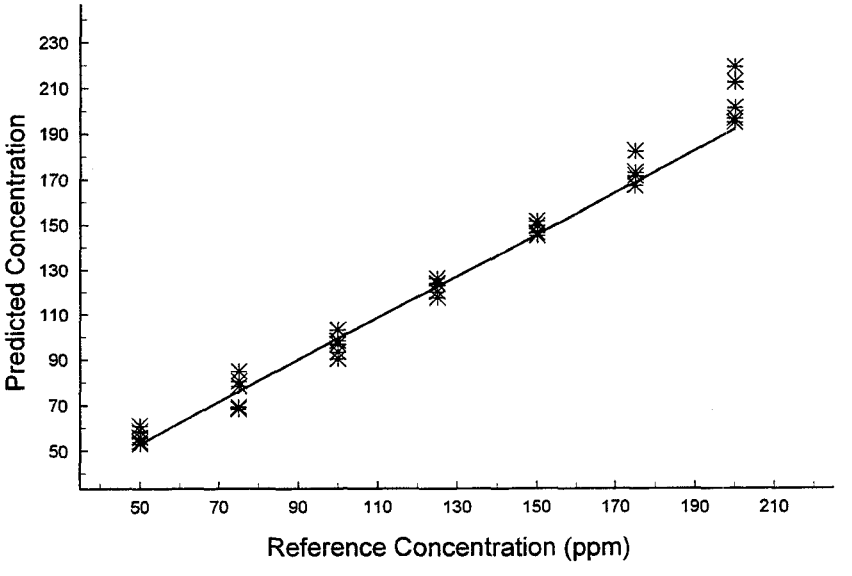

Figure 16. Reference versus predicted concentration for TCE based on PLS generated models from Set 3 (+) and Set 4 (x).

in generating all models were aqueous analyte concentrations. In general, partitioning of these organic analytes into silicone from aqueous solutions increases its relative concentration. This is the premise for this type of sensing format. However, this partitioning, as well as any co-solvency effects that could be occurring due to the high concentration of other analytes, has not been included in the generation of the calibration

models. Furthermore, as discussed earlier, a change in sample pathlength occurs when the relatively high index analytes diffuse into the cladding. This change, a nonlinear effect, also has not been incorporated into the models generated using either algorithm.

Conclusions

Linear chemometric algorithms, such as PLS and PCR, were used to model TCE, TCA, and toluene concentrations in mixtures between 20 and 300 ppm. Both algorithms performed comparably on the calibration sets with RMSEP's typically less than 10% of the concentration range examined. Both algorithms performed comparably on a given data set. For data sets obtained from aqueous solutions of toluene, TCE, TCA, and chloroform between 50 and 200 ppm, both algorithms performed comparably with RMSEPs again typically less than 10% of the concentration range. However, addition of PCE to the mixtures, while allowing comparable precision for predicting concentrations of TCE and TCA, reduced the precision for toluene and chloroform. This is due to spectral overlap of the analyte absorption features with the silicone cladding that undergoes significant variance due to the presence of the relatively high refractive index PCE. These results suggest the limitation for this type of sensing format due to the frequent presence of PCE in contaminated ground water sites.

References

1) Blaser, W. W.; Bredeweg, R. A.; Harner, R. S.; LaPack, M. A.; Leugers, A.; Martin, D. P.; Pell, R. J.; Workman, J., Jr.;, and Wright, L. G.; *Anal. Chem.* **1995,** *67*, 47R-70R.

2) Cline-Thomas, T.; Mills, W.; Trach, A.; *Field Screening Methods for Hazardous Wastes and Toxic Chemicals: Proceedings of the 1993 U.S. EPA/A&WMA Intl. Sym.*, **1993**, *2*, 771-782.

3) Hitoshi, U.; Nakamuro, K.; Moto, T.; Yasuyoshi, S.; *Proc. of the IWSA International Specialized conference on Advanced Treatment and Integrated Water System Management into 21st Centrury*, Osaka, Jpn, **1995**, *13*, 171-176.

4) Louch, D.; Motlagh, S.; Pawliszyn, J.; *Anal. Chem.*, **1992**, *64*, 1187-1199.

5) Okahata, Y.; Ebato, H.; *Anal. Chem.*, **1989**, *61*, 2185-2188.

6) Wohltjen, H.; *Sens. and Act.*, **1984**, *5*, 307-325.

7) Krska, R.; Kellner, R.; Schiessl, U.; Tacke, M.; Katzir, A.; *Appl. Phys Lett.* **1993**, *63*, 1868-1870.

8) Krska, R.; Rosenberg, E.; Taga, D.; Kellner, R.; Messica, A.; Katzir, A.; *App. Phys. Lett.* **1992**, *61*, 1778-1780.

9) Kvasnik, F.; McGrath, A.D.; *SPIE Chemical, Biochemical, and Environmental Sensors* **1989**, *1172*, 75-82.

10) Blair, D. S.; Burgess, L. W.; Brodsky, A. M.; *Anal. Chem.* submitted.

11) Blue, R.; Stewart, G.; *Int. J. Optoelec.* **1995**, *10,* 211-222.

12) DeGrandpre, M. D.; Burgess, L. W.; *App. Spec.* **1990**, *44*, 273-279.

13) DeGrandpre, M. D.; Burgess, L. W.; *ISA Transactions* **1988**, *28*, 71-77.

14) Burck, J.; Conzen, J. P.; Bechaus, B.; Ache, H. J.; *Sens. and Act. B* **1994**, *18-19*, 291-295.

15) DeGrandpre, M. D.; Burgess, L. W.; *Anal. Chem.* **1988**, *60*, 2582-2586.

16) Burck, J.; Conzen, J. P.; Ache, H. J.; *Fres. J. Anal. Chem.* **1992**, *342*, 394-400.

17) Burck, J.; Conzen, J. P.; Ache, H. J.; *Appl. Spec.* **1993**, *47,* 753-763.

18) Snyder, A. W.; Love, J. D.; *Optical Waveguide Theory*, Chapman and Hall: New York, 1983.

19) Harrick, N. J.; *Internal Reflection Spectroscopy*, Interscience: New York, 1967.

20) Kapany, N S.; *Fiber Optics*, Academic Press: New York, 1967.

21) Gloge, D.; *Appl. Optics* **1971**, *10*, 2252-2258.

22) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T.; *Numerical Recipes in Pascal: The Art of Scientific Computing*, Cambridge University Press: Cambridge, 1989.

23) Haaland, D. M.; Thomas, E. V.; *Anal. Chem.* **1988**, *60*, 193-1202.

24) Klunder G. L.;Burck, J.; Ache, H. J.; Silva, R. J.; Russo, R. E.; *Appl. Spectr.* **1994**, *48,* 387-41.

Appendix A

Data Set 1: Run Number	Trichloroethylene	Trichloroethane	Toluene
1	150	300	150
2	150	300	20
3	20	300	20
4	150	20	300
5	300	150	150
6	150	300	300
7	20	150	150
8	150	150	300
9	20	150	300
10	20	150	20
11	150	150	20
12	20	300	150
13	20	20	150
14	300	20	150
15	300	300	150
16	20	300	300
17	150	20	20
18	300	150	300
19	20	20	20
20	150	150	150
21	300	20	20
22	300	300	20
23	300	150	20
24	20	20	300
25	300	150	300
26	300	300	300
27	150	20	150

Data Set 2: Run Number	Trichloroethylene	Trichloroethane	Toluene
1	300	240	80
2	220	80	300
3	120	120	120
4	260	60	20
5	80	20	160
6	280	40	240
7	140	180	180
8	100	100	200
9	60	280	140
10	20	220	280
11	180	300	260
12	40	140	40
13	160	200	60
14	200	260	100
15	240	160	220

Data Sets 3 & 4: Run Number	Trichloroethylene	Trichloroethane	Toluene	Chloroform	PCE (in Set 4 only)
1	75	150	200	100	50
2	150	125	125	150	75
3	200	175	175	175	125
4	175	50	100	125	100
5	50	75	75	200	150
6	100	200	150	50	200
7	125	100	50	75	175
8	175	125	50	100	75
9	50	200	75	200	150
10	100	150	200	75	125
11	200	175	175	50	50
12	150	100	125	125	175
13	75	50	100	150	200
14	125	75	150	175	100
15	125	175	150	175	100
16	200	125	50	200	75
17	50	75	175	75	200
18	150	200	125	150	150
19	75	100	75	125	50
20	100	50	100	50	125
21	175	150	200	100	175
22	175	200	125	200	200
23	150	100	50	175	100
24	125	50	200	125	150
25	50	150	100	150	75
26	75	175	75	50	50
27	100	125	175	100	125
28	200	75	150	75	175
29	150	75	125	50	50
30	200	50	200	200	150
31	175	200	75	75	125

Data Sets 3 & 4: Run Number	Trichloroethylene	Trichloroethane	Toluene	Chloroform	PCE (in Set 4 only)
32	50	150	50	100	175
33	100	175	100	150	200
34	125	100	150	125	100
35	75	125	175	175	75