

(To be submitted to IEEE Sensors Proceedings by October 1, 2005)

Developing Sensor Activity Relationships for the JPL Electronic Nose Sensors Using Molecular Modeling and QSAR techniques

A.V. Shevade, M. A. Ryan, M. L. Homer, A. D. Jewell,
H. Zhou, K. Manatt, and A.K. Kisor

Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena CA 91109

Abhijit.Shevade@jpl.nasa.gov, Margaret.A.Ryan@jpl.nasa.gov, Margie.L.Homer@jpl.nasa.gov, April.D.Jewell@jpl.nasa.gov,
Hanying.Zhou@jpl.nasa.gov, Kenneth.S.Manatt@jpl.nasa.gov, Adam.Kisor@jpl.nasa.gov

Abstract— We report a Quantitative Structure-Activity Relationships (QSAR) study using Genetic Function Approximations (GFA) to describe the polymer-carbon composite sensor activities in the JPL Electronic Nose (ENose), when exposed to chemical vapors at parts-per-million (ppm) concentration levels. A unique QSAR molecular descriptor set developed in this work combines the default analyte property set (thermodynamic, structural etc.) with sensing film-analyte interactions that describes the sensor response. These descriptors are calculated using semi-empirical and molecular modeling tools. The QSAR training data set consists of 15-20 analyte molecules specified by NASA for applications related to Life Support and Habitation in space. Statistically validated The applicability of these QSAR models was also tested independently to predict the sensor activities for test analytes not considered in the training set.

Key Words: Multivariate analysis, Molecular modeling, Quantitative Structure-Activity Relationship (QSAR), Electronic nose, Environmental monitoring, Polymer composite,

I. INTRODUCTION

The JPL electronic nose (ENose) consists of polymer-carbon composite sensing films used to detect organic vapors and other environmental contaminants for Life Support and Habitation in space applications [1,2]. The measured sensor responses are conductivity changes in polymer-carbon composite films upon exposure to target vapors (or analytes) at parts-per-million (ppm) concentrations. Training an array for a given set of analytes and a given set of environmental conditions (temperature, pressure, and humidity) is time consuming; in addition, developing training sets and

calibration information may decrease the useful lifetime of the sensors. The ability to predict sensor responses accurately will be of great help in characterizing sensing materials. The JPL ENose polymers fall in four categories (i) Hydrogen bond basic (ii) Dipolar and hydrogen bond basic Moderate dipolarity (iii) weak hydrogen bonding and weakly dipolar (iv) weak or no hydrogen-bonding.

Previous modeling efforts [3-5] to predict the response of polymer-film based sensors, primarily on pure polymer films, use Linear Solvation Energy Relationships (LSER) and solubility parameters to obtain a good correlation between the calculated and measured responses. Modeling efforts in the past for polymer-carbon composite sensing films based on solubility parameters [6], have taken into account only the effect of polymer-analyte interactions and assumed that neither carbon nor analyte in the film play a role in sorbing analyte molecules or in contributing to the response of the film. This model of sensor response may not represent a complete picture of response in polymer-carbon composite sensors.

We report a Quantitative Structure-Activity Relationships (QSAR) study using Genetic Function Approximations (GFA) with a unique molecular descriptor set, to describe the activities of the JPL ENose polymer-carbon sensors. In our work, the sensor activity for a given analyte is defined as the coefficient, A_1 , which is correlated to the sensor response as $y=A_1x + A_2x^2$, where x is the concentration of analyte. A_2 is generally three to five orders of magnitude smaller than A_1 . The unique QSAR descriptor set combines the default analyte properties (structural, spatial, topological, conformational, and thermodynamic) with descriptors for sensing film-analyte interactions, which describes the sensor response. The modeled analyte descriptors are calculated

using QSPR techniques. The polymer-analyte interaction energies that describe sensor response are calculated using molecular modeling tools.

II. EXPERIMENTAL DETAILS

The JPL ENose array consists of 32 sensors made from 16 polymer-carbon composite films. As previously described, the training of the JPL ENose sensors is done by delivering measured concentrations of analyte with controlled humidity at a given sensor substrate temperature. Response data for each analyte and each sensor are fit to an equation of the form $y=A_1x + A_2x^2$, where x is the analyte concentration and y is the normalized change in resistance. To develop sensor activity model that is concentration independent, the QSAR studies uses coefficient A_1 as the sensor activity to be correlated with molecular descriptors. The poly(styrene-co-maleic acid) sensing film response is selected for this study. The choice of the current system is to further add to our previous understanding by developing a model that relates the sensor activity to molecular descriptors. The analytes and concentration ranges considered for this study are shown in TABLE I.

III. MODEL DEVELOPMENT

A. QSAR Descriptors: Analyte properties and Sensor response

The QSAR descriptors considered for the study include those that describe intrinsic analyte properties as well as sensor response.

The default analyte descriptors for QSAR studies [7] that describe the intrinsic analyte properties fall into Electronic, Spatial, Structural, Thermodynamic and Topological categories. Sensor response at a molecular level is described by interactions between the sensing film and the analyte. Descriptors for sensor response include interaction energies of the sensing film components (polymer and carbon-black) with the analyte molecules (target molecule and water). Also added to the default descriptor set, is the vapor pressure of the analytes at 300K. The rationale behind using vapor pressure as a descriptor, is based on functional dependence of the sensing film partition coefficient to the bulk analyte concentration (hence the vapor pressure).

The pair interaction energies considered to represent the sensor response descriptors are: polymer-analyte, CB-analyte, polymer-water, and CB-water, analyte-analyte and analyte-water. Depending on the type of interactions considered, the suffixes could be polymer (p), carbon black (cb), analyte (a), or water (w). For example interaction energy between the polymer (p) and the target analyte (a) is denoted by E_{pa} .

A combined descriptor set that includes the default analyte descriptors along with sensor response descriptors was used in the QSAR studies to correlate the sensor coefficients A_1 with the molecular descriptors.

B. QSAR Descriptor calculations

Default analyte descriptors were predicted by empirical and semi-empirical Quantitative Structure Property Relationships (QSPR) using the commercial software Cerius² [8] on a Silicon Graphics O2 workstation.

Details on the molecular models and forcefield used for the sensing film-analyte interaction calculations are discussed elsewhere [7,9]. The sensing film-analyte interaction energies were calculated using Monte Carlo simulation techniques. Interaction energy between the polymer and an analyte molecule, E_{pa} , was calculated by fixing the polymer structure in space and sampling the analyte molecule around the polymer, then averaging the energy calculated over all the random analyte configurations generated around the polymer. In this study we used 10^5 configurations. E_{pw} , E_{p-cb} , E_{cb-a} , E_{cb-w} , E_{cb-cb} , E_{aa} , and E_{aw} descriptors were calculated similarly. The BLENDS module in the Cerius² software performs the calculations based on the methodology described above.

TABLE I: ANALYTE LIST AND CONCENTRATION RANGE TESTED IN PARTS-PER-MILLION (PPM) FOR ENOSE OPERATION (760 TORR, 23 °C). DATA ARE TAKEN AT A CONSTANT HUMIDITY OF 5000 PPM WATER (APPROXIMATELY 18% RELATIVE HUMIDITY).

Analyte	Concentration tested
	Low - High (ppm)
1. Acetone	64 - 600
2. Ammonia	6 - 60
3. Chlorobenzene	3 - 30
4. Dichloromethane	10 - 150
5. Ethanol	200 - 6000
6. Isopropanol	30 - 400
7. Xylenes (mixed)	33 - 300
8. Tetrahydrofuran	13 - 120
9. Trichloroethane	7 - 200
10. Acetonitrile	1 - 25
11. Ethylbenzene	20 - 180
12. Freon113	15 - 500
13. Hexane	15 - 150
14. Methyl ethyl ketone	15 - 150
15. Methane	1600 - 50000
16. Methanol	6 - 100
17. Toluene	5 - 50
18. Benzene	10 - 100
19. Indole	25 - 450
20. Dichloroethane	10 - 100

C. QSAR equation: model development and sensor activity representation

The first step in the QSAR model development process is to investigate the number of terms (N_{term}) and the functional forms (linear or linear-quadratic or spline) to be used in the QSAR equations. Based on our previous experience [7], we have chosen to go with linear-quadratic equation with 3-4 terms. The reliability and significance of the developed QSAR models are determined by using statistical parameters such as correlation coefficient r^2 and F .

The response of a polymer carbon sensing film to a given analyte molecule is based on how the sensing film components (polymer and carbon black) in the polymer-carbon composite interact with the analyte molecule. Therefore, the QSAR equation that we have chosen to represent a given sensor is selected from a set of cross-validated equations generated by the GFA algorithm. The selected equation is the statistically most significant one (largest r^2 value) of the equation set which also contains the polymer-analyte (E_{pa}) descriptor

IV. RESULTS AND DISCUSSION

The polymers selected for the current investigation is poly(styrene-co-maleic acid) (Figure 1), which falls in the Hydrogen bond basic category. The P_{ka} of polystyrene-co-maleic acid at 298 K was calculated [10] as 5.19 and 3.12. These two values correspond to contributions by the two carboxylic acid groups in the maleic acid monomer of the polystyrene-co-maleic acid polymer dissociating H^+ ions.

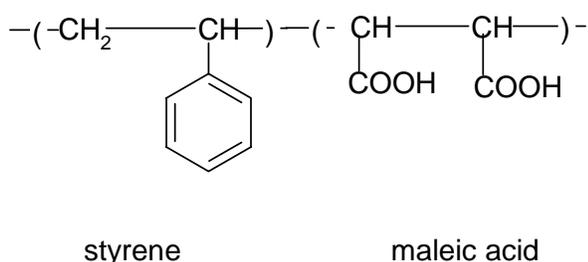


Figure 1: Poly(styrene-co-maleic acid) monomer unit

We have used a three term linear-quadratic form for the QSAR study. After cross-validation, the statistically significant equation for the poly(styrene-co-maleic acid) containing the descriptor E_{pa} was:

$$\begin{aligned} \text{Calculated activity} = & 1.00E-06 VP^2 \\ & - 1.48 E_{aw} \\ & + 0.39 E_{pa} \end{aligned}$$

$$(r^2=0.88, F=50.6)$$

The term VP, refers to the vapor pressure of the analyte. The strong interaction energy between the water and the two carboxylic groups in the polymer results in dominance of the analyte-water interaction energy term, hence the appearance of the E_{aw} in equation (1). A plot of QSAR calculated versus experimental sensor activities (A_1 coefficient) is shown in Figure 2. Equation (1) was used to predict the sensor coefficient A_1 for the test analytes: benzene, dichloroethane and indole. As seen in Figure 2, similar to the previous case, the model works satisfactorily for the benzene and dichloroethane and indole.

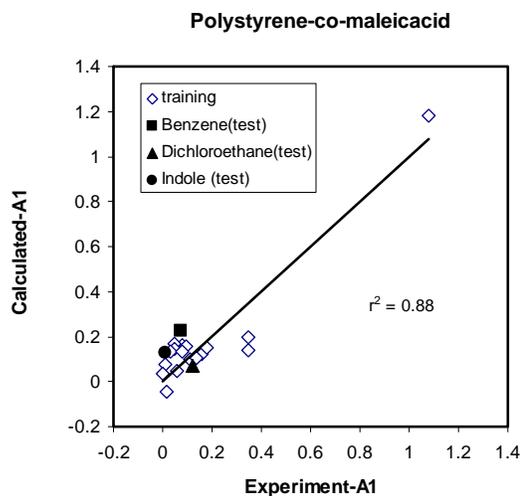


Figure 2: A plot of QSAR calculated vs. experimental sensor activity for the training and test analyte set using combined descriptor set. The calculated values for both the training and test analyte set are obtained using Equation (1). The r^2 value refers to the correlation between the calculated vs. the experimental values obtained for the training data set.

It should be noted that we do not expect to calculate the exact value of the coefficient A_1 using the QSAR approach described here. It is likely that there are effects influencing sensor response which are not accounted for in the models of interaction energy or in the physico-chemical properties of the analyte. For example, any addition of pathways for electronic conduction, such as ions in the polymer matrix, would result in a decrease in resistance, but such a decrease would not be accounted for in either the calculated interaction energies or in the default descriptor set.

V. CONCLUSIONS

Experimental data for a polystyrene-maleic acid carbon black sensor was correlated using QSAR with intrinsic analyte properties and molecular interaction energy terms. The model developed showed good correlation for the entire analyte set as well as analyte subsets. The descriptors that predict the polymer-carbon sensor response indicate that the polymer-analyte interaction is not the only important interaction to consider. In addition to predicting sensor response, it may be possible to elucidate the sensing mechanisms using this approach. The approach will be extended to other polymer composite sensors used in the JPL ENose system.

ACKNOWLEDGMENT

The research reported in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration. The JPL ENose development program is supported by Advanced Environmental Monitoring and Control, Life Support and Habitation, ESMD, NASA. The authors thank Prof. Anton Hopfinger and Dr. David Rogers for their valuable suggestions.

REFERENCE

- [1] M. A. Ryan, H. Zhou, M. G. Buehler, K. S. Manatt, V. S. Mowrey, S. P. Jackson, A. K. Kisor, A.V. Shevade and M. L. Homer, *IEEE Sensors Journal*, **4(3)**, 337 (2004).
- [2] M. A. Ryan, A. V. Shevade, H. Zhou and M. L. Homer, *MRS Bulletin*, **29**, 714 (2004).
- [3] J. W. Grate, S. J. Patrash, S. N. Kaganove, M. H. Abraham, B. M. Wise and N. B. Gallagher, *Anal. Chem.*, **73(21)**, 5247 (2001).
- [4] A. Hierlemann, E. T. Zellers and A. J. Ricco, *Anal. Chem.*, **73(14)**, 3458 (2001).
- [5] M. P. Eastman, R. C. Hughes, G. Yelton, A. J. Ricco, S. V. Patel, M. W. Jenkins, *J. Electrochem. Soc.*, **146(10)**, 3907 (1999).
- [6] W.A. Goddard III, T. Cagin, M. Blanco, N. Vaidehi, S. Dasgupta, W. Floriano, M. Belmares, J. Kua, G. Zamanakos, S. Kashihara, M. Iotov and G.H. Gao, *Comput. Theo. Poly. Sci.*, **11(5)**, 329 (2001).
- [7] A.V. Shevade, M. A. Ryan, M. L. Homer, A. D. Jewell, H. Zhou, K. Manatt, and A.K. Kisor, "Correlating
- [8] Cerius² v 4.2, Accelrys Inc., San Diego, California, USA (2000).

Polymer-Carbon Composite Sensor Response with Molecular Descriptors, submitted to the JCES.

- [9] A. V. Shevade, M. A. Ryan, M. L. Homer, A. M. Manfreda, H. Zhou and K. Manatt, *Sens. and Actuat. B: Chemical*, **93**, 84 (2003).
- [10] pKa program, Advanced Chemistry Development (ACD)