2016

Sensor-Based Estimation of BTEX Concentrations in Water Samples Using Recursive Least Squares and Kalman Filter Techniques

Karthick Sothivelr
Marquette University

Florian Bender
Marquette University, florain.bender@marquette.edu

Fabien Josse PhD
Marquette University, fabien.josse@marquette.edu

Edwin E. Yaz
Marquette University, edwin.yaz@marquette.edu

Antonio J. Ricco
Stanford University

Electrical and Computer Engineering Faculty Research and Publications/College of Engineering

Sensor-Based Estimation of BTEX Concentrations in Water Samples Using Recursive Least Squares and Kalman Filter Techniques

Karthick Sothivelr
Marquette University, Milwaukee, WI, US

Florian Bender
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin, United States

Fabien Josse
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin, United States

Edwin E. Yaz
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, Wisconsin, United States

Antonio J. Ricco

Department of Electrical Engineering Stanford, University Stanford, California, United States

Abstract:

This work investigates sensor signal processing approaches that can be used with a sensor system for direct on-site monitoring of groundwater, enabling detection and quantification of BTEX (benzene, toluene, ethylbenzene and xylene) compounds at µg/L (ppb) concentrations in the presence of interferents commonly found in groundwater. A model for the sensor response to water samples containing multiple analytes was first formulated based on experimental results. The first signal processing approach utilizes only RLSE (recursive least squares estimation) whereas the second, a two-step processing technique, utilizes both RLSE and bank of Kalman filters for the estimation process. The estimation techniques were tested using actual sensor responses to contaminated groundwater samples. Results indicate that relatively accurate concentration estimates (within ±15–23% for benzene) can be obtained in near-real time using these techniques. The two-step processing technique gave more accurate results. This approach allows the use of a single sensor, even for multiple analyte detection and quantification.

SECTION I. Introduction

Most sensor systems utilize a single sensor parameter, such as the equilibrium frequency shift, for analyte identification and quantification, thus requiring the use of a sensor array to achieve selectivity [1]. However, there may be additional information available in the total sensor response that—if properly extracted—will enable the use of a smaller sensor array or even a single sensor device for analyte identification and quantification. In that case, an efficient sensor signal processing unit will be a key component of the sensor system. This work investigates an in-situ sensor system for long-term groundwater monitoring for detection of accidental releases of fuel and oil, e.g. from underground storage tanks (USTs). Accidental releases from USTs often result in contamination of groundwater [2]. Contaminated groundwater samples usually contain many chemically similar organic compounds, thus requiring advanced sensor system design and signal processing to identify and quantify the analyte(s) of interest in the sample.

Groundwater monitoring at active UST sites is important to safeguard human health and the environment. BTEX compounds (benzene, toluene, ethylbenzene and xylenes) are good indicators of fuel and oil releases [3]. The monitoring of benzene is particularly important because of its carcinogenicity. At present, USTs are only monitored in 2–3 year intervals and the groundwater samples collected at the monitoring wells must be shipped to a laboratory for off-site analysis; this process can take many days or weeks. Thus, the current practice is too time-consuming, labor intensive and costly to permit frequent monitoring of UST sites. Therefore, it is of great importance to develop an efficient sensor technology for long-term monitoring of groundwater that can rapidly analyze and quantify the analytes of interest in the groundwater.

In past work, the authors have demonstrated successful use of SH-SAW (shear-horizontal surface acoustic wave) devices coated with various partially selective polymer thin films for the quantification of BTEX compounds directly in the aqueous phase [3]. Response time constants and sensitivities have previously been measured for each combination of BTEX compounds and sensor coating [3], and the use

of estimation theory-based techniques to rapidly identify and quantify binary mixtures of BTEX analytes has been investigated [4]. In this work, a signal processing approach is investigated that uses the response time constants and sensitivities for each coating-analyte pair, coupled with customized signal processing based on estimation technique, to detect and quantify BTEX compounds in the presence of interferents typically found in groundwater. Specifically, the approach uses a recursive least-squares estimation (RLSE) algorithm with or without bank of Kalman filter (KF). The approach enables effective detection and quantification of BTEX compounds, even in samples containing multiple chemically similar analytes, with greater reliability, improved chemical selectivity, and shorter data processing time.

SECTION II. Modeling of the Sensor Response

In order to detect and quantify BTEX compounds in the presence of interferents using estimation theory-based techniques, it is necessary to accurately model the SH-SAW sensor response to the samples. The sensor response model for samples containing multiple analytes follows directly from the single-analyte (and binary mixture) sensor response model which has been discussed in detail in a previous publication [4]. For the sensor response to samples of multiple analytes, two critical assumptions were made. First, it is assumed that polymer absorption of the mixture obeys Henry's law. Henry's law implies that the concentration of the mixture in the coating at any time t is the sum of the concentrations of each individual analyte, Ci(t), as it would be measured in a single-analyte response. Based on experimental observations, Henry's law is valid for analyte concentrations in the range of 0 to 50 ppm (mg/L) depending on the analyte and the sorbent coating [3]. The process of analyte absorption can then be assumed to be first order (similar to the model of single analyte sensor response). Second, the frequency shift due to the mixture at any time t is the sum of the frequency shifts due to each analyte in the mixture at that time. As a result of these two assumptions, the sensor response to a mixture of n analytes can be represented by the following equations:

$$\cdot\, C_i(t) = -\frac{1}{\tau_i} C_i(t) + \frac{K_{p-w}}{\tau_i} C_{amb,i}(t) \text{ and}$$
$$\Delta f(T) = -\Sigma_{i=1}^{n} a_i C_i(t),$$

(1a)(1b)

where $C_{amb,i}(t)$ is the ambient analyte concentration at time t, t, τ_i is the response time constant for a given analyte/coating combination, K_{p-w} is the polymer-liquid partition coefficient for a given analyte/coating combination, $\Delta f(T)$ is the frequency shift observed at time t, and a_i is the steady-state (or equilibrium) signal, which is a function of the sensor platform, the sensor coating, and the analyte. The subscript $i = 1,2J, \ldots n$ refers to each analyte in the mixture.

The total number of analytes in groundwater varies and can be large. Therefore, to use (1) for the detection and quantification of BTEX compounds in the presence of an unknown number of interferents, it is assumed that the sensor response due to these interferents does not dominate the response due to the target analytes. From experimental observations, it can be inferred that these interferents will either have slower response time constants or lower sensitivities than the target analytes. This inference holds if appropriate sorbent coatings are selected. Therefore, one approach to take into account any number of interferents in the mixture is to model all the detectable interferents as a single analyte with a response time constant longer than that of all BTEX analytes, provided their physicochemical properties-

primarily size and molecular weight in this case—are appropriate. Also, note that ethylbenzene and the three xylenes are chemical isomers that have similar characteristic response time constants (for many coatings) which permits representation of the response due to all of them by a single exponential term. For this analysis, using (1), we can assume $n = 4$ where the subscript $i = 1,2,3$ represents the response due to the BTEX compounds and the subscript $i = 4$ represents the response due to the interferents in the mixture.

For (1) to be used with the estimation theory-based techniques, the model is normalized, discretized and converted into the state-space form:

$$y_k = \Delta f_k = \Sigma_{i=1}^4 \alpha_i x_k^{(i)} + w_k \text{ and}$$

$$\begin{bmatrix} x_{k+1}^{(1)} \\ x_{k+1}^{(2)} \\ x_{k+1}^{(3)} \\ x_{k+1}^{(4)} \end{bmatrix} = \begin{bmatrix} 1-S_1 & 0 & 0 & 0 \\ 0 & 1-S_2 & 0 & 0 \\ 0 & 0 & 1-S_3 & 0 \\ 0 & 0 & 0 & 1-S_4 \end{bmatrix} \begin{bmatrix} x_k^{(1)} \\ x_k^{(2)} \\ x_k^{(3)} \\ x_k^{(4)} \end{bmatrix} + \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{bmatrix} U_k$$

(2a)(2b)

where $x_k^{(i)}$ represents the normalized concentration of absorbed analyte at time t, α_i is the steady-state frequency shift, S_i is the absorption rate constant (i.e. $S_i = \frac{T}{\tau_i}$ where T is the sampling period), and U_k represents the input to the system. Typically, U_k is assumed to be a step input to represent the step change in the concentration for the transition from clean (filtered) water to the sample (for $t < 0, C_{amb,i}(t) = 0$ and for $C_{amb,i}(t) = C_{max}$. However, in practice, it is observed that the sample flow in a tube is laminar, resulting in an effective non-step-like concentration profile [5]. Due to these constraints, precautions have been taken in the experiments to achieve a nearly step-like concentration profile.

SECTION III. Results and Discussion

In this work, estimations were performed either by using RLSE alone (one-step processing) or by using both RLSE and a bank of KF (two-step processing). In the two-step processing, an initial estimation of the concentration is performed using RLSE, which is then used as input for the second step using a bank of KF to refine the estimation results. The bank is formed as a parallel combination of KFs, each initialized with different frequency shifts, α_i, with the range of αi, based on the frequency shifts associated with the extracted concentrations from RLSE. Note that these techniques are independent of the initial values of the unknowns; therefore, any concentration range of BTEX compounds can be analyzed. The results obtained using both one-step and two step processing are shown and discussed below. The state space model of (2) was used in the estimation process using RLSE and bank of KF. The steady-state frequency shift, α_i, is estimated so that the corresponding concentration of each BTEX compound in the sample can be extracted using the sensitivities extracted from single-analyte experiments. The proposed estimation theory-based techniques were examined using measured responses of SH-SAW sensors coated with either 0.6 μm PECH or 0.8 μm PIB to BTEX-containing samples (i.e., aqueous solutions of light non-aqueous phase liquid (LNAPL) or samples of contaminated groundwater). The results obtained using both one-step processing and two step processing are compared to BTEX concentrations

measured independently using a GC-PID (gas chromatograph photoionization detector). Note that the BTEX concentrations determined using GC-PID are subject to an average error of ±7%. A representative measurement of the sensor response using a 0.6 μm-thick PECH coated device to LNAPL sample in groundwater is shown in Fig. 1 together with the estimated response using the two-step processing. Fig. 1 shows very good agreement between the measured data and the estimated sensor response. The estimated concentrations are 443 ppb (7% difference to GC-PID measurement) for benzene, 760 ppb (18%) for toluene, and 120 ppb (60%) for ethylbenzene and xylenes. Note that the absolute difference between estimated and measured concentrations for ethylbenzene and xylene is only 45 ppb notwithstanding the high percentage error.
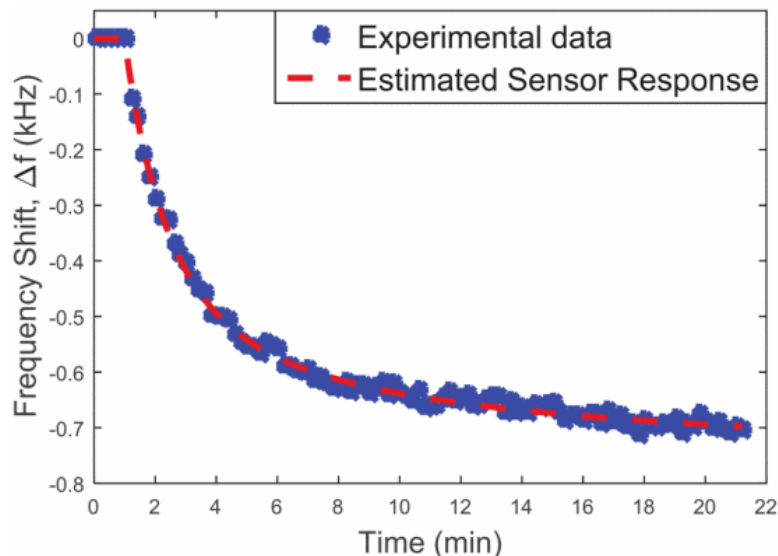

Figure 1. Measured response of a SH-SAW sensor coated with 0.6 μm PECH to a LNAPL sample in groundwater containing 475 ppb benzene, 646 ppb toluene and 75 ppb ethylbenzene and xylenes, and an unknown number and concentration of interferents. Also shown is the estimated sensor response curve obtained from two-step processing.

Fig. 2 summarizes estimated BTEX concentrations from various measurements using the coated device and both one step and two-step processing techniques. The two-step processing produces more accurate estimates compared to the one-step processing, an improvement most pronounced for benzene, which is the major target of this study due to its carcinogenicity. The same trend is observed for the estimation results obtained using data collected using a SH-SAW sensor coated with 0.8 μm PIB with relative percentage error of 23% for benzene, 34% for toluene and 138% for ethylbenzene and xylenes using the two-step processing. The error is largest for ethylbenzene/xylenes, probably due to the simplifying assumption of modeling the combined response of the other interferents in the mixture as a single analyte ($i = 4$). Another possible source of inaccuracy in ethylbenzene/xylenes estimates is the low concentration range tested for these compounds, where the signal noise limits the accuracy of the estimated concentrations.

SECTION IV. Conclusions
This study presents two estimation-theory-based techniques, one-step processing based on RLSE and two-step processing based on RLSE and bank of KF. Both techniques are capable of detecting and

quantifying BTEX compounds in water in the presence of interferents using only one coated sensor. Based on the estimated results, it can be concluded that the two-step processing will yield more accurate results. While this increased accuracy is achieved at the expense of slightly longer processing time, a good estimate of BTEX concentrations can still be obtained in near-real-time. The estimated concentrations for benzene using two-step processing fall within ±15-23% of the concentrations measured using GC-PID, even without any knowledge of the initial concentration range of the analytes. Note that the accuracy of the estimation results depends on the reproducibility of the steady-state sensitivities and time constants. Finally, the results obtained also validate the formulated multi-analyte sensor-response model and all the assumptions made for the analysis in the present work.
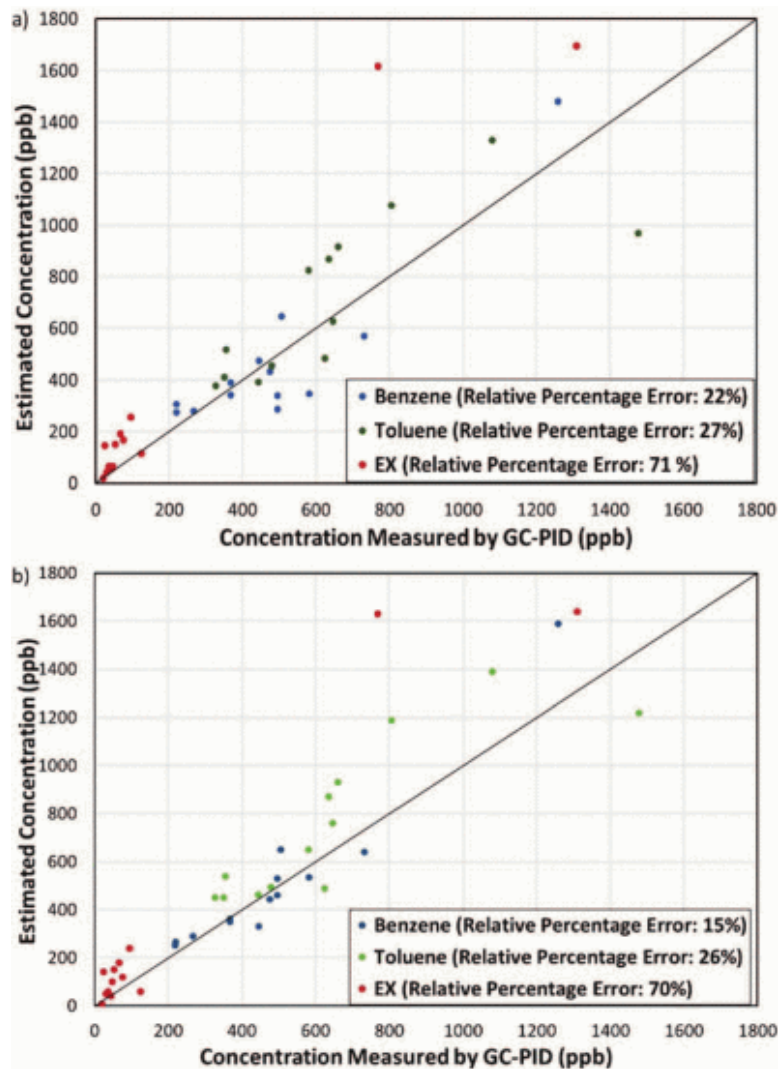

Figure 2. Estimated BTEX concentrations using a) one-step processing and b) two-step processing vs. Concentrations measured by GC-PID for a SH-SAW sensor coated with 0.6 μm pech. The legends show the average relative percentage error between the estimated and the measured concentrations. EX stands for ethylbenzene and xylenes; the diagonal line represents the ideal case (estimated concentration = measured concentration).

References

1. J. S. Cooper et al., "Quantifying BTEX in aqueous solutions with potentially interfering hydrocarbons using a partially selective sensor array", Analyst, vol. 140, pp. 3233-3238, March 2015.
2. US Environmental Protection Agency: Underground Storage Tanks (USTs), [online] Available: https://www.epa.gov/ust.
3. F. Bender et al., "Quantification of Benzene in Groundwater Using SH Surface Acoustic Wave Sensors", The 14th International Meeting on Chemical Sens, pp. 473-476, 2012.
4. K. Sothivelt et al., "Online Chemical Sensor Signal Processing Using Estimation Theory: Quantification of Binary Mixtures of Organic Compounds in the Presence of Linear Baseline Drift and Outliers", IEEE Sensors J., vol. 16, pp. 750-761, Feb. 2016.
5. G. Taylor, "Dispersion of Soluble Matter in Solvent Flowing Slowly through a Tube", Proc. R. Soc. London Ser. A, vol. 219, pp. 186-203, August 1953.