Int. J. Nonlinear Anal. Appl. 14 (2023) 1, 2753–2758
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.29387.4146

Modeling the UV-visible spectroscopic data of an aromatic hydrocarbon mixture to solve the problem of gas mixtures

Muthana Alboedam[a,*], A.A. Al-Rubaiee[b]

[a]Department of Clinical laboratory science, University of Kerbala, Iraq
[b]Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq

(Communicated by Ehsan Kozegar)

Abstract

The information data obtained from the spectra of the sensors were combined, which allowed us to build a matrix including thousands of spectrum-simulating gas mixtures, taking into account the presence of different components in several concentrations. Signal process technique and experimental design were utilized all the way for simulating spectral datasets. Particularly, the purpose of the simulation is for database creation for the evaluation of the final detection system for mixtures of gases (aromatic hydrocarbons). The dependable components of a fluorescent read platform are a light source for fluorescence excitation and for isolating the filtered fluorescent light from the excitement light and image detector. However, this technique has several limitations because of the need to utilize many chemo-sensor elements at the same time, which make titration complicated. This is needed to search for another way to overcome complications. In our study, certain filters were used to solve the problem of fluorescence spectroscopy.

Keywords: (UV-Visible) spectra, Aromatic hydrocarbons, filter, noise
2020 MSC: 11F72

1 Introduction

Aromatic compounds are widely used in advanced industries. Therefore, it is indispensable to define its air concentration, particularly in the industrial, and densely populated regions, because it is harmful in parts of billion. They cause components like naphthalene, benzene, xylene, and toluene, which cause unusual harm to human health, such as: (anaemia, infertility, and several types of cancer), it also damages the environment.

Despite many health problems caused by these aromatic compounds, we noted the production of these compounds growing worldwide. Over the past 20 years, sensors have been used to analyze chemical compounds in the environment in all fields of industrial and medical aspects. Various sensors have already been developed and used in many applications such as medicine, biochemistry, security systems and food packaging. In order to individually view and produce optical sensor materials, one must be able to calculate spectral properties [7, 17, 18]. These materials also exhibit functional properties typical of sensual substances under controlled with increasing conditions. Optical prospectors set up on the foundation of a fluorescence detector usually consist of a sensor material and a sensors program. So, the sensor substances include indicator particles attached to polymers, or other forms of matrix [16].

*Corresponding author
Email addresses: muthana.j@uokerbala.edu.iq (Muthana Alboedam), dr.rubaiee@uomustansiriyah.edu.iq (A.A. Al-Rubaiee)

Received: October 2022 *Accepted:* January 2023

2 Materials and methods

2.1 Sensor spectra

Groups of aromatic hydrocarbons are prepared in the laboratory. concentration compositions include gas supplies for naphthalene (0.2 g/min), toluene (10.1 g/min), benzene (44.0 g/min) at various temperatures, for naphthalene (50 C), toluene (100 C), and gasoline (110 C). Spectral characteristics are recorded directly. The flow rate is 200 ml/min, we have tested more than 1300 samples.

2.2 Spectra import

The spectrum is stored in many databases, and it is recorded in several devices that have different properties. So an algorithm was developed for importing spectral data with different wave number domains, resolutions, and input file formats. All spectra imported from the sensor have been transformed into widespread coordination and structure. To best reduce the noise from the database spectra, they were imported by using all the original marks, thus using the highest possible resolution.

The next step was to convert the noise-reduced signals to obtain unified sets of the measured spectrum at a stationary concentricity (1-ppb) in the spectral range of (1000-2500 cm^{-1}). At this ending, the smoothed spectrums are operated by a Matlab function that re-samples the spectrum at the desirable resolve into user-defined spectra windows. Primarily, to gain an out-put that resembles the presumptive spectrum obtained from the device, the essential spectrum is convolved using a Gaussian function with an average of one, and a standard deviation derived from all width at half maximum (FWHM.), as the equation:

$$\sigma = \frac{FWHM}{2\sqrt{2\ln(2)}}. \tag{2.1}$$

The outcoming signal is sampled with the desirable wave number averages by shifting a second-order polynomial interpolation to three successive points, simultaneously.

3 Results and discussion

3.1 Median Filter

The median filter is nonlinear filter often utilized for removing noise. (Fig. 1). shows the result of spectrum processing by this filter.

Advantages: does not affect step and sawtooth functions; well suppresses single impulse noise (random noise emissions of readings and misses); the existence of a fast 2D implementation [9, 10].

Disadvantages: median filtering - suppression of Gaussian noise is less effective than that of linear filters; 2D processing leads to more significant signal attenuation [14].

3.2 Savitsky-Golay smoothing method

Savitzky-Golay anti-aliasing filters are commonly utilized to "smooth out" a noisy signal with a high bandwidth (no-noise). They have also known as polynomial digital smooth filters, or least squares smooth filters. Savitzky Golay filters act better in many applicants than optimum FIR average filters, which slop to filter height frequency contents at the noise. Savitzky Golay filters are a high effect at keeping the height frequency components of the signal but are less effective with suppressing noise.

The procedures depend on established mathematical processes that were popularized by Abraham Savitsky and Marcel J. E. Golet [2, 3], their publication tables involved the convolution coefficients of different polynomials, and set up the sizes in 1964. The method has been extended to handle 2D and 3D data.

SG smoothing works in the subwave length range, including neighboring wave numbers, constructs a polynomial with the ordinal wave numbers as the independent variable, and fits the polynomial coefficients using the principle of least squares regression [4, 6]. In the process of polynomial approximation, spectral data from neighboring wave numbers were embedded in coefficients. The coefficient of each polynomial term will be a linear combination of the spectral data in the subband; the order term will become the value of the smoothed spectrum of the smoothing derivative of the order at the central point; the linear combination coefficients are called SG smoothing coefficients [8].

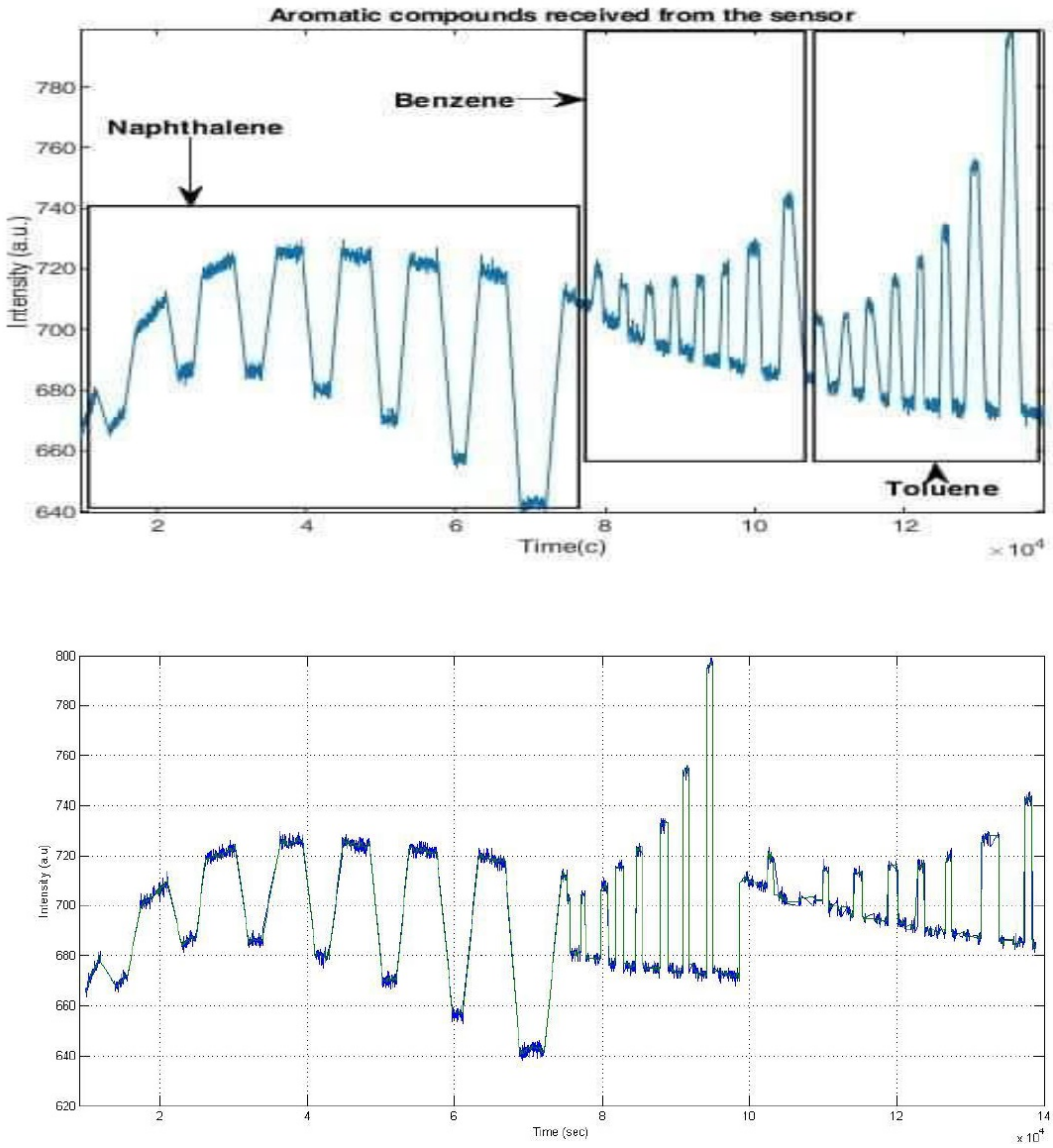

Figure 1: Original data and data processed by the median filter

The appearance of monochromatic light on a sample is characterized by some phenomena like reflected, scattered or absorbed. In general cases, the scattered process competes with the absorption processes. When the radiation is absorbed, the particle goes into a less exciting electronic state. A particle returning to the ground state can be either completely illuminated or carried away by the light radiations with (lesser periodicity). These radiations are known as luminescence (luminosity). Through mathematical models for the recording signals, the signal received from the sensor can be used by some of the included scientific models that form the signal of the spectrometer, the self-ignition origin, and noise this is mean all types of noise. Since the nature of noise filtering in signals is diverse, we apply a median filter at the preliminary stage to remove impulse noise (Fig. 2).

The median filter is high effect than linear filters in conditions, which the variations in signal amounts are largest compare with the variance in Gaussians-noise. This process because of the fact that it gives a lower RMS error value to the output signal than to the silent input signal compared to optimal linear filters. This allows it to be widely used to remove outliers in data arrays [11, 12, 13]. Its main feature is that it does not react well to indicators that stand out sharply against the background of neighboring ones. The output signal (y_k) of a moving median filter of width $(2n+1)$ for the current sample k is formed from the input time series $(..., x(k-1), xk, x(k+1), ...)$

$$y_k = Me(x_{k-n}, x_{k-n+1}, ..., x_{k-1}, x_k, x_{k+1}, ..., x_{k+n-1}, x_{x+n} \tag{3.1}$$

where $(x_1, ..., x_m, ..., x_{2n+1}) = x_{n+1}, x_m$ elements of the variation series i.e. ranked in ascending order of values x_m:

$$x_1 = \min(x_1, x_2, ..., x_{2n+1}) \leq x_1 \leq x_2 ... \leq x_{2n+1} = \max(x_1, x_2, ..., x_{2n+1}).$$

The average filter width $(2n+1)$ is chosen to suppress a sample n wide pulse. The (Savitsky Golay) filtered method is best than the mean filter because it preserves spectrum features, like peak width and height. The essential useful of this process is to support the regional extremes, which usually equalized by other averaging processes [19]. The Savitzky-Golay filter equation is shown in (Equation (3.2)):

$$S_i^* = \sum_{j=-n}^{n} c_j * S_{i+j} (i = n+1, N-n) \tag{3.2}$$

Here (S_i) are samples of the filtered signal, (c_j) are the (Savitsky Golay) coefficients, and (n) is the half width of the smoothing window. The Savitzky Golay coefficients are counted using the formula (Equation (3.3)):

$$c_j = C_{1,j} (j = 1, 2n+1) \tag{3.3}$$

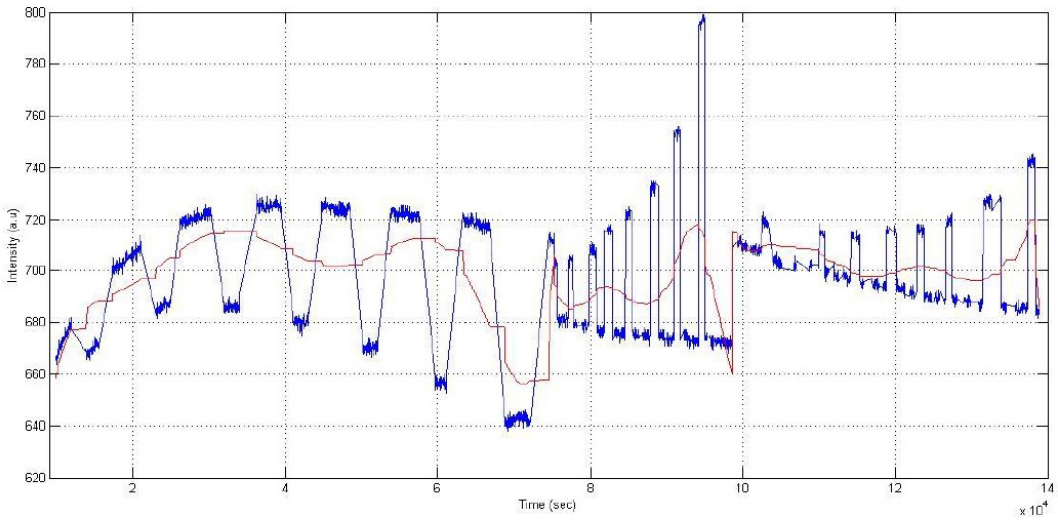

Figure 2: Initial data processed from Savitsky-Golay

3.3 Noise suppression

Various noise structures influence the spectrum of the sensor, and as well perfectly reasonably vary from the noise structures of the measured spectra by using the device that was developed. For this cause, it is needful for separating useful spectroscopic information, which will likewise be located in the signals of the novel instrumentation, from the noise introduced by the specific instrument utilized to record each spectrum. This aim was achieved utilizing wave-transform (WT) signal processing techniques because of the (WT's) ability for displaying the analyzed signal in both the origin and the corresponding frequency scope at the same time. As well as, the use of different wavelets to disband the signal into the (WT.) location allows the using of the high number of representations, through which one can choose the extremely efficient one [1].

The signal-to-noise rate (SNR.) indicates how much the dynamic range of the signal is greater than the dynamic range of the noise.

The larger the SNR value, the better the signal quality. Taking the maximum values as signal/noise indicators is not entirely relevant due to possible random outliers that are not characteristic of the signal [5, 15].

$$SNR = \frac{\sigma^2_{Signal}}{\sigma^2_{Noise}} \tag{3.4}$$

$$ENL = \frac{\mu_t^2}{\sigma_t^2} \tag{3.5}$$

where σ is the standard deviation.

Each gas has an individual set of peaks. For example, naphthalene, benzene and toluene have the following characteristic peaks listed in table 1, 2, 3.

In this work, the most distinct individual peaks were taken as the main components. The reference amplitude of these peaks was calculated for various gas spectra. The amplitudes of these peaks after filtering were also calculated.

Table shows the amplitudes of the peaks after background fluorescence has been removed by the algorithms. Peak amplitudes after processing procedures are compared with the reference ones. Deviation from the reference value, expressed in fractions, are presented in tables

The (Savitsky – Golay) filter possess a smaller error scattered for the considered characteristic peaks comparing with the medium filter, so it's very easy for correcting this spread [14].

Table 1: Peak amplitudes for the Naphthalene spectrum

	$0.1185 * 10^5$	$0.2098 * 10^5$	$0.3018 * 10^5$	$0.3938 * 10^5$	$0.4528 * 10^5$	$0.5595 * 10^5$	$0.6494 * 10^5$	$0.7472 * 10^5$
Signal Before Processing	681.8	731.6	724.6	728.2	728.5	725.5	722.8	714.8
Median filter	679.8	710.5	723	726	724.4	723.1	717.6	713.9
SG smoothing	677.2	696.6	714.3	708.7	701.9	711.8	678.3	691.4

Table 2: Peak amplitudes for the spectrum of Benzene

	$0.7886 * 10^5$	$0.8208 * 10^5$	$0.8596 * 10^5$	$0.894 * 10^5$	$0.9301 * 10^5$	$0.9645 * 10^5$	$0.1001 * 10^5$	$0.1047 * 10^5$
Signal Before Processing	723.3	718.5	716	717.5	718.1	722	729.9	745.6
Median filter	720.8	715.2	713	716.1	716.8	718.5	728.3	743.7
SG smoothing	709.4	707.2	698.9	698.9	701.6	698.2	698.1	720

Table 3: Peak amplitudes for the spectrum of Toluene

	$1.088 * 10^5$	$1.124 * 10^5$	$1.152 * 10^5$	$1.189 * 10^5$	$1.223 * 10^5$	$1.252 * 10^5$	$1.294 * 10^5$	$1.345 * 10^5$
Signal Before Processing	705	706	711	718	725	735	756	799
Median filter	703	703.8	707.2	716.2	723.1	732.9	753.9	796.6
SG smoothing	694.2	685.3	689.8	693.7	689.9	678.8	705	713

4 Conclusion

Programmable sampling of spectral readings, provided by the use of spectrometers with random spectral processing, provides an improved method of reconstructing the concentrations of gas mixture components from the spectrometric data of the sensor used. The method is based on obtaining and processing spectral data from the sensor. It is possible to effectively solve the problem of quantitative analysis of gas mixtures, especially if they contain substances with similar spectra.

In this work, we faced the problem of modeling the spectroscopic data of a chemical sensor and creating a new database, particularly after removing the background of fluorescence and autofluorescence of mixing of gases number. The starting point is to study the operation of the sensor and obtain raw data from it, as well as how the spectra taken in the ultraviolet and visible spectrum of aromatic hydrocarbons (naphthalene, benzene and toluene) are physically formed, which must be properly processed in order to obtain the same characteristics with terms of accuracy etc. In our study, we took into account the influence of various methods of spectral processes on the original contents (aromatic-compounds), spectrum characteristics of the compounds, and characteristic-peaks.

References

[1] D.W. Ball, *Field guide to spectroscopy*, SPIE Publication, Bellingham, 2006.

[2] N. Brown and P. Bladon, *Spectroscopy and structure of (1, 3-diketonato) boron difluorides and related compounds*, J. Chem. Soc. A: Inorganic, Physical, Theor. (1969), 526–532.

[3] P. Cadusch, M. Hlaing, S. Wade, S. McArthur and P. Stoddart, *Improved methods for fluorescence background subtraction from Raman spectra*, J. Raman Spectroscopy **44** (2013), 1587–1595.

[4] T.T. Cai, D. Zhang and D. Ben-Amotz, *Enhanced chemical classification of Raman images using multiresolution wavelet transformation*, Appl. Spectroscopy **55** (2001), no. 9, 1124–1130.

[5] M.A. Choma, M.V. Sarunic, C. Yang and J.A. Izatt, *Sensitivity advantage of swept source and Fourier domain optical coherence tomography*, Optics Express **11** (2003), no. 18.

[6] J.D.A. Espinoza, V. Sazhnikov, S. Sabik, D. Ionov, E. Smits, S. Kalathimekkad, G. Van Steenberge, M. Alfimov, M. Pośniak, E. Dobrzyńska and M. Szewczyńska, *Flexible optical chemical sensor platform for BTX*, Proc. Eng. **47** (2012), 607–610.

[7] C.V. Gazarov, V.E. Pozhar and V.N. Zhogun, *Acousto-optical spectrometer for air pollution monitoring*, CIS Selected Papers: Optic. Monitor. Envir. SPIE **2107** (1993), 143–146.

[8] R. Heinrich, A. Popescu, A. Hangauer, R. Strzoda and S. Höfling, *High performance direct absorption spectroscopy of pure and binary mixture hydrocarbon gases in the 6–11 μm range*, Appl. Phys. B **123** (2017), no. 223, 1–9.

[9] M. Jalali-Heravi and H. Parastar, *Assessment of the co-elution problem in gas chromatography-mass spectrometry using non-linear optimization techniques*, Chemom. Intell. Lab. Syst. **101** (2010), no. 1, 1–13.

[10] Y.M. Jung, *Principal component analysis based two-dimensional correlation spectroscopy for noise filtering effect*, Vibrat. Spectroscopy **36** (2004), no. 2, 267.

[11] R. Kengne-Momo, P. Daniel, F. Lagarde, Y. Jeyachandran, J. Pilard, M.J. Durand-Thouand, G. Thouand, *Protein interactions investigated by the Raman spectroscopy for biosensor applications*, Int. J. Spectroscopy **2012** (2012).

[12] A.-I.Z. Khalaf, M. Alboedam, H.J. Abidalhussein and A.-Z.S. Hassan, *The role of blood proteins and nucleic acids in the detection of multiple Myeloma based on Raman spectroscopy*, EurAsian J. BioSci. **14** (2020), no. 1, 1955–1963.

[13] A.-I.Z. Khalaf, M. Alboedam, H.J. Abidalhussein and A.-Z.S. Hassan, *Detecting levels amino acids for proteins of dierent for patients with myeloma and comparing them using a portable Raman spectrometer*, EurAsian J. BioSci. **14** (2020), no. 1, 2029–2036.

[14] J. Luo, K. Ying, P. He and J. Bai, *Properties of Savitzky–Golay digital differentiators*, Digital Signal Process. **15** (2005), no. 2, 122–136.

[15] V. Mazet, C. Carteret, D. Brie, J. Idier and B. Humbert, *Background removal from spectra by designing and minimising a non-quadratic cost function*, Chemom. Intel. Lab. Syst. **76** (2005), no. 2, 121–133.

[16] A. Mirzaei, J.H. Kim, H.W. Kim and S.S. Kim, *Resistive-based gas sensors for detection of benzene, toluene and xylene (BTX) gases: a review*, J. Mater. Chem. C **6** (2018), no. 16, 4342–4370.

[17] U. Platt, D. Perner and H.W. Pätz, *Simultaneous measurement of atmospheric CH2O, O3, and NO2 by differential optical absorption*, J. Geophys. Res.: Oceans **84** (1979), no. C10, 6329–6335.

[18] V.I. Pustovoit and V.E. Pozhar, *Long-path optical spectral AOTF-based gas analyzer*, Instrument. Air Poll. Glob. Atmosph. Monitor. SPIE **4574** (2002), 174–178.

[19] M.A. Rahman, M.A. Rashid and M. Ahmad, *Selecting the optimal conditions of Savitzky–Golay filter for fNIRS signal*, Biocyb. Biomed. Eng. **39** (2019), no. 3, 624–637.