# BTEX compounds identification by means of gas sensors arrays

*Marek* Badura[1,*], *Andrzej* Szczurek[1], and *Kamil* Banaszkiewicz[1]

[1]Wrocław University of Science and Technology, Faculty of Environmental Engineering, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

**Abstract.** BTEX compounds can cause a threat to environment and human health. For this reason measurement devices are needed for rapid identification of such pollutants. The paper presents the results of recognition of mixtures of benzene, toluene, ethylbenzene and xylenes in humid air by means of two gas sensor arrays and linear discriminant analysis. Measurements were conducted during stabilization/solidification processes of contaminated soils. High classification ratios were obtained for both arrays (I: 88%–94%; II: 94%–96%). Improvement of identification was achieved when two copies of every sensor were included in analysis – classification rates reached 97.1–100%.

## 1 Introduction

Hydrocarbons from BTEX group (i.e. benzene, toluene, ethylbenzene and xylenes) are considered as one of the most dangerous pollutants in the environment. Those contaminants could be found in groundwater and soil [1, 2] and are also classified as hazardous air pollutants [3, 4]. BTEX presence in ecosystem is related to the use of solvents [5] and petroleum products [6, 7] and to the emission from vehicles [8].

BTEX compounds have very similar structure, but some of their properties are quite diverse. For example, their volatility is distinct and this causes different emission characteristics and behaviour in the environment [9]. Also the impact on living beings is not the same for every compound in that group, but their adverse health effects and toxicity are well documented [10].

Taking into account all environmental and health problems caused by BTEX compounds there is an increasing interest in developing instruments capable to detect them and measure their concentrations. The most common method is based on sampling and laboratory analysis with gas chromatography (GC), available in different configurations [5, 7, 8, 11]. However, that approach is time consuming and nowadays rapid measurements could be done by means of portable or mobile devices [12]. Unfortunately, solutions based on GC are still very expensive, thus lower cost instruments are popular. Handheld devices like flame ionization detectors (FIDs) or photo-ionization detectors (PIDs) are very helpful in fast assessments [13], but they are non-specific, so they cannot identify BTEX compounds. Promising method for screening measurements is based on differential ion

---

[*] Corresponding author: marek.badura@pwr.edu.pl

mobility spectrometry [14], however the usage of radioactive ion source in those instruments is problematic.

Safer and cheaper solution is based on gas sensor arrays. Partially selective semiconductor sensors coupled with pattern recognition algorithms could be especially useful for such tasks [15–17].

The aim of this work was the use of sensor arrays for identification of benzene, toluene, ethylbenzene and xylenes in humid air. The paper focuses on the choice of sensors for classification of such air mixtures. Sensors selection is crucial in case of multidimensional arrays and could improve recognition properties of those devices [17].

Measurements were conducted during stabilization/solidification (S/S) processes of contaminated soils. In S/S processes contaminants are immobilized by means of mixing soils with different hydraulic binders. It has been shown that S/S actions pose a risk of releasing BTEX vapours [18, 19], increasing the health danger for workers and create a threat to environment. In that case BTEX monitoring system seems to be necessary.
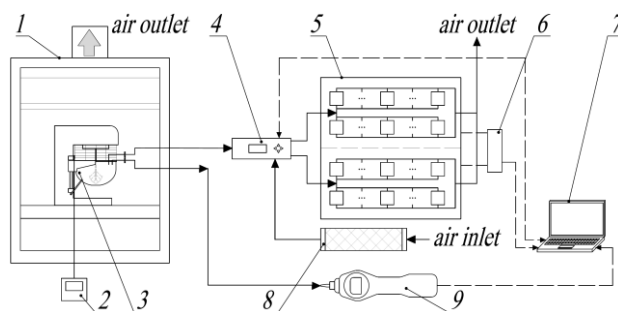
## 2 Experimental

### 2.1 Experimental setup

#### 2.1.1 Installation for stabilization/solidification processes

Experiments were carried out on a laboratory-scale installation dedicated to neutralization of contaminated soils. The experimental setup is shown in Fig. 1.

S/S processes were performed in a rector of cement mortar Tecnotest B205/X5, placed in a fume hood. Sampling tubes with dust filters were mounted above reactor's bowl and gas samples were transported via PTFE tubing to measuring devices. Temperature and relative humidity in vicinity of the reactor were measured by means of a probe of AR236/2 datalogger (APAR, Poland).



**Fig. 1.** Experimental setup: 1 – fume hood; 2 – temperature and humidity datalogger; 3 – cement mortar reactor; 4 – gas flow controller; 5 – sensor arrays (two blocks); 6 – voltage supplier and signals measurement unit; 7 – personal computer; 8 – filter set; 9 – photo-ionization detector.

#### 2.1.2 Characteristic of photo-ionization BTEX detector

Photo-ionization detector PhoCheck TIGER (Ion Science Ltd, UK) was used in experiments to estimate the emission level of BTEX compounds.

This non-specific instrument was calibrated against a isobutylene before each measurements series. In case of benzene, toluene, ethylbenzene or xylenes measurements

an appropriate response factor was selected from instruments' gas table (factory set coefficient). Data was logged in 1-second intervals and sent to computer after each test.

### 2.1.3 Characteristic of gas sensor arrays unit

Two blocks of sensing arrays were used in this study. Each block consisted of the same number and models of Taguchi Gas Sensors (Figaro Engineering Inc., Japan): TGS800, TGS822, TGS823, TGS825, TGS826, TGS830, TGS832, TGS842, TGS2180, TGS2600, TGS2602, TGS2620, TGS2104 and two copies of TGS2201. TGS2201 contained two independent sensing layers, so every array was equipped with 17 sensing elements.

Sensors were mounted inside individual flow-type chambers, made of aluminium. Chambers were connected with system of PTFE tubes for transport of gases.

Voltage supplier was used to power the sensors and a set of reference resistors was applied for measuring output signals. Signals from digital converter were sent to computer and saved as voltage variations on the load resistors. Data acquisition was conducted with time resolution of 1 second.

Supply of gases was possible through the use of gas flow controller. This module consisted of 2 diaphragm pumps, 2 mass flow controllers and a set of electromagnetic valves enclosed in a RACK-type case. This configuration allowed aspiration of gas samples from S/S reactor or suction of zero air. Zero air was generated by purifying ambient air in filter filled with activated carbon and drying it in filter with silica gel.

Flow control unit was also able to direct gas mixtures or zero air to sensor arrays. The device was controlled by computer application.

## 2.2 Methodology of the stabilization/solidification processes

### 2.2.1 Measurement series

Experiments were conducted in four measurement sessions. Each session was dedicated to one pollutant from BTEX group.

### 2.2.2 Soil samples preparation

Soil samples (200 g) were placed in containers and contaminated with 0.1 ml of liquid hydrocarbon. P.a. grade benzene, toluene, ethylbenzene or isomeric mixture of xylene were used. Closed containers were shaken for 120 seconds in an overhead shaker and after that they were used in S/S processes.

### 2.2.3 Binders characteristics

Portland cement CEM I 42.5 R and various additives were adopted in this study. The following samples were prepared: 1) 100 g Portland cement + 10 g textile cord, 2) 100 g Portland cement + 1 g activated carbon, 3) 100 g Portland cement + 10 g textile cord + 1 g activated carbon, 4) 100 g Portland cement.

In every measurement session zero sample was also prepared. In that case contaminated soil was not mixed with hydraulic binders. Every sample type (including zero sample) was prepared and tested three times in every measurement series.

### 2.2.4 Stabilization/solidification processes

Every S/S process started with simultaneous introducing of soil and binding medium into clean reactor's bowl. After that two main stages of the process took place:
1 – mechanical mixing of components (the homogenization phase),
2 – introducing diluent water to initiate the hydration of cement components and further mixing (the hydration phase).

Each stage lasted 180 seconds and all materials were mixed with a constant rotational speed. The resulting mixtures were packed in containers.

## 2.3 Measuring procedure of sensor arrays

Measuring procedure of arrays was based on alternating sensors exposure to gas samples and zero air. Exposition in every phase lasted 30 seconds. This mode of operation has been applied to reduce the so-called "memory effect" [20], resulting from sensors sensing mechanism.

In this study exposition of both arrays to a gas mixture from reactor was not simultaneous. While sensors from first array (Block I) were placed in a gas sample stream, sensors from second array (Block II) were purged with a stream of zero air. This procedure allows sampling reactor's gases continuously. Scheme of a operation algorithm is presented in Fig. 2. The constant flow rate of 3 l/min was kept in every phase.
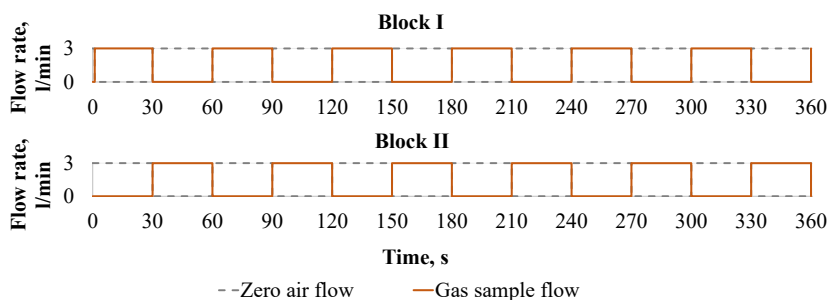


**Fig. 2.** Scheme of operation of gas sensor arrays.

## 2.4 Data analysis procedure

### 2.4.1 Classification method for BTEX compounds identification

Identification of mixtures of benzene, toluene, ethylbenzene or xylenes in humid air was performed using linear discrimination analysis (LDA). LDA is a well known pattern recognition technique used for supervised classification [21].

One-against-all approach was employed for this study. The class that included selected compound was discriminated from class with all other compounds. In that way four two-class problems were examined: (1) mixtures with benzene vs. mixtures without benzene, (2) mixtures with toluene vs. mixtures without toluene, (3) mixtures with ethylbenzene vs. mixtures without ethylbenzene, (4) mixtures with xylenes vs. mixtures without xylenes.

### 2.4.2 Construction of feature vectors

Gas sensors used in the measurements characterize with poor selectivity, i.e. they respond to many reducing gases. In case of multi-element arrays responses of individual sensors to the same gas mixture could be very different. A characteristic pattern ("fingerprint") of mixture could be created in that way [16].

Patterns in form of feature vectors were considered for the purpose of this study. Mean value of sensor signal from the gas sample exposition phase (30-second mean) was used as a feature.

Basic feature vectors were constructed from different number and combinations of sensors in arrays. In case of each array dimension of feature vector varied from $N = 1$ (feature vector based on signal from only one sensor) to $N = 17$ (feature vector based on signals from all sensors in array). The number of possible combinations was given by binomial coefficient and all of the possibilities were taken into consideration in the analysis. Sensor arrays were treated in two ways: A) as independent devices, B) as one device.

In case A) feature vectors were constructed only on the basis of features from particular block and data analysis was conducted for every array separately as described above. Hence, every feature vector was related to 30-second sampling of gas mixture from reactor.

In case B) two 30-second mean values of signals of particular sensor model from appropriate phases of both arrays exposition were used to create a set of features. Those sets of features were used to build feature vectors. So, in this case feature vectors had dimension from $N = 2$ (feature vector based on signals from two copies of particular sensor model) to $N = 34$ (feature vector that included signals from all sensors in both arrays). In other words, case B) included two copies of every sensor model in analysis. Because of the shift of exposure phases between arrays data vectors were related to 60-second sampling of gases.

### 2.4.3 Construction of training and test sets

Every measurement session consisted of triplicate measurements of samples with 6 exposition phases to gas mixtures in every one of them. The total number of exposition phases was 360 for every array and this was also the number of data vectors.

All vectors constructed from features were labelled with information about BTEX compounds. For the purpose of two-class problems, labels with abbreviations of compounds were used, for example: "B" (from benzene) and "TEX" (from toluene, ethylbenzene, xylenes).

The entire dataset was divided into training set and test set. It was assumed that for training purposes only data from zero samples measurements will be useful. This assumption was based on results from previous experiments [18, 19], where samples without binders were characterized with the highest emission of organic compounds during S/S processes. In this way, a large range of BTEX compounds concentrations could be obtained for training.

Taking into account that zero samples were measured three times in every session, BTEX identification possibilities were checked also three times. In every test data from one repetition of zero sample measurement was used for training and the rest from the pool of signals was used for testing purposes. Therefore, 24 data vectors were used for training and 336 vectors for testing in every validation approach.

The identification performance was evaluated by means of classification rate. This index was calculated as a percentage of correctly classified test data vectors among

the overall number of vectors from test set. The mean value from triplicate validation procedure was used for final assessment of recognition possibilities.

All calculations were performed in MATLAB environment.

## 3 Results and discussion

All experiments were conducted in laboratory with following conditions of temperature and humidity (averaged values): 1) 22°C and 55% RH in benzene measurements, 2) 23°C and 48% RH in toluene measurements, 3) 28ºC and 60% RH in ethylbenzene measurements, 4) 27°C and 40% RH in xylenes measurements. It should be noted that those conditions were not stable during every S/S test. Increase in relative humidity was observed especially after introduction of water in second stage of S/S operation.

Concentrations of BTEX compounds, detected by photo-ionization device, were in ranges: 1) 0.1–405 ppm in benzene measurements, 2) 0.2–526 ppm in toluene measurements, 3) 0.1–155 ppm in ethylbenzene measurements and 4) < 0.1–99.2 ppm in xylenes measurements. The widest ranges of concentrations were recorded for zero samples, as expected. However, BTEX emission behaviour was not repeatable for every sample – slightly different maximum concentrations and time shifts of occurrence of concentrations peaks were observed.

Maximum values of BTEX classification rates in relation to number of sensor models used for creation of feature vectors are presented in Fig. 3. It can be seen that data vectors based on signals from all sensors ($N = 17$) in most cases gave worse results than vectors with smaller dimensions.

Taking into account all classification problems, the best results for Block I (Fig. 3a) were obtained with combinations of features from $N = 3$ to $N = 8$ sensor models. Efficiency of benzene recognition was better than 93.2% for such combinations and reached 94.4% ($N = 4$). In case of toluene the level of around 91% was achieved for combinations created from $N = 6$ to $N = 8$ features. Xylenes were identified with highest rate of 91.2–92.4% with combinations based on $N = 3$ to $N = 5$ sensors. Ethylbenzene classification was the most difficult – classification ratio at the level of 88% was gained.

It should be also mentioned that different sets of models gave the highest results in BTEX identification tasks. For example, in benzene recognition TGS842, TGS2620, TGS2104 and TGS2201 proved to be the best, while xylenes identification was most effective with set consisting of TGS832, TGS2180 and TGS2201 (the second copy). Thus the necessity to choose the most appropriate sensors combination to particular task was showed. In case of second array (Fig. 3b) the best identification results were obtained with data vectors with dimensions from $N = 3$ to $N = 9$. It was similar with Block I, but somewhat higher ratios were observed for the undertaken tasks. In case of toluene 96.3% recognition efficiency was reachable with only $N = 3$ sensors. Benzene recognition was possible at the level of 96%, but with combination of $N = 8$ sensor models. Ethylbenzene mixtures were distinguishable easier from the rest compounds than in first array – even 95.2% was reached (for $N = 4$). Only a little worse result was documented for xylene isomers – 94.6% with $N = 6$ and $N = 7$ sensors in combination. Just like in case of Block I, different sensors models were necessary to carry out each discrimination task. Moreover, there were differences between models appropriate for particular tasks in each array. For example, the best combination in Block II for benzene samples identification did not contain any of TGS models selected for Block I. This can result from two things: 1) there was a time shift between exposition stages of arrays, so composition of sampled mixtures was different and features from sensors signals were different too, 2) there is a lack in reproducibility of manufacture of sensors, so each copy has its own characteristics. This results show that every sensor device should be treated individually and sensors selection is

crucial in such case. Fig. 3c presents the best classification ratios for the scenario when two copies of every sensor model were included in calculations. In general, the connection of signals from two sensor arrays gave better results than previously described situation. Toluene mixtures were perfectly recognized with combinations based on $N = 4$ to $N = 6$ sensor models. The 100% identification possibilities were also observed for ethylbenzene, starting from $N = 2$ to $N = 9$ sensor models combinations. Very high classification rate was also attained for benzene samples – 99.4% and 99.7% for $N = 4$ and $N = 5$ sensor models, respectively. Slightly worse identification rate was only noticed for xylene samples (97% level for $N = 4$ to $N = 7$) and it may be due to the fact that three isomers of that compound were present in collected samples, resulting in very variable composition of gas mixtures. As in previous cases, different sets of sensor models were useful in particular classification tasks. In addition, some sensor models in the best designated combinations were not present in combinations that were chosen for Block I or Block II. Moreover, despite the fact that the idea of doubled copies of sensor models gave generally higher identification possibilities, duplication of sensors was not beneficial in every type of combination.
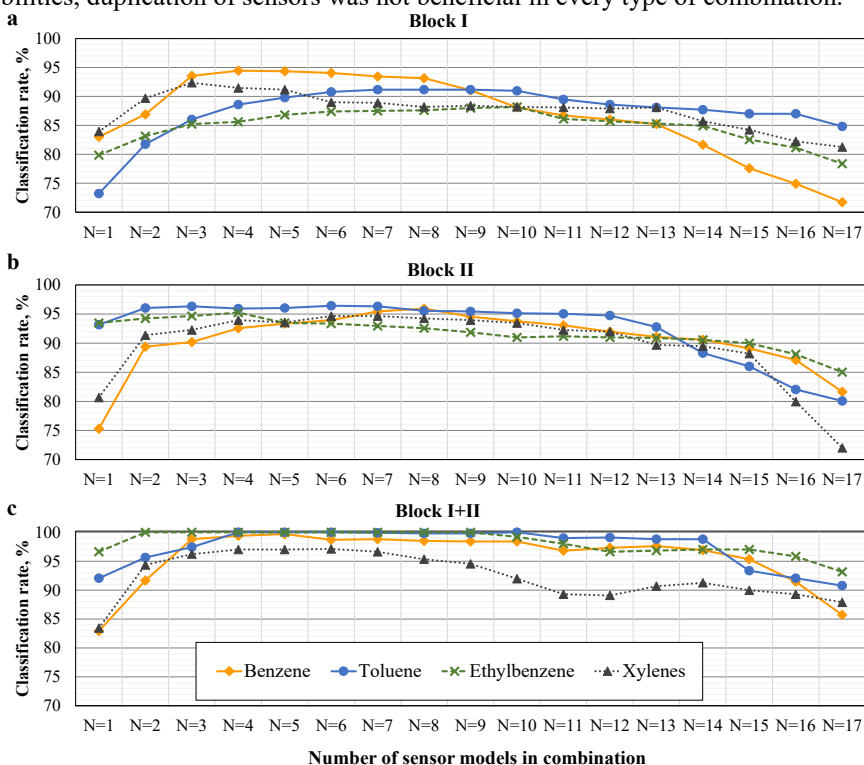


**Fig. 3.** The best classification rates of BTEX compounds as a function of number of sensor models in feature vectors combinations: a) and b) the case of two independent sensor arrays, c) the case of sensor device with doubled sensor models.

## 4 Conclusions

Two sensor arrays (Block I and Block II) were adopted for measurements of BTEX emissions from stabilization/solidification processes and it has been revealed that selection of sensors is necessary to obtain high classification rates. When arrays were treated independently the information about composition of gas samples was available every 30 seconds and the highest recognition ratios were at the level of 88% to 94% for Block I

and 94% to 96% for Block II, depending on the classification task. Improvement of classification efficiency was noticed when copies of each sensor model from both arrays were used for feature vectors formation. Classification rates equal 100% were reached in case of toluene and ethylbenzene samples, 99.7% in case of benzene and 97.1% for xylenes mixtures. However, in that case, 60-second sampling was required. The results show that BTEX compounds identification could be done by near real-time monitoring system based on semiconductor gas sensors and pattern recognition technique.

# References

1.  P. Panagos, M. Van Liedekerke, Y. Yigini, L. Montanarella, J. Environ. Public Health **2013**, 158764 (2013)

2.  J. Pinedo, R. Ibáñez, J.P.A. Lijzen, Á. Irabien, J. Environ. Manage. **130**, 72 (2013)

3.  U.S. Environmental Protection Agency, https://www.epa.gov/haps/initial-list-hazardous-air-pollutants-modifications [access: 30.01.2018]

4.  E. Durmusoglu, F. Taspinar, A. Karademir, J. Hazard. Mater. **176**, 870–877 (2010)

5.  E.M. Martins, P.F. de Sá Borba, N.E. dos Santos, P.T. dos Reis, R.S. Silveira, S.M. Corrêa, Environ. Monit. Assess. **188**, 608 (2016)

6.  E. López, M. Schuhmacher, J.L. Domingo, Environ. Sci. Pollut. Res. **15**, 278 (2008)

7.  S.M. Correa, G. Arbilla, M.R.C. Marques, K.M.P.G. Oliveira, Atmos. Pollut. Res. **3**, 163 (2012)

8.  Z. Jiang, B. Grosselin, V. Daële, A. Mellouki, Y. Mu, Sci. Total Environ. **574**, 1659 (2017)

9.  A. Serrano, M. Gallego, J.L. González, Environ. Pollut. **144**, 203 (2006)

10. International Agency for Research on Cancer, http://monographs.iarc.fr/ENG/Classification/latest_classif.php [access: 30.01.2018]

11. Y. Kerchich, R. Kerbachi, J. Air Waste Manag. Assoc. **62**, 1370 (2012)

12. Y.Y. Zhou, J.F. Yu, Z.G. Yan, C.Y. Zhang, Y.B. Xie, L.Q. Ma, Q.B. Gu, F.S. Li, Environ. Monit. Assess. **185**, 3037 (2013)

13. C.E. Devine, P. Bennett, K. Synowiec, S. Nelson, R. Mohler, M. Einarson, Groundwater Monit. Remediat. **34**, 95 (2014)

14. A. Szczurek, M. Maziejuk, M. Maciejewska, T. Pietrucha, T. Sikora, Sens. Actuators B Chem. **240**, 1237 (2017)

15. A. Szczurek, M. Maciejewska, Talanta **64**, 609 (2004)

16. A. Szczurek, M. Maciejewska, Ł. Bodzoj, B. Flisowska-Wiercik, IEEE Sens. J. **10**, 924 (2010)

17. A. Szczurek, M. Maciejewska, B. Flisowska-Wiercik, Talanta **83**, 916 (2011)

18. K. Banaszkiewicz, T. Marcinkowski, Environ. Prot. Eng. **40**, 57 (2014)

19. K. Banaszkiewicz, T. Marcinkowski, E3S Web Conf. **22**, 00009 (2017)

20. A. Szczurek, M. Badura, K. Banaszkiewicz, M. Maciejewska, T. Marcinkowski, Environ. Eng. Sci. **31**, 533 (2014)

21. R.G. Brereton, *Chemometrics for pattern recognition* (John Wiley & Sons, Ltd, UK, 2009)