Artificial Neural Network-Based Classification of Volatile Organic Compounds for Indoor Air Quality Control

Georgi Georgiev[1], Zvezditza Nenova[2] and Stefan Ivanov[3]

Abstract – **The volatile organic compounds (VOCs) are chemical contaminants which are subject to indoor air quality control. The results of measurement of VOCs concentration in test chamber by metal-oxide gas sensors are presented. An artificial neural network (ANN) approach for the classification of acetone, ethanol, trichloroethylene, dibutyl phthalate, xylene-o and benzene is proposed.**

Keywords – **volatile organic compounds, artificial neural network.**

I. Introduction

The levels of many biological, chemical, and particulate contaminants are monitored at indoor air quality control. According to a scores laboratory centers and organization which regulate eligible and dangerous for human health pollution levels, the indoors, in which the problem of air quality is considered, are offices, schools, industrial and commercial buildings and hospital facilities [1-5]. The volatile organic compounds are some of the most common chemical pollutions. The levels of specific contaminant are controlled in offices, schools and homes. The total volatile organic compounds (TVOCs) are monitored in indoors with industrial and commercial activities. These compounds are not subject to measurement in hospitals. There are different sources of VOCs as adhesives, varnishes, cosmetics, cleaners, furniture, carpets, curtains, floor and ceiling coverings, office machines, etc. [5-7]. The effect of the volatile organic compounds on the people can cause a variety of health problems and illnesses [8, 9].

The problem of identification of gases is solved by using of different types of discriminant analysis as Kernel Discriminant method [10], the method of k-nearest neighbors [11], ANN methods, etc. An ANNs are widely used because of the achieving of higher classification accuracy [12-14].

The paper presents ANN-based classification of some of the most common volatile organic compounds – acetone (A), ethanol (E), trichloroethylene (T), dibutyl phthalate (D), xylene-o (X) and benzene (B), which are subject of control in at least two types of indoors.

II. Experimental procedure

The experimental equipment, shown in Fig. 1, consists of test gas chamber, measuring module AS-ML of AppliedSensor company and three types of gas sensors – AS-MLV for detection of substance of VOCs, AS-MLC for carbon monoxide (CO) and AS-MLN for nitrogen dioxide (NO2) [15]. The sensor module control, data acquisition and communication are realized by personal computer (PC) and serial interface RS-232. The parameters which are subject of measurement are resistances *Rs* of the sensing elements as marked with *RsV*, *RsC* and *RsN*, respectively. The variation of gas concentration leads to change of these parameters. The injecting of the tested compounds in the chamber is performed

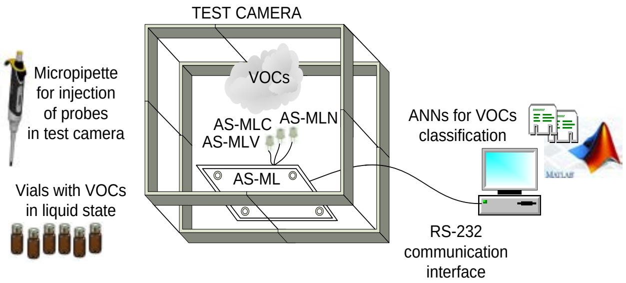


Fig. 1. Experimental equipment

by micropipette SVA – 100 model of Accumax company [16].

In Table I are given the measurement ranges of the investigated gases.

TABLE I
MEASUREMENT RANGES OF VOCS

Gas	Measurement ranges, ppm
acetone (A)	5 - 25
ethanol (E)	3 - 19
trichloroethylene (T)	1.2 – 7.2
dibutyl phthalate (D)	0.4 – 2.4
xylene-o (X)	0.88 – 5.28
benzene (B)	1.17 – 7.02

[1]Georgi Georgiev is with the Faculty of Electrical Engineering and Electronics at Technical University of Gabrovo, 4 H. Dimitar Str., Gabrovo 5300, Bulgaria, E-mail: givanow@abv.bg.

[2]Zvezditza Nenova is with the Faculty of Electrical Engineering and Electronics at Technical University of Gabrovo, 4 H. Dimitar Str., Gabrovo 5300, Bulgaria, E-mail: z_nenova@yahoo.com.

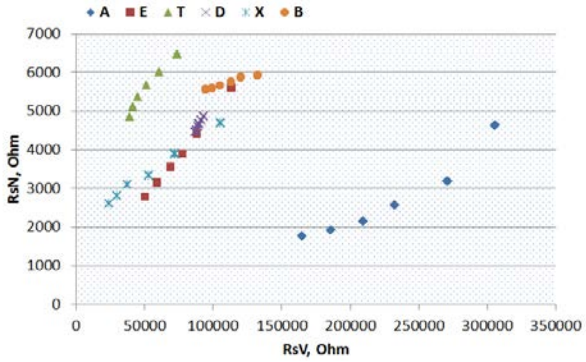
[3]Stefan Ivanov is with the Faculty of Electrical Engineering and Electronics at Technical University of Gabrovo, 4 H. Dimitar Str., Gabrovo 5300, Bulgaria, E-mail: st_ivanov@abv.bg.

Each compound is injected separately in the test chamber as the obtaining of a wanted gas concentration in ppm is based on the evaporation of a pre-calculated amount of a liquid compound in μl [17]. The minimum and the maximum values of the measurement ranges correspond to the lowest and the highest gas concentrations which are created into the chamber.

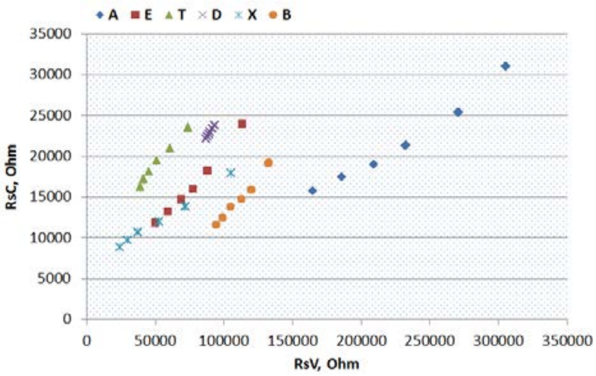
III. RESULTS AND DISCUSION

3.1. Rs measurements

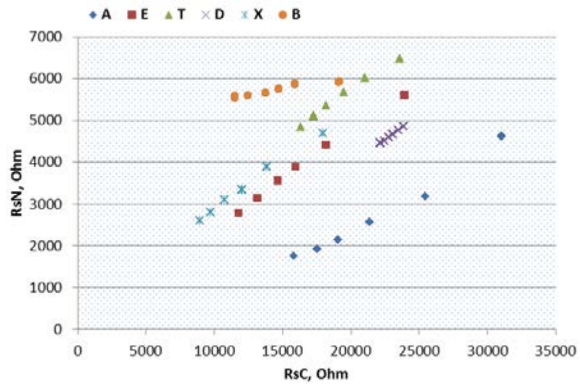
Two-dimensional (2D) and three-dimensional (3D) presentations of the measured resistance of the sensors are shown in Fig. 2 and Fig. 3.



a)



b)



c)

Fig. 2. 2D presentation of: a) RsN=f(RsV), b) RsC=f(RsV) and c) RsN=f(RsC)

The 2D representation shows overlap between the groups B with E, D with E for sensors AS-MLN and AS-MLV and E with X for AS-MLC and AS-MLV. Although in use of the sensors AS-MLC and AS-MLN overlap of groups doesn't observe, the samples are closely located for groups as B and T, X and E. Therefore for the VOCs classification the parameters *RsV*, *RsC* and *RsN* of the three sensors are used.

3.2. ANN classification

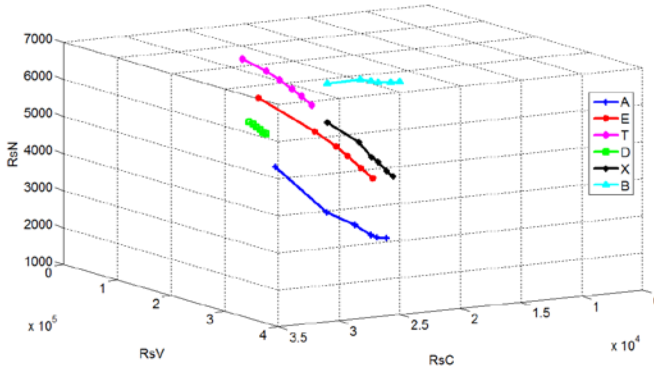


Fig. 3. 3D presentation of RsV, RsC and RsN

The instrumentation of artificial neural networks is widely used for the processing of tasks with varying grades of complexity and dimension of the input and output parameters [18-21]. The advantage of back-propagation learning algorithm is in the process of updating the network weights during each iteration to minimize the error between the expected and the calculated result of the network output [22, 23].

To process the recognition task of the analyzed compounds neural network with one input layer, one hidden layer and output layer is proposed. ANN is built using Matlab R2009b.

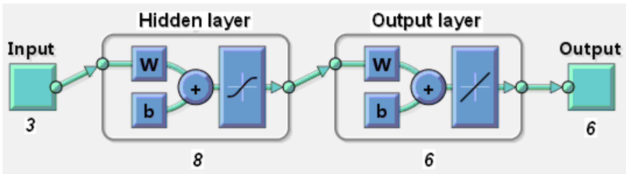


Fig. 4. Artificial neural network architecture

The neural network architecture is presented in Fig. 4. In the hidden layer tangent-sigmoid function is used, and in the output layer - a linear activation function. The input data comprises a group of 90 samples for each gas, i.e. total 540 samples. The dataset is normalized in range of [-1, +1] and divided into training set (60%), validation set (20%) and test set (20%). In Table II are presented the network input parameters, the target output parameters and the parameters of network training. For each of the examined VOCs a definite code combination of the output parameters having 0 and only one 1 is assigned. The group presence is marked by the position of 1.

Tests with different number of hidden neurons are conducted during the training of the network. Optimal network architecture is achieved with eight neurons when a minimum mean squared error from 0.0033 is obtained. In Fig. 5 the change of the mean squared error depending on the number of epochs for ANN training is shown. The validation and test curves have a very similar nature of amendment. There is not an increasing amendment of the mean squared error in respect to the process of ANN testing, which indicates that an overfitting of neural network has not occurred [24]. The minimum gradient magnitude is reached at epoch 8.

TABLE II
INPUT, TARGET AND LEARNING NETWORK PARAMETERS

Input parameters	
Sensor	**Resistance**
AS-MLV	RsV
AS-MLC	RsC
AS-MLV	RsN
Target parameters	
VOCs	**Code combination**
acetone	1 0 0 0 0 0
ethanol	0 1 0 0 0 0
trichloroethylene	0 0 1 0 0 0
dibutyl phthalate	0 0 0 1 0 0
xylene-o	0 0 0 0 1 0
benzene	0 0 0 0 0 1
Learning network parameters	
show	25
epoch	1000
lr (learning rate)	0.05
goal	0.01
min_grad	1e-05

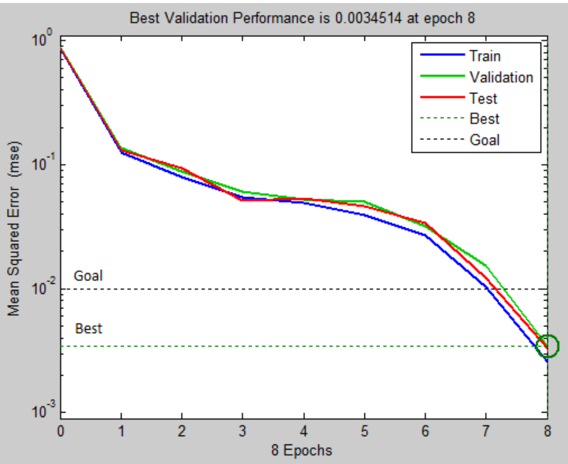


Fig. 5. Change of the mean squared error

In Table III the parameters of linear regression relations between network and target outputs – the slope *m* and y-intercept *b* of the best linear regression and the correlation coefficients *R* are presented.

TABLE III
REGRESSION PARAMETERS RELATING TARGETS TO NETWORK OUTPUTS

Output	Slope of the best linear regression *m*	Y-intercept of the best linear regression *b*	Correlation coefficient *R*
1	0.9875	-0.0086	0.9986
2	0.9266	0.0406	0.9879
3	0.9473	0.0384	0.9894
4	1.0047	-0.0079	0.9981
5	0.8891	-0.0060	0.9886
6	0.8914	0.0018	0.9953

The confusion matrix, which defines the correct and incorrect classifications, is given in Fig. 6. The matrix is formed on the base of the test set from 108 samples and shows a correct classification of all considered groups (classes) by ANN proposed.

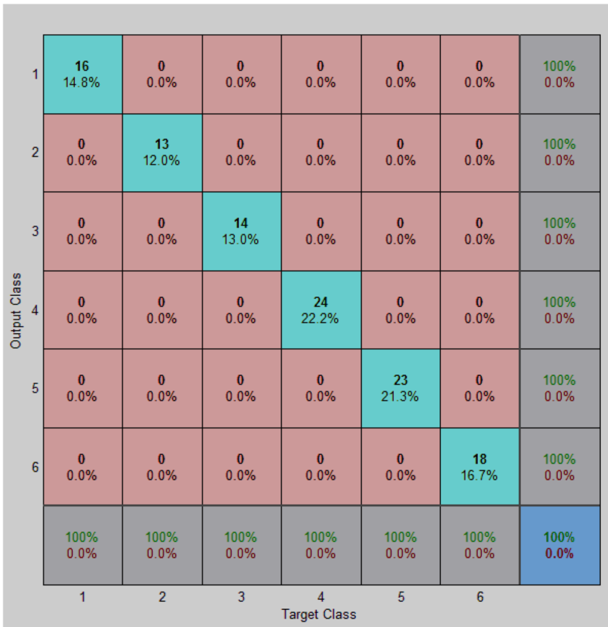


Fig. 6. Confusion matrix for the classes 1-A, 2-E, 3-T, 4-D, 5-X and 6-B

The absolute network error for the samples of the test set for each output group (class) is shown in Fig. 7. The error is calculated as a difference between the network output values and the target values for each of the test samples and varies in the range from -0.2349 to 0.2446. The smallest errors are observed for the fourth group (D) and the greatest – for the fifth group (X).

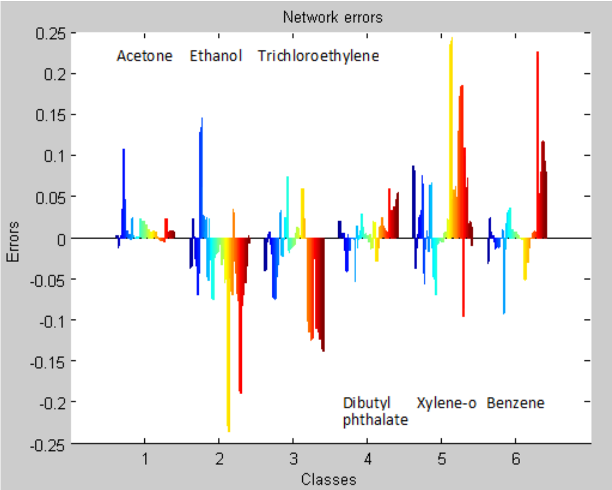


Fig. 7. Change of the network error

The obtained results confirm the possibility of using the specified sensors and ANN for classification of VOCs at indoor air quality control.

IV. Conclusion

The identification of volatile organic compounds at indoor air quality control is a major step to ensure a health environment with minimum contamination levels, to improve the quality of study at schools, to provide a comfortable environment at homes. The trained ANN and the relevant gas sensors could be successfully used in the computer based systems for monitoring of VOCs contamination in a different category of indoors.

References

[1] World Health Organization – Regional Office for Europe, “WHO Guidelines for Indoor Air Quality: Selected Pollutions”, 2010. http://www.euro.who.int.

[2] Air Quality Sciences, “The Office Equipment Industry’s Guide to Managing Product Emissions”, 2003. http://www.aqs.com.

[3] M. Freitas, et al.*, “Indoor Air Quality in Primary Schools”, Environmental Health and Air Pollution Case Studies, DOI: 10.5772/17609, pp.361-384, 2011.

[4] California Environmental Protection Agency, Ventilation and Indoor Air Quality in New Homes. California Energy Commission, PIER Collaborative Report, CEC-500-2009-085, 2009. http://www.calepa.ca.gov.

[5] G. Georgiev, Z. Nenova, “Control of Indoor Air Pollutions”, International Scientific Conference “UNITECH’12”, Gabrovo, Bulgaria, vol. II, pp.371-376, 2012.

[6] International Association of Certified Home Inspectors, “Air Quality in the Home”, 2012. http://www.nachi.org.

[7] A. Pickett, M. Bell, “Assessment of Indoor Air Pollutions in Homes with Infants”, International Journal of Environmental Research and Public Health, 8, pp.4502-4520, 2011.

[8] United States Environmental Protection Agency, “An Introduction to Indoor Air Quality”, 2012. http://www.epa.gov.

[9] California Environmental Protection Agency, “Indoor Air Quality and Personal Exposure Assessment Program”, 2003. http://www.arb.ca.gov.

[10] L. Zhang, F. Tian, “A new kernel discriminant analysis framework electronic nose recognition”, National Center for Biotechnology Information, Anal. Chim. Acta, 833, pp.48-55, 2014.

[11] S. Belhaouari, R. Abaza, “Gas Identi cation by Using of a Cluster-K-Nearest Neighbor”, International Conference of Machine Learning and Computing, vol.3, pp.123-127, 2011.

[12] P. Chang, J. Shih, “The Application of Back Propagation Neural Network of Multi-channel Piezoelectric Quartz Crystal Sensor for Mixed Organic Vapours”, Tamkang Journal of Science and Engineering, vol. 5, No. 4, pp. 209-217, 2002.

[13] J. Shih, H. Hsu, “Multi-Channel Surface Acoustic Wave (SAW) Sensor Based on Artificial Back Propagation Neural (BNP) Network and Multivariable Linear Regression Analysis (MLR) for Organic Vapours”, Jornal of Chinese Chemical Society, 2007, 54, p.401-410, 2007.

[14] C. Chen, J. Shih, “Multi-Channel Piezoelectric Quartz Cristal Sensor with Printipal Componenten Analysis and Back-Propagation Neural Network for Organic Pollutant from Petrochemical Plants”, Journal of the Chinese Chemical Society, 55, pp.979-993, 2008.

[15] AppliedSensor, “Quick Reference Manual – Test Kit AS-ML Sensor Components”, 2009. http://www.appliedsensor.com.

[16] Accumax, “Accumax SMART series of pipette”. http://www.accumaximum.com.

[17] A. Uyanik, N. Tinkilic, “Preparing Accurate Standard Gas Mixtures of Volatile Substances at Low Concentration Levels”, Chem. Educator, 4, pp.141-143, 1999.

[18] A. Srivastava, “Detection of volatile organic compounds (VOCs) using SnO_2 gas-sensor array and artificial neural network”, ELSEVIER Sensors and Actuators B 96, pp.24-37, 2003.

[19] S. Benedetti, S. Mannino, A. Sabatini, G. Marcazzan, “Electronic nose and neural network use for classification of honey”, INRA/DIB-AGIB/ EDP Sciences, Apidologie 35, pp.1-6, 2004.

[20] A. Yamazaki, T. Ludermir, “Classification of Vintages of Wine by an Artificial Nose with Neural Networks”, Centre of Informatics Federal University of Pernambuco, Brazil, 2001.

[21] S. Sharma, “Implementation of artificial neural network for odor identification using e-nose”, National Conference of Computational Instrumentation “NCCI 2010”, CSIO Grandigarh, India, 2010.

[22] R. David, L. Sovan, D. Ioannis, J. Jean, L. Jacques, S. Aulagnier, Artificial neural networks as classification method in the behavioural sciences, ELSEVIER Behavioural Processes 40, pp.35-43, 1997.

[23] G. Zhang, “Neural Network for Classification: A Survey”, IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews”, vol.30, №4, pp.451-462, 2009.

[24] Mathworks, “Analyze Neural Network Performance After Training”, 2014. http://www.mathworks.com.