sciendo

DOI: 10.2478/eces-2025-0009 ECOL CHEM ENG S. 2025;32(2):171-188

Zbigniew SUCHORAB[1*], Agnieszka MALEC[1], Łukasz GUZ[1], Monika SAK[1]
Dariusz MAJEREK[2], Thirumalini SELVARAJ[3] and Grzegorz ŁAGÓD[1]

ESTIMATION OF MICROBIAL VOC CONCENTRATION OF FUNGAL ORIGIN USING GAS SENSOR ARRAY AND ARTIFICIAL INTELLIGENCE

Abstract: Microbial volatile organic compounds (mVOC) are indicators of the mycological load of buildings. It is an unfavourable phenomenon and has a negative impact on the health of residents, as it is associated with the problem of Sick Building Syndrome and buildings affected by it should be subject to repair processes. In order to diagnose the presence of mould fungi in the air of closed rooms, diagnostic tests should be carried out, which are usually time-consuming and require large financial outlays. The basic diagnostic methods include microbiological methods, associated with the cultivation and counting of fungal colonies. The second type of tests includes chemical tests, including gas chromatography, which allows the detection of chemical substances that are markers of the presence of fungi (including mVOC). This article proposes the use of an array of MOS gas sensors (the so-called electronic nose) as a tool for detecting mycological infestation of rooms. The signals obtained from the gas sensor array were then subjected to advanced analysis consisting in the development of predictive models for estimating concentrations. For this purpose, deterministic models based on linear regression without multidimensionality reduction were developed, and then artificial intelligence tools were used - Multilayer Perceptron (MLP) and Support Vector Machine (SVM). The analyses carried out confirmed the effectiveness of artificial intelligence models for predicting fungal marker concentrations and smaller concentration estimation errors compared to deterministic models based on linear regression.

Keywords: mould, electronic nose, artificial intelligence, gas chromatography, microbial volatile organic compounds

Introduction

In the interior of buildings, there may be several hundred species of fungi, bacteria and other microorganisms. These include the following types of fungi: *Cladosporium sphaerospermumn*, *Penicillium chrysogenum*, *Aspergillus niger*, *Aspergillus versicolor*, *Alternaria alternata*, *Stachybotrys chartarum* [1, 2]. They constitute one of the groups of biological factors commonly found in both the work environment and the communal environment. Many species of fungi have an allergenic, toxic or infectious effect on

[1] Faculty of Environmental Engineering and Energy, Lublin University of Technology, ul. Nadbystrzycka 40B, 20-618 Lublin, Poland, email: z.suchorab@pollub.pl, aga_malec@interia.pl, l.guz@pollub.pl, monika.garbacz@interia.pl, g.lagod@pollub.pl, ORCID: ZS 0000-0001-8658-864X, ŁG 0000-0002-5024-4738, GŁ 0000-0002-0621-7222
[2] Faculty of Mathematics and Information Technology, Lublin University of Technology, ul. Nadbystrzycka 38, 20-618 Lublin, Poland, email: d.majerek@pollub.pl, ORCID: 0000-0002-0035-7187
[3] CO_2 Research and Green Technologies Centre, Vellore Institute of Technology, Vellore, 632014 Tamil Nadu, India, email: p.thirumalinj@yahoo.in, ORCID: 0000-0002-3431-6025
[*] Corresponding author: z.suchorab@pollub.pl

humans [3]. Exposure to these microorganisms may lead to many adverse and serious health effects, e.g. allergies, mycoses or toxic reactions [4]. Most fungi are microscopic organisms, invisible to the naked eye (often referred to by mycologists as microscopic fungi) or visible in the form of blooms [5]. The growth of fungi is facilitated by high relative air humidity (optimum above 70 %) and appropriate ambient temperature (optimum is 16 ℃ - 35 ℃). Therefore, their development depends on the climate and season [6]. The most harmful species of moulds include *Aspergillus* species, which are commonly present in buildings with very damp wall surfaces. This genus contains several toxic species, the most important of which are *Aspergillus pasiticus*, *Aspergillus flavus*, *Aspergillus fumigatus* and *Aspergillus niger*. The aflatoxins produced by these species are among the strongest mycotoxins. These are some of the most toxic substances with particularly harmful properties for the liver, brain, kidneys and heart, and are also highly teratogenic [7]. Another type of fungi that is commonly found in buildings is *Stachybotrys chartarum*. Fungi belonging to the genus *Stachybotrys chartarum* grow intensively in conditions of high humidity. They are mesophilic organisms, developing in the range from 18 °C to 35 °C, and also in a much wider range [8].

Indoor environment contamination with fungi is mainly caused by the presence of Microbial Volatile Organic Compounds (mVOC), which are a subgroup of volatile organic compounds (VOC) - considered carcinogenic - mycotoxins, fungal spores and mycelium fragments [9]. These substances are secondary metabolites formed in the fermentation process and are volatile due to their physicochemical properties [10]. Compounds from the mVOCs group with a characteristic mouldy odour indicate the presence of filamentous fungi in buildings. Generally, higher concentrations of mVOCs are observed in rooms, mainly closed most often due to improperly functioning ventilation [11]. The toxins considered to be the most harmful are produced by fungi of the genera *Stachybotrys chartarum*, Fusarium and *Aspergillus versicolor* [12]. The most dangerous mycotoxins produced by fungi include ochratoxin a (OT), zearalenone (ZEN), aflatoxins (AF), trichothecenes, and fumonisins (F) [13]. Other toxins produced by fungi are ketones, alcohols, esters, terpenes, and sulphur compounds, including the following substances [14]: ketones: 2-heptanone, 2-pentanone, and 3-octanone; alcohols: 2-methyl-1-butanol, 3-methyl-1-butanol, 2-methylpropanol, 3-octanol, 1-octen-3-ol, 1-hexanol, 1-pentanol, 2-methylisoborneol, and geosmin; terpenes and sesquiterpenes: limonene and pinene; furans: 3-methylfuran; sulphur: 296 compounds; and hydrocarbons: alkanes, olefins, dienes, and trienes [12, 15-17]. Microbial volatile organic compounds mVOC are formed during both the primary and secondary metabolism of microorganisms as by-products. Currently, about 400 such compounds and about 350 species of fungi that produce them are known.

Microbiological and chemical methods are used to assess the degree of mycological contamination of rooms. Microbiological methods for assessing indoor air quality can be divided into the traditional (mycological) method and molecular methods. Mycological assessment involves macroscopic and microscopic observation of collected and cultured microorganisms, determining humidity in the samples taken, and assessing the efficiency of ventilation and other sanitary installations. This method is based on detecting the presence of mould fungi in air samples or swabs, calculating the number of colonies, and determining the species or genus. Traditional techniques include [18]: 1) Koch's sedimentation method, 2) the cascade method, which is based on the flow of contaminated air through small openings. Molecular methods, which are among the most accurate, but

also the most expensive methods of measuring microbiological hazard. Molecular methods for detecting microbial infections include the PCR (Polymerase Chain Reaction) method, which is based on the amplification of a DNA fragment by heating and cooling the sample in laboratory conditions and the NGS (Next Generation Sequencing) method, which is performed on the entire DNA or RNA isolated from the building material and allows for a complete description of the entire community without the need for culturing [19, 20]. Microbiological methods are time-consuming, require qualified laboratory personnel and equipment. Molecular tests, on the other hand, are characterised by expensive analyses.

Chemical methods provide a lot of information about the presence of microorganisms. They are most often used to estimate the metabolic activity and, consequently, the toxic potential of a microbial population for a given substrate. The possibilities are as follows: 1) detection of chemical components that make up microbial cells, such as components that make up mycelium cells in the case of fungi (chitin and ergosterol), adenosine triphosphate (ATP) which is an energy-producing molecule, and cell wall polysaccharides (β-d-glucan). The amount of these components can be related to the number of microorganisms or correlated with the type of microbial species [21]; 2) detection of chemical compounds produced by microorganisms, such as nitric oxide, various toxins (mVOC, mycotoxins, endotoxins, etc.) and other metabolites taken up from substrates. Using this indirect method, it is possible to assess the metabolic (or biological) activity of microorganisms and, as a result, provide estimates of the microbial population. This method is mainly used to collect information on the amount of potentially harmful compounds (volatile compounds or metabolites on substrates) and to identify the pathogenic potential of the tested environment and the resulting health risks [22]. Chemical methods mainly allow for the identification of selected markers that are products of metabolites, such as mycotoxins or volatile organic compounds. Among all the methods, the most popular are [23] gas chromatography mass spectrometry and high-performance liquid chromatography. Both methods require a complex and time-consuming sample preparation process [24].

An alternative to the methods discussed earlier are the techniques based on gas sensor arrays, the so-called electronic noses (e-noses). The principle of operation of the e-nose is based on a simplified model of a biological olfactory analyser. The sensor array consists of a plate on which several or several dozen different sensors are placed that react selectively to the presence of different compounds or groups of chemical compounds [25]. The main element of the electronic nose is the sensor array. Stimulation of its sensitive elements in specific types of odour detection and identification systems leads to the generation of specific signals, similarly to the stimulation of odour receptors in the olfactory organ. After applying an appropriate method of processing the received signals and analysing the data, it is possible to generate unique responses that are characteristic for individual odour samples, just as fingerprints are unique to a given person [26]. Recognition of a given gas sample is possible not by stimulating the entire sensor array, but each of the sensors to a different degree. As a result of such stimulation, a characteristic set of responses from individual sensors is created [27]. Analysis of the obtained response vector allows the classification of a given gas sample into the set of response vectors known earlier.

The aim of this work is to demonstrate the possibility of using multi-sensor arrays built from low-selective metal oxide semiconductor (MOS) gas sensors for rapid detection of the threat of biological contamination of rooms caused by fungi by estimating the concentrations of selected chemical compounds.

Materials and methods

The research was conducted in the Lublin province in residential buildings constructed between 1960 and 2005. The list of the tested rooms is presented in Table 1.

Table 1

List of facilities examined during the research

No.	Room type	Year of construction	Microbial contamination
1	Non-residential - teaching room	2001	No
2	Residential - room	2009	No
3	Non-residential - teaching room	2001	No
4	Non-residential - laboratory	2001	No
5	Residential premises - basement	1996	Yes, visible mould growths on walls and pipes, additionally, a sharp, unpleasant smell
6	Residential premises	1967	Yes, a sharp, unpleasant smell of mould is noticeable
7	Non-residential - teaching room	2001	No
8	Residential premises	1967	No
9	Residential premises	1963	Yes, a sharp, unpleasant smell of mould is noticeable
10	Residential premises - basement	1963	Yes, there is a sharp, unpleasant smell of mould, additionally, visible efflorescence on the walls and ceiling
11	Residential premises	1968	Yes, a sharp, unpleasant smell of mould is noticeable
12	Residential premises	1968	Yes, there is a sharp, unpleasant smell of mould. additionally, visible efflorescence on the walls and ceiling

These were both public utility buildings (marked with the following numbers: 1, 3, 4, 7) and the residential buildings (numbers: 2, 5, 6, 7, 9, 10-12). None of the buildings were of an industrial or medical nature, because such facilities could contain additional sources of volatile organic compounds due to the nature of the premises, which could negatively affect the measurements conducted. In the facilities, both entire residential premises (6, 8, 9, 11, 12) and individual rooms (5, 10) were examined. The buildings were characterised by the fact that they were all constructed using the masonry technology. The walls in all facilities were covered with cement-lime plasters. None of the rooms examined were equipped with air conditioning. Only one of them (No. 4) had organised mechanical ventilation system and all the other rooms were ventilated using gravity ventilation. The external finish of the buildings did not indicate the presence of damage that could potentially affect the dampness of the building. The facilities were located in urban and rural development. The selection of rooms for testing was based on the following selection criteria: 1) diversity of microbial load, 2) initial visual and sensory assessment, including an interview with the users of the rooms.

The research was conducted during winter period 2019/2020 and included chemical tests to identify chemical substances present in the air in the tested rooms, as well as tests using a MOS sensor array.

Chemical analyses

Chemical tests were carried out using the chromatographic method. The substance concentration values obtained in this way were of reference nature and allowed the development of both deterministic models and those based on supervised machine learning, allowing their estimation using a sensor array.

Due to the fact that mVOCs constitute a large group of compounds, the chemical tests concerned 8 previously selected substances. These were: 1) cumin, 2) styrene, 3) ethylbenzene, 4) α-pinene, 5) 1oct-3ol, 6) 2m-1butanol, 7) 2-heptanone and 8) limonene.

The equipment of the research station for chromatographic studies included:

- gas chromatograph with a GC-MS TraceUltra - PolarisQ, Thermo mass spectrometer with an Equity 5MS chromatographic column from Supelco, length 30 m, internal diameter 0.25 mm, phase thickness 0.25 µm.
- SPME Sulpeco CAR fibres, phase thickness df 75 µm (3 pieces).

The microextraction technique on SPME fibres was used to study the concentrations of chemical compounds. Supelco CAR fibres were used, which were placed for 60 minutes in the tested room to absorb fungal metabolites in the air. The fibre was placed in the middle of the room with free access to air on all sides of the fibre. After absorption, the fibre was transported to the laboratory at a temperature of 4 °C, where it was subjected to desorption. After desorption, the fibres were analysed by gas chromatography coupled with a GC-MS TraceUltra - PolarisQ, Thermo mass spectrometer. The helium flow through the column was 1.2 mL/min with a purity of 99.999 %. As a result of the chromatographic analysis, chromatograms were obtained, on the basis of which the concentrations of compounds classified as mVOC volatile organic compounds produced by fungi were determined.

Analyses using MOS array

Array configuration

The research was carried out using an array equipped with MOS gas sensors. The measuring device (so-called electronic nose) is a device of our own design developed at the Lublin University of Technology [28]. It is an array consisting of 17 sensors sensitive to inorganic substances (TGS-2600-B00, TGS-4161, TGS-2442-B02, TGS-825-A00, TGS-821) and organic substances (TGS-2610-D00, TGS-2611-E00, TGS-2620-C00, TGS-813-A00, TGS-830, TGS-832-A00). The manufacturer of the sensors is FIGARO [29]. The detection ranges for individual substances are in the range of 1 ppm - 10000 ppm, which ensures good detection of chemical compounds even at low concentration levels. The configuration of the array used for the tests is presented in Table 2, and the actual array with visible individual sensors is shown in Figure 1.

a) b)

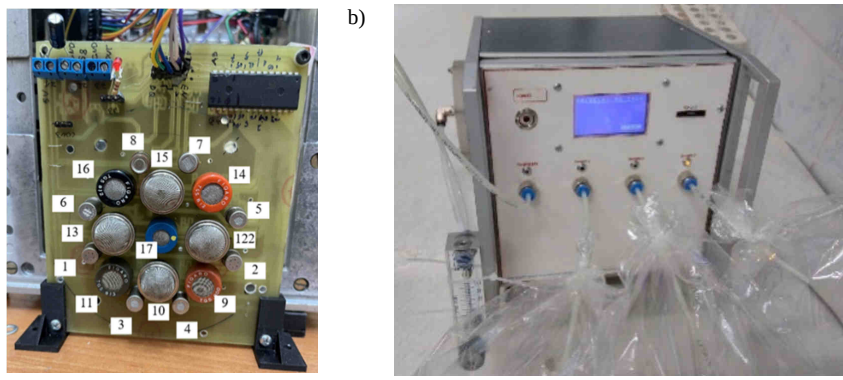

Fig. 1. Device used in experiment: a) array with visible sensors, b) electronic nose

Table 2

A list of MOS sensors mounted in the array used for testing

No.	Name	Detected substance	Additional detected substance	Detection range [ppm]
s1	TGS-2600-B00	General air pollutants	H_2, CO, cigarette smoke	1-30 (for H_2)
s2	TGS-2602-B00	Toxic gases (NH_3, H_2S, C_2H_5OH, C_6H_5-CH_3)	H_2	1-30 (for ethyl alcohol)
s3	TGS-2610-D00	Propane and butane (on carbon filter) (C_3H_8, C_4H_{10})	Ethyl alcohol, methane, isobutane	500-10000
s4	TGS-2611-E00	Methane (on carbon filter) (CH_4)	Ethyl alcohol, isobutane	500-10000
s5	TGS-2620-C00	Ethyl alcohol	Methane, isobutane	50-5000
s6	TGS-4161	Carbon monoxide(II)	-	350-10000
s7	TGS-2444	Ammonia	H_2S, NH_3, ethyl alcohol	10-100
s8	TGS-2442-B02	Carbon monoxide(II)	Ethyl alcohol, H_2, CO	30-1000
s9	TGS-800	General air pollutants	CO, methane	1-1000 (for isobutane)
s10	TGS-825-A00	Hydrogen sulphide	Ethyl alcohol	5-100
s11	TGS-813-A00	Flammable gases	Ethyl alcohol, methane, isobutane	500-10000
s12	TGS-821	Hydrogen	Methane, ethyl alcohol, CO	10-5000
s13	TGS-823-A00	Solvent vapours	Isobutane, CO	50-5000
s14	TGS-812	Toxic and explosive gases (methane, isobutane, hydrogen, carbon monoxide(II))	-	100-5000
s15	TGS-830	Chlorofluorocarbons	-	100-3000
s16	TGS-832-A00	Chlorofluorocarbons	Ethyl alcohol	100-3000
s17	TGS-2106	Exhaust gases, exhaust gases for diesel engine	-	0.1-10 (for NO_2)

Air sampling

The cycle of a single measurement in the room lasted 15 minutes and included a phase of rinsing the sensors with clean air for 5 minutes and a phase of sampling for the next 10 minutes. Before each measurement in the room, the array was rinsed with clean air.

For this purpose, tedlar bags were used, which were filled in the laboratory with synthetic air from a cylinder without hydrocarbons ($C_nH_m \leq 0.1$ ppm) before starting the field measurements. The air from the bags was passed through the array before starting the actual measurement for a period of 5 minutes. After rinsing the sensors (zero measurement), the device automatically switched to taking air samples from the tested room. In order to eliminate interference, the rooms were left without people for the measurement period. The measurement results were saved on the memory card as averages from 5 seconds. The last 2-minute recording is used for signal analysis in each case, which is characterised by a stabilised signal generated by the individual sensors, as shown in Figure 2, where the responses of the individual sensors are represented by lines marked with different colours.

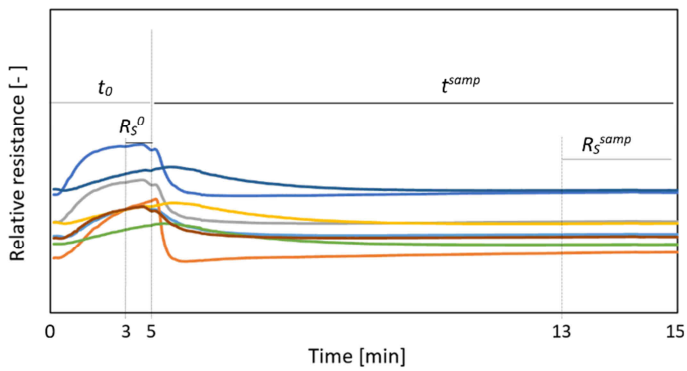

Fig. 2. Example of a graph with the distribution of signals from the array (graph only for sensors with a clear response to the tested sample)

In order to minimise the effects of changes in the sensor characteristics over time, caused by their so-called "poisoning" [30], the so-called relative resistance is determined. For each sensor, the relative resistance R_w [-] was determined according to the following formula [31]:

$$R_w = {R_S}^{samp} / {R_S}^0 \tag{1}$$

where: ${R_S}^{samp}$ - stable sensor resistance value for the tested sample [kΩ], ${R_S}^0$ - baseline (sensor resistance for zero air) [kΩ]. The ${R_S}^0$ baseline was assumed to be the average resistance value of a given sensor from the final phase of flushing the sensors with zero air before measurements.

Data analysis methods

Linear regression without dimension reduction

Linear regression is a statistical approach used to model the relationship between a dependent variable and one or more independent variables. In this study, the method was employed in a multiple regression framework to predict the concentration of chemical compounds based on sensor array readings. This model assumes a linear relationship between the predictors (sensor readings) and the response variable (chemical concentration). The regression model is mathematically represented as:

$$C = \beta_0 + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \ldots + \beta_{17} S_{17} \tag{2}$$

where: β_0 - β_{17} - structural parameters of the model, S_1 - S_{17} - readings from MOS array sensors, C - estimated concentration value of a given substance [$\mu g/cm^3$].

To construct the model, a dataset comprising sensor readings and their corresponding compound concentrations was used. The parameters of the model were estimated using the method of least squares, which minimises the sum of squared differences between the observed and predicted values of C. This optimisation process ensures that the model provides the best linear fit to the data.

The model's construction involved training the regression algorithm on a subset of the data to estimate the parameters. During this process, the algorithm computed the optimal values of β parameters to reduce the residual errors. The residual errors are the differences between the actual values and those predicted by the model. Once trained, the model was applied to a separate subset of the data to verify its predictive performance.

Linear regression without dimensionality reduction was chosen for its interpretability and its ability to utilise the complete set of sensor data. This approach maintains all the sensor readings as predictors, allowing the model to capture the full spectrum of information provided by the sensor array. Despite its simplicity, this method offers valuable insights into the relationships between sensor responses and compound concentrations, making it a fundamental tool in environmental data analysis.

Artificial Intelligence models

Due to the complexity of the relationship between input and output data, the use of linear models may not fully reflect reality. For this reason, the use of models based on artificial intelligence seems to be advisable. Two methods of machine learning technique were used Multilayer Perceptron (MLP) and Support Vector Machine (SVM).

The Multilayer Perceptron (MLP) is a type of artificial neural network used in this study to model the non-linear relationships between sensor array readings and the concentrations of chemical compounds. MLPs are particularly well-suited for capturing complex patterns in data, making them a robust choice for analysing multi-sensor measurements. An MLP consists of three primary components: an input layer, one or more hidden layers, and an output layer. Each layer is composed of neurons that are interconnected, with weights assigned to these connections.

In this study, the MLP architecture included an input layer with 17 neurons, one for each sensor in the MOS array. A single hidden layer with 20 neurons was employed, and the output layer consisted of a single neuron, representing the predicted concentration of a specific compound. Each hidden neuron applied a non-linear activation function, such as the rectified linear unit (ReLU), to introduce non-linearity and enable the network to learn complex dependencies between the input and output. The MLP was trained using supervised learning, where the input consisted of sensor readings, and the output was the corresponding concentration of a chemical compound determined through reference measurements. During training, the network adjusted the weights of the connections to minimise the error between the predicted and actual concentrations. This error was quantified using a loss function, such as mean squared error (MSE). Optimisation was performed using backpropagation combined with a gradient descent algorithm, which iteratively updated the weights to reduce the loss. A separate MLP model was constructed for each compound, allowing the network to specialise in predicting the concentration of individual chemicals. The use of MLPs was critical in capturing the non-linear interactions between sensor responses and compound concentrations, which cannot be effectively

modelled with linear approaches. By leveraging hidden layers and non-linear activation functions, the MLP was able to process complex sensor data and produce accurate predictions.

The Support Vector Machine (SVM) is another advanced machine learning method used in this study to analyse the relationship between sensor array readings and chemical compound concentrations. SVMs are particularly effective for modelling non-linear relationships and are based on the concept of finding the optimal hyperplane that separates data points in a high-dimensional space. In this regression context, known as Support Vector Regression (SVR), the goal is to find a function that approximates the data within a specified margin of tolerance.

For this study, the SVM algorithm utilised the sensor readings as input features and predicted the concentration of a specific compound. A kernel function, such as the radial basis function (RBF), was applied to transform the input data into a higher-dimensional space where the relationships between sensor signals and concentrations could be more easily captured. The kernel function enables SVM to handle non-linear dependencies by mapping the input features into a feature space where linear regression can be performed effectively. The SVM training process involved selecting support vectors - data points that lie closest to the decision boundary or hyperplane. These support vectors are critical in defining the model, as they determine the optimal regression function that fits the data while maintaining generalisation. The model parameters, such as the penalty term and kernel parameters, were optimised through a grid search and cross-validation to achieve the best balance between accuracy and complexity. SVMs were chosen for their ability to model non-linear relationships while maintaining robustness against overfitting, especially in high-dimensional datasets like those derived from sensor arrays. Unlike MLPs, which require iterative weight adjustments across multiple layers, SVMs focus on identifying the most critical data points, making them computationally efficient and highly interpretable. By leveraging the flexibility of kernel functions, SVM provided precise predictions of chemical concentrations, even in cases with complex patterns of sensor responses.

Both MLP and SVM methods complemented each other in this study. While MLPs excelled in capturing intricate, high-dimensional patterns through neural network architectures, SVMs offered a robust, kernel-based approach to model non-linear dependencies with fewer hyperparameters to tune. These methods collectively provided a comprehensive framework for analysing the sensor data and estimating the concentrations of volatile organic compounds.

Results

The results obtained from the studies included concentrations of substances recognised as fungal metabolites determined using a chromatographic method, as well as multidimensional resistance readings of MOS sensors recorded on the electronic nose device.

Data from gas chromatography

Raw data from chromatographic measurements are so-called chromatograms, which allow for quantitative determination of concentrations of individual substances by measuring peak areas for given retention times. Chromatograms obtained by the GC-MS TraceUltra - PolarisQ were automatically interpreted using Excalibur software [32]. Table 3

contains the results of chromatographic tests for selected chemical substances for the tested, in both series, for all rooms. All the substances tested belong to the mVOC group, which means they can be fungal metabolites. The literature on the subject does not provide critical concentration values of these compounds for which microbial contamination or its absence can be determined. Available scientific publications provide only the results of qualitative analysis, i.e. only the occurrence or absence of a given substance in the tested rooms [33, 34]. Table 4 contains descriptive statistics for individual compounds identified by GC-MS in all studied rooms and buildings.

Table 3

Concentrations of selected substances determined by gas chromatography in individual rooms

No.	cumin	styrene	ethyl-benzene	α-pinene	1oct-3ol	2m-1butanol	2-heptanone	limonene
	[μg/dm^3]							
1	0	0.0151	0.0032	0.0092	0	0	0.0018	0.0015
3	0.0008	0.0090	0.0091	0.0064	0.0017	0	0.0005	0.1452
4	0	0.0100	0.0137	0.0154	0	0	0.0012	0.0202
5	0	0.0038	0.0074	0.0087	0	0.0417	0	0.0105
6	0.0004	0.0020	0.0023	0.0062	0.0005	0.0082	0	0.0019
7	0.0007	0.0072	0.0259	0.1809	0	0.0113	0	0.0180
8	0.0003	0.0044	0.0049	0.0123	0	0.0394	0	0.0077
9	0.0003	0.0028	0.0079	0.0161	0	0	0	0.0025
10	0.0016	0.0150	0.0154	0.0961	0.0032	0.0142	0.0062	0.0283
11	0.0004	0.0051	0.0044	0.1032	0	0	0.0010	0.0207
12	0.0002	0.0027	0.0019	0.0558	0	0	0	0.0139
13	0.0005	0.0075	0.0070	0.1505	0	0	0.0020	0.0274

Table 4

Descriptive statistics for individual compounds identified in all studies

Desc. stat.	cumin	styrene	ethyl-benzene	α-pinene	1oct-3ol	2m-1butanol	2-heptanone	limonene
	[μg/cm^3]							
Min	0.0000	0.0020	0.0019	0.0062	0.0000	0.0000	0.0000	0.0015
Max	0.0016	0.0151	0.0259	0.1809	0.0032	0.0417	0.0062	0.1452
Median	0.0004	0.0062	0.0072	0.0158	0.0000	0.0000	0.0003	0.0160
Q1	0.0002	0.0036	0.0041	0.0091	0.0000	0.0000	0.0000	0.0064
Q3	0.0006	0.0093	0.0103	0.0979	0.0001	0.0120	0.0014	0.0224
Mean	0.0004	0.0071	0.0086	0.0551	0.0005	0.0096	0.0011	0.0248
SD[*]	0.0004	0.0043	0.0066	0.0598	0.0010	0.0147	0.0017	0.0374

[*] *SD* - standard deviation

Because the selected substances belong to compounds of high volatility, the presented statistical parameters are in the low numerical range. The highest median value was characterised by limonene (0.0160 μg/cm^3) and) α-pinene (0.0158 μg/cm^3), while the lowest median values were characterised by 1-oct-3-ol and 2m-1butanol (0 μg/cm^3).

Data from MOS array

The most important issue from the point of view of the stated goal of this work was research using multi-sensor arrays, therefore, analyses of air quality were performed in real objects characterised by various degrees of microbial contamination using the multi-sensor

method and the chemical method as a reference. For all tested rooms, both infected and not biologically infected, Table 5 presents descriptive statistics of relative responses R_w (relative resistance according to formula (1)) of individual sensors together with substances to which they are selective, while Figure 3 illustrates the range of sensor variability using the 1-st and 3-rd quartile (Q1, Q3) and the median. The largest range of variability was found for sensors s5 (TGS-2620-C00), s13 (TGS-823-A00) and s16 (TGS-832-A00). According to Figaro's technical specifications, these are weakly selective sensors that respond to a number of different chemical substances, among which compounds from these groups are also found in the category of fungal metabolites.

Table 5

Descriptive statistics for relative responses R_w of particular MOS array sensors

Sensor	Type	Min [-]	Max [-]	Q1 [-]	Median [-]	Q3 [-]	Mean [-]	*SD* [-]
s1	TGS-2600-B00	0.391	0.758	0.484	0.556	0.596	0.545	0.083
s2	TGS-2602-B00	0.057	0.448	0.103	0.218	0.332	0.227	0.123
s3	TGS-2610-D00	0.355	0.675	0.415	0.480	0.553	0.481	0.086
s4	TGS-2611-E00	0.392	0.758	0.483	0.555	0.596	0.545	0.083
s5	TGS-2620-C00	0.057	0.695	0.306	0.460	0.579	0.435	0.174
s6	TGS-4161	0.057	0.117	0.076	0.080	0.085	0.081	0.011
s7	TGS-2444	0.057	0.456	0.096	0.205	0.323	0.214	0.122
s8	TGS-2442-B02	0.095	0.525	0.153	0.186	0.202	0.221	0.118
s9	TGS-800	0.194	0.468	0.336	0.400	0.425	0.381	0.065
s10	TGS-825-A00	0.057	0.419	0.119	0.183	0.287	0.206	0.106
s11	TGS-825-A00	0.057	0.467	0.096	0.219	0.327	0.223	0.122
s12	TGS-813-A00	0.095	0.603	0.159	0.178	0.405	0.262	0.150
s13	TGS-821	0.237	0.849	0.387	0.491	0.635	0.513	0.151
s14	TGS-823-A00	0.027	0.999	0.063	0.499	0.646	0.390	0.308
s15	TGS-812	0.277	0.670	0.488	0.541	0.585	0.523	0.083
s16	TGS-830	0.253	0.890	0.452	0.590	0.765	0.587	0.175
s17	TGS-832-A00	0.058	0.999	0.999	0.999	0.999	0.919	0.228

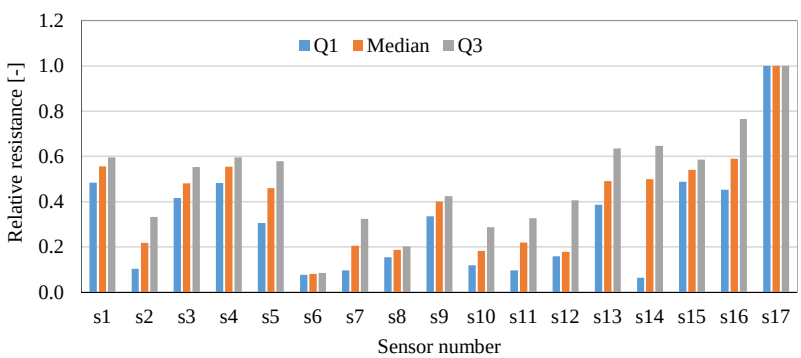

Fig. 3. Basic statistics of the relative response (R_w) of the individual MOS array sensors used in the research

The above presented Figure shows that sensors s6 (TGS-4161 for detecting carbon dioxide) and s17 (TGS-2106 for detecting exhaust gases, waste gases from diesel engines) do not respond to the tested air sample. The changes in the carbon dioxide level are mainly influenced by the number of people in the room as well as the number of plants and indirectly by the temperature (higher breathing frequency). The state of mould in the rooms was stabilised over a longer period of time, so it did not affect the response of the sensors during the 15-minute measurement. No presence of exhaust gases or waste gases was detected either because there were no heating boilers installed in the rooms and the rooms were not located in the immediate vicinity of the streets. This may also explain the poor response of the *S*8 sensor dedicated mainly to detecting carbon monoxide(II), which is mainly a product of the incomplete combustion process.

Discussion

Multiple regression model without dimensionality reduction

By using multiple regression without dimensionality reduction, deterministic models were developed that allow for estimating the concentrations of individual substances using a mathematical equation described by the general formula (2). In order to present the predictive capabilities of the array and this method of multidimensional signal analysis, Table 6 presents the model fit measures on the training set and the test set. They are basic statistics that can be used to assess the quality of the estimation of concentrations of individual substances. They include the coefficient of determination (R^2), the root mean square error (*RMSE*) and the mean absolute error (*MAE*).

Table 6

Goodness of fit measures of multiple regression models without dimensionality reduction on training and test sets

Substance	Training set			Test set		
	R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*
-	-	[µg/cm^3]	[µg/cm^3]	-	[µg/cm^3]	[µg/cm^3]
Cumin	0.850	0.000	0.000	0.807	0.000	0.000
Styrene	0.836	0.002	0.001	0.770	0.002	0.002
Ethylbenzene	0.793	0.003	0.002	0.775	0.003	0.002
α-pinene	0.717	0.030	0.024	0.690	0.028	0.024
1oct-3ol	0.442	0.043	0.026	0.574	0.051	0.033
2m-1butanol	0.478	0.011	0.008	0.491	0.013	0.009
2-heptanone	0.951	0.000	0.000	0.921	0.000	0.000
Limonene	0.588	0.025	0.018	0.479	0.027	0.020
Average value	0.707	0.014	0.010	0.688	0.016	0.011

To assess the model fit, the basic parameters used were the coefficients of determination (R^2), the average values of which are similar to each other, but in the case of most of the tested substances they show higher values for the training set (average value of $R^2 = 0.71$ for the training set) than for the test set (average value of $R^2 = 0.69$) (exceptions are 1oct-3ol and 2m-1butanol). On the other hand, the average *RMSE* values obtained on the test set (0.016 µg/m^3) are slightly higher than those corresponding to the training set (0.014 µg/m^3), which is in line with expectations and, similarly to regression models with dimensionality reduction, indicates the absence of the so-called "model overfitting"

phenomenon, i.e. a situation in which the fit measures on the training set are much better than on the test set.

AI models

Similarly to multiple regression models basic statistics that can be used to assess the quality of the estimation (R^2, *RMSE* and *MAE*) are presented in Tables 7 (SVM models) and 8 (MLP models).

Table 7

Fitting SVM models on training and test sets

Substance	Training set			Test set		
	R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*
Unit	-	[µg/cm^3]	[µg/cm^3]	-	[µg/cm^3]	[µg/cm^3]
Cumin	0.994	0.000	0.000	0.976	0.000	0.000
Styrene	0.996	0.000	0.000	0.986	0.001	0.001
Ethylbenzene	0.995	0.001	0.001	0.99	0.001	0.001
α-pinene	0.997	0.005	0.005	0.986	0.008	0.006
1oct-3ol	0.999	0.006	0.005	0.997	0.007	0.006
2m-1butanol	0.997	0.001	0.001	0.981	0.003	0.002
2-heptanone	0.994	0.000	0.000	0.99	0.000	0.000
Limonene	0.996	0.003	0.003	0.897	0.012	0.005
Average value	0.996	0.002	0.002	0.975	0.004	0.003

Models based on artificial intelligence allow for obtaining even greater fit coefficients to the analysed data, as evidenced by the high value of the R^2 coefficient, which is on average 0.996 and 0.976 in the case of SVM analysis and 0.901 and 0.876 in the case of ANN analysis, respectively for the training and test sets in both methods. SVM models are much better fitted to the data, which results from the fact that they allow for the estimation of nonlinear dependencies.

Table 8

Fitting MLP models on training and test sets

Substance	Training set			Test set		
	R^2	*RMSE*	*MAE*	R^2	*RMSE*	*MAE*
Unit	-	[µg/cm^3]	[µg/cm^3]	-	[µg/cm^3]	[µg/cm^3]
Cumin	0.083	0.001	0.001	0.08	0.001	0.001
Styrene	0.864	0.002	0.001	0.822	0.002	0.002
Ethylbenzene	0.932	0.002	0.001	0.919	0.002	0.001
α-pinene	0.998	0.002	0.002	0.989	0.005	0.004
1oct-3ol	0.864	0.002	0.001	0.822	0.002	0.002
2m-1butanol	0.578	0.001	0.001	0.457	0.002	0.001
2-heptanone	0.993	0.005	0.003	0.988	0.009	0.006
Limonene	0.995	0.003	0.002	0.985	0.005	0.003
Average value	0.788	0.002	0.002	0.758	0.004	0.003

The results of the study demonstrate the efficacy of using a multi-sensor array equipped with MOS sensors to estimate the concentrations of volatile organic compounds (mVOC) of fungal origin in indoor environments. These findings underline the potential of combining traditional chemical analysis techniques with advanced statistical and machine learning models for environmental monitoring.

The gas chromatography-mass spectrometry (GC-MS) analysis provided a reliable reference for compound concentrations. This method allowed for precise quantification of selected mVOC compounds, offering a robust benchmark for validating the predictive models. However, the labour-intensive and time-consuming nature of GC-MS analysis limits its practicality for continuous or widespread monitoring. The study highlights the importance of developing alternative methods, such as sensor-based systems, to address these limitations.

In order to quantitatively compare individual methods of signal analysis from gas sensor arrays, the average values of the obtained statistics were compared. It was found that methods based on linear multiple regression allow for obtaining average values of the coefficient of determination R^2 at the level of 0.707 for the training set and 0.688 for the test set. Artificial intelligence methods allowed for obtaining these parameters of 0.788 and 0.758, respectively, in the case of MLP algorithms and 0.996 and 0.975 in the case of SVM algorithms. This indicates that artificial intelligence models best reflect the examined relationships. This is confirmed by the *RMSE* and *MAE* errors, which in the case of multiple regression for the training and test sets are (*RMSE*) 0.014 µg/cm^3 and 0.016 µg/cm^3, respectively, while in the case of both machine learning techniques these values are only 0.002 µg/cm^3 and 0.004 µg/cm^3. In the case of *MAE* errors, these values are 0.01 and 0.011 µg/cm^3 for multiple regression and 0.002 µg/cm^3 and 0.003 µg/cm^3 for both machine learning methods, respectively.

The obtained measurement errors of both methods of estimating concentrations using artificial intelligence methods are also very promising when referring them to the threshold values of individual substances. Although in the case of 1oct-3ol threshold concentration is 0.00524 µg/cm^3 [34, 35] and the obtained errors for the SVM model are 0.006 µg/cm^3 (*RMSE*) and 0.005 µg/cm^3 (*MAE*) and are greater than this limit value, in the case of MLP models these errors are much lower and equal to 0.002 µg/cm^3 and 0.001 µg/cm^3, respectively, which confirms their effectiveness in estimating the concentrations of this substance. As for 2m-butanol, the acceptable concentration of which is 0.0025 µg/cm^3 [36], it should be noted that the *RMSE* and *MAE* errors of both data analysis methods are 0.001 µg/cm^3 and are smaller than the threshold value. As for 2-heptanone, with an acceptable concentration of 0.00467 µg/cm^3 [34, 35] the estimation errors depend on the algorithm used, in the case of using the SVM network they are smaller than the adopted method resolution, while in the case of the MLP method they are less favourable and close to this threshold concentration. Finally, in the case of limonene with acceptable concentration of 0.0025 µg/cm^3 [37], the *RMSE* and *MAE* estimation errors using the SVM method are equal to 0.003 µg/cm^3 and are slightly higher than the threshold value, similarly to the case of the MLP method, where they are 0.003 µg/cm^3 and 0.002 µg/cm^3, respectively, which means that they are slightly smaller than the maximum acceptable values and are sufficiently effective in estimating the concentrations of this substance in the air. In the case of traditional regression model, the estimation error values were several times higher than the threshold values (except for 2-heptanone), which means that they are not effective in estimating the concentrations of these substances.

Linear regression models, while straightforward and interpretable, showed moderate predictive capabilities. The coefficients of determination (R^2) and error metrics (*RMSE*, *MAE*) indicated that these models captured the overall trends in sensor responses but struggled with more complex, non-linear dependencies. The simplicity of linear regression

makes it a valuable baseline method, but its limited ability to account for intricate relationships underscores the need for more sophisticated approaches.

The application of artificial intelligence models - Multi-Layer Perceptrons (MLP) and Support Vector Machines (SVM) - yielded significantly improved results compared to linear regression. MLP models demonstrated their strength in learning complex, non-linear patterns within the data. By leveraging a hidden layer with non-linear activation functions, the MLP models effectively captured the relationships between the 17 sensor responses and the concentrations of the chemical compounds. The model specialisation for individual compounds further enhanced prediction accuracy. Similarly, the SVM models exhibited high predictive performance, particularly due to their use of kernel functions, such as the radial basis function (RBF), which enabled the capture of non-linear relationships in a transformed feature space. SVM models were particularly robust against overfitting and performed consistently across various compounds. Their focus on support vectors - critical data points in the dataset - enhanced computational efficiency while maintaining accuracy.

The comparison between traditional linear models and AI-based approaches highlights the advantages of modern machine learning methods in handling complex datasets. The SVM and MLP models not only outperformed linear regression in terms of predictive accuracy but also provided greater flexibility in modelling non-linear relationships. This is particularly important for the application of sensor arrays, where sensor responses may interact in complex ways due to the presence of multiple compounds and environmental factors. The results also demonstrate the potential of electronic noses equipped with MOS sensors as a practical tool for rapid and non-invasive monitoring of indoor air quality. By incorporating AI models, these systems can achieve near-real-time estimation of mVOC concentrations, offering a valuable alternative to traditional chemical analysis methods. However, the study also emphasises the need for continuous calibration and validation of sensor-based systems to account for sensor drift and environmental variability over time.

While the AI models provided excellent results, their performance is contingent on the quality and diversity of the training data. The study relied on a dataset collected under specific conditions, and the generalisability of the models to different environments or sensor configurations remains an area for further research. Additionally, the computational requirements of training AI models, particularly MLPs, may pose challenges for deployment in resource-constrained settings.

Future work could focus on enhancing the robustness and scalability of the sensor-based system. This may include integrating adaptive algorithms to compensate for sensor drift, expanding the dataset to include a wider range of environmental conditions, and exploring hybrid approaches that combine the strengths of linear and AI models. Furthermore, the inclusion of additional sensors or novel materials with higher specificity could improve the detection and quantification of mVOC compounds.

Conclusion

In conclusion, the study demonstrates that combining MOS sensor arrays with AI-based models can provide a reliable and efficient solution for monitoring microbial contamination in indoor environments. This approach bridges the gap between the precision of chemical analysis and the practicality of real-time monitoring, paving the way for more accessible and scalable air quality assessment systems.

The use of artificial neural networks allows for precise and quantitative determination of factors negatively affecting indoor air quality with an acceptable level of accuracy. This is evidenced by the average values of the coefficient of determination R^2 of models estimating the concentrations of fungal metabolites equal to 0.996 in the case of the SVM network and 0.788 in the case of the MLP network, and *RMSE* and *MAE* errors equal to 0.002 µg/cm^3 in both the *MAE* network and the SVM method.

The *RMSE* and *MAE* estimation errors obtained using machine learning models (mainly MLP), unlike deterministic model, are in most cases smaller than the threshold concentration values of these substances in the air, which means that these methods are effective in detecting their presence and thus predicting mycological contamination of these rooms.

The conducted research has shown that it is possible to estimate fungal markers in the indoor air of buildings in order to quickly detect the threat of microbial contamination; however, due to the nature of the substances analysed and the very large impact of external factors, additional research is required to further develop methods for the rapid detection of biological contamination of buildings.

Acknowledgements

This work was financially supported within the authors' scientific units under subvention for a Scientific Disciplines program.

References

[1] Sánchez Espinosa KC, Rodríguez Davydenko S, Rojas Flores TR, Fernández-González M, Almaguer M. Xerophilic and cellulolytic fungi in the indoor air of houses in Havana. Int Biodeter Biodegr. 2024;188:105730. DOI: 10.1016/j.ibiod.2024.105730.

[2] Kuske M, Rubio R, Romain AC, Nicolas J, Marco S. Fuzzy k-NN applied to moulds detection. Sens Actuator B Chem. 2005;106:52-60. DOI: 10.1016/j.snb.2004.05.066.

[3] Heung LJ, Wiesner DL, Wang K, Rivera A, Hohl TM. Immunity to fungi in the lung. Semin Immunol. 2023;66:101728. DOI: 10.1016/j.smim.2023.101728.

[4] Peccia J, Kwan SE. Buildings, beneficial microbes, and health. Trends Microbiol. 2016;24(8):595-7. DOI: 10.1016/j.tim.2016.04.007.

[5] Fathy RM, Daigham GE, Mahfouz AY. High-performance carboxymethyl cellulose stabilized Au-Ag nanoparticles using gamma rays for management of the airborne fungus Aspergillus sydowii AGH-5 OP593090. Process Biochem. 2025;152:58-74. DOI: 10.1016/j.procbio.2025.02.014.

[6] Moslem MA, Yassin MA, El-Samawaty AER, Sayed SRM. New toxigenic Penicillium species associated with apple blue mold in Saudi Arabia. Fresenius Environ Bull. 2011;20(12):3194-8. Available from: https://www.scopus.com/record/display.uri?eid=2-s2.0-84856304709&origin=recordpage.

[7] Isaksson T, Thelandersson S, Ekstrand-Tobin A, Johansson P. Critical conditions for onset of mould growth under varying climate conditions. Build Environ. 2010;45:1712-21. DOI: 10.1016/j.buildenv.2010.01.023.

[8] Sessa R, Di Pietro M, Schia Voni G, Santino I, Atieri A, Pinelli S, et al. Microbiological indoor air quality in healthy buildings. New Microbiol. 2002;25:51-6. Available from: https://pubmed.ncbi.nlm.nih.gov/11837391/.

[9] Chen X, Li F, Liu C, Yang J, Zhang J, Peng C. Monitoring, human health risk assessment and optimized management for typical pollutants in indoor air from random families of university staff, Wuhan City, China. Sustainability. 2017;9:1115. DOI: 10.3390/su9071115.

[10] Schenkel D, Lemfack MC, Piechulla B, Splivallo R. A meta-analysis approach for assessing the diversity and specificity of belowground root and microbial volatiles. Front Plant Sci. 2015;6:707. DOI: 10.3389/fpls.2015.00707.

[11] Gallego E, Roca FJJ, Perales JFF, Sánchez G, Esplugas P. Characterization and determination of the odorous charge in the indoor air of a waste treatment facility through the evaluation of volatile organic compounds (VOCs) using TD-GC/MS. Waste Manage. 2012;32:2469-81. DOI: 10.1016/j.wasman.2012.07.010.

[12] Oliveira IS, Galdino da Silva A Jr, Souza de Andrade CA, Oliveira MDL. Biosensors for early detection of fungi spoilage and toxigenic and mycotoxins in food. Curr Opin Food Sci. 2019;29:64-79. DOI: 10.1016/j.cofs.2019.08.004.
[13] Gallo A, Catellani A, Ghilardelli F, Lapris M, Mastroeni C. Review: Strategies and technologies in preventing regulated and emerging mycotoxin co-contamination in forage for safeguarding ruminant health. Animal. 2024;18(suppl. 2):101280. DOI: 10.1016/j.animal.2024.101280.
[14] Eggleston PA, Bush RK. Environmental allergen avoidance: An overview. J Allergy Clin Immunol. 2001;107:403-5. DOI: 10.1067/mai.2001.113673.
[15] Zhao Y, Chen D, Duan H, Li P, Wu W, Wang X, et al. Sample preparation and mass spectrometry for determining mycotoxins, hazardous fungi, and their metabolites in the environment, food, and healthcare. TrAC Trends Analyt Chem. 2023;160:116962. DOI: 10.1016/j.trac.2023.116962.
[16] Wilkins K, Larsen K. Variation of volatile organ ic compound patterns of mold species from damp buildings. Chemosphere. 1995;31(5):3225-36. DOI: 10.1016/0045-6535(95)00184-A.
[17] Sham NM, Ahmad NI, Pahrol MA, Leong YH. Fungus and mycotoxins studies in hospital environment: A scoping review. Build Environ. 2021;193:107626. DOI: 10.1016/j.buildenv.2021.107626.
[18] Przystaś W, Zabłocka-Godlewska E, Nelaniuk-Wolny E. A comparison of sedimentation method and active sampler analysis of microbiological indoor air quality - case study. Ecol Chem Eng S. 2023;30(1):37-48. DOI: 10.2478/eces-2023-0009.
[19] Mendell MJ, Adams RI. The challenge for microbial measurements in buildings. Indoor Air. 2019;29(4):523-6. DOI: 10.1111/ina.12550.
[20] Sadowska A, Nynca A, Ruszkowska M, Paukszto L, Myszczynski K, Orlowska K, et al. Transcriptional profiling of porcine granulosa cells exposed to 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin. Chemosphere. 2017;178:368-77. DOI: 10.1016/j.chemosphere.2017.03.055.
[21] Gutarowska B. Metabolic activity of moulds as a factor of building materials biodegradation. Pol J Microbiol. 2010;59:119-24. Available from: https://pubmed.ncbi.nlm.nih.gov/20734757/.
[22] Verdier T, Coutand M, Bertron A, Roques C. A review of indoor microbial growth across building materials and sampling and analysis methods. Build Environ. 2014;80:136-49. DOI:10.1016/j.buildenv.2014.05.030.
[23] Wei GL, Zeng EY. Gas chromatography-mass spectrometry and high-performance liquid chromatography-tandem mass spectrometry in quantifying fatty acids. TrAC Trends Anal Chem. 2011;30:1429-36. DOI: 10.1016/j.trac.2011.05.005.
[24] Fornal E, Parfieniuk E, Czeczko R, Bilinska-Wielgus N, Frac M. Fast and easy liquid chromatography-mass spectrometry method for evaluation of postharvest fruit safety by determination of mycotoxins: Fumitremorgin C and verruculogen. Postharvest Biol Technol. 2017;131:46-54. DOI: 10.1016/j.postharvbio.2017.05.004.
[25] Guz L, Sobczuk H, Suchorab Z. Odor measurement by using a portable device with semiconductor gas sensors array. Przem Chem. 2010;89(4):378-81. Available from: https://www.scopus.com/record/display.uri?eid=2-s2.0-77953499748&origin=recordpage.
[26] Bonah E, Huang X, Aheto JH, Osae R. Application of electronic nose as a non-invasive technique for odor fingerprinting and detection of bacterial foodborne pathogens: A review. J Food Sci Technol. 2020;57:1977-90. DOI: 10.1007/s13197-019-04143-4.
[27] Zhai Z, Liu Y, Li C, Wang D, Wu H. Electronic noses: From gas-sensitive components and practical applications to data processing. Sensors. 2024;24(15):4806. DOI: 10.3390/s24154806.
[28] Suchorab Z, Frąc M, Guz Ł, Oszust K, Łagód G, Gryta A, et al. A method for early detection and identification of fungal contamination of building materials using e-nose. PLoS One. 2019;14(4):e0215179. DOI: 10.1371/journal.pone.0215179.
[29] Figaro USA, Inc. MOS gas sensors technical information. Available from: https://www.figarosensor.com.
[30] Gutierrez-Osuna R. Pattern analysis for machine olfaction: A review. IEEE Sens J. 2002;2:189-202. DOI: 10.1109/JSEN.2002.800688.
[31] Distante C, Leo M, Siciliano P, Persaud KC. On the study of feature extraction methods for an electronic nose. Sens Actuator B Chem. 2002;87:274-88. DOI: 10.1016/S0925-4005(02)00247-2.
[32] Sawoszczuk T, Syguła-Cholewińska J, del Hoyo-Meléndez JM. Optimization of headspace solid phase microextraction for the analysis of microbial volatile organic compounds emitted by fungi: Application to historical objects. J Chromatogr A. 2015;1409:30-45. DOI: 10.1016/j.chroma.2015.07.059.
[33] Polizzi V, Adams A, Malysheva SV, De Saeger S, Van Peteghem C, Moretti A, et al. Identification of volatile markers for indoor fungal growth and chemotaxonomic classification of Aspergillus species. Fungal Biol. 2012;116(9):941-53. DOI: 10.1016/j.funbio.2012.06.001.
[34] Schleibinger H, Laussmann D, Bornehag CG, Eis D, Rueden H. Microbial volatile organic compounds in the air of moldy and mold-free indoor environments. Indoor Air. 2008;18(2):113-24. DOI: 10.1111/j.1600-0668.2007.00513.x. PMID: 18333991.

[35] NIOSH, National Institute for Occupational Safety and Health, 2006.
[36] National Academies of Sciences, Engineering, and Medicine, Spacecraft Maximum Allowable Concentrations for Selected Airborne Contaminants: Volume 3. The National Academies Press, Washington, DC 1996, DOI: 10.17226/5435.
[37] Nielsen E, Ladefoged O, Søborg I. Evaluation of health hazards by exposure to d-Limonene and proposal of a health-based quality criterion for ambient air. Division of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, 2013. ISBN: 9788793026339.