Metrol. Meas. Syst., Vol. 27 (2020) No. 1, pp. 167–179
DOI: 10.24425/mms.2020.131716

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.pg.gda.pl

EVALUATION OF THE NO_2 CONCENTRATION PREDICTION POSSIBILITY BASED ON STATIC AND DYNAMIC RESPONSES OF TGS SENSORS AT CHANGING HUMIDITY LEVELS

Paweł Kalinowski, Łukasz Woźniak, Grzegorz Jasiński, Piotr Jasiński

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, G. Narutowicza 11/12, 80-233 Gdańsk, Poland (✉ *pawkalin@pg.edu.pl*, +48 58 347 2509, *lukwozni@pg.edu.pl, grzegorz.jasinski@eti.pg.edu.pl, piotr.jasinski@eti.pg.edu.pl*)

Abstract

The commercially available metal-oxide TGS sensors are widely used in many applications due to the fact that they are inexpensive and considered to be reliable. However, they are partially selective and their responses are influenced by various factors, *e.g.* temperature or humidity level. Therefore, it is important to design a proper analysis system of the sensor responses. In this paper, the results of examinations of eight commercial TGS sensors combined in an array and measured over a period of a few months for the purpose of prediction of nitrogen dioxide concentration are presented. The measurements were performed at different relative humidity levels. PLS regression was employed as a method of quantitative analysis of the obtained sensor responses. The results of NO_2 concentration prediction based on static and dynamic responses of sensors are compared. It is demonstrated that it is possible to predict the nitrogen dioxide concentration despite the influence of humidity.

Keywords: TGS sensors, partial least squares, nitrogen dioxide, dynamic measurements.

1. Introduction

Electronic devices which are able to detect gases, commonly known in the literature as electronic noses, have found a wide range of potential applications. For example, in medicine to the detection of diseases [1], in industry to monitoring quality of food or drinks [2] or in environmental monitoring [3–5]. Their advantages, *e.g.* the possibility of working in the real-time operation mode, compact size or low cost, resulted in electronic gas-analysing systems becoming an attractive alternative to other gas-analysing systems, like gas chromatographs. However, these devices still have their limitations, which need to be overcome.

A typical electronic gas-analysing device consists of four main elements, namely, a gas delivery subsystem, an array of gas sensors, data acquisition and power supply circuits and data analysis software. Such systems usually employ feature vectors created based on the measured

responses, which are given as inputs for pattern recognition algorithms. There are reports, *e.g.* [6–8], in which the features associated with the power supply system of sensors are used. This technique is called temperature modulation. There are also reports in which the features for gas recognition are associated with the data acquisition system, *e.g.* fluctuation-enhanced gas sensing [9, 10] or electrochemical measurement techniques [11]. Another approach used for obtaining additional information about the type and concentration of a measured gas uses a gas-delivery subsystem and is associated with the changes of gas flow rate in the measurement chamber. The application of different gas flow patterns enables to obtain unique shapes of the sensor responses which provide additional features for the recognition algorithms [12].

The main drawback of electronic gas-analysing systems is associated with the properties of gas sensors [13]. For example, gas sensors tend to drift [14]. The lack of long-term stability is one of the crucial problems and impose the necessity of frequent recalibration of the gas-recognition system. The influence of humidity is one of the sources of this instability and the methods used for its mitigation are the topic of a number of reports [15–19]. Due to the fact that TGS sensors are sensitive to many factors, the information obtained from static measurements can be insufficient for reliable gas detection. There are reported several techniques of dynamic measurements of gas sensors that are used to improve the possibility of gas detection [20, 21]. In general, they are based on changing the measurement conditions by changing the gas flow rate in the measurement chamber. One of the procedures of dynamic measurements of a sensor array is called the stop-flow mode of operation [22]. It consist of three stages. During the first stage the target gas is delivered to the measurement cell. In the next phase the gas flow is stopped. Finally, in the last stage of measurement, the cell is cleaned with a reference gas, *e.g.* synthetic air. This technique also enables to extract additional information hidden in the shape of the obtained acquisition. In the study presented here, all measurements were performed using the stop-flow mode of sensor operation.

There is a need for reports in which the gas-recognizing system performance is examined over a longer period of time. In this study, we demonstrate the results of evaluation of the responses of eight commercially available sensors combined in an array to predict the nitrogen dioxide concentration measured at different humidity levels. The dynamic measurements performed using the stop-flow technique had been performed for three months. The data obtained during the initial successive weeks were used for system calibration and the data collected during further measurements were used for validation. *Partial Least Squares* regression (PLS) [23] was chosen as a method of creating a multivariate calibration model. Four approaches to the preparation of regression models were compared. Namely, the regression models were calculated using single-value resistances of the sensors, using raw and normalized acquisitions obtained during dynamic measurements and using the features extracted from acquisitions with and without pre-processing.

2. Experimental

2.1. Measurement stand

The measurements were carried out on a sensor array consisting of two measurement cells (denoted as #1 and #2) connected in series with, in total, eight commercially available semiconductor TGS sensors (four sensors in each cell) and two temperature sensors (one for each cell). A scheme of the measurement stand is presented in Fig. 1 and a detailed description of the measurement cells can be found in [24]. Temperature inside cell #2 was adjusted by a heating cable wrapped around the measurement chamber #2, controlled by a Fuji PXR4 temperature con-

Metrol. Meas. Syst., Vol. 27 (2020), No. 1, pp. 167–179
DOI: 10.24425/mms.2020.131716

troller. This made it possible to perform measurements at a constant temperature of the chamber. Temperatures inside measurement chambers #1 and #2 (stabilized with the heating cable) were equal to (43 ± 2)°C and 47°C, respectively. The sensors used in the array along with a list of their catalogue data are shown in Table 1. The sensor heaters were supplied with a constant voltage of 5 V or 6.2 V (TGS 2106) using HAMEG 8040 power supply. The responses of sensors were automatically measured using a Keithley 2700 multichannel multi-meter and software designed for data acquisition. The resistance was measured using the 4-probe method. The gas-delivery system consisted of three Brooks GF Series mass flow controllers connected to the PC with an RS 485 interface. For measurements at changing relative humidity levels, one of the controllers was connected to an air bubbler. The gas profiles were programmed with MEDSON software. The total gas flow in the measurement cell was constant and set to 100 cm^3min^{-1}. Two humidity sensors (SY-HS-220) were used for monitoring the humidity level. One of the humidity sensors was placed at the inlet of the measurement chamber and the second was placed at the outlet of the measurement chamber to ensure obtaining a required humidity level. For further analysis, only the readings from the sensor placed at the inlet were selected.

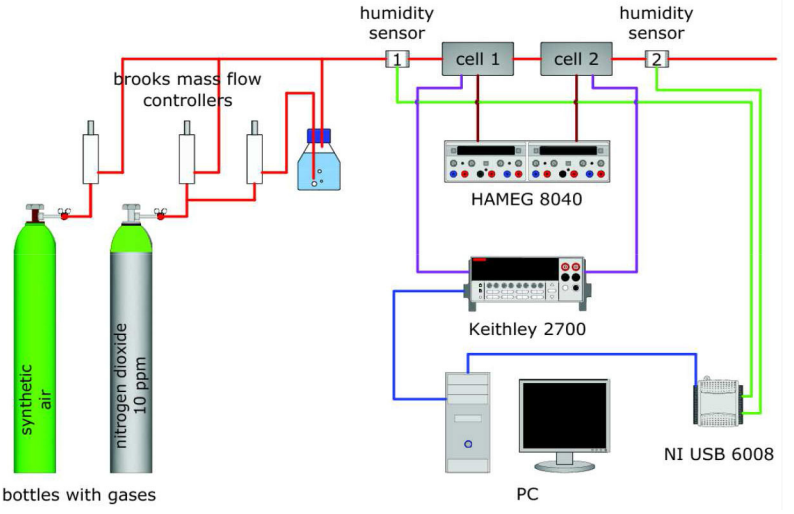

Fig. 1. The measurement stand.

Table 1. Composition of an array.

Number	Sensors	Target gases	Concentrations of target gases*
S_1 S_2	TGS 826 #1 TTGS 826 #2	NH_3	> 5 ppm
S_3 S_4	TGS 825 #1 TGS 825 #2	H_2S	> 3 ppm
S_5 S_6	TGS 2106 #1 TTGS 2106#2	NO_2	0.1~10 ppm
S_7	TGS 880 #2	Air pollutants	–
S_8	TGS 2600 #1	Air pollutants	1~30 ppm H_2

*according to the catalogue data of sensors

3. Data analysis

3.1. Measurement procedure for dynamic measurements

The air pollutant selected for the measurements as a target gas was nitrogen dioxide. Despite the fact that only one type of sensor (TGS 2106) was designed for the detection of this compound, all used sensors changed their output resistance in the presence of the target gas due to their partial selectivity. As the reference gas, synthetic air was used. All gases were provided by Linde Gaz. Concentrations of nitrogen dioxide for the measurements were equal to 25 ppm, 50 ppm, 75 ppm and 100 ppm. All nitrogen dioxide concentrations were measured in dry air and at 10%, 20%, 30%, 40% and 50% relative humidity levels. Before performing the dynamic measurements, the sensor responses were measured at all NO_2 concentrations, all relative humidity levels, as well as different flow rates to characterize TGS sensors. One cycle of the stop-flow procedure took 1 hour. The exposure to the target gas had lasted for 7 min, then the gas flow had been stopped for another 7 min. For the remaining 46 minutes the clean air had flowed through the measurement chambers. Originally [22], all stages in the stop-flow mode of operation lasted 7 minutes, but, for the purpose of cleaning, the air-on stage was extended. The resistance values were measured every 30 seconds, thus a single acquisition contained 120 values. All measurements taken for further analyses were acquired with a total flow rate of 100 $cm^3 min^{-1}$.

The measurements were carried out over a period of three months. The most frequently occurring data for calibration and validation sets were selected randomly from the available dataset. However, this work was intended to demonstrate the behaviour of sensors over a longer period of time and to examine stability of the calibrated system, which is closer to an electronic gas-recognition system working in real conditions. Therefore, the acquisitions obtained in the first successive weeks were composed into a calibration set, while the remaining ones were used as a validation set. In total, 79 acquisitions for each sensor were obtained. The initial 51 acquisitions obtained in the first two months were treated as a calibration set. The remaining 28 ones were used to validate the performance of sensors. The obtained multivariate set was not large enough to use popular machine learning techniques, such as Artificial Neural Networks or Support Vector Machine. Therefore, the statistical linear regression method PLS was selected, since it does not require a large calibration dataset [25].

There were two main goals of the analysis performed using the dynamic responses of TGS sensors presented in this work. The first was to find the optimal calibration model for prediction of the nitrogen dioxide concentration. For this purpose the Partial Least Squares Regression method was employed and tested with various feature vectors obtained from sensor responses. The results were compared with a PLS model calibrated using only single-value resistances of eight sensors, as well as with a PLS model obtained from all points of acquisitions from dynamic measurements. Such values were derived from the dynamic acquisitions as the resistance response of sensors after seven minutes of exposure to the target gas. The second goal was to create a calibration model which was insensitive to changes in humidity levels. This goal was achieved by using techniques of pre-processing data obtained from the dynamic measurements of sensors and feature extraction.

3.2. Feature extraction and data pre-processing techniques

The analysis of the dynamic responses of TGS sensors was based on an assumption that the shape of acquisition obtained during the dynamic measurements contained information about the concentration of the measured target gas. By calculating the number of parameters (features) it is possible to preserve this information about gas concentration and reduce the volume of

Metrol. Meas. Syst., Vol. 27 (2020), No. 1, pp. 167–179
DOI: 10.24425/mms.2020.131716

data used as the input vector for the calibration model. Additionally, through appropriate data processing, such as normalization of the acquisitions and monitoring of the current humidity level, drift caused by the influence of humidity can be compensated. In this work, the features for the calibration set were calculated based on both – normalized and raw acquisitions. The acquisitions were normalized according to [26].

Various features can be extracted from the response curves obtained during dynamic measurements [21]. Some of them are taken directly from the acquisition, *e.g.* the minimum or maximum value of sensor response, based on which amplitude or relative amplitude can be calculated. Some features are connected with the dynamic properties of sensors, *i.e.* quasi time-constants (a selected part of the initial exposure to the target gas), or the features which represent the total information about the reaction on the sensing layer of sensors, *i.e.* the surface under the measured curve. Some features which can be calculated, such as kurtosis or skewness, are measures which describe the shape of the measured curve. Skewness is the indicator of the curve asymmetry, while kurtosis describes its flatness. These parameters are usually used in the context of statistical distribution shapes, but in this paper we use them as features for gas concentration recognition. They were calculated from each obtained dynamic acquisition. A total of five features were selected to prepare the feature vectors, according to [21], namely: amplitude (A), surface under the curve (S), skewness (Sk), kurtosis (K) and full width at half maximum ($T_{1/2}$). All of the features described were calculated automatically using software written by the authors in Matlab. Matlab was also used in the preparation of the calibration model.

3.3. Partial Least Squares

The PLS method was used to create a multivariate calibration model. This linear supervised method requires two datasets – a matrix of measured samples **X** and a corresponding matrix of responses **Y**. The PLS method models a relationship between these two matrices. In some cases, instead of a matrix of dependent variables, a vector of one dependent variable is considered. In this work, there is used a vector of dependent variables **Y** which contains the values of NO_2 concentrations. The method determines a set of latent variables (LVs) in such a way that the LVs explain both the variance of **X** (*i.e.* a set of feature vectors described above or a set of vectors of sensor resistances) as well as the correlation with **Y**. As a result, a matrix of predictors is obtained, which estimates the gas concentration. Another advantage of using this method is the possibility of achieving a reduction in data volume since only in the first few latent variables there is included most information from the original data.

3.4. Analysis scenario and model verification

In Fig. 2a flowchart of the performed analysis is presented. There are four prepared sets of vectors: V_1 composed of eight values of TGS single-value resistance responses and information about the measured RH level; V_2 – composed of the acquisitions from all sensors (120 measurement points ×8 sensors + information about the measured RH level); V_3 – composed of the normalized acquisitions from all sensors (120 measurement points ×8 sensors + information about the measured RH level); V_4 – containing 41 elements (5 features ×8 TGS sensors + information about the measured RH level). The features in this case were calculated based on raw acquisitions, without normalization. The last prepared set of feature vectors; V_5 – contains 41 elements (5 features ×8 TGS sensors + information about RH levels). The features were calculated based on normalized acquisitions. The calibration sets contain 51 observations collected in different NO_2 concentrations and RH levels over a few weeks. Thus, dimensions of the resulting

multivariate calibration sets were equal to 51 (number of observations) ×8 for vector V_1, 51 ×921 for vectors V_2 and V_3, 51 ×41 for vector V_4 and 51 ×41 for vector V_5. The validation sets contain 28 observations for each set of prepared vectors collected within a few weeks after the calibration data were obtained.

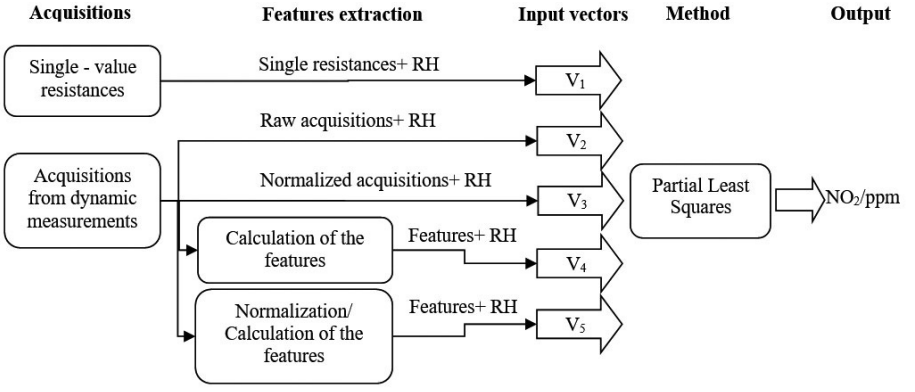

Fig. 2. A flowchart of analysis scenario.

In order to analyse the significance of sensors and features, the length of the feature vectors was reduced by omitting responses of selected sensors or calculated features from the input vector. Namely, the responses of single consecutive sensors and their combinations were iteratively excluded from the dataset. In each case, the calibration was performed (including optimization of the parameters of PLS method by selecting the number of LVs, which provide the lowest value of prediction error) and the prediction error was calculated based on the validation data. The excluded sensor or combination of sensors was rejected permanently if the value of obtained *Root Mean Squared Error of prediction* (RMSEp) increased or did not change. In the other case, the sensor responses were preserved. Only the combination of sensor responses which gave the lowest prediction error was preserved. For input vectors V_4 and V_5 further reduction of dimension of the datasets were performed by omitting the features or their combinations in a similar way that in the case of eliminating a sensor. The optimal number of LVs used for preparation of the final prediction models was selected independently for each analysis scheme, based on the observation of the value of prediction errors. In each case, the value of variance in data explained by the latent variables was higher than 96% .

4. Results and discussion

4.1. TGS sensor responses to various NO_2 concentrations and changes in relative humidity level

In Fig. 3a, the normalized responses of seven TGS sensors to 100 ppm NO_2 concentrations for all measured relative humidity levels are presented. The response of the sixth sensor, namely TGS2106#2 (S_6) is not presented in the figure. During validation measurements it was noticed that this sensor was damaged, thus it was rejected from further analyses. The normalized responses N_{NO_2} shown in Fig. 3a were calculated by the subtraction of the response in the presence of synthetic air (the baseline) from the sensor resistance response at 100 ppm of NO_2, and dividing the difference by the baseline. It can be interpreted as a measure of sensor sensitivity that occurs

Metrol. Meas. Syst., Vol. 27 (2020), No. 1, pp. 167–179
DOI: 10.24425/amms.2020.131716

when sensors are exposed to the presence of NO_2. It can be seen that the normalized response N_{NO_2} differs not only among sensors, but also varies for each measured relative humidity level. Some of sensors, *e.g.* S_4 and S_7, are characterized with a lower variability of response, while *e.g.* the changes of responses of sensors S_3 or S_8 are significantly higher. It can be seen that sensors belonging to the same type (*e.g.* S_3 and S_4) behave differently in the presence of nitrogen dioxide concentrations. It results from the fact that one of sensors of the same type was new, whereas the other was used before in other experiments. In Fig. 3b an example of responses of two selected sensors, namely S_1 and S_2, to the NO_2 concentrations at 20% RH and 40% RH is shown. Both sensors have similar sensitivities, however their resistance responses are different.

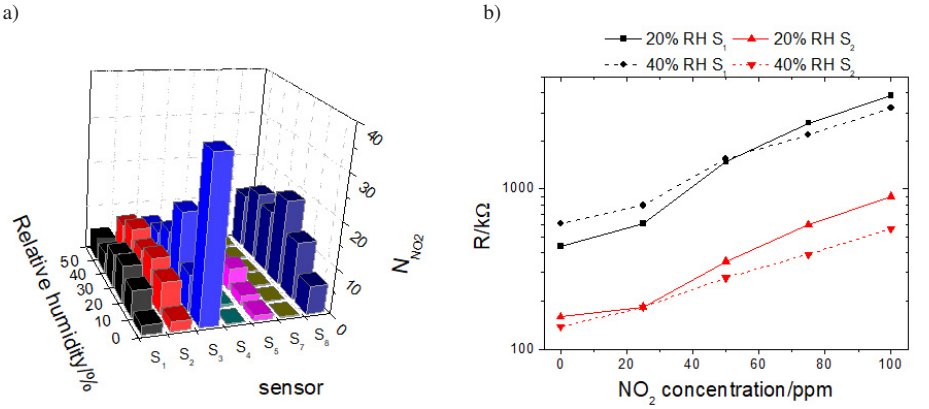

Fig. 3. Normalized responses of TGS sensors to 100 ppm NO_2 at different relative humidity levels (a). An example of resistance responses of S_1 and S_2 sensors to NO_2 concentrations at 20% RH and 40% RH (b).

Figure 4a presents the normalized responses of seven TGS sensors to the changes in relative humidity for all of the measured NO_2 concentrations. The parameter N_{RH} was calculated by the subtraction of the response in dry air (the baseline) from the sensor resistance response in the presence of 50% RH, and dividing the difference by the baseline. The interpretation of N_{RH} is similar to that in the case of parameter N_{NO_2}, presented earlier. However, in this case, the

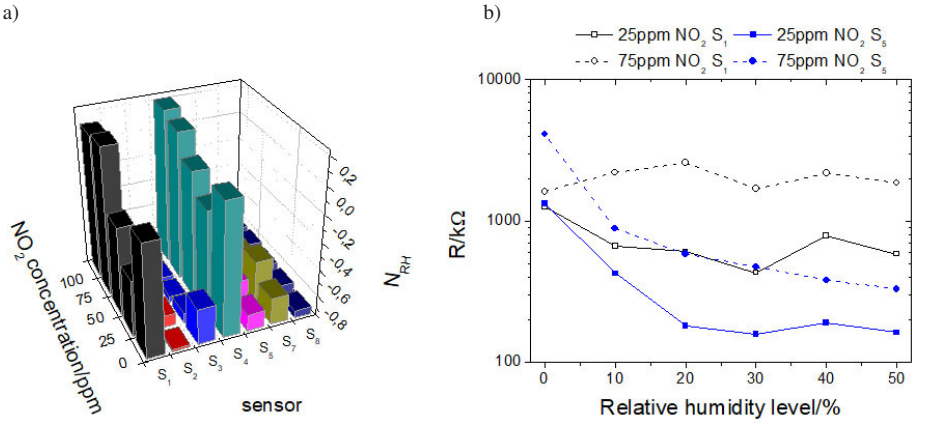

Fig. 4. Normalized responses of TGS sensors to 50% RH at different NO_2 concentrations (a). Examples of resistance responses of S_1 and S_5 sensors in the presence of different humidity levels at 25 ppm and 75 ppm NO_2 (b).

variability of sensor responses is associated with the changes in relative humidity level. In the case of parameter N_{RH} it is expected that it should have negative values due to the fact that, when used in an array, the responses of TGS semiconductor sensors decrease in the presence of humidity. It may be observed that in some cases (*e.g.* in the case of sensor S_1 – TGS 826#1) the calculated values are positive for selected concentrations of NO_2. This is presented in detail in Fig. 4 b), in which there are shown the dependencies of two sensors, namely S_1 and S_5, on humidity. It may be observed that for S_5 there is a visible trend observed in the case of two presented nitrogen dioxide concentrations. The characteristics of S_1 behave differently in the presence of humidity. Namely, there is no visible monotonic trend.

4.2. Dynamic measurements of TGS sensors

An example of single-sensor dynamic responses (TGS 2600#1) obtained using the stop-flow technique is shown in Fig. 5. Three stages, namely the gas-on, gas-stop and air-on are marked. The responses were measured in the presence of 50 ppm and 100 ppm of dry nitrogen dioxide. The characteristic shape of the acquisitions can be seen. Given an assumption that all sensors provide unique responses, the calculated features should have different values, which will be sufficient to obtain a prediction of nitrogen dioxide concentration.

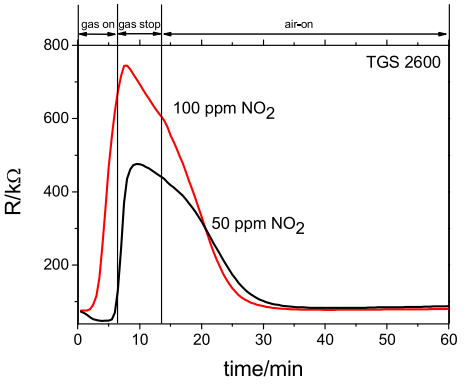

Fig. 5. Sensor S_8 responses in the stop-flow mode to 50 ppm and 100 ppm of NO_2.

In Fig. 6a the results of measurements performed using the stop-flow method for five levels of relative humidity are presented. An example of drift caused by the presence of humidity is shown. The shape of acquisitions are preserved; however, the baseline of the sensor responses changes in such a way that it decreases with an increase in humidity level. Fig. 6b shows the normalized acquisitions, based on which a set of vectors V_3 and features for vector V_5 were calculated. The normalization revealed that there are some differences, *e.g.* in amplitude of the acquisitions, despite the fact that the presented results were measured in the presence of the same concentration of NO_2. It is not clear whether those differences were caused by the influence of humidity or by another drift-causing factor. It can also be seen that after normalization the baseline drift was compensated and some characteristic features, *e.g.* slopes of the acquisitions, are practically the same in all presented cases. Therefore, it is expected that by extracting features from the normalized acquisitions it would be possible to achieve a calibration model for prediction of NO_2 concentration independent of the influence of humidity.

Metrol. Meas. Syst., Vol. 27 (2020), No. 1, pp. 167–179
DOI: 10.24425/amms.2020.131716

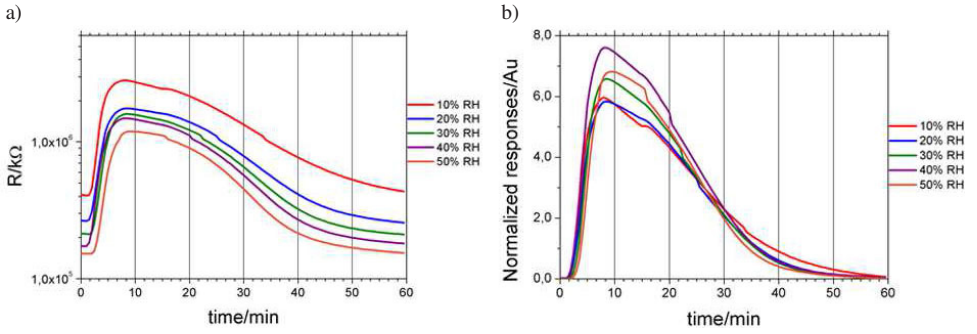

Fig. 6. Sensor S_2 dynamic responses to 100 ppm of NO_2 at different humidity levels (a). Normalized dynamic responses of sensor S_2 to 100 ppm of NO_2 at different humidity levels (b).

4.3. Calibration using PLS regression

A. *Regression based on a set of single resistances*
The first approach was based on a set of vectors V_1, which are composed of a single resistance. In the first step, the responses from seven sensors (S_6 – TGS 2106#2 was omitted) were taken for the analysis. All further analyses were performed without using sensor S_6. The goal was to achieve the minimal value of RMSE. Additionally, the information about RH levels was added to the dataset to improve the results. The calculations showed that the errors obtained with models enhanced using information about the measured RH levels were almost the same as the errors obtained without enhancement; further analyses were performed using a set of vectors without information concerning the relative humidity level. The goal of a further analysis using a set of vectors V_1 was to improve the results of the prediction by reducing dimension of the vectors. Namely, the successive sensor responses from vector V_1 were eliminated and RMSE was calculated. The obtained results showed that the information from all sensors gives the best results by means of the minimal RMSE value. Further elimination of resistance responses of individual sensors as well as their combinations does not improve the final results and often causes an increase in RMSE.

B. *Regression based on dynamic acquisitions*
A similar analysis scenario was employed with the use of a set of vectors V_2 which contained raw acquisition data obtained from all sensors using the stop-flow method. The results showed that the achieved predictions were significantly better in the case of using the acquisitions from dynamic measurements than in the previously presented case of using a set of vectors V_1. The lowest error was obtained with the use of five TGS sensors, with S_1, S_2 and S_6 eliminated. The achieved error was equal to 8.25 ppm. Dynamic responses of sensors S_1 and S_2 were not reproducible enough to provide satisfactory results, thus the elimination of their responses caused the most significant reduction in the RMSE value. Moreover, it can be seen in Fig. 5 that dynamic responses from sensor S_2 were different for the same concentration of target gas when measured at different relative humidity levels. This was also the reason justifying the elimination of responses of that sensor from the analysis.

C. *Regression based on normalized dynamic acquisitions*
In this case, a set of vectors V_3, containing the normalized dynamic acquisitions, was used for calibration with the PLS method. The lowest error was obtained with the use of six TGS sensors, with S_1 and S_6 eliminated. The achieved value of prediction error, equal to 8.89, was comparable with that of case B, and significantly lower than obtained in case A. However, it

was shown that the normalization is not obligatory in the case of PLS analysis. A possible reason of errors can be associated with the problem which was shown in Fig. 5. The sensor responses measured in the same conditions are different, and it is revealed in the values of RMSE.

D. *Regression based on features obtained from non-normalized acquisitions*
Similarly to the previous case, the PLS model calibrated using a set of vectors V_4 enabled to significantly reduce the RMSE value compared with the PLS model based on a set of vectors V_1. The elimination of specific sensors showed that the best results were achieved using the features from five sensors, with S_1, S_3 and S_6 eliminated. This is different from the previous case, where sensor S_2 was removed from the feature vector. This time the contribution from sensor S_2 was significant enough to preserve its response as part of the reduced features of vector V_4. The possibility of improving the results was examined by the elimination of specific features from vector V_4. In the case of eliminating amplitude (A), surface under the curve (S) or full width at half maximum ($T_{1/2}$) from vector V_4, the values of prediction errors increased. Thus, those features should not be eliminated from the set of vectors V_4. The error decreased when skewness (Sk) and kurtosis (K) were eliminated. The best resulting set of features, based on which the PLS model for prediction of nitrogen dioxide concentration was calibrated, contains amplitude (A), surface under the curve (S) and full width at half maximum ($T_{1/2}$) for five selected sensors. From the original composition of 41 features included in vector V_4, after reduction, there remained 15 elements (3 features × 5 sensors).

E. *Regression based on features obtained from normalized acquisitions*
The last analysis scenario was employed with the use of a set of vectors V_5 which contained features calculated from normalized acquisitions obtained with the stop-flow method. The purpose of normalization of the data was to eliminate the baseline shift caused by humidity and to examine whether it was possible to improve the results. The achieved prediction ability was comparable to the case of vectors V_4. The lowest value of RMSE was obtained with the use of four TGS sensors (S_2, S_3, S_4 and S_8). It is the smallest number of used sensors amongst the analysed scenarios. Moreover, the achieved error was the second lowest comparing with the best case for a set of vectors V_2. The elimination of features associated with specific sensors from vector V_5 revealed that the RMSE did not change significantly when amplitude (A) was eliminated from vector V_5; however, the error changed its value from 9.24 ppm before reductions to 9.28 ppm after elimination of amplitude. Therefore, this feature was preserved in vector V_5. An increase of the error value is also observed when full width at half maximum ($T_{1/2}$) is eliminated. Only one feature, surface (S), was rejected due to the highest decrease in the error value – from 9.24 ppm to 9.09 ppm. The final composition of the reduced vector V_5 contains 16 elements (4 features ×4 sensors).

4.4. Summary of analysis

The summary of the performed analyses is presented in Table 2. Since the error in the best case is rather high, in a practical application it would make a serious problem. However, for the purposes of this examination this factor is of secondary importance. Responses from commercially available TGS sensors are influenced by many factors, *e.g.* humidity, effects of aging or poisoning, so that their metrological properties vary in time. The results obtained showed significant advantages of applying dynamic measurements. This can be seen in Fig. 7a, where distributions of prediction values of concentrations of NO_2 calculated by the PLS model compared with the actual ones for the worst (with the use of a set of vectors V_1) and best (with the use of a set of vectors V_2) cases are shown. It can be seen that the results obtained from dynamic acquisitions form less spread

Metrol. Meas. Syst., Vol. 27 (2020), No. 1, pp. 167–179
DOI: 10.24425/mms.2020.131716

distributions and thus the errors obtained with the PLS are smaller. The acquisitions obtained using the stop-flow method contained more information about the gas concentration that could be extracted in the form of features. The PLS regression calibrated using whole acquisitions as well as a set of features provided a better prediction ability compared with the approach based on the use of only single resistance responses of the sensors.

Table 2. Prediction errors obtained in all analysis scenarios.

Analysis case	RMSE/ppm	Eliminated sensors	Eliminated features
A	18.78	TGS 2106#2	–
B	8.25	TGS 826 #1 TGS 825 #1 TGS 2106#2	–
C	8.89	TGS 826#1	–
C	9.56	TGS 826 #1 TGS 825 #1 TGS 2106#2	skewness, kurtosis
D	9.09	TGS 826 #1 TGS 2106 #1 TGS 2106#2 TGS 880 #2	surface

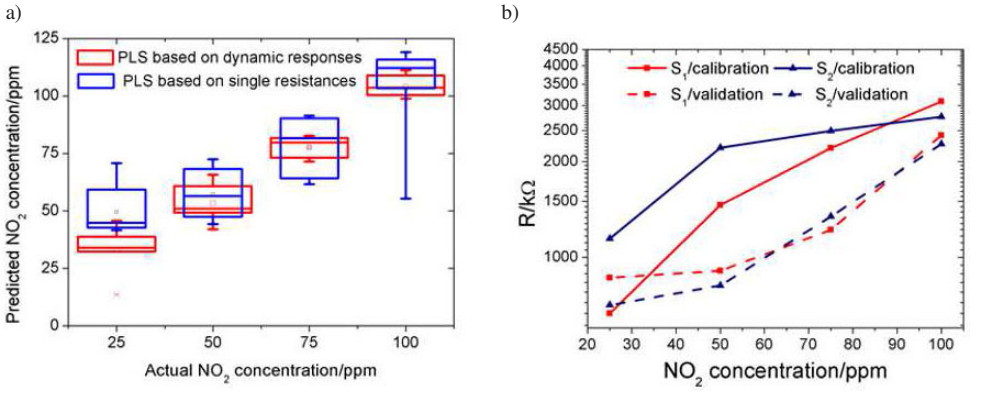

Fig. 7. Distribution of predicted concentrations of NO_2 vs. actual concentration for PLS (a). Instability of the responses of TGS 826#1 and TGS826#2 measured in the presence of 10% RH. Validation measurements were performed one month after calibration measurements (b).

High error values were also caused by the procedure of gathering the data for the purposes of calibration and validation. In many literature reports concerning gas recognition or prediction of gas levels, the data are collected continuously. It is well known that without a significant amount of time between gathering data for the training and validating sets, the interfering factors such as the aging effect and other drift-causing effects will not occur. In this study, the first validation measurements were carried out after at least a two-week interval after the end of collecting the calibration set. Further measurements were used without the recalibration process. An example of instability of characteristics of TGS sensors is shown in Fig. 7b, where responses of two used TGS 826 sensors are plotted. The concentration values measured a month after the calibration

give totally different resistances. This proves that taking single-value resistances for calibration generates massive errors and should not be used in gas-analysing systems.

5. Conclusions

In this work, an array of commercially available TGS sensors was examined. The ability of predicting the nitrogen dioxide concentration value was checked for varying relative humidity levels over a period of three months. The results of analysis of the behaviour of sensors in the presence of the target gas for different relative humidity levels as well as different low gas flow rates were shown. Four approaches to calibration using the PLS method were presented and compared. The best prediction ability was achieved when the PLS regression was performed based on dynamic measurements using the stop-flow technique. The feature extraction enabled to significantly reduce the dataset volume preserving the information about the nitrogen concentration from dynamic measurements. The approach where static measurements were used as a feature vector shows the worst prediction ability and should not be used in gas recognition systems. It was also shown that complementing the feature vectors with additional information about measured humidity levels does not improve the results.

Acknowledgements

This work is partially supported by Statutory Funds of Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology as well as by the National Centre for Research and Development, project LIDER No. 22/103/L-2/10/NCBiR/2011 „Multi-sensor system for measuring air pollutants". Authors would like to thank Mr. Igor Sulim for his help during measurements.

References

[1] Guo, D., Zhang, D., Li, N., Zhang, L., Yang, J. (2010). A novel breath analysis system based on electronic olfaction. *IEEE Trans. Biomed. Eng.*, 57, 1–11.

[2] Peris, M., Escuder-Gilabert, L. (2009). A 21st century technique for food control: Electronic noses. *Anal. Chim. Acta.*, 638, 1–15.

[3] Capelli, L., Sironi, S., Del Rosso, R. (2014). Electronic Noses for Environmental Monitoring Applications. *Sensors.*, 14, 19979–20007.

[4] Kalinowski, P., Strzelczyk, A., Wozniak, L., Jasinski, G., Jasinski, P. (2014). Determination of toxic gases based on the responses of a single electrocatalytic sensor and pattern recognition techniques. *Meas. Sci. Technol.*, 25, 25101.

[5] Mikołajczyk, J., Bielecki, Z., Stacewicz, T., Smulko, J., Wojtas, J., Szabra, D., Lentka, Ł., Prokopiuk, A., Magryta, P. (2016). Detection of gaseous compounds by different techniques. *Metrol. Meas. Syst.*, 23(2), 205–224.

[6] Yin, X., Zhang, L., Tian, F., Zhang, D. (2016). Temperature Modulated Gas Sensing E-Nose System for Low-Cost and Fast Detection. *IEEE Sens. J.*, 16, 464–474.

[7] Gwiżdż, P., Brudnik, A., Zakrzewska, K. (2015). Hydrogen detection with a gas sensor array – processing and recognition of dynamic responses using neural networks. *Metrol. Meas. Syst.*, 22(1), 3–12.

[8] Wozniak, L., Kalinowski, P., Jasiński, G., Jasiński, P. (2016). Determination of chlorine concentration using single temperature modulated semiconductor gas sensor. *Proc. SPIE.*, 10161, 101610U.

[9] Kotarski, M., Smulko, J. (2009). Noise measurement setups for fluctuations enhanced gas sensing. *Metrol. Meas. Syst.*, 16(4), 457–464.

[10] Macku, R., Smulko, J., Koktavy, P., Trawka, M., Sedlak, P. (2015) Analytical fluctuation enhanced sensing by resistive gas sensors. *Sens. Actuators, B.*, 213, 390–396.

[11] Jasinski, G., Jasinski, P., Chachulski, B., Nowakowski, A. (2005). Lisicon solid electrolyte electrocatalytic gas sensor. *J. Eur. Ceram. Soc.*, 25, 2969–2972.

[12] Jasinski, G., Strzelczyk, A., Kalinowski, P., Wozniak, L. (2016). Techniques of acquiring additional features of the responses of individual gas sensors. *Proc. SPIE.*, 10161, 101610P.

[13] Barsan, N., Koziej, D., Weimar, U. (2007). Metal oxide-based gas sensor research: How to? *Sens. Actuators B.*, 121, 18–35.

[14] Korotcenkov, G, Cho, B.K. (2013). Engineering approaches for the improvement of conductometric gas sensor parameters: Part 1. Improvement of sensor sensitivity and selectivity (short survey). *Sens. Actuators B.*, 188, 709–728.

[15] Delpha, C., Siadat, M., Lumbreras, M. (1999). Humidity dependence of a TGS gas sensor array in an air-conditioned atmosphere. *Sens. Actuators B.*, 59, 255–259.

[16] Sohn, J.H., Atzeni, M., Zeller, L., Pioggia, G. (2008). Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares. *Sens. Actuators B.*, 131, 230–235.

[17] Korotcenkov, G., Cho, B.K. (2011). Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey). *Sens. Actuators B.*, 156, 527–538.

[18] Romain, A.C., Nicolas, J. (2010). Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview. *Sens. Actuators B.*, 146, 502–506.

[19] Massok P., Loesch, M., Bertrand, D. (1995). Comparison between two Figaro sensors (TGS 813 and TGS 842) for the detection of methane, in terms of selectivity and long-term stability. *Sens. Actuators B.*, 25, 525–528.

[20] El Barbri, N., Duran, C., Brezmes, J., Cañellas, N., Amírez, J.L., Bouchikhi, B., Llobet, E. (2008). Selectivity enhancement in multisensor systems using flow modulation techniques. *Sensors.*, 8, 7369–7379.

[21] Brudzewski, K., Ulaczyk, J. (2009). An effective method for analysis of dynamic electronic nose responses. *Sens. Actuators B.*, 140, 43–50.

[22] Maciejewska, M., Szczurek, A., Ochromowicz, Ł. (2009). The characteristics of a "stop-flow" mode of sensor array operation using data with the best classification performance. *Sens. Actuators B.*, 141, 417–423.

[23] Wold, S., Sjöström, M., Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. Chemom, *Intell. Lab. Syst.*, 2001, 109–130.

[24] Jasinski, G., Kalinowski, P., Wozniak, L. (2016). An electronic nose for quantitative determination of gas concentrations. *Proc. SPIE*, 10161, 101610.

[25] Chen, Q., Zhao, J., Liu, M., Cai, J., Liu, J. (2008). Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms. *J. Pharm. Biomed. Anal.*, 46, 568–573.

[26] Palit, M., Tudu, B., Bhattacharyya, N., Dutta, A., Dutta, P.K., Jana, A., Bandyopadhyay, R., Chatterjee, A. (2010). Comparison of multivariate preprocessing techniques as applied to electronic tongue based pattern classification for black tea. *Anal. Chim. Acta.*, 675, 8–15.