chemistry proceedings

MDPI

Proceedings

Drift compensation of the electronic nose in the development of instruments for out-of-laboratory analysis †

Anastasiia Shuba [1,*], Tatiana Kuchmenko [1], Dariya Menzhulina[2]

[1] Voronezh State University of Engineering Technologies; tak1907@mail.ru
[2] Voronezh Children's Clinical Hospital named after N.N. Burdenko; darinochka@mail.ru
* Correspondence: shuba1nastya@gmail.com; Tel.: +79204031622
† Presented at the title, place, and date.

Abstract: A technique for evaluating and compensating for the drift of eight mass-sensitive sensor array in an open detection cell was developed to take into account the influence of external factors (temperature, changes in the chemical composition of the background) for out-of-laboratory analysis of biosamples when long-term monitoring of health state humans and animals. The effective way to compensate for the sensor signal drift when the sorption properties of the sensitive coatings change during their long-term intensive operation is the daily internal standardization of the system. Distilled water is proposed as a standard for the biosamples based on the water matrix (blood, exhaled breath condensate, urine, etc.). Internal standardization is based on daily calculating the specific sensor signals by dividing the sensor signals for biosample on the corresponding averaged values from 3-5 measurements of standard. The stability of the sensor array operation is estimated using the theory of statistical process control (exponentially weighted moving average control charts) based on the specific signal of the sensor array. The control limits for the statistic quantity of the central tendency for each sensor and the whole array, variations of the sensor signals were determined. The average time to signal and run lengths for statistically substantiated monitoring of the electronic nose stability were calculated. Based on the analysis of tendency and variations in sensor signals for three months of operation, a technique was drawn up to control the stability of the sensor array for the out-of-laboratory analysis of biosamples. This approach is successfully verified by classifying the results of the analysis of blood and water samples obtained for this period. The proposed technique can be introduced into the software algorithm of the electronic nose, which will increase the correctness of decision-making during long-term monitoring of the health state of humans and animals.

Keywords: piezoelectric sensor array; drift ; electronic nose; stability; statistical process control chart; blood.

Citation: Shuba, A.; Kuchmenko, T.; Menzhulina, D. Drift compensation of the electronic nose in the development of instruments for out-of-laboratory analysis. *Chem. Proc.* **2021**, *3*, x. https://doi.org/10.3390/xxxxx

Published: date

1. Introduction

When designing a new methodology of routine analysis of biological samples, great attention is paid to reducing systematic error linked to the changing of properties of the measuring devices. The control of external conditions (temperature and humidity) is determined by the nature of the working element of the device. Thus, in gas analyzers of sorption type, the signal is defined by the interaction of modifiers phases with vapors and gases in the pre-sensory space. In reusable sensors, stability of sorption surfaces, defining processes of «sorption-desorption», is the main requirement. Since the temperature changes the volatility of substances, their and sorbents sorption properties, operating at the normal temperature, relative humidity for the sensors of sorption type is a critical factor. The development of approaches to ensure the stability of measurements during long-term operation is relevant and paramount for portable and mobile analytical

systems providing the time minimization from the moment of sample selection to the measurement, which is vital for biosamples analysis.

It is well-known that the drift for sensor signals influencing classification solutions and quantitative determination of components in a mixture. A short-term drift of basic lines of the piezoelectric sensor of sorption type is caused by malfunctions of electric schemes and abrupt change of external conditions of the experiment. As a rule, such mistakes are fixed immediately during the measurement and discarded as outliers. A long-term drift basically relates to the change in sensitive coatings of sensors during operation. Currently, two ways to decrease the drift impact of a sensor array on analysis results are applying extremely stable sensitive sensor coatings [1] and mathematical algorithms of drift compensation, increasing accuracy when classifying samples [2]. The mathematical algorithms based on multivariate analysis to drift correction of sensor signals are difficult to implement in practice for out-of-laboratory analysis since they require additional computational costs, software change or an application of additional programs, and operator's modeling skills. The most critical thing at mathematical processing implementation is valid information selection, obtained during a long time, for a single measurement but not for a data set. A goal of the investigation, nature of biosample, and operation conditions define the drift compensation approach's choice.

In this work, we present a technique for drift compensation of the signal of mass-sensitive piezoelectric quartz sensors during operation in the open detection cell and frontal input of the gaseous phase over biosamples with significant water content. The technique consists of the piezoelectric sensor signals correction in portable e-nose during a lengthy routine analysis of water and blood samples combined with a simple algorithm for assessing system instability to increase quality measurement and decrease errors of the first and second order in classification. This technique can be used for real samples analysis, including biological ones (blood, secretions, mucus, exhaled breath condensate, perspiration, urine, saliva) with minor modification of device software for out-of-laboratory analysis.

2. Methods and analysis

Gaseous phase analysis above distilled water and blood samples was conducted on the gas analyzer «Diagnost-Bio-8» (LLC «SenSino», Kursk, Russia) using «frontal analyte input» mode [3]. The sensor array consists of 8-piezoelectrical quartz BAW-type resonators with 10,0 MHz basic oscillation frequency with coatings from solid nanostructured sorbents («Living system» set): Sensors 1, 8 – carboxylate carbon nanotubes of different mass, marked in the tables and text as MCNT1, MCNT2, Sensors 2, 7 –phases of nitrate of zirconium oxide of different mass, Zr1, Zr2, Sensor 3 – Dicyclohexano-18-Crown-6, DCH18C6, Sensors 4, 5 – biohydroxyapatite phases of different mass, HA1, HA2, Sensor 6 – polyethylene glycol succinate, PEGsc. The manufacturing of sensors is described in work [4]. The basic sensor line in the array remained stable (± 1 Hz) during an 80- second blank measurement before sample analysis.

The blood or water (volume of 0,5 cm^3) was placed onto a glass Petri plate and then covered by a detection cell of the device. The measurement mode was combined 80 s was sorption and 120 s spontaneous desorption [3].

For analysis of work stability and drift correction of sensor array signal, there was chosen 3-month period (October to December 2019) of device operating (656 measurements, among them 75 water samples and 31 blood samples of somatically healthy people, who had indicators of general and biochemical analysis of blood within normal limits). As the training set of the samples, there were selected sensor array data for 64 samples (19 blood samples and 45 water samples) for the first one and a half months of operation (October-November). Sensor array data for 42 samples (12 blood samples and 30 water samples) were selected as the test set of samples for the next period of operation (November-December). During the experiment, external factors have variation: room temperature change from 20 to 25 °C, the humidity was 45-55%, the slight changes of

room smell (background) per day corresponded to the schedule of disinfection and ventilation of the laboratory. Measurements were not performed during the disinfection and ventilation and 90 minutes after them. As the original sensor data for each sample, special software determines maximum sensor responses ($\Delta F_{max,i}$, Hz), which correspond to 80 seconds of measurement (sorption time).

We suggest the correction of the sensor signals in an array based on daily internal standardization. The distilled water was chosen as the standard. Therefore we calculated the specific signals $\bar{F}_i$ via formula:

$$\overline{F_i} = \frac{\Delta F_{max,i}(for\ sample)}{\Delta \bar{F}_{max,i}(for\ water)}, \quad (1)$$

where $\Delta F_{max,i}$, Hz is values of the original signal of i-th sensor at biosample analysis, $\Delta \bar{F}_{max,i}$ is average signal of i-th sensor for 3 – 5 analysis of water samples on the same day.

For specific signals calculated for water samples, we applied a statistical process control method, namely Exponentially Weighted Moving Average (EWMA) Control Chart for uni- and multivariate data [5] for assessing the stability of sensor array. We investigated for sensor array the statistic quantity T_i^2, for each sensor – parameter z_i, and exponentially weighted mean square (EWMS) error s_i during three months of operating. Before the computation of statistic criteria, the obtained specific signals were standardized by the average value equal to 1,00 for all the sensor signals and standard deviation for each sensor according to data for the first 25 operation days.

As an alternating method of drift compensation for sensor signals, component correction by principal component analysis (CC-PCA) is employed to remove a part of the information, which describes sensor signal drift, corresponding to the first principal component, from the data matrix.

As variables for designing classification models, we used original maximum sensor responses and specific signals. The results of water and blood gaseous phases analysis by sensor array were classified into two classes ("water" and "blood") by linear discriminant analysis with the preliminary processing by principal component analysis (PCA-LDA) with significance level 0.05 (CAMOSoftware Unscrambler v.10.0.0, Oslo, Norway) to assess the effectiveness of drift compensation techniques.

3. Results and Discussion

The significant factors determining the drift of sensor signals in the device with an open detection cell are the changes of sorption properties of sensitive coatings during operation and external factors of the experiment (temperature, humidity, air composition in the laboratory (background)). Frequently, these factors act together, and it is hard to estimate and predict the contribution of each factor to the total signal change. The other internal factors such as the malfunctions in the oscillating scheme of the device, the change in the features of utilized quartz and electrodes, the problems linked to signal transmission into software for our devices varies sensor signals no more than 5% [9]. The algorithm of multivariate data processing for drift compensation of sensor signals is limited by the pilot experiment conditions for model design and cannot always contain all the factor combinations influencing the signals of the system.

A daily internal standardization may perform a simple account of the external factors that influence sensor signals during long-term operation. As the standard, we offer to use the substance dominating by its content in analyzed samples. For biosamples such as blood, urine, exhaled breath condensate, saliva, distilled water is an appropriate standard. For three months of the experiment, the value of the original signals of the sensors in water vapor varied from 20 to 27%, moreover, the dispersions of original sensor signals were inhomogeneous according to Cochran's C-test. Application of the daily standardization reduces the variation to 12 %, and dispersions of all signals become homogeneous. Average specific signals of all sensors throughout the total operation period are equal to 1.00. The implementation of the specific signals considers the changes in

biosamples, sorption properties of sensor coatings, depending on external factors, and reduces the errors of sample classification. However, more accurate results of classification could be obtained using information about the stability of sensors.

3.1. Stability assessment of piezoelectric sensor array by specific signals.

Application of multivariate EWMA control charts by specific signals discovers only the abrupt change in external conditions (measurement of a contaminated standard – operator's blunder). Therefore, detailed information about stability sensor array describes EWMA control charts for individual signals of the sensor array (Fig. 1).

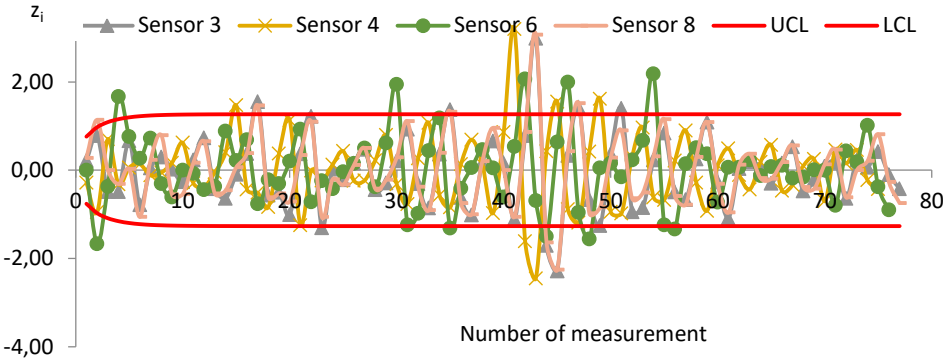

Figure 1. EWMA control chart for specific signals of four sensors from the array. The solid red lines indicate UCL and LCL for 3rd sensors in the array.

According to multiple excesses of control limits, there was established high instability of sensor 6 with the PEGSb film because of the excellent water impact on the properties of polymer films. The control limits for the sensor with the DCH18C6 film (sensor 3) have the most significant values (Table 1). It is explained by the nature of the modifier that consists of solvated crown-ether macromolecules. Such a structure is less stable and varies more depending on external conditions.

Table 1. Maximum values of control limits for parameter z_i for each sensor: upper «+» and low «-».

MCNT1	Zr1	DCH18C6	HA1	HA2	PEGSb	Zr2	MCNT2
± 1.01	± 1.24	±1.27	±1.25	±1.22	±1.06	±1.18	±1.22

Parameters z_i for the sensors with hypersensitive nanostructured films of carboxylated nanotubes (sensor 8) and hydroxyapatite (sensor 4) are almost always within the control limits. The exception is the day of synchronous change of z_i parameters for all sensors and the following 3 days (41-51 measurements), which is connected to the changes of coating properties during the sorption of the contaminated standard vapors and its subsequent reactivation by purging with laboratory air and repeated measurements of pure standard samples. The data of the biosample analysis obtained these days could be unreliable and contribute to error in the classification model.

For the final decision about exclusion of the sensor signals from classification data and assessing the sensor regeneration level after poisoning, it will be required to analyze the changes of signal variability. The evaluation of the sample generalized variance for sensor array is not informative since it values throughout the total studied period were either negative or extremely low (less than 10^{-60}). Therefore, we will consider EWMS errors s_i of specific sensor signals (Figure 2). On the day of the abrupt change in external conditions (measurements of the contaminated sample of the standard, measurements № 41-43), variation for all the sensors and zi exceeded the upper limit control (UCL=1.45). Therefore, when a synchronous increase of control limits by parameters z_i and s_i for all sensors, it is necessary to terminate measurements, verify compliance with the conditions

for the measurement procedure, and reactivate the sensor array's coatings. When calculating EWMS for each subsequent day after the carried-out reactivation of coatings, EWMS $s_0 = 1$ is accepted as the initial point for all the sensors. In the following 8 days (measurements № 44-65), EWMS of all sensors, except sensors 1 and 8 with carbon nanotubes coatings, exceeded UCL as well. It might indicate sorption quality changes of hydrophilic modifiers after sorption of contaminated standard vapor and their gradual recovery when measuring the pure standard or slight inertia of the surface reactivation process of nanostructured mineral coatings.

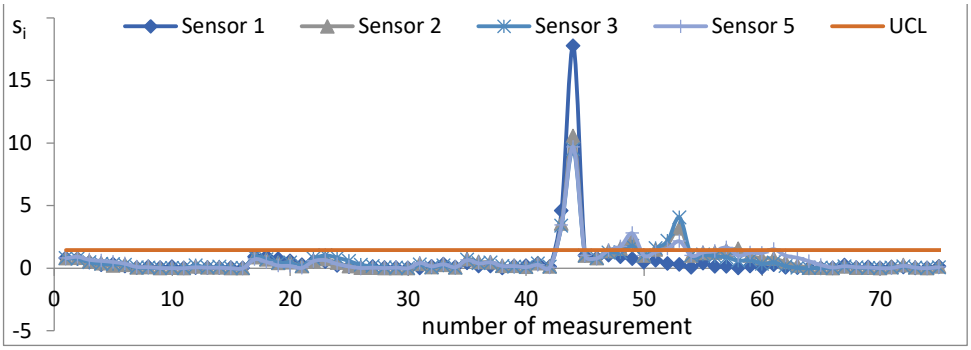

Figure 2. Control Chart for EWMS errors of four sensors

Average run length (ARL) and time to signal (ATS) were calculated according to EWMA theory. ARL is equal 3 times to detect 3 σ shifts in signals and ATS is 2,5 hours. After a long break or recovering the sorption properties of coatings ARL is 5 times, ATS is 1,5 hours.

General recommendations for sensor data analysis of biosamples accounting of its stability were formulated as a scheme that could be implemented in the device's software to monitor stability sensor array with minimal control by the operator, which is essential for out-of-laboratory analysis. The proposed approach is implemented for biosample analysis, for example, blood analysis.

3.2. Blood and water samples classification

To ensure the orthogonality of the data for the discriminant analysis, we performed the principal component analysis of original and specific sensor signals. For further analysis, we used the first 4 principal components with the 98 % explained variance, which allows us to consider all the information from the sensor array while sample classification. At the first stage, we estimated the possibility of water and blood sample classification by original signals of sensors. The accuracy of this classification model for the training set of samples was 90, 4% (Table 2).

Table 2. Accuracy and correctness of sample division into 2 classes by PCA-LDA models for original and specific sensor signals.

Variables	Accuracy of PCA-LDA model for training set, %		Correctness of classification of blood samples from test set, %	
	1[1]	2	1	2
Original signals, ($\Delta F_{max,i}$, Hz)	90.4	96.2	0	0
Specific signals, ($\overline{F_i}$)	94.2	95.2	33.3	80.0
CC-PCA	90.4	96.2	8	10

[1] 1 – for model designing, we have used the whole data array, 2 – the samples and sensor data have been excluded according to the stability evaluation of sensor array operation.

The misclassification of blood samples to "water" class is observed. It signifies that during the analysis of the gaseous phase over blood samples, sample matrix influences on sensors, and over time differences of the gaseous composition in samples become less noticeable for sensors. The greatest value for accuracy and correctness of blood sample classification belongs to the model constructed by specific sensor signals that accounted for sensor stability information (Table 2). Consequently, drift compensation of sensor signals using specific signals with combination sensor stability monitoring by EWMA control charts can be applied with the routine sensor array operation in the laboratory and out-of-laboratory analysis. Furthermore, the correct classification of samples considering the impact of the dominant component allows us to suggest that such an approach can provide a more accurate classification of slight differences in the composition of gaseous phases of blood with pathologies.

4. Conclusion

We have proposed a fast and efficient compensation method for sensor signal drift based on daily standardization (dividing sensor signals for biosamples into corresponding average signals for standard samples (3 times measured during the day)). The distilled water was suggested as a standard for blood samples and other biofluids (urine, perspiration, exhaled breath condensate, saliva). EWMA control charts were applied for sensor array stability monitoring as an additional module in the device software for routine analysis. The most stable coatings from the studied sensor array when measuring water and blood samples are carboxylated carbon nanotubes. The effectiveness of applying for drift compensation by daily standardization with a combination of sensor stability monitoring was proven by significantly raising blood samples' classification accuracy. Similarly, it is possible to compensate for the drift of sensor signals when analyzing other biosamples using appropriate standards. The additional parameters of sorption or features of sensor output curves after standardization could be used to improve the correctness of classification during the long-term operation of the sensor array.

Author Contributions: Conceptualization, T.K.; methodology, T.K. and A.S.; validation, D.M. and A.S.; investigation, D.M.; data curation, A.S.; writing—original draft preparation, D.M.; writing—review and editing, A.S.; supervision, T.K.; project administration, T.K.; funding acquisition, T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Russian Science Foundation, grant number 18-76-10015..

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by Ethics Committee of Voronezh State University of Engineering Technologies (protocol code № 2 from 21 February 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patients to publish this paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Tolentino, M.A.K.P; Albano, D.R.B; Sevilla, F.B. Piezoelectric sensor for ethylene based on silver(I)/polymer composite. *Sensors and Actuators B-Chemical* **2018**, *254*, 299–306.
2. Di Carlo, S.; Falasconi, M. Drift Correction Methods for Gas Chemical Sensors in Artificial Olfaction Systems: Techniques and Challenges. In *Advances in Chemical Sensors*, 1st ed.; Wen Wang; IntechOpen, 2011; pp. 305-326.
3. Shuba, A.; Kuchmenko, T.; Umarkhanov, R.; Chernitskiy, A. Portable E-nose for Diagnostic of Inflammation and Diverse Variation in Health Status of Humans and Animals. In Proceedings of the Fifth International Conference on Advances in Sensors, Actuators, Metering and Sensing (ALLSENSORS 2020), Valencia, Spain, 21–25 November 2020; pp. 56–62.
4. Kuchmenko, T.; Shuba, A.; Umarkhanov, R.; Lvova, L. The New Approach to a Pattern Recognition of Volatile Compounds: The Inflammation Markers in Nasal Mucus Swabs from Calves Using the Gas Sensor Array. *Chemosensors* **2021**, *9*, 116.
5. Montgomery, D. *Introduction to Statistical Quality Control,* 6th ed.; Wiley: Jefferson City, USA, 2009; pp. 419–424; 509–518.