

Drift compensation of gas sensor array data by Orthogonal Signal Correction

M. Padilla^{1*}, A. Perera¹, I. Montoliu¹, A. Chaudry², K. Persaud³ and S. Marco¹

¹ Departament d'Electrònica. Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain/ Institut de Bioenginyeria de Catalunya (IBEC), Baldiri i Reixac 13, 08028 Barcelona, Spain

² Protea Ltd, 11 Mallard Court, Mallard Way, Crewe Business Park, Crewe, Cheshire, CW1 6ZQ, UK

³ School of Chemical Engineering and Analytical Science, The University of Manchester, PO Box 88, Sackville St, Manchester, M60 1QD, UK

Abstract

Drift is an important issue that impairs the reliability of gas sensing systems. Sensor aging, memory effects and environmental disturbances produce shifts in sensor responses that make initial statistical models for gas or odor recognition useless after a relatively short period (typically few weeks). Frequent recalibrations are needed to preserve system accuracy. However, when recalibrations involve numerous samples they become expensive and laborious. An interesting and lower cost alternative is drift counteraction by signal processing techniques. Orthogonal Signal Correction (OSC) is proposed for drift compensation in chemical sensor arrays. The performance of OSC is also compared with Component Correction (CC). A simple classification algorithm has been employed for assessing the performance of the algorithms on a dataset composed by measurements of three analytes using an array of seventeen conductive polymer gas sensors over a ten month period.

Keywords: gas sensor array, drift, Orthogonal Signal Correction, Component Correction, cross-validation, electronic nose, Data shift.

1. Introduction

Gas sensor arrays are potentially economic and faster alternatives for gas analysis and aroma evaluation compared to conventional analytical instruments such as gas chromatographs. Many different technologies based on different principles for gas sensing are available. These include Metal Oxide (MOX), Quartz Microbalances (QMB), Conductive Polymers (CP), Surface Acoustic Waves (SAW), and electrochemical sensors among others. While solid-state sensors offer very big advantages in terms of miniaturization, response times and power consumption, those sensors show poor selectivity. Since the seminal paper of Persaud and Dodd [1], it has been known that improved selectivity may be achieved by combining different sensors with partial and overlapping sensitivities and pattern recognition systems.

While a large number of successful applications for gas sensor arrays have been published, these are largely laboratory based and practical applications of chemical sensor arrays in real commercial applications are limited. This may be attributed to a number of reasons, such as insufficient sensitivity (or limit of detection too high), lack of selectivity, and other problems. However, from the authors' point of view, the main reason is the lack of stability over time and the cost of recalibration. This paper addresses the application of different methods to improve the stability over time of sensor arrays from a signal and data processing perspective.

Gas sensor drift consists of a random temporal variation of the sensor response when it is exposed to the same analytes under identical conditions. This problem is generally considered to be due to sensors aging [5], but it has also been attributed to a variety of sources, like environmental factors such as humidity variations [2], system sampling non-

specific adsorption [3,4], variations on flow rate, thermo-mechanical degradation and poisoning among others [6,7]. All of these factors can modify both the baseline and the sensitivity of the sensors in the array in different ways, depending on sensor technology.

Therefore, for operation over long periods, the ability of the instrument to recognize analytes is degraded, since statistical models built in the calibration phase become useless after a short period of time, in some cases weeks or few months. After that time, the instrument must be completely re-calibrated, which is a time-consuming, laborious and expensive task, to ensure that the predictions remain valid. The working hypothesis in this work is that system stability may be improved using proper multivariate data processing techniques.

Several methods have been reported in the literature in order to improve stability over time by modifications in sensor technology and design [8-10] or by the use of different sensor operation modes [11,12]. On the other hand, signal processing methods for drift counteraction are based on different approaches: univariate [13,14] or multivariate [15-17], linear [18,19] or non-linear [15-17], adaptive [15,16] or not, based in reference samples [14] or based in component removal [18,19]. Univariate corrections include basic baseline corrections [13], or more complex per-sensor correction by means of a calibration sample [14]. These methods, however, do not take profit of correlated drift effects. For this purpose, other multivariate methods have been proposed. Non-linear multivariate adaptive algorithms like Self Organizing Maps (SOM) [15,16], multiple SOM (mSOM) [17] and system identification theory [15,18] have also been proposed in the past. However, fully unsupervised adaptive SOM methods have problems in case of overlapping classes, since the reference pattern updating may follow the wrong class. Techniques based on system identification theory predict sensors outputs by means of dynamical models for every sensor from the rest of the sensors' response, assuming the sensors behave very similarly. Other adaptive approaches include the use of sliding window wavelet decomposition based feature extraction for drift detection and compensation, sometimes post-processed by Principal Component Analysis or Fuzzy Adaptive Resonance Theory networks (ARTMAP) based algorithms [19-21]. Data processing by wavelets decomposes data into multiple time scales. Since drift is a slow process, it is better captured by the coarsest scales, while noise and other events appear in the finest scales. Provided that the separation of drift from real responses is feasible, this procedure improves posterior classification or regression tasks. Linear methods like Component Correction (CC) based on Principal Component Analysis (PCA) [22,23], or Canonical Correlation Analysis (CCA) and Partial Least Squares (PLS) [24] have also been reported to provide good results. In particular, Component Correction has received considerable attention in the community. However, Component Correction methods assume that all chemical classes behave in the same way in the presence of drift and this is not always the case (as it will be seen in the present study).

Our primary goal is to demonstrate that Orthogonal Signal Correction (OSC) is suitable for drift compensation. OSC is a linear technique that removes components orthogonal to a variable Y indicative of the data structure, which is correlated to the data. This method is a very common tool used in spectroscopy to correct spectra by removing systematic non-relevant information, such as baseline variation, but it is not so commonly used in the field of chemical sensors, although we find some examples in literature [25,26] where it is used for removing local variance. This technique will be compared to Artursson's Component Correction (CC) method. The effectiveness of the proposed algorithms will be evaluated on an experimental dataset composed by measurements of several analytes using an array of conductive polymer sensors. The time duration of experiments has been 10 months.

A secondary goal of this paper is to study the impact of the training set size on the ability of the methods for drift compensation. To be of practical interest, drift counteraction should

be effective with only a reduced set of training samples, spanning a limited time duration.

A key point in this paper is the validation methodology. Very common validation techniques, like k-fold cross-validation, random subsampling, bootstrap and leave-one-out, which abound on the literature, totally neglect the influence of drift. In those techniques validation and training samples are interleaved along the time axis. Then future evolution of the sensor responses is modeled by the classifier and consequently provides overoptimistic results. Of course, this is not representative at all of real operation conditions. The authors would like to emphasize that those validation techniques should never be used in conditions where drift is present. Instead, in this work validation samples are always in the future of training samples.

2. Theory

2.1. Component Correction

Artursson proposed in 2001 the Component Correction method (CC) [22]. It is a signal processing technique based on a Principal Component Analysis (PCA) decomposition of a reference class data subset that is later used to correct undesired variance from the rest of the dataset. This method assumes that the reference class is representative of the entire population. Therefore variation found in the reference class will also be present in the rest of the dataset. This means that, if the variance in the reference class is due to drift effects, drift will be also removed from the complete dataset.

2.2. Orthogonal Signal Correction (OSC)

Wold et al. firstly introduced orthogonal Signal Correction (OSC) for its use on NIR spectra correction [27]. The main idea is to remove variance not correlated to variables in a vector (or matrix) Y , which contains some extra data information. This is done by constraining the deflation of non-relevant information of X , so that only information orthogonal to Y should be removed. The inclusion of the condition of *orthogonality to Y* ensures that the signal correction removes as little information as possible.

After Wold's paper a number of OSC-like algorithms that tried to improve the original OSC method were published [28-34]. Comparisons among them can be also found in literature [35,36]. In this work, the version applied is based on the Wise implementation of the algorithm [37]. This first searches for a direction of maximum variance of the data X using PCA. The scores vector t corresponding to this first principal component is then orthogonalized with respect to the information matrix Y , in order to obtain a new scores vector nt not correlated with Y that captures the highest possible amount of variance of X . A Partial Least Squares (PLS) step between initial data X and nt , with a suitable number of latent variables (LV), gives scores $T1$ and loadings $P1$ vectors that contain the information not related to Y . For the rest of the paper, we will refer to this step as the inner PLS. The number of inner PLS latent variables, it is calculated in this implementation from an specification of the variance explained in the X -block. We refer to this as the OSC tolerance and it will be given in percentage values. In a final step $T1$ is again orthogonalized with respect to Y and $P1$ is updated. The final $T1$ and $P1$ correspond to first OSC factor and its products are then removed from original data. To obtain a second OSC factor the complete treatment is then applied on corrected data, therefore for n OSC factors corrected data X_{OSC} is given by:

$$X_{OSC} = X - \sum_{i=1}^n T_i P_i' \quad (1)$$

3. Experimental

3.1. The dataset

The dataset was from Osmetech plc (Cheshire, UK). Three different analytes (ammonia, propanoic acid and n-butanol), at different concentrations levels, were periodically measured over 10 months with an array of 17 conductive polymer sensors, the total number of samples in the dataset being 3415. The concentration and number of samples for each analyte are shown in table 1. We consider each analyte as a class. Additionally, we define a group as a particular analyte at a given concentration. Hence 3 classes and 8 groups are present in the dataset.

For every sample, the full sensor response to a sampling transient is recorded. That is, the sensor array is initially exposed to clean air. Subsequently, the analyte at the desired concentration is introduced in the sensor chamber for 185s. Finally, clean air is introduced again. Every transient signal lasts for 200s at a sampling frequency of 1 Hz. All waveforms are baseline corrected so that starting baseline in all samples is common at time 0s. Figure 1 shows an example of waveform for one sensor and three classes. To apply the proposed drift counteraction techniques, the data matrix is organized into a two-way matrix with dimensions: *number of samples* \times (*sensors* \times *transient time*) or 3415 samples \times 3400 variables.

At every sample interval of the transient signals, the response of the sensor array produces a characteristic pattern for every class. Figure 2 shows the patterns of three different classes at time 185s. of the transient waveform. Traditionally only one point of a transient signal is considered for every sensor, this point usually corresponds to the maximum value of the signal (in our case around instant 185s in every transient signal). When introducing the complete transient waveform in the data matrix, dynamic information is also being considered [38]. It is well known that transient information provides additional discriminatory information.

3.2. Methodology

The proposed model validation procedure has been specifically conceived to illustrate the capabilities of the different algorithmic solutions towards drift rejection. The secondary goal is to ascertain the behavior of the algorithms when a small sample training dataset is encountered. This is usually the case since, from a practical point of view; calibration costs have to be reduced to a minimum.

As figure of merit, the performance of a classifier in time-ordered validation subsets is evaluated. The classifier consists of a dimensionality reduction step using Principal Component Analysis, followed by a k-NN classifier. While the input dimensionality is 3400, the output dimensionality is limited to the number of Principal Components (PCs) that capture most of the X data variance. In this space a k-NN classifier is used with $k=3$ nearest neighbors. The dataset has been classified in three classes, corresponding to the three chemical species.

3.2.1 Validation methodology

In order to carry out rigid validation, algorithm optimization has been restricted to the use of calibration set information. Final testing was performed with data subsets never used for algorithm building and optimization. To test the effectiveness of the drift counteraction techniques, the complete dataset (3415 samples), was divided in 10

sections (or subsets) of 342 samples each approximately. All samples are ordered in time. A scheme of the validation methodology is shown in figure 3. The first subset consists of samples measured during the first 15 days of experiments. We will refer to this first section as the calibration set.

In a first step, the algorithms parameters are optimized by using an internal cross-validation within the calibration set. The last quarter of the calibration set is used for inner validation. The model is built with a random selection of 66% of the remaining calibration set samples (first $\frac{3}{4}$). By repeating this process 10 times we evaluated the robustness of the models towards the particular selection of training data. Graphics of results provide the corresponding error bars.

Once the algorithms were optimized, stability over time evaluation was done by assessing the performance with the remaining nine data subsets ordered in time (figure 3).

To evaluate the performance of drift correction for smaller training sets, ten additional data subsets were built with random samples from the original calibration set. The sizes of these subsets range from 10% to 100% of its size. Therefore, the smallest calibration subset contains 34 samples and the largest one 342. These ten subsets have been used as training sets for building the algorithms models. Please note that in this section, no further optimization of the algorithms inner parameters is done. Hence, these parameters stay fixed as the size of the training set varies.

3.2.2 Algorithm optimization

To select the best internal parameters, we propose to use as a figure of merit the Fisher Ratio: ratio between inter-group and intra-group variance of the cross-validation dataset.

In OSC, two internal parameters must be optimized; the number of OSC components and the number of internal PLS latent variables. Here, instead of the number of inner PLS latent variables, we use the equivalent amount of captured variance in the X-block (or tolerance) of the inner PLS relation. This is the way original Wise's algorithm is implemented [37].

On the other hand, in the PCA-CC method the number of principal components and a reference group have to be selected as well. The criterion used to select the reference group was based on visual inspection of the score plots in the calibration set. This group should have an inner covariance structure similar to most of the remaining groups in the dataset. The reference group is modeled with a Principal Components decomposition that captures a high amount of the group inner variance in the calibration set. In fact, removing many components in PCA-CC is risky, since some of those components may have information useful for posterior classifications. This is the case when drift direction is parallel to discriminant direction. Because of these factors only three components were removed by PCA-CC.

4. Results and discussion

An exploratory analysis by PCA displays the initial distribution of the classes. Figure 4 shows the PCA scores plot of the calibration set (solid symbols). In the figure, it can be seen that some classes apparently mix, there is scattering due to drift and also additional intra-class variability. The direction of the main dispersion component in all classes is quasi parallel, except ammonia 0.05%. Some groups, ammonia 0.01% and 0.02% and both concentration levels of n-butanol, present dispersion over mainly only one direction. Main dispersion directions are pointed out by an arrow for every type of gas. A thick arrow shows the direction of the displacement of the eighth validation set (non solid symbols) with respect to the calibration set. Furthermore, in this data subset sensors are very

correlated since only ten principal components capture more than 99% of the total variance, and the first two PCs capture about 90%.

A posterior data subset of 342 samples, measured between day 97 and day 118, has also been projected on the same PCA subspace. Scores of this projection are shown in the same figure in non-solid symbols. The distribution of the groups in this latter subset is similar to the one in the first subset; only ammonia 0.05% shows an important deviation in its main dispersion directions. However, the whole subset is completely shifted from the location of the first data subset. Therefore, additional to variability and scattering due to local noise/drift, there is a long-term drift effect, which displaces and changes the variance within the data set structure.

Figure 5 shows the maximum of the transient for sensor one for the total duration of the experiment. In this figure, the effect of drift and intra-class variability in the measurements is clearly seen, and also a clear correlation between the different traces can be observed. The irregular shape of these curves may be due to sensors aging and also the environmental changes that happened during the 10 months of measurements, since conductive polymer sensors are strongly perturbed by temperature and humidity.

4.1. Fitting PCA-CC and OSC parameters

In figure 6, the Fisher ratio is plotted for different numbers of removed components and tolerances for OSC and PCA-CC. Fisher ratio measures the ratio between mean values of the distances of groups among themselves and the groups compactness:

$$FR = \frac{\sum_{i,j=1, i \neq j}^N |c_i - c_j|}{\sum_{i=1}^N mx_i}$$

where $c_{ij} = |c_i - c_j|$ is the distance between the mean centers of groups i and j , and mx_i is the mean of the distances of the N_k elements of group i to its center c_i :

$$mx_i = \left(\frac{1}{N_k} \right) \sum_{k=1}^{N_k} |x_k - c_i|, \text{ with } c_i = \left(\frac{1}{N_k} \right) \sum_{k=1}^{N_k} |x_k|$$

Therefore, high values are desired since they would mean small groups and high separation among them. The proposed figure of merit helps to avoid over fitting, however it does not guarantee the best choice for drift rejection in the long term, since the cross-validation samples are very close in time to the training samples.

Figure 6 shows that the Fisher Ratio is higher for OSC than for CC, suggesting that the former will outperform the later in the final classification task.

For the OSC, we observe a continuous increase in the Fisher Ratio with the number of extracted components at multiple levels of the tolerance value. It may be surprising that this component removal does not saturate fast, but it has to be considered that this preprocessing is carried out in an input space featuring high dimensionality, and the graph only explores removal of up to 16 components. On the other hand, the figure of merit shows more sensitivity to the training samples (larger error bars) when the number of OSC components increases. This is probably due to an overfitting to the actual samples used. Additionally, the tolerance value chosen has a higher impact in the results for higher number of OSC components, while it has little influence for a reduced number of components. For further analysis, two cases are considered one and eight OSC components both with a high tolerance, 99.0% (other high tolerances around 99% could also have been chosen).

Regarding PCA-CC method, a reference group must be chosen. The firsts loadings of a

PCA on this group are subtracted from all data, thus this group must be representative of all groups. In figure 4, it has been seen that all groups present main dispersion directions that are nearly parallel, except for ammonia 0.05%. This direction is very well defined for n-butanol and ammonia 0.01 and 0.02% but ammonia 0.05%, and propanoic acid presents more dispersion components. Therefore, either n-butanol or ammonia 0.01 or 0.02%, can be a good reference. Finally we have selected n-butanol 1%. Figure 6 shows Fisher ratio for PCA-CC computed with two groups; ammonia 0.05% and n-butanol 1%. Best performance is obtained when using n-butanol 1% as a reference. In fact, when selecting a reference group the dispersion of which is not representative (ammonia 0.05%), figure 6 shows that CC does not improve Fisher Ratio, but may even worsens it. This result confirms the choice of the reference group (n-butanol 1%) from the visual inspection of the PCA scores plot in figure 4. For both reference groups, the Fisher Ratio saturates sooner than in the OSC. Finally, figure 6 also shows that the PCA-CC models are less sensitive to the particular choice of training samples than OSC models (especially for higher number of removed components). Three Principal Components have been removed, since they capture a high amount of inner variance (about 95%) in the reference group.

Figure 7 shows one removed OSC component and only the first removed component in PCA-CC for the sake of a better visualization. An exploration on the shape of the loadings for the PCA of the reference group helps to understand where the drift appears (figure 7). This may have some impact on later instrument optimization.

It is interesting to observe, that the first principal component for PCA-CC drift correction almost equally weights all the sample intervals in transient time. Regarding sensors, sensor 11 and 12 contribute to drift slightly more than other sensors in this component. This component also shows high values (peaks) at points where the transient signal changes abruptly, corresponding to the instant when the analyte comes into the sensors chamber and when dry air cleans the chamber (at 2s and 185s approx.). It seems to reflect some jitter in this point and a better synchrony between signal acquisition and the chemical sampling system is needed.

First OSC component is similar to first PCA-CC component but with emphasis on the last part of the transient. It also confirms that sensors 11 and 12 are less stable than the others. As in PCA-CC, in general the sensors that contribute the most are 11, 12 and 17. Also the highest components values (peaks) are located at the same transient instant, like in PCA-CC, showing that transition times contain most of the variance, thus they are more affected by drift.

4.2. Data distribution of corrected data

A PCA scores plot for the calibration set and a posterior test set has been shown on figure 4. Such test set consists on samples measured between day 97 and day 118. PCA scores plot for corrected test set with each method are shown on figures 8 and 9. In these figures, data from the test subset are projected on the PCA model (non solid symbols) built from corrected calibration set (solid symbols) like in figure 4.

Distribution of data corrected by OSC (8 component and 99% tolerance, figure 8) shows little dispersion in its eight groups, which are well separated, the better the ones corresponding to the highest concentration of every analyte. Groups belonging to the same class are located along one direction in the figure, therefore a data set distribution consist of clusters along three different directions, one for each class, with decreasing concentrations towards the center of the data distribution. Corrected test set is displaced from the training set, presents higher dispersion than training data and shows a small change on the orientation of the branch corresponding to every chemical specie. All these three effects are due to drift that OSC has not modeled.

Drift corrected data by means of PCA-CC (3 components, figure 9) also shows a distribution along three main directions, with different orientation according to the chemical species. It presents a very small cluster for the reference group in the training set and other class of the same corresponding chemical specie. Analytes whose initial training set distribution (figure 4) does not have parallel components to the ones of the reference group, like ammonia 0.05%, show larger variance. Due to drift, corrected validation data is also slightly displaced from training data and it causes a change in the orientation of some class directions. The final eight groups are in general larger than in OSC plots and poorer in separation.

It can also be noticed that the sparsest clusters in figures 8 and 9 are the ones corresponding to ammonia. It has already been observed in figure 4, that ammonia is the class whose covariance structure has changed the most from first subset to seventh, due to the effect of long term drift that both signal processing methods are not able to correct.

4.3. Time stability

In figure 10 the performance of PCA-CC and OSC drift counteraction techniques are plotted in the form of classification rates (CR) for corrected and non-corrected data. Dataset dimensionality has been reduced to 10 PCs, capturing about 99% of variance, previous to the k-NN classifier. CR is calculated over nine test sets ordered in time and models are built with 308 samples from calibration set. Results show that both methods systematically improve the classification rate. Obviously, the correction is not perfect, and the CR still decreases with time.

OSC-8 (8 components) method outperforms OSC-1 (1 component) and PCA-CC providing higher CR values along the first month of measurements (first validation set). After first month, OSC-1 clearly outperforms the other two methods until nearly day 100. OSC-8 outperforms PCA-CC until around day 100. Later, all methods become comparable, although PCA-CC presents the more unstable behavior. Regarding sensitivity to training set samples, in general PCA-CC shows the smallest value and OSC-1 presents a lower error bar than OSC-8. This sensitivity was also observed in the Fisher Ratio figure (figure 6). We can conclude that OSC-1 is in general the most suitable method for drift correction. However, if the data is very close in time to training data OSC-8 is the most effective one. Up to 100 days both OSC methods outperform PCA-CC.

The main reason of the different performance of both methods is the way they calculate the components of variance to be removed. The main component obtained by PCA-CC has an orientation given by the direction of the largest variance of a reference group and, if more components are to be calculated, these ones are orthogonal to it. This makes that the removed components may be parallel to important directions for posterior classification, since it is not possible to control their direction. On the other hand, PCA-CC only takes samples from one group of reference to build a model, unlike OSC methods, which takes samples from all groups. PCA-CC reference group must be representative of all groups; therefore the more similar the groups are the better this method performs. As a result PCA-CC depends strongly in the selection of the reference.

Removed components in OSC are calculated considering information from all groups. The information regarding sample distribution is in matrix Y , which contains the sample memberships to each of the 8 groups. Y is a matrix with dimensions *number of samples* \times *number of groups* and binary elements; a '1' at row i and column j means that sample i belongs to group j . The condition of orthogonality assures that extracted components of variance are not parallel to important directions. Furthermore, OSC components are not forced to be orthogonal among them. These facts mean that OSC gives very good results for samples very close in time to the training set (little affected by drift), outperforming PCA-CC until around day 100. Later it degrades smoothly along the time resulting in a

better and more stable behavior than non-corrected data, and even more stable than PCA-CC.

For both methods, selection of parameters is a critical step. More components or more strict tolerance (in the case of OSC) would result in a better-fitted model to training samples but in a loss of ability of generalization to correct posterior data affected by drift. This fact can be observed on comparing OSC-1 and OSC-8 results: OSC-8 outperforms OSC-1 only on correcting samples from first validation set.

In fact, in both cases the main difficulty resides in estimating from a short period (calibration set duration) the directions of drift. For the method to be effective, the system should exhibit at least stability in the statistical properties of the variance structure. If the structure of the noise/drift variance changes in time beyond the calibration phase, obviously the methods have no capability to re-adapt.

On the other hand, OSC classification results show higher uncertainty than PCA-CC. We attribute this increased variance to the fact that it takes samples randomly from the eight groups in the training set, not only one as in PCA-CC. In consequence, the identified sources of variance may vary depending on the particular samples that are included in the training set

The results indicate that this preprocessing is also useful to obtain better results also in times close to the training set. Methods are effective not only correcting drift, but also other sources of variance.

4.4. Effect of the training set size

From a practical point of view, for these methods to be useful, they have to be able to correct data with a limited number of calibration samples. Figure 11 shows the performance of both methods as a function of the training set size for the first validation set. All curves present better results with increasing number of training samples. However, it is remarkable that very good results can be obtained with very few training samples. The number of training samples in figure 11 corresponds to samples belonging to all classes, thus valid for OSC and the raw data case. However, PCA-CC only uses training samples belonging to one group; therefore, the effective number of training samples is 1/8th of what is shown in horizontal axis. It is remarkable that PCA-CC is able to improve the results only with 10 training samples from the reference class.

PCA-CC classification rates stabilize with fewer training samples than OSC. The reason for this is the need of OSC to contain several samples from all classes to build a model. Again, PCA-CC only needs several samples from one class.

The apparent advantage for PCA-CC regarding the number of samples needed (1/8 of the total calibration set) is not absolutely true, because samples from the rest of the classes are also used to build the classifier. However, in the current scenario OSC uses all the information available to build the model.

Nevertheless a different scenario may be clearly envisioned, that of using additional future samples for recalibration. In PCA-CC recalibration samples come only from a single reference analyte and concentration, while OSC would require samples for all classes and conditions. Of course, the use of a single substance and not the whole set of analyte species would be a clear advantage for choosing PCA-CC in this second scenario.

5. Conclusions

Lack of stability over time or drift, is a main drawback for the use of chemical sensor array based instruments.

In this work, a drift compensation technique based on Orthogonal Signal Correction

(OSC) has been proposed and compared to Component Correction (CC) method.

The results clearly show that the application of these pre-processing techniques greatly improves the data distribution resulting in smaller clusters with better separation and consequently better discrimination. Additionally, and from an interpretational point of view, the analysis of the loadings of the removed components help to identify the sources of unwanted variability. It may help to identify particularly unstable sensors, or areas of the transient signal that are not stable. In this particular dataset, results show that times very close to gas switching in the sampling system are rather unstable.

Results show that OSC outperforms PCA-CC for a limited period of time (about 100 days in present example), while later on the advantage is not clear. Complex OSC models do a better correction of variance for shorter times, but they degrade faster than simpler OSC that remain stable for a longer time.

It is also important to remark that both methods are relatively robust regarding small calibration set sizes and perform rather well with a reduced calibration set. OSC results show a higher variance than PCA-CC, but this may be a result of the number of components chosen in the models. On the other hand, PCA-CC needs a smaller training set and a single chemical species. This advantage may turn into disadvantage if the reference class is not properly chosen.

Since gas sensor arrays systems are typically plagued with stability problems, the authors would like to emphasize the importance of using validation methodologies that use test samples interspersed in time outside of the training set and acquired subsequent to the training set.

In summary, these two drift counteraction techniques provide better performance over posterior classifiers or regression methods by removing unwanted variance from the data. Although improving the performance of any posterior data processing, they do not completely solve the problem of drift. While recalibration of the instrument is still necessary, the time between recalibrations can be extended.

6. Acknowledgments

This work was partially funded from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 216916: Biologically inspired computation for chemical sensing (NEUROCHEM).

7. References

- [1] Persaud y Dodd. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature*, 299 (1982) 352-355
- [2] R. Ionescu, A. Vancu, A. Tomescu, Time-dependent humidity calibration for drift corrections in electronic noses equipped with SnO₂ gas sensors, *Sensors and Actuators B-Chemical*, 69 (2000) 283.
- [3] C. Di Natale, E. Martinelli, A. D'Amico, Counteraction of environmental disturbances of electronic nose data by independent component analysis, *Sensors and Actuators B-Chemical*, 82 (2002) 158.
- [4] S. Roussel, G. Forsberg, P. Grenier, V. Bellon-Maurel, Optimisation of electronic nose measurements. Part II: Influence of experimental parameters, *Journal of Food Engineering*, 39 (1999) 9.
- [5] R.K. Sharma, P.C.H. Chan, Z.N. Tang, G.Z. Yan, I.M. Hsing, J.K.O. Sin, Investigation of stability and reliability of tin oxide thin-film for integrated micro-machined gas sensor devices, *Sensors and Actuators B-Chemical*, 81 (2001) 9.
- [6] P. Mielle, Managing dynamic thermal exchanges in commercial semiconducting gas sensors, *Sensors and Actuators B-Chemical*, 34 (1996) 533.

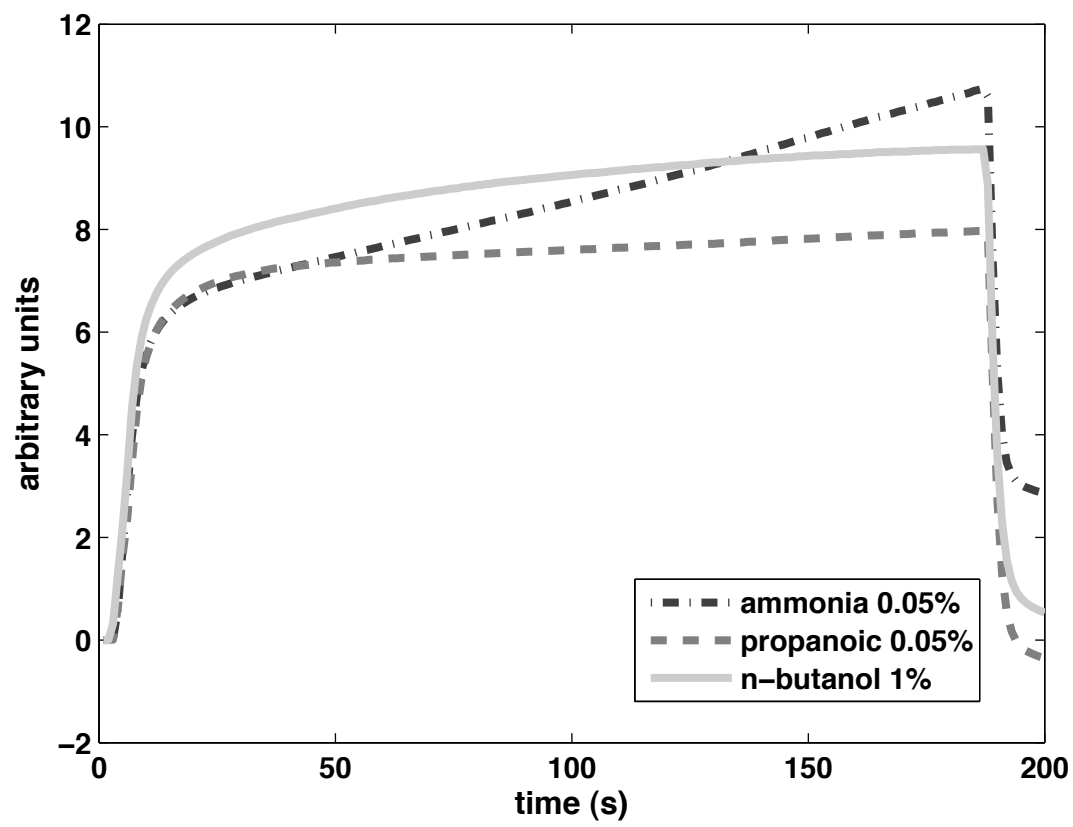
- [7] B. Ruhland, T. Becker, G. Muller, Gas-kinetic interactions of nitrous oxides with SnO₂ surfaces, *Sensors and Actuators B-Chemical*, 50 (1998) 85.
- [8] M. Burgmair, M. Zimmer, I. Eisele, Humidity and temperature compensation in work function gas sensor FETs, *Sensors and Actuators B-Chemical*, 93 (2003) 271.
- [9] R. Paris, S. Pawel, R. Herzer, T. Doll, P. Kornetzky, R.P. Gupta, G. Eranna, *Proceedings of IEEE Sensors 2002. First IEEE International Conference on Sensors (Cat. No.02CH37394)*
- [10] G. Muller, A. Friedberger, P. Kreis, S. Ahlers, O. Schulz, T. Becker, A MEMS toolkit for metal-oxide-based gas sensing systems, *Thin Solid Films*, 436 (2003) 34.
- [11] Gramm, A. Schutze, High performance solvent vapor identification with a two sensor array using temperature cycling and pattern classification, *Sensors and Actuators B-Chemical*, 95 (2003) 58
- [12] M. Roth, R. Hartinger, R. Faul, H.E. Endres, Drift reduction of organic coated gas-sensors by temperature modulation, *Sensors and Actuators B-Chemical*, 36 (1996) 358.
- [13] E.L. Hines, E. Llobet, J.W. Gardner, Electronic noses: a review of signal processing techniques, *IEEE Proceedings-Circuits Devices and Systems*, 146 (1999) 297
- [14] J.E. Haugen, O. Tomic, K. Kvaal, A calibration method for handling the temporal drift of solid state gas-sensors, *Analytica Chimica Acta*, 407 (2000) 23.
- [15] M. Holmberg, F. Winquist, I. Lundstrom, F. Davide, C. DiNatale, A. Damico, Drift counteraction for an electronic nose, *Sensors and Actuators B-Chemical*, 36 (1996) 528.
- [16] S. Marco, A. Ortega, A. Pardo, J. Samitier, Gas Identification with Tin Oxide Sensor Array and Self-Organizing Maps: Adaptive Correction of Sensor Drifts *IEEE Transactions on Instrumentation and Measurement*, 47 (1998) 316.
- [17] M. Zuppa, C. Distanto, P. Siciliano, K.C. Persaud, Drift counteraction with multiple self-organising maps for an electronic nose, *Sensors and Actuators B-Chemical*, 98 (2004) 305.
- [18] M. Holmberg, F.A.M. Davide, C. DiNatale, A. Damico, F. Winquist, I. Lundstrom, Drift counteraction in odour recognition applications: lifelong calibration method, *Sensors and Actuators B-Chemical*, 42 (1997) 185.
- [19] H. Ding, H.F. Ge, J.H. Liu, High performance of gas identification by wavelet transform-based fast feature extraction from temperature modulated semiconductor gas sensors, *Sensors and Actuators B-Chemical*, 107 (2005) 749.
- [20] H. Ding, J.H. Liu, Z.R. Shen, Drift reduction of gas sensor by wavelet and principal component analysis, *Sensors and Actuators B-Chemical*, 96 (2003) 354.
- [21] E. Llobet, J. Brezmes, R. Ionescu, X. Vilanova, S. Al-Khalifa, J.W. Gardner, N. Barsan, X. Correig, Wavelet transform and fuzzy ARTMAP-based pattern recognition for fast gas identification using a micro-hotplate gas sensor, *Sensors and Actuators B-Chemical*, 83 (2002) 238.
- [22] T. Artursson, T. Eklov, I. Lundstrom, P. Martinsson, M. Sjostrom, M. Holmberg, Drift correction for gas sensors using multivariate methods, *Journal of Chemometrics*, 14 (2000) 711.
- [23] O. Tomic, T. Eklov, K. Kvaal, J.E. Haugen, Recalibration of a gas-sensor array system related to sensor replacement, *Analytica Chimica Acta*, 512 (2004) 199.
- [24] R. Gutierrez-Osuna, Drift reduction for metal-oxide sensor arrays using canonical correlation regression and partial least squares, in: *Electronic Noses and Olfaction, 2000: Proceedings of the 7th International Symp. On Olfaction and Electronic Nose*, Brighton, UK, July 20-24, 2000. Institute of Physics Publishing, p. 147.
- [25] O. Tomic, H. Ulmer, J.E. Haugen, Standardization methods for handling instrument related signal shift in gas-sensor array measurement data, *Analytica Chimica Acta*, 472 (2002) 99.
- [26] J.E. Carlson, P.E. Martinsson, Exploring interaction effects in two-component gas mixtures using orthogonal signal correction of ultrasound pulses, *Journal of the Acoustical Society of America*, 117 (2005) 2961.
- [27] S. Wold, H. Antti, F. Lindgren, J. Ohman, Orthogonal signal correction of near-infrared spectra, *Chemometrics and Intelligent Laboratory Systems*, 44 (1998) 175.

- [28] T. Fearn, On orthogonal signal correction, *Chemometrics and Intelligent Laboratory Systems*, 50 (2000) 47.
- [29] J. Sjöblom, O. Svensson, M. Josefson, H. Kullberg, S. Wold, An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra, *Chemometrics and Intelligent Laboratory Systems*, 44 (1998) 229.
- [30] J.A. Westerhuis, S. de Jong, A.K. Smilde, Direct orthogonal signal correction, *Chemometrics and Intelligent Laboratory Systems*, 56 (2001) 13.
- [31] Li, A. J. Morris, and E. B. Martin, Orthogonal signal correction: algorithmic aspects and properties, *Journal of Chemometrics*, 16 (2002) 556
- [32] R. N. Feudale, H. Tan, S. D. Brown, Piecewise orthogonal signal correction, *Chemometrics and Intelligent Laboratory Systems*, 63 (2002) 129 – 138
- [33] J. Trygg, O2-PLS for qualitative and quantitative analysis in multivariate calibration, *Journal of Chemometrics*, 16 (2002) 283
- [34] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), *Journal of Chemometrics*, 16 (2002) 119
- [35] O. Svensson, T. Kourti, J.F. MacGregor, An investigation of orthogonal signal correction algorithms and their characteristics, *Journal of Chemometrics*, 16 (2002) 176.
- [36] H. C. Goicoechea and A. C. Olivieri, A comparison of orthogonal signal correction and net analyte preprocessing methods. Theoretical and experimental study, *Chemometrics and Intelligent Laboratory Systems*, 56 (2001) 73.
- [37] B.M. Wise, N.B. Gallagher, <http://www.eigenvector.com/MATLAB/OSC.html>.
- [38] J. Samitier, J.M. Lopez-Villegas, S. Marco, L. Camara, A. Pardo, O. Ruiz, J.R. Morante, A new method to analyse signal transients in chemical sensors, *Sensors and Actuators B-Chemical*, 18 (1994) 308.

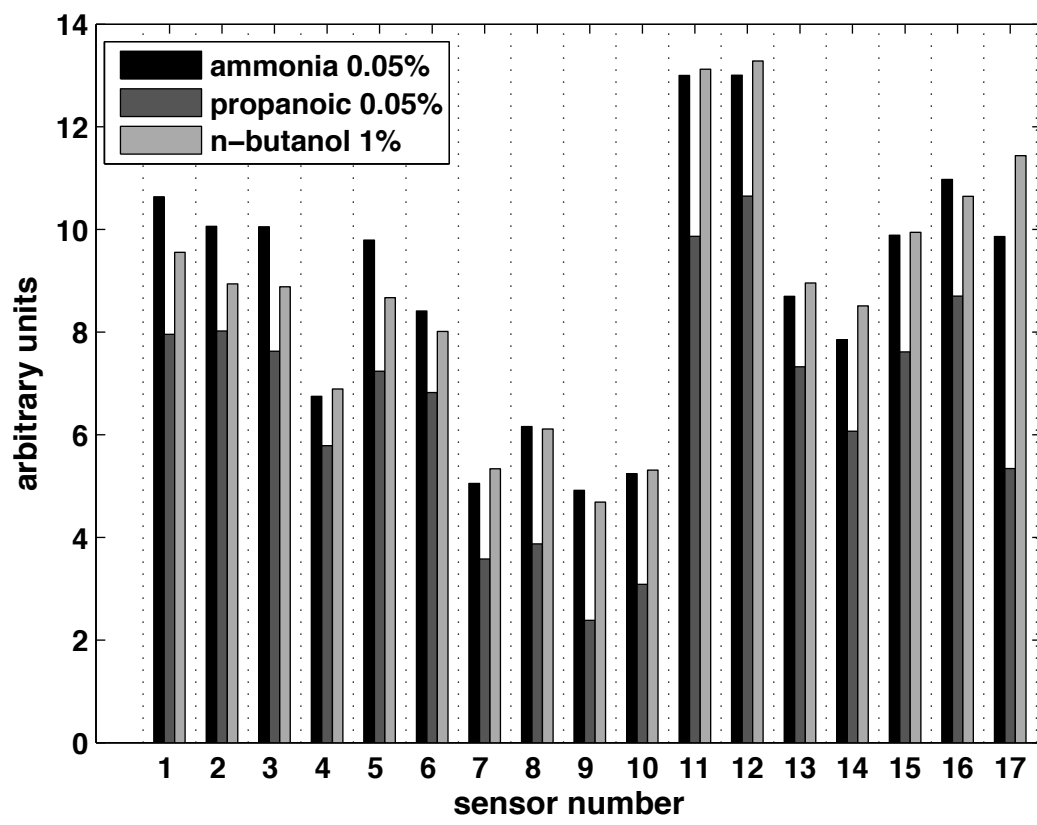
Table 1. Measured compounds, concentrations and number of samples in the dataset.

Analyte	Samples	Concentration level
Ammonia	447	0.01%
	452	0.02%
	307	0.05%
Propanoic Acid	458	0.01%
	457	0.02%
	423	0.05%
n-Butanol	446	0.01%
	425	1.00%

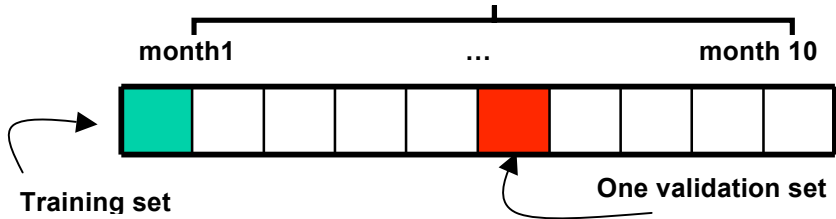
Figure(s)



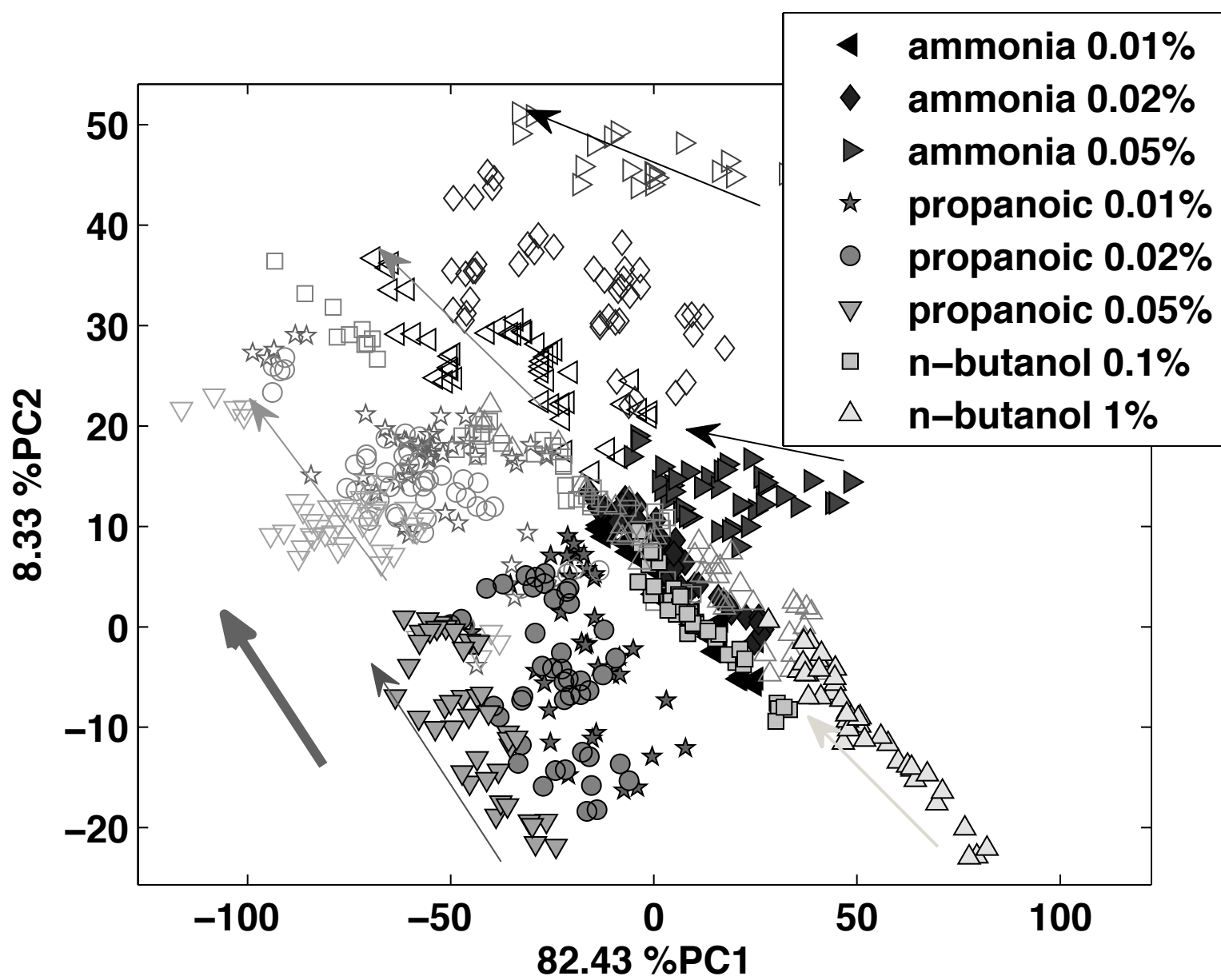
Figure(s)



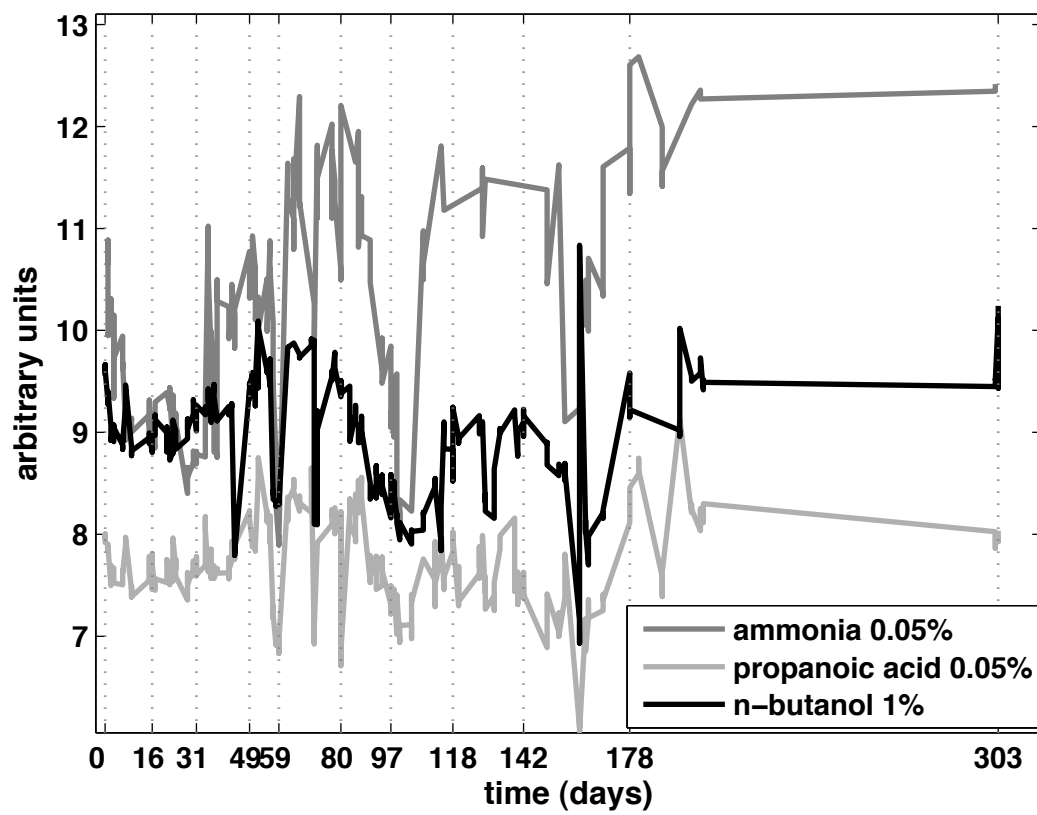
9 Validation sets



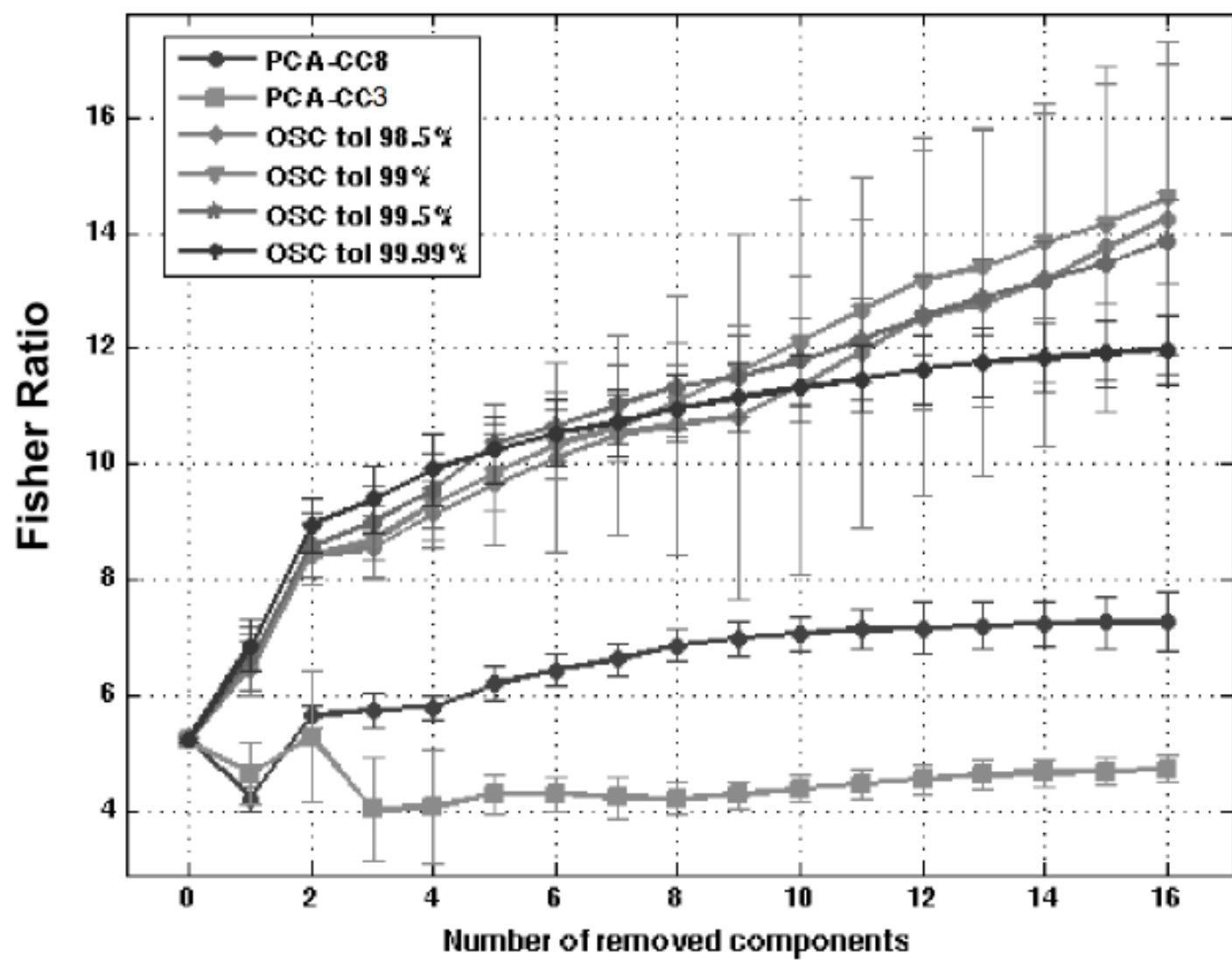
Figure(s)

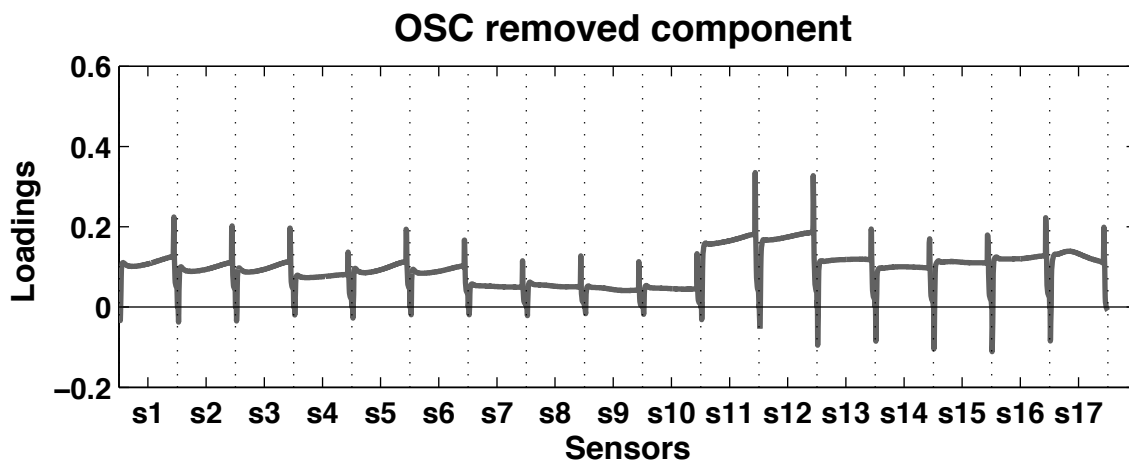
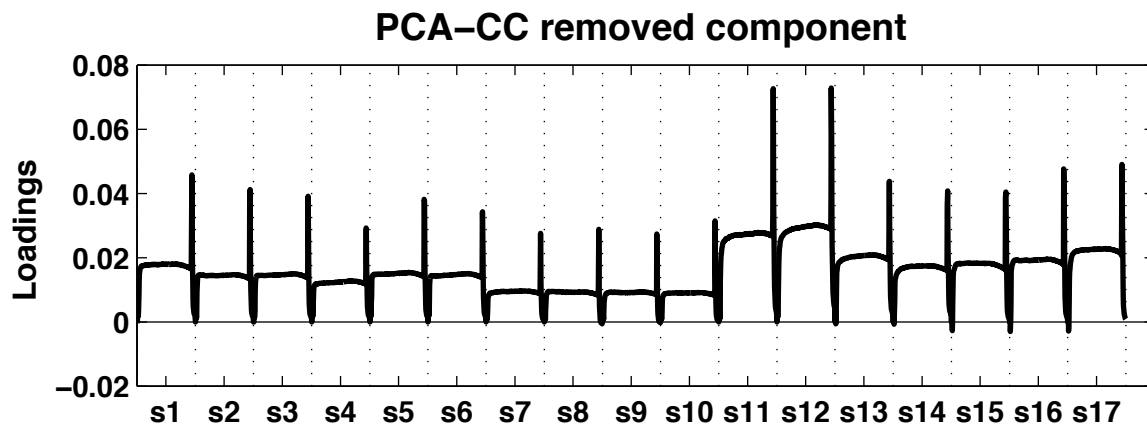


Figure(s)

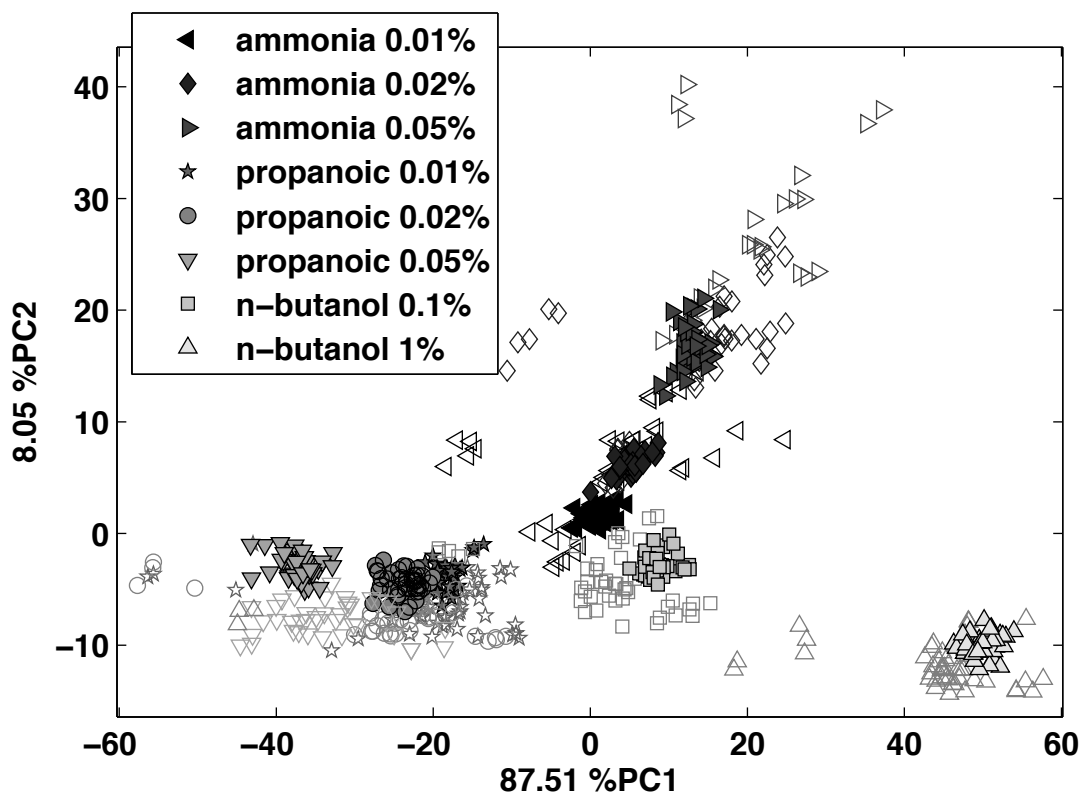


Figure(s)

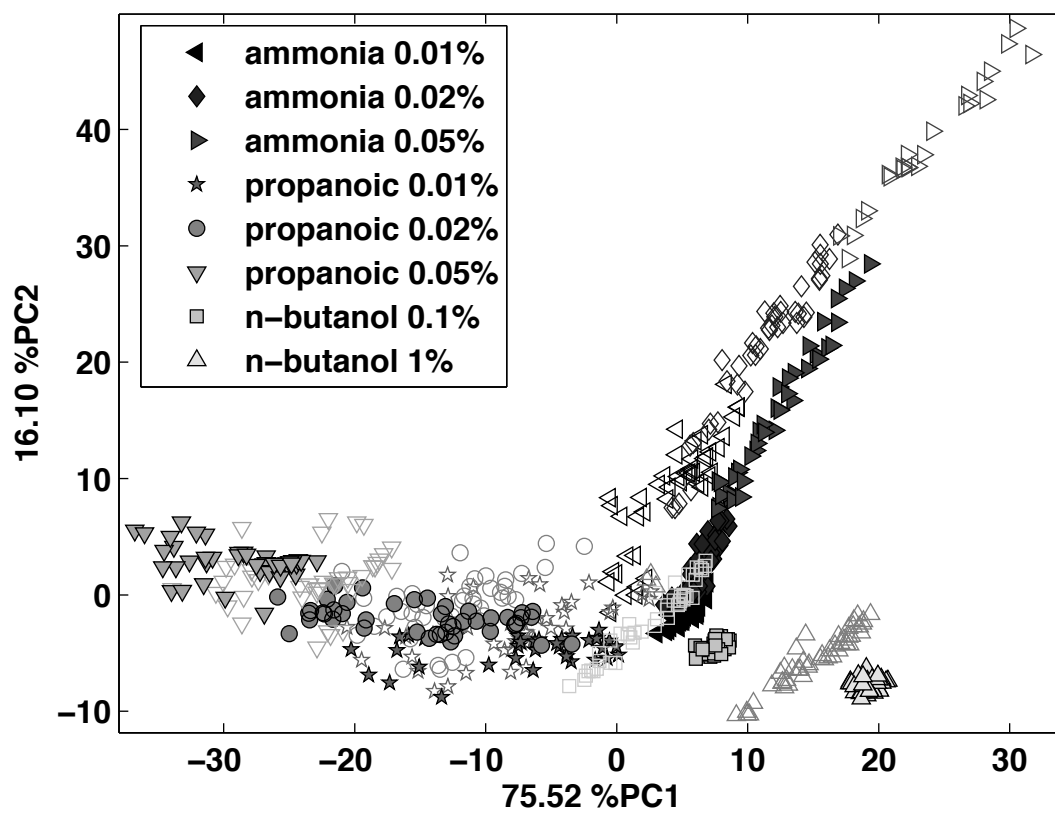




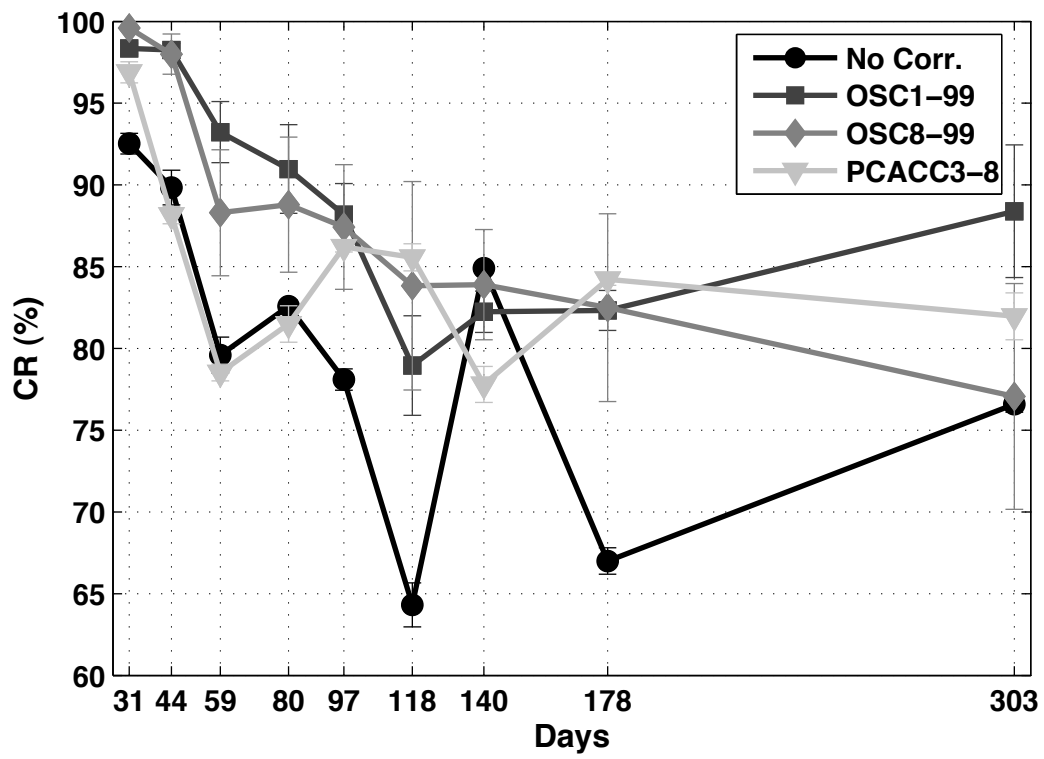
Figure(s)



Figure(s)



Figure(s)



Figure(s)

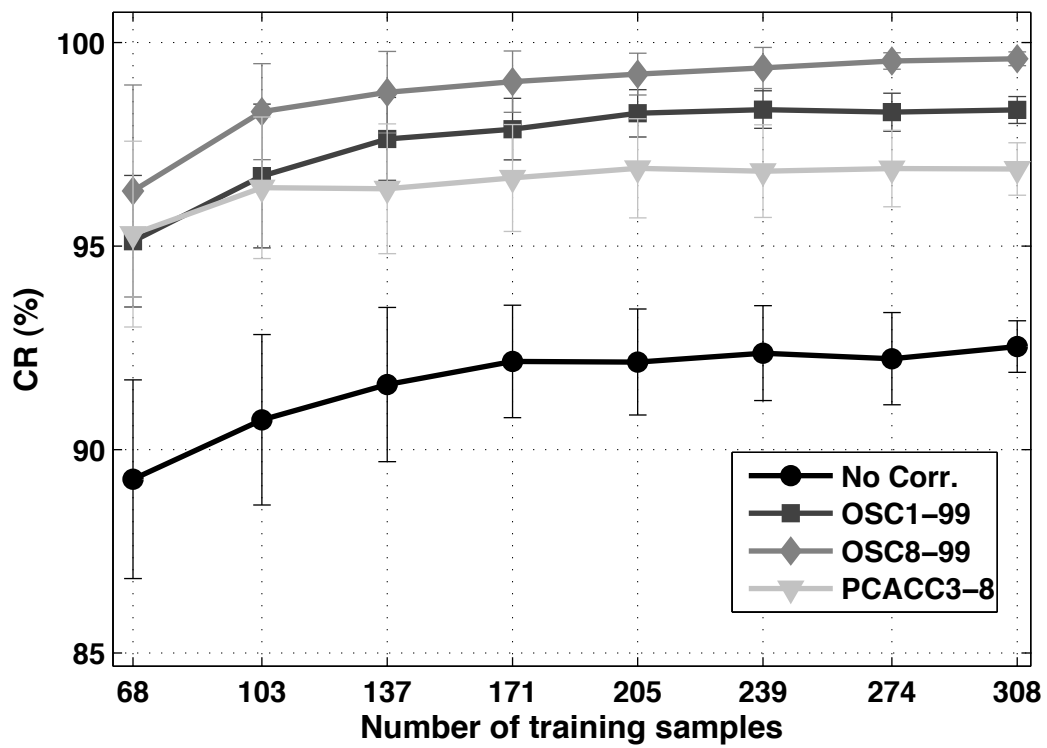


Fig.1 Transient response of sensor 1 to three analytes.

Fig.2 Examples of patterns corresponding to 17 sensors and samples of all classes at the maximum time point of the transient signal (instant 185s).

Fig.3. Scheme of time stability evaluation.

Fig.4 PCA scores of the calibration set (solid symbols) and eighth test set (measurements between day 97 and day 118, in non solid symbols).

Fig.5 Responses of sensor 1 to three analytes along the time.

Fig.6 Fisher Ratio vs. number of OSC and PCA-CC removed components and internal OSC tolerances. PCA-CC is computed with two different groups of reference; ammonia 0,05% (group 3) and n-butanol 1% (group 8). Error bars represents sensitivity to training samples.

Fig.7 First data removed component by PCA-CC and OSC.

Fig.8 PCA scores of the calibration set (solid symbols) and eighth test set (non solid symbols) corrected by OSC-8 (8 removed components).

Fig.9 PCA scores of the calibration set (solid symbols) and eighth test set (non solid symbols) corrected by PCA-CC (3 removed components, group of reference 8).

Fig.10 Classification rates along the time for drift correction by PCA-CC and OSC methods and non-corrected data (No Corr). Calibration set contains 308 samples.

Fig.11 Classification rates for corrected first test subset by PCA-CC and OSC methods and non-corrected data (No Corr) vs. number of samples in the

calibration set.