

On-line Event Detection by Recursive Dynamic Principal Component Analysis and Gas Sensor Arrays under drift conditions

A. Perera¹, N. Papamichail², N. Bârsan², U. Weimar², S. Marco¹

¹Department of Electronics, Universitat de Barcelona
Martí Franqués 1, 08028-Barcelona-Spain. e-mail: santi@el.ub.es

²Institute of Physical Chemistry, University of Tübingen
Auf der Morgenstelle 8, 72074-Tübingen-Germany

Abstract

Leakage detection is an important issue in many chemical sensing applications. Leakage detection by thresholds suffers from important drawbacks when sensors have serious drifts or they are affected by cross-sensitivities. Here we present an adaptive method based in a Dynamic Principal Component Analysis that models the relationships between the sensors in the array. In normal conditions a certain variance distribution characterizes sensor signals. However, in the presence of a new source of variance the PCA decomposition changes drastically. In order to prevent the influence of sensor drifts the model is adaptive and it is calculated in a recursive manner with minimum computational effort. The behavior of this technique is studied with synthetic signals and with real signals arising by oil vapor leakages in an air compressor. Results clearly demonstrate the efficiency of the proposed method.

Keywords

Drift, gas sensors, event detection, electronic nose, PCA, ADPCA, adaptive

INTRODUCTION

A substantial number of chemical sensing applications can be classified as event detection problems. Problems like leakage detection, filter contamination or spoilage, chemical change detection are all types of applications that share fundamental properties. In some high pressure air industrial applications is necessary to maintain a level of quality in the line. Some of these applications include pharmaceutical industry, food industry, or hospitals. The residual oil content of compressed air in food (mineral water, ice-cream production, etc) and pharmaceutical applications must be monitored and kept to the lowest levels for the customer protection. Unfortunately, there is a high probability that an air compressor delivers some oil to the pressure line, producing a non-desired contamination of the rest of equipments that use that pressure for their regular daily work. This oil delivery is controlled by a proper filter installed in the pressure delivery outlet, protecting the line. Nevertheless, in some circumstances, a breakage of the filter can occur, delivering oil content and producing the contamination of the line.

Smart Gas sensor arrays (sometimes referred as electronic noses) are suitable to be used for sensing the quality of the air in the high pressure line. This approach favors an online analysis and monitoring approach that allows for continuous monitoring of the correct functioning of the compressor filter. A gas sensor array based system provides a pre-emptive alarm as soon as malfunctioning occurs. Abnormal functioning can be considered as a strange deviation of the chemical composition of the line pressure.

The early detection of filter breakages has a strong impact on the cost of the application. The time elapsed since the failure is pro-

duced until the line is closed determines the degree of contamination of the line and thus the cost of the line cleaning.

The CMOSSens European project (CMOSSens, IST-1999-10579) aimed the development and characterization of a CMOS based microhotplate gas sensor for different applications. CMOS based technology offers small size, low power consumption, low costs for mass production and possibility of automated coating. The sensitive layer is integrated with CMOS compatible circuitry allowing on board signal acquisition, digital heater temperature control, on chip function tests and on-board signal evaluation. Among the different goal applications within this project, Residual Oil Analysis (ROA) was focused in the early detection of separator breakages and aging. The use of CMOS technology would permit the integration of a sensor array in the same IC and fast temperature modulation over the sensing material [1,2].

Nevertheless, although semiconductor based sensors have good sensitivity for air quality applications, they suffer from severe drift in the framework of long time applications. This instability leads to poor reproducibility of the sensor responses [3,4].

A side effect of sensor drift is the need of a periodic recalibration of the sensor array, which itself produces a strong economic impact. Some methods have been proposed for drift correction, although the most efficient are based in a periodic recalibration of the model [5], and using multivariate linear models [6]. Some other authors proposed techniques based on adaptive classifiers based on self-organizing maps [7,8,9].

Few methods have been proposed for on-line system monitoring drift correction. In this paper presents an adaptive algorithm based in Dynamic Principal Component Analysis that models the regular behavior of the sensor array [10,11]. The variance analysis nature of the algorithm permits the detection of new chemical components in the dynamic headspace; the adaptive nature of the algorithm provides drift rejection to the detection.

This paper is structured as follows, in section II and III an introduction to the theory used in this work is given, section IV describes a description of the experimental setup, section V compares the method proposed with other solutions, section VI presents the results with experimental data and finally section VII ends with some conclusions.

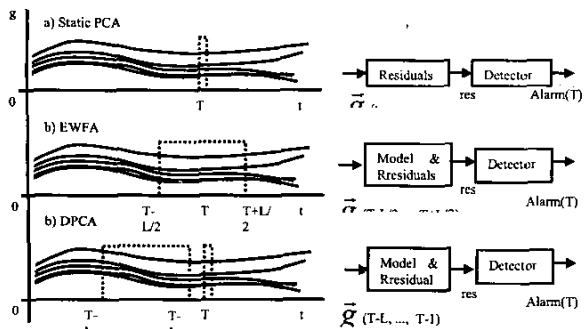


Figure 1 Schematic view of the three models explored in this work. a) Static PCA, b) EWFA and c) Adaptive Dynamic PCA

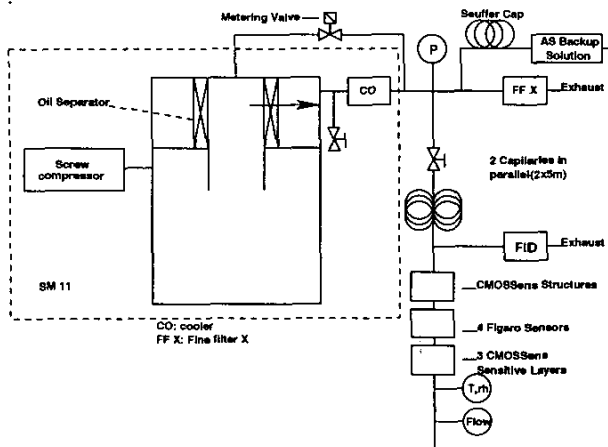
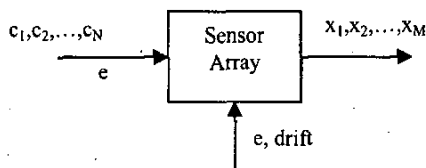


Figure 2. Schematic test setup for ROA application..

THEORY

The problem statement can be formulated as the detection of changes in the chemical composition of a target atmosphere in such a way that long term drift does not influence the performance of the device. A sensor array is actually a transducer that provides the following transformation,



where a map is done from the number of chemical species and temperature ($c_1 \dots c_N, T$) onto a higher dimensionality space given by the number of features extracted from the sensors ($x_1 \dots x_M$). The sensor array will respond to a number of components in the input. Ideally, if we assume that the sensor array responds linearly to the chemical species concentration, the number of different sources of variance in the input space (chemicals, temperature and humidity) will be structurally reflected in the data distribution at the output of the sensor array. For instance,

variations in only two components within the $\{c_i, T, HR\}$ set will produce a distribution in the feature space of only two components of variance. This distribution will obviously form a two dimensional subspace in feature space. The combined effect of the non-linearity of the gas sensor responses, and the noise influence in the system, will produce a regularization of the subspace to a higher dimension. Notwithstanding, this effect on the sensor array extracted features can be assumed small if the variations of temperature and humidity and the variations in the chemical composition of the headspace are assumed locally small.

On the other hand, gas sensor drift will have a different effect on the distribution of the data in feature space. In our example with variation of only two components, the two-dimensional cloud in feature space will evolve. This evolution produces changes in both, the mean position of the samples distribution in this subspace (as drift affects the baseline and response to a given chemical) and the orientation of this distribution in the subspace (as the sensibility of the sensor array is also modified by drift effects). The practical effect of the drift is that invalidates practically any response level based trigger detector for long term operation. In this paper we propose an approach for the detection of new chemicals based on an adaptive variance analysis, namely Adaptive Dynamic Principal Component Analysis, or ADPCA

We will compare the performance of this method with Evolving Window Factor Analysis, or EWFA, static PCA and sensor trigger level and its behavior under heavy drift conditions will be explored.

ADAPTIVE DYNAMIC PRINCIPAL COMPONENT ANALYSIS

Principal Components Analysis (PCA), also known as Karhunen-Loève expansion in information theory [12,13,14], is a analysis of variance technique very common in chemometrics that has been used, very extensively, for visualization purposes of gas sensor array distribution data. PCA is an unsupervised projection that computes an orthogonal subspace that is oriented in such a way that this given subspace explains most of the variance given by the dataset. PCA subspace can be computed by finding the eigenvectors of the covariance distribution of the data.

$$\Sigma_y = A^T \Sigma_x A = \Lambda$$

where Λ is a diagonal matrix, Σ_x is the covariance of the data in the feature space, Σ_y is the covariance of the data in the projected subspace. A is a projection matrix such that Σ_x becomes a diagonal matrix Λ in the projected space.

In the example exposed in the introduction where we had N sources of variance in the data, the subspace obtained would have N orthogonal eigenvectors, $\{a_i; i = 1, \dots, N\}$. With these subspace representation we would be able to explain most of the data,

$$\hat{x} = \sum_{i=1}^N y a_i$$

with a low error.

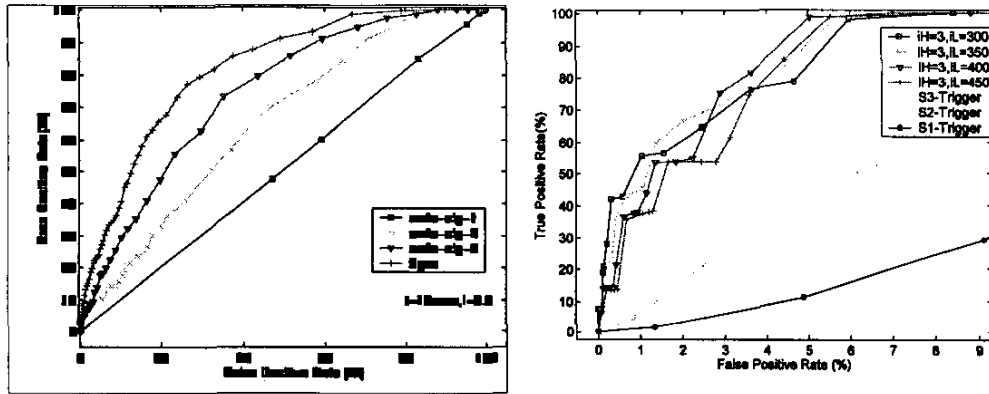


Figure 3. Left: ROC curves corresponding to EWFA filters compared with ADPCA. Right: ROC curves comparing different ADPCA designs and sensor levels.

In an online application like early leak detection, we are provided by continuous stream of data, instead of a finite set of samples. For overcoming drift effects, the central idea of the Adaptive version of PCA is that APCA performs an analysis of variance only in a time-window set. The model captures the variance of the data only in this time-window. If within the time-window there are only N sources of variance (only small variations in humidity and temperature, for instance), the APCA model will capture N meaningful principal components.

The application of this method to chemical event detection is based in the idea that the appearance of a new component in contact to the sensor array will produce a new source of variance. The dynamic nature of the ADPCA is given when past samples are considered when building the features of the covariance matrix. By this way, an order h and length L data matrix will be constructed by

$$X_{h,L} = \begin{pmatrix} x_t^T & \cdots & x_{t-h}^T \\ \vdots & \ddots & \vdots \\ x_{t-L}^T & \cdots & x_{t-L-h}^T \end{pmatrix}$$

where t is the current time, n is the window length to be considered and h is the lag to be added in the feature set. x_{t-L-h}^T represents a row vector with the readouts of the sensor array in the instant $(t-L-h)$, different vectors at different lags are added together to form the new feature vector. As shown in Figure 1, this matrix is build as a sliding window form as new data is acquired from the sensor array.

The natural detection of new sources of variance can be done by monitoring the eigenvalues of each variance analysis at each time-window step. When data matrix is build without lag ($h=0$) and eigenvalues are monitored, we recover a very similar technique named Evolving Factor Window Analysis (EFWA). This technique captures effectively the new sources of variance Evolving Window Factor Analysis and has its roots on the Meader's EFA (Evolving Factor Analysis, see [15]). EWFA [16] is a version of EFA with a fixed size window instead of an increasing size window. However, the monitoring of the eigenvalues is not appropriate to early chemical detection as it contains a natural delay. There will be a noticeable change in the eigenvalue when the window captures an important portion of data containing the new component.

In this paper we propose to achieve a much faster response by measuring the residuals of the incoming sample versus an ADPCA model computed iteratively on each incoming sample. The residuals are computed as the distance of the sample to the ADPCA subspace, which can be computed as

$$q_j^2 = (x_i - \sum_{j=1}^N y_j a_j)^T (x_i - \sum_{j=1}^N y_j a_j).$$

Where, the coefficients a_j form the A projection matrix computed from X , x_i is the incoming sample, and y_i is the projected sample in the PCA subspace (score vector). The trigger levels and other parameters of the ADPCA filters are analyzed by the Receiver Operating Characteristic (ROC), where the quality goal is to minimize the rate of false alarms (false positives) and the rate of non-detected events (false negatives). A false positive is an indication of an event or fault when this has not actually occurred. A false negative is obviously when the event or fault is not detected but has occurred. The ROC curve is the plot of the true alarm rate versus the false alarm rate when a parameter controlling the detector sensitivity is varied, for extended information please see [17, 18]. ROC plots are computed in the following form, a given trigger level is defined for the output of a given filter. The number of False Positives (FP, event captured in absence of real event) and the number of True Positives (TP, events captured when a real event is produced) is plotted for each filter for different trigger levels. When the trigger is low, most of real events are correctly captured (TP \rightarrow 100%) but a considerable number of False Positives is generated (FP \rightarrow 100%). On the other side, when trigger levels are high, a few number of false events are triggered (FP \rightarrow 0%) but the efficiency of capturing real events is decreased as well (TP \rightarrow 0%). The optimum ideal filter would present zero false positives and 100% of true positives. As a general rule a better filter is the one that presents a convex curve closer to the upper-left corner. A detector that does not provide discrimination would draw a diagonal curve from bottom-left to upper-right coordinates.

EXPERIMENTAL SETUP

Data for ROA application was generated at IPC (Institute of Physical Chemistry, University of Tübingen) with an experimental setup that involved screw-type compressor (Kaeser SM11) conveniently modified.

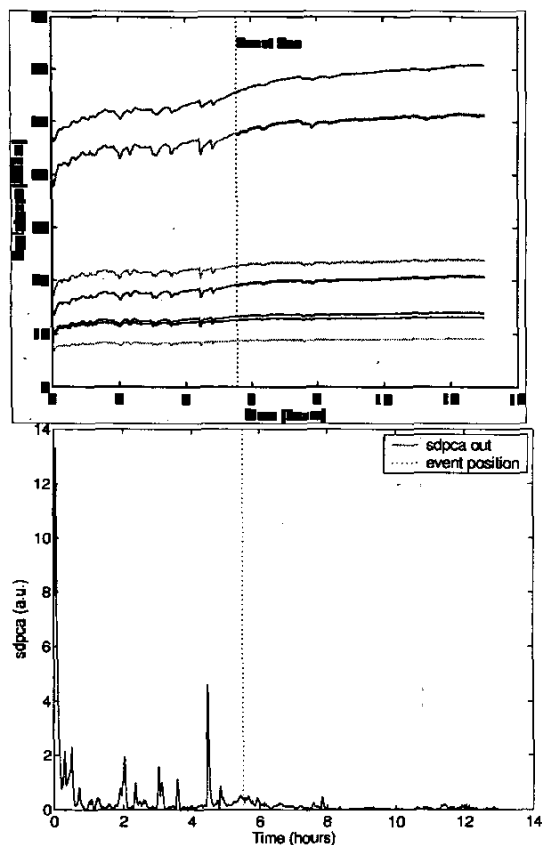


Figure 4. Top: 12-hour sensor array signals readout. A small synthetic event is produced around hour 6. Bottom: Static PCA residuals plot on the same data.

A by-pass valve was mounted in parallel to the oil separator, as shown Figure 2. The compressor generates 700 kPa overpressure and delivers a flow of 1 m³/min. The sensor chamber is placed after a stainless steel capillary of five meter length and an inner diameter of 180 μm. This capillary reduces the 700kPa and generates a flow of 50 ml/min. The maximum overpressure found after the capillary is of 50kPa. This flow reaches the sensor chambers that hold four commercial sensors (FIGARO) and three different CMOSens sensors with different dopands, named S1, S2, S3. The sampling frequency for each sensor readout is established to 5 seconds. The MOX sensor chamber is followed by commercial sensors for temperature, humidity (Testo 601) and flow (Sensorion AFM1400). Hydrocarbons level is measured by a Flame Ionization Detector (FID) detector. Additionally, a pressure gauge is installed downstream the oil separator bypass. The setup is done in such a way that permits the forced and controlled contamination of the pressure line. This is done by the bypass valve shown on the upper left side. This valve permits the introduction of controlled oil doses into the main air flow. The maximum oil dose is around 150mg/m³. Unfortunately, valve was not able to provide accurate and reproducible oil doses, yielding to non repetitive injections that ranged from 34 to 75mg/m³. However, the typical dose provided with a single valve turn corresponded approximately to a quantity of oil of 35mg/m³. A new oil separator provides a purity of less than mg/m³. For proper operation within the lifetime of the separator, this value should not rise above 30 mg/m³. Further details about the experimental setup and measurements can be found in [19]. First examination of the sensor signals showed clear drift effects partly correlated with temperature. The sensor signals also show some spikes that were attrib-

uted to water drops at the end of the capillary. These signal variations are a potential source to mask oil leakages.

Algorithm characterization

In the following sections we want to evaluate the ROC sensibility provided by different algorithms. For that goal a number of small events should be generated in order to measure the discrimination power of each algorithm. Unfortunately it is not possible to generate small doses with present setup. For these reason a two-step approach is taken for the evaluation of the performance of ADPCA. First, random small events will be generated over continuous real data stream like shown in Figure 4 in order to evaluate sensibility of the methods. On a second step, ADPCA detection will be tested on long term data under drift conditions and real events. First, artificial abrupt changes are generated in 12 hour real data. We compare ADPCA, EWFA, Static PCA and level-triggered signals over single sensors conductance. A plot of 12 hour continuous data is shown in Figure 4. In order to understand the behavior of different approaches we generate an artificial event on the 12-hour signals, which is shown around the sixth hour in the same figure. This event is generated virtual events are generated from the following form:

$$I_{add} = I_{max} S_s (1 - e^{-t/\tau})$$

where I_{max} is the intensity of the event, τ is the time constant of the event and S_s is the sensibility of the sensor, which varies randomly (and independently) for each sensor and event within the range $[-\sigma_s, 0]$, where σ_s is the sensor variance in the whole acquisition (here this has been computed over a 12 hours period).

Virtual events with intensity smaller than σ_s will be difficult to detect since they are on the order of magnitude of the sensor variability under normal conditions. In this analysis, we have processed the resistance signals so events are characterized by a decrease of the signals. In the synthetic experiments one hundred events at $I_{max}=1$ were generated in each simulation whose position was randomly located over normal operation signals. These kinds of events are randomly generated on the data in a non-overlapping manner. In the analysis performed in this work, we have used a constant number of components for ADPCA, $N=2$, under the assumption that only two main components of variance will be present in the data as oscillations of temperature and humidity

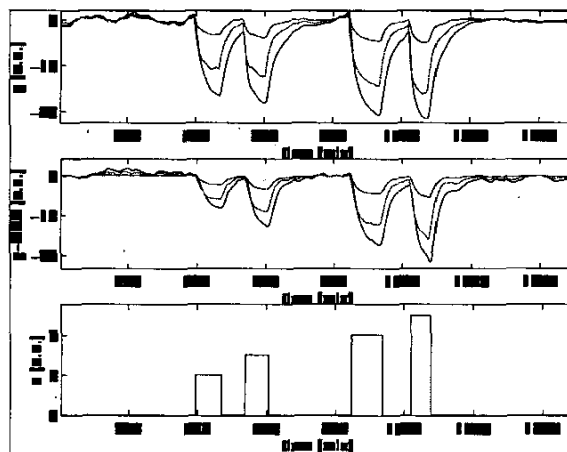


Figure 5 Top: Real three sensor data stream with four oil doses of increasing magnitude. Middle: Result of the ARX model adjusted from real data. Bottom: Input signal representing the oil dose magnitude applied to the line

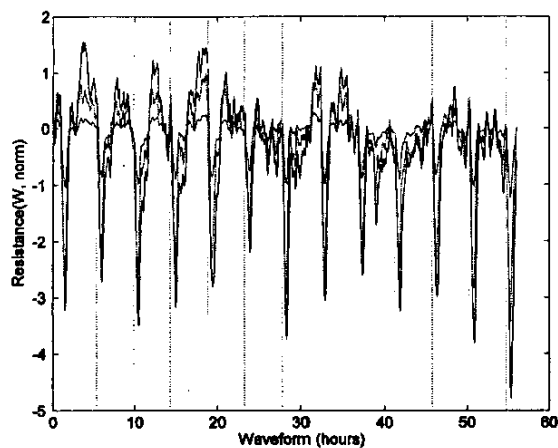


Figure 6. 80 hours of data generation from ARX model. Events are produced every 4.5 hours

From the time being, only the three sensors corresponding to the CMOSsens sensing layers will be used for the generation of the results. The three main eigenvalues of EWFA and ADPCA are compared in ROC plot shown in Figure 3(left).

As expected, the different eigenvalues (in decreasing order) retrieved by EWFA give increasing recognition rates. The first eigenvalue (diagonal in the ROC of Figure 3 left) has no recognition power as it captures the evolution of the data probably due to small fluctuations in temperature and humidity. Nevertheless the third eigenvalue provides the highest quality ROC curve. This fact confirms, to a certain degree, that the first two components of variance captured by the array are not related to the problem, and strengthens the hypothesis that ADPCA should consider only two principal components. Static PCA did not provide meaningful results with low-intensity events, defined as $I_{\max} < 1$ (as shown in Figure 4(bottom)), while ADPCA outperformed clearly EWFA in capturing low intensity data events. These results can be compared with the selection of the sensor conductance as a detector. The ROC plot produced by the sensors will provide the optimum level trigger in conductance. As shown in Figure 3 (right), the sensors itself produce poor levels of event recognition. This fact suggests that the combined information of different sensors is a better approach than the use of a single sensor, and effectively detect new events.

Long term data analysis

The experimental setup depicted in the previous section permits to dose arbitrary amount of oil content into the pressure line. The oil is in form of aerosol and controlled by needle a valve mounted in bypass to the oil separator. The ideal dataset we expect to have is continuous stream of sensor readout with controlled oil doses evenly generated with time. The amount of events should be rather high in order to have a statistical meaningful filter design and still have data for validation purposes.

Unfortunately, the procedure of generating data with real events is complex, slow and expensive, and impedes the creation of long term data with a representative amount of events on it. In order to overcome this situation an alternative approach is taken in order to find the optimum filter design, suitable for real data analysis. The basis of the approach is to design a controlled experiment with real data and real oil dosing and build a model of this data. This model capture the random oscillation found in the signal and the effect of the oil dose. This model will permit arbitrary generation of long-term sensor signals and the optimization of the detec-

tor with real-like events. We used a linear ARX ([20, 21] model for capturing the sensor array evolution with increasing doses of oil content. The output of the array will be given by

$$A_0 y(T) + A_1 y(T-1) + \dots + A_n y(T-n) = B_0 u(t) + B_1 (t-T) + \dots + B_m (t-T) + e(T)$$

where A_i is a (3x3) matrix as we have three sensors and B_j is a three dimensional vector. It is found that the model presented a good correspondence with the experimental data provided by choosing $n = 30$ and $m = 4$. The experimental data and the input signal used in the modeling are shown in the Figure 5 (top, bottom). Once the model is adjusted a stream is generated by the ARX model from a random number generator, representing a white noise source. The expression for the data generation can be found from previous equation isolating $y(t)$ and considering that $e(t)$ is a three dimensional vector random number generator. Although this model reproduces qualitatively the behavior of the sensor array, it is only a coarse model that provides a linear approximation to a non-linear system. Nevertheless, the model will permit the generation of long term data and provide qualitative results. The result of a realization of the filter is shown in Figure. The virtual oil doses are generated evenly each four and a half hours. Data is generated with a heavy component of low frequency noise that reproduces, to some extent, drift observed in the sensor array. Over a signal of 450 hours the filter was optimized with help of ROC curves generated by exploring $h = \{1,3,5\}$ and L between 10 and 140 seconds in steps of 10 seconds. The optimum filter is found at $h=3$ and $L=70$ secs.

DRIFT EFFECT AND DISCUSSION

In order to test the robustness of the method, a linear drift component was added in the ARX sensor generation signal. The events produced are of low amplitude $u=1$, that represents a value of 37mg/m^3 of oil dose. The drift component added was deliberately high and its slope was equivalent of six times the sensor variance every 25 hours. The performance of ADPCA on the signals without drift component is shown in Figure 7. The figure shows the original signal a) and the sensor array signal under heavy drift in b). The output of ADPCA over heavy drift signals in d) has lowered slightly its output but still able to detect clearly the events. Actually ADPCA captures not only the event produced (a new chemical is in contact with the sensor array) but when this chemical is removed from the headspace. This effect is seen as the second peak just after detecting the oil dose. The results of ADPCA under heavy drift provide a new approach based in sliding window variance analysis. The variance analysis captures the principal components over a time-window slice captures. These sources of variance are usually assigned to small variations of humidity and temperature. When a new chemical is present in the headspace, the residuals of the first incoming samples will be high as the new component is not captured by the time-window history. On the other hand the adaptive nature of ADPCA provides drift robustness as drift is typically a long time constant process. As ADPCA will find sources of variance only in the time-window segment of length L , the PCA subspace will adapt slowly following the evolution of the drift effect. These reasons explain that a residual analysis of an adaptive variance model provides good results at detecting new chemicals in the headspace, while rejecting slow drift variations of the sensor array calibration. Nevertheless a number of issues are still to be explored. The modeling of the sensor array is a simple linear model and may not reflect the intrinsic non-linear behavior of metal oxide sensors. On the other hand, ADPCA would trigger an alarm when any new chemical appears in the headspace.

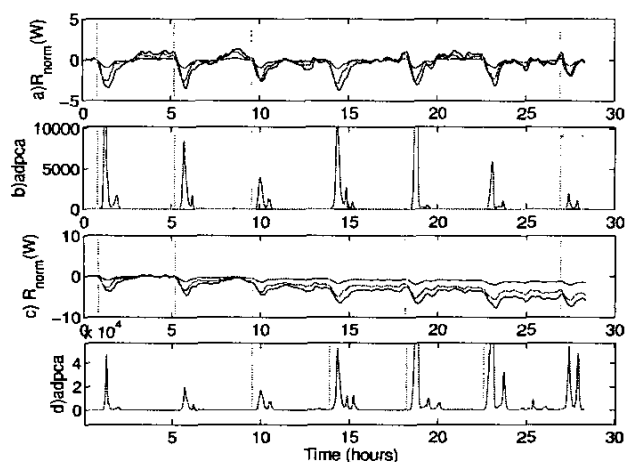


Figure 7. From top to bottom: a) original signal b) ADPCA based detector for original signal. c) sensor array signals under heavy drift, d) resulting ADPCA output

It is known that in ROA applications they may appear relatively high concentrations of CO that could trigger the ADPCA alarm. Further methods should be considered for the correct discrimination of the oil aerosols from other compounds present at the air intake of the compressor. These methods should be based in the adaptive nature of ADPCA in order to maintain a reasonable drift rejection effect.

CONCLUSIONS

In this paper the use of a novel adaptive approach based in sliding window variance analysis namely Adaptive Dynamic Principal Component Analysis is proposed for online gas array applications. This method has been compared against Evolving Window Factor Analysis and the static version of PCA with favorable results. The use of signal processing methods like ADPCA for online chemical detection provides much better and reliable results than the use of conductance level-based detection. Furthermore, ADPCA has proven to be very robust against heavy drift thanks to its adaptive nature when the filter is properly design.

ACKNOWLEDGEMENTS

This work was supported by the EU Information Society Technologies project, IST-1999-10579.

BIBLIOGRAPHY

- [1] Barrettino, D.; Graf, M.; Zimmermann, M.; Hagleitner, C.; Hierlemann, A.; Baltes, H.; A micro-hotplate-based monolithic CMOS gas sensor array *Circuits and Systems*, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on , vol. 4 , May 25 - 28, (2003) 852 -855
- [2] Graf, M.; Barrettino, D.; Taschini, S.; Hagleitner, C.; Hierlemann, A.; Baltes, H. Monolithic metal-oxide microsensor system in industrial CMOS technology *Micro Electro Mechanical Systems*, 2003. MEMS-03 Kyoto. IEEE The Sixteenth Annual Int. Conference on , 19-23, (2003) 303 -306
- [3] H. Abe, S. Kanaya, Y. Takahashi, S. Sasaki, Method for Monitoring Sensor Stability in a Multiple Semiconductor Gas-Sensor System, *Anal. Chim. Acta*, 225, 89-96(1989)
- [4] W. Göpel, K.D. Schierbaum, Definitions and typical examples W. Göpel, T.A. Jones, M. Kleitz, I. Lundström, T. Seiyama (Eds.), *Chemical and Biochemical Sensors, Part I*, vol. 2, VCH, Weinheim, 1992, p. 1.

- [5] J.E. Haugen, O.Tomic, and K. Kvaal, A calibration method for handling the temporal drift of solid-state gas sensors. *Anal. Chim. Acta*, vol. 407, pp 23–39, 2000
- [6] T. Artursson, T. Eklöv, I. Lundström, P. Martensson,, M. Sjöström and M. Holmberg, Drift correction for gas sensors using multivariate methods, *J. Chemom*, vol 14, no 5-6 pp 711-723 (2000)
- [7] Martin Holmberg, Fredrik Winquist, and Ingemar Lundström, Fabrizio Davide, Corrado DiNatale and Arnaldo D'Amico. Drift counteraction for an electronic nose, *Sensors and Actuators B*, vol. 36, no. 1-3, pp. 528-535 (1996)
- [8] Marco, S.; Ortega, A.; Pardo, A.; Samitier, J. Gas identification with tin oxide sensor array and self-organizing maps: adaptive correction of sensor drifts, *Instr. and Meas.*, *IEEE Trans. on* , vol. 47 no. 1 (1998) 316 -321
- [9] C. DiNatale, F. Davide, and A. D'Amico, "A self-organizing system for pattern classification: Time varying statistics and sensors drift effects", *Sensors and Actuators B*, vol 26-27, 1-3, pp237-241, 1995
- [10] Ku W.; Storer R.H.; Georgakis C. Disturbance detection and isolation by dynamic principal component analysis, *Chemometrics and Intelligent Laboratory Systems Volume 30, Issue 1, (1995)*, Pages 179-196
- [11] Evan L. Russell, Leo H. Chiang and Richard D. Braatz, Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis *Chemometrics and Intelligent Laboratory Systems* vol. 51, no. 1 , 8 (2000) 81-93
- [12] J. E. Jackson. Quality control methods for several related variables. *Technometrics*, 1(4):359–377, 1959
- [13] M. Kendall. *Multivariate Analysis*. Charles Griffin&Co.1975
- [14] E. Oja. *Subspace Methods of Pattern Recognition*. Research Studies Press, Letchworth, Hertfordshire UK, 1983
- [15] M. Maeder. Evolving factor analysis for the resolution of overlapping chromatographic peaks. *Anal. Chem*, 527–530, 1987.
- [16] H. Keller and D. Massart. Peak purity control in liquid chromatography with photodiode-array detection by a fixed size moving window evolving factor analysis. *Analytica Chimica Acta*, 246:379–390, 1991.
- [17] K. H. Zou, W. J. Hall, and D. E. Shapiro. Smooth non-parametric receiver operating characteristic(roc) curves for continuous diagnostic tests. *Statistics in Medicine*, 16:2143–2156, 1997.
- [18] F. Provost and T. Fawcet. Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions. In *Proc. of the 3rd Int. Conf. on Knowledge Discovery and Data Mining*, 1997
- [19] Nikos Papamichail, Nicolae Barsan, Udo Weimar Monitoring of oil aerosol contamination in pressurised air with SnO₂-based thick film sensors in real life conditions, *Int. Symp. On artificial olfaction and electronic nose, ISOEN '03*, 2003, Riga Latvia
- [20] L. Ljung. *System Ident. Toolbox*. The Mathworks 2003
- [21] Pardo, A.; Marco, S.; Samitier, J.; Nonlinear inverse dynamic models of gas sensing systems based on chemical sensor arrays for quantitative measurements, *Inst and Meas.*, *IEEE Trans. on* , vol. 47 no. 3 , June 1998 Page(s): 644 -651