Drift compensation of gas sensor array data by Orthogonal Signal Correction

M. Padilla*, A. Perera†, I. Montoliu†, A. Chaudry Õ, K. Persaud‡ and S. Marco†

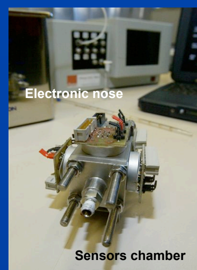
† Sistemas d'Instrumentació i Communicacions, Departament d'Electrònica, Universitat de Barcelona, Martí i Franquès 1, 08028Barcelona, Spain

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

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

INTRODUCTION

Electronic noses **are a faster, cheaper (but less accurate) and portable alternative to conventional instruments for gas analysis. They mainly consist on an array of sensors and a pattern recognition unit.**



Drift **is an important problem that affects strongly chemical gas sensors. Sensors ageing and environmental disturbances produce changes in sensors responses that make initial statistical models for gas or odor recognition become useless after a relatively short period of time (weeks).**

Periodically recalibrations of instruments that use this kind of sensors, like *electronic noses*, are necessary but they are expensive and laborious. An interesting and cheap alternative is drift counteraction by signal processing techniques.

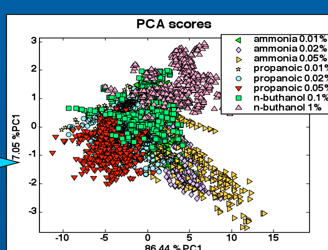
GOAL

In this work a drift compensation method based on Orthogonal Signal Correction (OSC) [1] for an array of gas sensors is proposed. A simple classification algorithm, kNN, will be employed for validating it on a dataset composed by measurements of three compounds at different concentrations, using a 17 sensor array during 10 months. OSC performance will be also compared with Component Correction technique [2].

THE DATASET

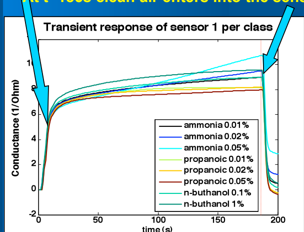
3415 samples of 8 classes (3 analytes at several concentrations measured by 17 gas polymer sensors during 10 months).

The 8 classes are very mixed and scattered due to the effect of drift and intra-class variability.

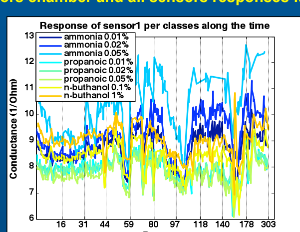


Every sample consist on a transient signal; sensor response measures during 200s:

- **At t=0s the analyte is introduced into the sensors chamber producing an abrupt rising of the sensor response**
- **At t=185s clean air enters into the sensors chamber and all sensors responses fall.**



Transient responses are characteristic of every interaction sensor – analyte (information about the chemical and physical reactions between the analyte and the sensors' sensitive layer).

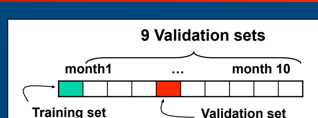


Great irregular variation along the time for all classes. Signals correspond to t=185s of transient responses of all samples.

EVALUATION OF THE TIME STABILITY

-Division of the dataset into 10 sets of 342 samples each; 311 samples from first set as training set, remaining are validation subsets.

-Classification Rates (CR) on 3 types of gas for every validation set by means of PCA (7 PCs) + kNN (3 nearest neighbors).



http://isp.el.ub.es

METHODS

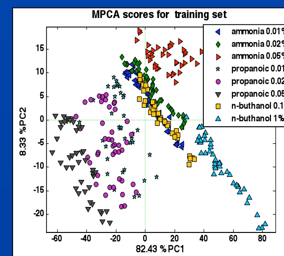
Component Correction (CC) [2]

- **Main goal: to extract from data X a number of components that capture most of the variance of a reference (representative) class.**
- **It supposes all classes evolves in the same way under drift conditions (not true in general).**
- **Components generated by Principal Components Analysis (PCA):**

$$X^k = T^k P^{k\prime} + E$$ PCA of the class of reference k

$$T = X P^k$$ Projection of all data on PCA-k subspace

$$X^c = X - T P^{k\prime}$$ Data correction X^c



PCA scores of training set. N-buthanol is chosen as *reference class*.

Orthogonal Signal Correction (OSC) [1]

- **Main goal: to extract from data X a number of components orthogonal to an information matrix Y that capture most of the data variance.**
- **Several OSC different versions: Wold, Sjöblom, Fearn, Wise, …, and several similar methods: DOSC, DO, OPLS, O2PLS, …**

Wise algorithm (Matlab, [3]):

for i=1:*NOSC* (=OSC components)

1. PCA(*X*) → *p*
2. Repeat until *t* stable: calculate *t*, orthogonalize *t* with respect to *Y* and update *p*
3. PLS(*X, t, NPLS*) → *w* (NPLS or tolerance)
4. calculate *t*, orthogonalize *t* with respect to *Y* and update *p*
5. extract component from data *X*: $X = X - tp'$ → $\begin{cases} np(:,i) = p \\ nt(:,i) = t \\ nw(:,i) = w \end{cases}$

end

New samples: $x^c = x - x\dfrac{nw}{np'nw}np'$

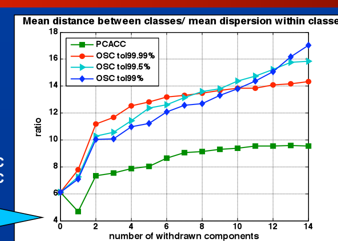
RESULTS

Selecting parameters

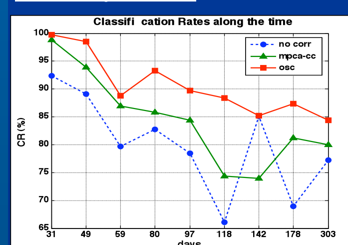
OSC tends to overfit => avoid max. tolerance, ok from 6 to 8 components and 99% tol. Components always orthogonal to information!

PCACC: try not to remove too many components => we may also remove information!

Fisher ratio for optimal choose of OSC/PCACC parameters and tolerance versus the number of OSC components for several tolerances.



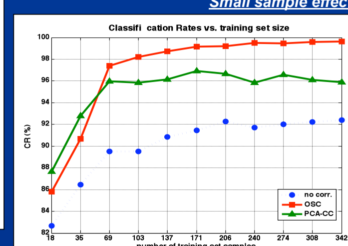
Time stability evaluation



CR for non corrected and drift corrected data by PA-CC, OSC along the time.

OSC: better performance and less oscillations

Small sample effect



CR for non corrected and drift corrected data by PA-CC, OSC vs number of training samples (of all classes in black, of reference class in green).

Better performance with increasing training set size, PCA-CC outperforms OSC for smaller training set size => OSC needs much more training samples to perform well

CONCLUSIONS

- **Drift compensation methods clearly improve time stability along a large period of time. Specially *Orthogonal Signal Correction* shows very good results.**
- **For tuning OSC and PCACC parameters the Fisher ratio of sample distribution is proposed.**
- **OSC and PCACC performance are compared:**
  - **OSC correction outperforms PCACC, but PCACC needs a much smaller training set,**
  - **stability is achieved with relatively few training set samples.**
- **Data preprocessing by these techniques improve the instrument performance during months. The problem of drift is not completely solved but time between recalibrations can be extended.**

REFERENCES

[1] S. Wold, H. Antti, F. Lindgren and J. Öhman, "Orthogonal signal correction of near-infrared spectra", Chemometrics and Intelligent Laboratory Systems 44 (1998):175-185

[2] T. Artursson, T. Eklöv, I. Lundström, P. Mårtensson, M. Sjöström and M. Holmberg, "Drift correction for gas sensors using multivariate methods", J. Chemometrics, no. 14, 711–723, 2000.

[3] Wise BM, Gallagher NB. www.eigenvector.com/MATLAB/OSC.html [May 2005].