

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/323908023>

Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors

Article in *Analytica Chimica Acta* · March 2018

DOI: 10.1016/j.aca.2018.03.005

CITATIONS

30

READS

213

2 authors:



Javier Burgués

ScioSense Germany GmbH

23 PUBLICATIONS 183 CITATIONS

[SEE PROFILE](#)



Santiago Marco

University of Barcelona

255 PUBLICATIONS 3,745 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Sampling and detection of security threats [View project](#)

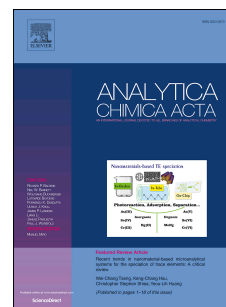


MOX sensor dynamics [View project](#)

Accepted Manuscript

Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors

Javier Burgués, Santiago Marco



PII: S0003-2670(18)30370-2

DOI: [10.1016/j.aca.2018.03.005](https://doi.org/10.1016/j.aca.2018.03.005)

Reference: ACA 235802

To appear in: *Analytica Chimica Acta*

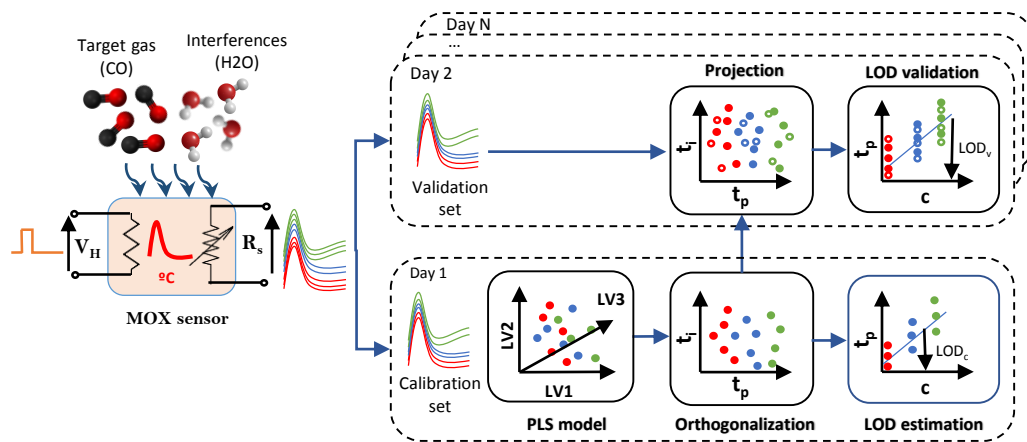
Received Date: 28 August 2017

Revised Date: 2 March 2018

Accepted Date: 5 March 2018

Please cite this article as: J. Burgués, S. Marco, Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors, *Analytica Chimica Acta* (2018), doi: 10.1016/j.aca.2018.03.005.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Multivariate estimation of the limit of detection by orthogonal partial least squares in temperature-modulated MOX sensors

Javier Burgués^{1,2}, Santiago Marco^{1,2}

¹Department of Electronic and Biomedical Engineering, Universitat de Barcelona, Martí i Franqués 1, 08028-Barcelona, Spain

²Signal and Information Processing for Sensing Systems, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, 08028 Barcelona Spain

Abstract: Metal oxide semiconductor (MOX) sensors are usually temperature-modulated and calibrated with multivariate models such as Partial Least Squares (PLS) to increase the inherent low selectivity of this technology. The multivariate sensor response patterns exhibit heteroscedastic and correlated noise, which suggests that maximum likelihood methods should outperform PLS. One contribution of this paper is the comparison between PLS and maximum likelihood principal components regression (MLPCR) in MOX sensors. PLS is often criticized by the lack of interpretability when the model complexity increases beyond the chemical rank of the problem. This happens in MOX sensors due to cross-sensitivities to interferences, such as temperature or humidity and non-linearity. Additionally, the estimation of fundamental figures of merit, such as the limit of detection (LOD), is still not standardized in multivariate models. Orthogonalization methods, such as Orthogonal Projection to Latent Structures (O-PLS), have been successfully applied in other fields to reduce the complexity of PLS models. In this work, we propose a LOD estimation method based on applying the well-accepted univariate LOD formulas to the scores of the first component of an orthogonal PLS model. The resulting LOD is compared to the multivariate LOD range derived from error-propagation. The methodology is applied to data extracted from temperature-modulated MOX sensors (FIS SB-500-12 and Figaro TGS 3870-A04), aiming at the detection of low concentrations of carbon monoxide in the presence of uncontrolled humidity (chemical noise). We found that PLS models were simpler and more accurate than MLPCR models. Average LOD values of 0.79 ppm (FIS) and 1.06 ppm (Figaro) were found using the approach described in this paper. These values were contained within the LOD ranges obtained with the error-propagation approach. The mean LOD increased to 1.13 ppm (FIS) and 1.59 ppm (Figaro) when considering validation samples collected two weeks after calibration, which represents a 43% and 46% degradation, respectively. The orthogonal score-plot was a very convenient tool to visualize MOX sensor data and to validate the LOD estimates.

Keywords: Metal Oxide Sensors; Partial Least Squares; Orthogonal Projection to Latent Structures; Maximum Likelihood Principal Component Regression; Limit of Detection; Temperature Modulation;

1. Introduction

Gas sensor technology is evolving towards miniaturized devices with low power consumption that can be integrated in smartphones and wearable devices [1–6]. With millions of units sold worldwide each year, metal oxide semiconductor (MOX) is the most commercially successful gas sensor technology. MOX sensors have been used to solve problems in a number of fields: energy [7], food

[8,9], automotive [10–12], environmental monitoring [13–16], safety and security [17–21] or biomedicine [22,23]. MOX sensors offer high sensitivity to many volatile organic compounds (VOCs), low fabrication cost, miniaturization potential thanks to micro-electromechanical systems (MEMS) technology and are very robust devices which can last for decades. However, after more than 50 years in the market, MOX sensors still suffer from very important limitations, namely poor stability [24,25], low selectivity [26], slow recovery time [27,28] and large tolerances in specifications for identical devices [29–31]. Among those, the lack of selectivity and stability (temporal drift) might be considered the most important ones.

The transduction mechanism of MOX sensors consists on measuring resistance changes resulting from chemical interactions between the gas species and the metal oxide surface. Clifford-Tuma [32] found that the response of a SnO_2 based Taguchi-type sensor in the presence of CO, H_2O and CH_4 follows a power law which contains products between several gas concentrations and quadratic terms. This non-linear response is due to competitive and synergistic interactions among gases that make the effect of the target gas can be either masked or enhanced by the presence of other gases (i.e. “matrix-effect”). The sensor response is also cross-sensitive to environmental factors, such as temperature, humidity or pressure. The temperature of the sensing surface affects the rate at which the chemical reactions occur. There is usually an optimum working temperature (in terms of sensitivity) for each target gas but only limited selectivity can be achieved by working at a constant temperature. Selectivity to a given analyte can be increased by modulating the operating temperature with a periodic heating power waveform [33]. The resulting multivariate response patterns capture the selectivity-temperature dependence characteristic of the target gas plus the effect of the interferences. The variables of the response pattern are highly correlated because the shape of the heating waveform smoothly varies the sensor temperature (except for the temperature transitions).

Using a set of calibration samples with known concentration of the target gas and varying contributions of potential interferences, a multivariate calibration model is built to predict the concentration of the target gas in future samples. Temperature-modulated MOX sensors are usually calibrated with Partial Least Squares (PLS) [34] models, which can deal with the highly collinear response patterns. The good performance of PLS from a theoretical point of view has been discussed by Helland [35,36]. PLS works optimally when the measurement errors are independent and identically distributed (iid). Because temperature-modulated MOX sensors usually present heteroscedastic and correlated measurement errors (i.e. non-iid), maximum likelihood methods, such as maximum likelihood principal components regression (MLPCR) [37], should in principle outperform PLS. MLPCR takes into account the measurement error structure for building the multivariate model. However, in a comparison between MLPCR and PLS, M.S. Reis et al. [38] have found that PLS seems to be more effective than MLPCR to reject heteroscedasticity. In fact, PLS prediction errors are close to those found by MLPCR in most scenarios. Reis attributes the good results of PLS to the effective way of this method to find a low dimensional predictive space, onto which the regressors are projected prior to being used for predicting the response. This projection is an effective filter that minimizes the effect of noise in X , regardless of the structure of the noise. The authors of this study also claim that the use of orthogonal projections benefit PLS for heteroscedastic data, when compared to maximum likelihood projections used by MLPCR. However, there are no references in the available literature regarding the suitability of MLPCR for the calibration of temperature-modulated MOX sensors. One contribution of this paper is the comparison between

PLS and MLPCR in modelling temperature-modulated MOX sensor data, in a concentration range near the limit of detection (LOD).

After the calibration model is built, it is common to report figures of merit, such as sensitivity, selectivity, precision or LOD, that characterize the chemical measurement process. the LOD can be considered the most relevant figure of merit for a myriad of applications requiring the detection of low concentrations of the target gases [10,11,13,39]. Despite of its relevancy, the estimation of the LOD in multivariate calibration is not standardized (see a recent review in [40]). The main issue is the lack of agreement between two alternative methods to estimate the LOD. The LOD is a consequence of uncertainty in analyte prediction at the blank level. As it is pointed out by Olivieri in [41], this uncertainty in prediction can be obtained either (i) analytically, by propagating the errors of the calibration concentrations and measured signals or (ii) empirically, by carrying out predictions on validation samples.

In the error-propagation approach, the uncertainties of the measured signals (both calibration and test samples) and the calibration concentrations are propagated through the model to estimate the variance in the predicted concentration of an unknown sample. Then, a sample-specific LOD can be computed from the prediction interval of blank samples. The most complete and general formula for the multivariate prediction error was proposed by Faber and Kowalski [42]. It incorporates the above sources of uncertainty plus the potential bias introduced by lack of model adequacy to the test sample. When a preprocessing method such as, Orthogonal Signal Correction (OSC) or Net Analyte Signal (NAS) is applied, Olivieri [41] shows that a correction needs to be applied to this formula to account for the uncertainty introduced by the filtering matrix. The uncertainty of the measured signals and the model parameters (due to variance in the calibration data) take the form of covariance matrices [43], which introduce complexity into the formulas. To simplify the covariance matrices to scalar values, it is assumed that the errors are iid. However, this assumption is unrealistic in MOX response patterns because the variables are highly correlated, and the errors of the variables close to the temperature transitions tend to be higher than at smoother temperature areas. Moreover, the formula assumes that real data is a contribution of additive signal and random noise and, as we previously discussed, MOX sensor response is highly non-linear. A practical limitation of using this formula is that the variance of the errors must be estimated a priori (e.g. from replicates), which is costly and, sometimes, impractical. The Faber and Kowalski formula [42] was tested in Montecarlo simulations [44], yielding unexpected results. Nonetheless, Allegrini et al. [45] developed a LOD definition based on the formulations of Faber and Kowalski [42], particularized for PLS models. The result is a range of LOD values: from the lowest (i.e. LOD_{min}), corresponding to the lowest blank leverage, to the largest (i.e. LOD_{max}), corresponding to the highest blank leverage. The LOD range clearly reflects that the composition of the sample in multivariate calibration affects the LOD estimate. However, the detection decision is slightly complex when the predicted concentration c of the test sample falls within the LOD range (i.e. $LOD_{min} < c < LOD_{max}$). In this case, it is unclear whether the decision to be made should be “detected” or “not detected”. Allegrini et al. [45] proposes to recompute a specific LOD for the test sample, approximating the leverage of the test sample by that of a blank. Then, the predicted concentration is compared against this single LOD value.

In the empirical approach, the uncertainty in prediction is estimated through a univariate “surrogate” variable computed on the calibration or (ideally) validation samples. The use of cross-

validation for error estimation is a well-accepted practice and has been recommended among others by Efron and Gong [57]. The well-accepted univariate LOD formula [46] is then applied to the pseudounivariate calibration graph that relates the univariate surrogate variable to the analyte concentration. Examples of surrogate variables are the net analyte signal (NAS) [47], the predicted concentrations [48] and the scores of the first component in principal component analysis (PCA) [49,50]. Among these, the approach based on the NAS is supported by the IUPAC [51]. The NAS of a certain sample is (a vector) defined as the part of the gross signal which can be uniquely attributed to the target analyte. The norm of the NAS (a scalar) can be used as the surrogate variable. However, there are different approaches to compute the NAS [47,52–55] which produce slightly different outcomes [56]. This is an issue because the resulting LOD estimates might differ depending on which method was used to estimate the NAS. The use of the predicted concentrations as the pseudounivariate variable resorts to the mathematical proof that the LOD is invariant for linear transformations of the response variable [48]. The predictions given by the model are obtained by multiplying the measured signals by the regression vector, which is a linear operation. However, this ignores the uncertainty in the estimated model parameters resulting from the use of noisy calibration data. Finally, computing the LOD directly from the scores of the model [49,50] is a promising approach because there is a unique way to compute the scores and they reveal the underlying structure of the model. However, this approach is limited to those cases in which the input data can be sufficiently well represented by a single component. In temperature-modulated MOX sensors, usually more than one PLS component is necessary to properly fit the calibration data. Due to the non-linearities and noise in the sensor conductance patterns that are not relevant for predicting the concentration of the target gas, several PLS components are needed to model one hyperplane or “latent direction” that relates the predictors to the predictand. Thus, the percentage of variance captured by the model is distributed among several latent variables and the LOD formulas based on the first component cannot be used.

If the error of the instrumental signals is known a priori, the LOD formula based on error-propagation should in principle provide more accurate results than any empirical approach. However, it is not usually the case that measurement errors (including all potential sources of error) are known a priori. In this case, replicate measurements must be used to compute the error covariance matrix, prior to applying the error-propagation formula. In multivariate calibration, these “replicate” measurements must contain the potential interferences that the analytical system might find during operation. The empirical approach does not strictly require replicate measurements but the obtained LOD is an average value instead of a sample-specific LOD. In this case, it is unclear whether the resulting LOD truly represents the degree of uncertainty of prediction at the blank level. Nevertheless, if replicate measurements are available (including blank replicates), then the empirical approach can estimate the uncertainty at the blank level by computing the errors of the model (i.e. the variance of the surrogate variable) in the blank replicates. One research question that this paper will address is how different is the LOD computed analytically or computed empirically, when replicate measurements are available.

As the number of latent factors increases, not only the computation of the LOD becomes difficult but the interpretability of the model decreases [58]. A simpler model is advantageous to understand the underlying structure of the model by, for example, answering some of the following questions: How is the model induced from the data? Which variables are relevant for prediction? Which variables/observations are more affected by the interferences? Which observations can be

considered outliers? Is there any time-based trend in the observations? To address the interpretability issues of PLS models mentioned above, Trygg et. al. [59] proposed a new method called orthogonal projections to latent structures (O-PLS). In the O-PLS algorithm, the structured noise in the input signals is removed by an integral orthogonal signal correction (OSC) filter, prior to PLS modelling. This yields a more simpler and relevant model with the same predictive power as the original PLS model. An extension of O-PLS, called O2-PLS, was presented by the same authors [60]. The advantage of O2-PLS over its predecessor is that it can estimate the pure constituent profiles, in the case of multiple predictands. However, for a single predictand case O-PLS and O2-PLS are identical. Ergon [61] proved that the same results of O-PLS could be obtained by post-processing the PLS model using a similarity transformation (ST). The resulting algorithm was named PLS+ST. O-PLS has been extensively applied in a wide diversity of fields such as spectroscopy [62], metabolomics [63], bioinformatics [64] and process monitoring [65]. Surprisingly, 15 years after its invention (as of July 2017) we have not found any work applying O-PLS, PLS+ST or similar methods to chemical sensor data. An alternative method to simplify a PLS model into a two component model was proposed by Ergon [66]. The idea behind 2PLS (also referred as bi-orthogonal PLS in [67]) is to keep the first PLS latent variable (as it is considered by some authors the best estimate of the spectral profile of the target gas) and compress the remaining latent variables (which in theory model the noise and interferences) into a new variable, allowing a 2D representation of the model. The usage of 2PLS in the literature is very scarce and has been exclusively used in process monitoring [68].

We consider that O-PLS (equivalently PLS+ST) provide a more convenient representation than 2PLS because the first component of the simplified model is in the direction of the regression vector (instead of the first latent variable, as in 2PLS). The importance of the regression vector is that it models the part of the signal that is useful for prediction. Thus, the projection of the scores onto the first component describes the variations of \mathbf{X} related exclusively to \mathbf{y} . We will call the first component the predictive component. Taking advantage of this property, we propose a new method to estimate the LOD by applying the well-accepted univariate formulas [46] to the predictive component of the orthogonalized model. We consider this is an improvement over existing methods based on the first PCA component [49,50] because our approach ensures that the first component condenses all the variance related to the analyte concentration. This approach yields the same results as the IUPAC accepted multivariate formulas based on the NAS [51], when the NAS is computed through the regression vector [53]. It is also equivalent to the method proposed by Ortiz et. al [48] in which the surrogate variable are the concentrations predicted by the model. Compared to these methods, the main advantage of computing the LOD through the orthogonal model is that the role of the interferences and the target gas can be visualized, which is beneficial for interpreting the LOD.

An underlying assumption when building a calibration model is that the characteristics of the prediction samples are the same as for the calibration samples. This is usually the case in analytical instruments, such as spectrometers, which are highly selective and usually calibrated and operated in laboratory conditions. However, MOX sensors are calibrated in laboratory conditions and then are operated without supervision outside the lab: in domestic scenarios (e.g. CO toxic alarm), mobile platforms (e.g. smartphone, robot) or in environmental samples (e.g. monitoring traces of contaminants), among others. The greater variability of the environmental factors in such scenarios, the more complex composition of the gas mixtures or even device shift will introduce bias and variability in the predictions. Subsequently, the LOD obtained in calibration might not be

representative of the model performance in the target scenario. In this work we will show that the validation of the calibration model using future samples is essential to quantify the degradation of the model and report more realistic LOD estimates.

2. Materials and methods

To illustrate that orthogonal PLS (computed either by pre- or post-processing) is a very convenient tool to treat chemical sensor data, we designed a scenario of CO detection in the presence of random humidity, using MOX sensors with different degree of cross-sensitivity to environmental interferences. In this section we describe the dataset used in the experiments and how this dataset was used to build MLPCR and orthogonalized PLS models. Finally, we explain the estimation and validation of the LOD from the orthogonal model. Figure 1 shows a block diagram of the full process.

The analysis of the data was done using MATLAB R2009a (The Mathworks) and functions *pls* and *orthogonalizepls* from the PLS_Toolbox_8.0.2 (Eigenvector Research).

2.1. Dataset

For the experiments presented in this work we used the dataset described in detail in [69]. Therefore, only a brief summary is described in this section. Seven units of two commercial MOX sensors (SB-500-12 and TGS 3870-A04, provided by FIS and Figaro, respectively) were exposed to dynamic mixtures of CO and humid synthetic air. The heater voltage was modulated in the range 0.2-0.9 V in periods of 25 seconds (FIS) and 20 seconds (Figaro), following the manufacturer recommendations (0.9V for 5s, followed by 0.2V for 20s (FIS) or 15s (FIGARO)). The resulting multivariate sensor conductance g_s was measured using a voltage divider with a load resistor of 1 M Ω . The sensors were pre-heated for one week before starting the experiments.

Dynamic mixtures of CO and humid synthetic air were delivered from high purity gases in cylinders to a gas chamber, by means of a piping system and mass flow controllers (MFCs). Each experiment consisted on 100 measurements: 10 experimental concentrations uniformly distributed in the range 0 to 20 ppm and 10 replicates per concentration (see Fig. 2). Each replicate had a relative humidity randomly chosen from a uniform distribution between 15 and 75 % r.h. A temperature/humidity sensor (SHT75, from Sensirion) was used to measure the humidity. The grid of concentration vs. humidity was designed irregularly as we considered it was more representative of a real scenario. The samples were presented to the sensors in random order, using an exposure time of 15 minutes per sample. This exposure time ensured that the sensors reached stable responses and it was determined by the size of the gas chamber and the gas flow rates used. A single experiment lasted 25 hours (100 samples x 15 minutes/sample) and was replicated on 13 working days spanning a natural period of 17 days. The uncertainty in the concentrations generated by our system was derived through error propagation, taking into account all known sources of errors (bottle uncertainty, mass flow control reading errors, etc.). A maximum uncertainty of 0.12 ppm was associated to the lowest concentration (2.2 ppm), because the generation of this concentration requires the MFC which dilutes the CO in air to operate at the maximum flow rate (the error of the MFC increases with the flow rate).

In each experimental day, the sensor conductance signals g_s were baseline-corrected and logarithmically transformed. The baseline correction consists on dividing the sensor conductance g_s ($k\Omega^{-1}$) by the sensor conductance in air g_0 ($k\Omega^{-1}$), which was measured at the beginning of each experimental day at a random humidity and temperature. The logarithmic transformation of the response in this sensor model linearizes the response at low concentrations, which produces simpler and more accurate models [69]. The data matrix \mathbf{X} ($n \times m$) consists of $n = 100$ observations (samples) of the baseline-corrected logarithmic sensor conductance patterns, sampled at $m = 100$ uniformly distributed points (variables) along the heating cycle. It should be noted that the replicates of each sample were included in the matrix \mathbf{X} . The vectors \mathbf{y} ($n \times 1$) and \mathbf{h} ($n \times 1$) contain the true CO concentration (ppm) and the humidity content (% r.h.) of the 100 samples, respectively.

2.2. PLS models

Given a matrix \mathbf{X} ($n \times m$) of n samples and m predictors or variables and their corresponding concentration \mathbf{y} ($n \times 1$), the PLS algorithm [70] can be used to solve the multiple linear regression problem in the inverse formulation (Eq. 1), for the case where there are more variables than samples ($m > n$).

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (1)$$

Here, \mathbf{b} ($m \times 1$) is the vector of regression coefficients and \mathbf{e} is the residual error. PLS solves the inverse calibration problem by finding a reduced number of underlying latent factors, also called latent variables (LVs), that account for as much as possible variance in \mathbf{X} while also capturing the correlation to \mathbf{y} . The advantage of inverse calibration is that, because \mathbf{y} is modelled as a function of \mathbf{X} , no explicit knowledge of the potential interfering species is required (although they must be sufficiently represented in the calibration set). The latent space is defined by a set of p basis vectors or *loadings* \mathbf{P} ($m \times p$) and \mathbf{q} ($1 \times p$), which give information about the weights for each predictor in \mathbf{X} when calculating each latent variable. The coordinates of each sample in the latent space are the scores \mathbf{T} ($n \times p$). The score matrix is able to summarize \mathbf{X} and predict \mathbf{y} with small errors in \mathbf{E} ($n \times m$) and \mathbf{F} ($n \times 1$):

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (2)$$

$$\mathbf{y} = \mathbf{T}\mathbf{q}^T + \mathbf{F} \quad (3)$$

\mathbf{X} was mean-centered and standardized to unit variance prior to PLS modelling. The vector of responses \mathbf{y} was mean-centered. PLS models were built using the samples corresponding to the first experimental day, which we denote as the calibration set. The complexity of the models was optimized by performing a 5-fold cross-validation (CV). The 5-fold CV involves randomly splitting the data into five groups and fitting the model using four folds (i.e. using 80% of the samples). The remaining fold (i.e. 20% of the samples) is hold out as the test set to compute the root mean squared error (RMSE) (Eq. 4). The folds are made such that the 80/20% split is maintained in each concentration level.

$$RMSE = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / n} \quad (4)$$

where n denotes the size of the test set, and y_i and \hat{y}_i are the true and predicted value for test sample i , respectively. This process is repeated five times so that each fold is used exactly once as the test set. The optimum number of latent variables (LVs) is found by inspection of the average $RMSE$ of the five folds vs. number of LV curve. The final models were built using the selected number of LVs and the full calibration set.

2.2.1. Orthogonalization of the PLS models (PLS+ST)

The PLS models were post-processed with an orthogonalization step (PLS+ST). This produces simpler models with orthogonalized loadings and scores, which condense all the \mathbf{y} -variance in the first weight and loading. The first loading of the orthogonalized model is in the same direction of the regression vector for data in which the structured noise has been filtered out. The predictions of the new model are identical to the non-orthogonalized model but the loadings and weights have been rotated. The second loading of the orthogonalized model captures the orthogonal variation to \mathbf{y} , which is related to the interferences. There are two main approaches for PLS orthogonalization [59,71], which produce similar results. In [71], a similarity transformation matrix \mathbf{M} (Eq. 5) is used to remove the orthogonal variation of \mathbf{y} :

$$\mathbf{M} = \begin{bmatrix} 1 & \mathbf{0} \\ q_1^{-1} \mathbf{q}_{2:A} & \mathbf{I} \end{bmatrix} \quad (5)$$

where $\mathbf{0}$ is a matrix of zeros and \mathbf{I} is the identity matrix. Using \mathbf{M} , Eq. 3 is transformed into Eq. 6:

$$\mathbf{y} = [\mathbf{t}_1 \quad \mathbf{T}_{2:A}] \mathbf{M} \mathbf{M}^{-1} \begin{bmatrix} q_1 \\ \mathbf{q}_{2:A} \end{bmatrix} + \mathbf{F} = \mathbf{t}_1^* \mathbf{q}_1 + \mathbf{F} \quad (6)$$

where

$$\mathbf{t}_1^* = \mathbf{t}_1 + q_1^{-1} \mathbf{T}_{2:A} \mathbf{q}_{2:A} \quad (7)$$

\mathbf{t}_1^* is a rotation of the original \mathbf{t}_1 scores. From Eq. 6, it can be seen that \mathbf{t}_1^* has the same direction as \mathbf{y} .

Similarly, Eq. 2 is transformed into Eq. 8:

$$\mathbf{X} = [\mathbf{t}_1 \quad \mathbf{T}_{2:A}] \mathbf{M} \mathbf{M}^{-1} \begin{bmatrix} \mathbf{p}_1^T \\ \mathbf{P}_{2:A}^T \end{bmatrix} + \mathbf{E} = \mathbf{t}_1^* \mathbf{p}_1^T + \mathbf{T}_{2:A} \mathbf{P}_{2:A}^* + \mathbf{E} \quad (8)$$

where

$$\mathbf{P}_{2:A}^* = \mathbf{P}_{2:A} - q_1^{-1} \mathbf{p}_1 \mathbf{q}_{2:A}^T \quad (9)$$

is the rotated loading matrix for the components orthogonal to \mathbf{y} .

2.3. MLPCR models

Maximum likelihood principal component regression (MLPCR) [37] is a supervised regression technique that performs a principal component decomposition (PCA) of \mathbf{X} (i.e. Eq. 2), considering the structure of the error matrix \mathbf{E} . The final models are built by ordinary least squares (OLS) regression of the scores \mathbf{T} and the nominal analyte concentrations \mathbf{y} . In practice, the error matrix \mathbf{E} is unknown and must be estimated from a finite number of samples. In the general case, the error variance among all measurement errors can be described by the full covariance matrix $\mathbf{\Sigma}$ ($mn \times mn$). However, the MLPCR algorithm [72] cannot handle full covariance matrices because of computational and storage cost. Therefore, several assumptions are required to simplify the problem. First, it is assumed that correlation only exists within each row of the data matrix \mathbf{X} (i.e. correlation among variables), but not within the columns (i.e. samples). In the case of error covariance in the rows, the error variance $\mathbf{\Sigma}$ can be described in terms of n row error covariance matrices $\mathbf{\Sigma}_i$ of size $m \times m$:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & & \\ & \ddots & \\ & & \mathbf{\Sigma}_n \end{bmatrix} \quad (10)$$

Each row error covariance matrix $\mathbf{\Sigma}_i$ ($i = 1, \dots, n$) is estimated in practice from q replicates of sample i :

$$\hat{\mathbf{\Sigma}}_i = \frac{1}{(q-1)} \sum_{k=1}^q (x_{i,k} - \bar{x}_i)^T \cdot (x_{i,k} - \bar{x}_i) \quad (11)$$

Where $\hat{\mathbf{\Sigma}}_i$ is the estimated row error covariance matrix corresponding to the i th sample, $x_{i,k}$ is the k th replicate of sample i and \bar{x}_i is the average of the q replicates.

The second assumption of MLPCR is that the row error covariance matrix is similar for each sample. In this manner, a single row error covariance matrix of size $m \times m$, pooled across the different samples, describes the error variance:

$$\hat{\mathbf{\Sigma}}_{pooled} = \frac{1}{n} \sum_{i=1}^n \hat{\mathbf{\Sigma}}_i \quad (12)$$

$\hat{\mathbf{\Sigma}}_{pooled}$ is called “equal row error covariance matrix” and is the default error matrix in the MLPCR algorithm.

2.4. Limit of detection (LOD)

The LOD is defined by the International Union of Pure and Applied Chemistry (IUPAC) as the minimum concentration that can be reliably distinguished from blank samples, with stated probabilities of error [46]. In the IUPAC recommendations, the LOD is formulated in terms of hypothesis testing, interpreting the distribution of the LOD and the blanks as gaussian random variables:

$$P(\hat{L} > L_c | L = 0) = \alpha \quad (13)$$

$$P(\hat{L} < L_C | L = L_D) = \beta \quad (14)$$

Here L is a generic symbol that represents either net signal (blank-corrected) or analyte concentration, \hat{L} is an estimate of L , L_C is the threshold used for the detection decision (i.e. analyte present or not) and L_D is the limit of detection or LOD. α and β are the probabilities of false positives and false negatives, respectively. Solving the above equations for L_D yields:

$$L_D = z_{1-\alpha}\sigma_0 + z_{1-\beta}\sigma_D \quad (15)$$

where $z_{1-\alpha}$ and $z_{1-\beta}$ are the critical values of the normal distribution for significance level α and β , respectively, while σ_0 and σ_D are the true standard deviation of \hat{L} when $L = 0$ and $L = L_D$, respectively. In practice, $\alpha = \beta$ and the critical value of the Gaussian distribution is replaced by the one sided t-critical value ($t_{1-\alpha, \nu}$) for the chosen confidence level (α) and degrees of freedom (ν). Similarly, the true standard deviations σ_0 and σ_D are replaced by the corresponding estimates s_0 and s_D , respectively. Assuming homoscedasticity, s_D is replaced by s_0 and a simplified formula for the LOD can be obtained (Eq. 16):

$$L_D = 2t_{1-\alpha, \nu}s_0 \quad (16)$$

The only variable that we need to estimate in Eq. 16 is s_0 , the standard deviation of the estimated net signal or predicted concentration of a blank sample. The two competing methods to estimate the LOD, which are explained in detail in the following sections, differ in how they compute s_0 .

2.4.1. Pseudo-univariate LOD

In the pseudounivariate LOD approach, s_0 is computed from a “surrogate” univariate variable extracted from the multivariate model. Examples of surrogate variables are the net analyte signal, the predicted concentration in a set of calibration samples, or the first score of an orthogonalized PLS model, as in the present proposal. To compute the LOD, the surrogate variable y is related to the analyte concentration x by a linear regression model (Eq. 17):

$$y = \hat{B} + \hat{A}x \quad (17)$$

where \hat{B} and \hat{A} are estimates of the intercept and slope of the calibration curve, respectively.

It should be noted that Eq. 17 represents a direct calibration model and, thus, the meaning of y and x is interchanged with respect to inverse models (e.g. PLS). From linear regression [73], we know that the standard deviation of the predicted value of a blank sample is:

$$s_0 = s_{y,x} \sqrt{1 + \frac{1}{n} + \frac{\bar{x}^2}{\sum(x_j - \bar{x})^2}} = s_{y,x} \sqrt{1 + h_0} \quad (18)$$

Where $s_{y,x}$ is the standard error of regression, n is the number of calibration samples, x_j are the concentration of the calibration samples, \bar{x} is the mean calibration concentration and h_0 is the leverage of a blank sample (Eq. 19).

$$h_0 = \frac{1}{n} + \frac{\bar{x}^2}{\sum (x_j - \bar{x})^2} \quad (19)$$

We can see that h_0 penalizes calibration designs with low number of samples (h_0 grows when n decreases) or high calibration ranges (h_0 grows when the mean calibration concentration \bar{x} increases). Substituting Eq. 18 into Eq.16, yields the LOD in the signal domain (y_D):

$$y_D = 2t_{1-\alpha, v} s_{y, x} \sqrt{1 + h_0} \quad (20)$$

The LOD is usually reported in concentration units. The concentration (\hat{x}) corresponding to an observed gross signal (y) is estimated by:

$$\hat{x} = \frac{y - \hat{B}}{\hat{A}} = \frac{\hat{y}}{\hat{A}} \quad (21)$$

Where \hat{y} is the net (blank-corrected) signal. The LOD in the concentration domain is obtained by inserting Eq. 19 into Eq. 20:

$$x_D = \frac{y_D}{\hat{A}} \cdot \left(\frac{K}{I}\right) = 2t_{1-\alpha} \cdot \frac{s_{y, x}}{\hat{A}} \cdot \sqrt{1 + h_0} \cdot \left(\frac{K}{I}\right) \quad (22)$$

The correction factor $\left(\frac{K}{I}\right)$ accounts for the uncertainty in the regression parameters (\hat{A}, \hat{B}) [74] and it is defined in Eq. 22:

$$\frac{K}{I} = \frac{1 - r(B, A) \cdot (\sigma_{\hat{B}}/\sigma_0) \cdot [t_{1-\alpha, v} \cdot (\sigma_{\hat{A}}/A)]}{1 - [t_{1-\alpha, v} \cdot (\sigma_{\hat{A}}/A)]^2} \quad (23)$$

Where $r(B, A)$ is the correlation coefficient between the slope and intercept of the calibration line (Eq. 17), $\sigma_{\hat{B}}/\sigma_0$ is the relative standard deviation of \hat{B} (estimated intercept), σ_0 is the standard deviation of the blanks and $\sigma_{\hat{A}}/A$ is the relative standard deviation of \hat{A} .

$$r(B, A) = \bar{x} / \sqrt{\sum_{i=1}^n x_i^2 / n} \quad (24)$$

The interpretation of the correction factor $\left(\frac{K}{I}\right)$ is as follows. If the slope (A) is estimated with low uncertainty (i.e. $(\sigma_{\hat{A}}/A) \rightarrow 0$), the term in square brackets in Eq. 16 goes to zero, and both K and I are equal to 1, yielding $\left(\frac{K}{I}\right) = 1$. In this case, the LOD is not affected by the correction factor. At the other extreme, when the uncertainty in estimating A is very high (as $\sigma_{\hat{A}}/A$ approaches $1/t_{1-\alpha, v}$), the denominator (I) goes to zero, the ratio $\left(\frac{K}{I}\right)$ goes to infinity and x_D is unbounded. The correction factor includes the effect of the uncertainty of \hat{B} when estimating A . If A and B are estimated individually, the correlation coefficient $r(B, A)$ is zero and the ratio $\left(\frac{K}{I}\right)$ only depends on the uncertainty of \hat{A} . When A and B are estimated jointly, they will be negatively correlated and the numerator (K) in Eq. 16 will increase with the uncertainty of \hat{B} .

2.4.2. Multivariate LOD or error-propagation approach

In the error-propagation approach, the uncertainties of the calibration concentrations and the measured signals (both calibration and test samples) are propagated through the model to estimate the variance $\sigma_{\hat{y}}^2$ in the predicted concentration \hat{y} of an unknown sample (see equation (61) in Reference [42]):

$$\sigma_{\hat{y}}^2 = h_u \cdot (\sigma_{\Delta y}^2 + \|\mathbf{b}\|^2 \sigma_{\Delta x}^2) + \|\mathbf{b}\|^2 \sigma_{\Delta x}^2 \quad (25)$$

Where h_u is the unknown sample leverage for mean-centered data (Eq. 26), $\sigma_{\Delta y}^2$ is the variance of the calibration concentrations, $\|\mathbf{b}\|$ is the Euclidean norm of the PLS regression vector and $\sigma_{\Delta x}^2$ is the variance of the instrumental signals of the unknown and calibration samples.

$$h_u = \frac{1}{n} + \mathbf{t}_u^T \cdot (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{t}_u \quad (26)$$

Here n is the number of calibration samples and \mathbf{t}_u ($1 \times p$) and \mathbf{T} ($n \times p$) are the PLS scores of the unknown sample and the calibration set, respectively. The expression of Eq. 25 assumes that the measurement errors are iid with zero mean and constant variance, which is not usually the case in chemical sensing. To accommodate other error structures, Allegrini et. al [43] derived a general formula for the prediction error (see equation (11) and Table 1 (Case 2) in Reference [43]) :

$$\sigma_{\hat{y}}^2 = h_u \cdot (\sigma_{\Delta y}^2 + \mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b}) + \mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b} \quad (27)$$

where $\boldsymbol{\Sigma}_y$ and $\boldsymbol{\Sigma}_x$ are the error covariance matrices of the instrumental signals of the unknown and calibration samples, respectively. The estimated standard deviation of the predicted analyte concentration of a blank sample can be obtained by particularizing Eq. 27 for $y = 0$ and taking the square root in both sides of the equation:

$$s_0 = \sqrt{h_u \cdot (\sigma_{\Delta y}^2 + \mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b}) + \mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b}} \quad (28)$$

Inserting s_0 into Eq. 16 would give the LOD in the concentration domain. However, in contrast to univariate calibration, the fact that there is no single blank sample in PLS calibration will result in multiple values of h_0 , one per each blank sample [45]. The result is a range of LOD values: from the lowest (Eq. 27), corresponding to the lowest blank leverage $h_{0\min}$, to the largest (Eq. 28), corresponding to the highest blank leverage $h_{0\max}$:

$$x_{D\min} = 2t_{1-\alpha, v} \sqrt{\mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b} \cdot (1 + h_{0\min}) + h_{0\min} \sigma_{\Delta y}^2} \quad (29)$$

$$x_{D\max} = 2t_{1-\alpha, v} \sqrt{\mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b} \cdot (1 + h_{0\max}) + h_{0\max} s_{y_{cal}}^2} \quad (30)$$

The details on how to compute $h_{0\min}$ and $h_{0\max}$ are given in [45]. The relationship between the multivariate LOD range (Eqs. 29-30) and the pseudounivariate LOD (Eq. 22) can be clarified by noticing that $h_{0\min}$ in Eq. 29 is equal to the leverage of the blank sample h_0 in Eq. 19 [45].

Neglecting the uncertainty in the calibration concentrations in Eq. 29 (i.e. $\sigma_{\Delta y}^2 = 0$) and rearranging the terms inside the square root yields:

$$x_{Dmin} = 2t_{1-\alpha, v} \sqrt{\mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b}} \cdot \sqrt{(1 + h_0)} \quad (31)$$

Comparing Eq. 31 to Eq. 22, we can see that the terms $\sqrt{\mathbf{b}^T \boldsymbol{\Sigma}_x \mathbf{b}}$ and $\frac{s_{y,x}}{\hat{A}}$ are equivalent when the correction factor $\left(\frac{K}{I}\right)$ approaches unity. Both terms represent the propagation of the uncertainty in the input signals through the model. In the first case, the propagation occurs directly from the measurement error $\boldsymbol{\Sigma}_x$. The coefficients of the regression vector \mathbf{b} are used to weight the variables in $\boldsymbol{\Sigma}_x$, according to their importance for prediction. The resulting value is already in concentration units. In the second case, the input signals are first projected onto the regression vector to obtain the surrogate variable. The uncertainty of the surrogate variable is captured by $s_{y,x}$. The slope \hat{A} is used to convert $s_{y,x}$ to concentration units.

3. Results and Discussion

The dataset described in this section contains samples measured in 13 experimental days. The problem considered here is to calibrate two MOX sensor families with the objective of predicting the CO concentration and rejecting the humidity interference. First, the sensor conductance patterns and the measurement noise are analyzed. Then, PLS and MLPCR models are built using samples from the first day. The RMSECV and the optimum number of components of both models are compared. Then, PLS models were orthogonalized to simplify the visualization and to calculate the LOD using the pseudo-univariate approach. The resulting values are compared against the multivariate LOD interval. The samples from days 2-13 were used to validate the model and quantify its degradation in terms of LOD. First, a comprehensive analysis of the FIS device is presented and, then, the main results of the Figaro unit will be shown.

3.1. FIS SB-500-12

The response of one sensor unit to varying levels of carbon monoxide (CO) under nearly constant humidity (52 %r.h.) is shown in Fig. 3a. The x-axis represents a single heating cycle in which the sensor is driven at high temperatures for 5 seconds and then it is cooled down for 20 seconds. The logarithm is applied to the y-axis to facilitate the visualization of the response at low heating temperatures. Fig. 3b shows the sensor response to varying levels of humidity, for three concentration levels. As it can be seen, the sensor response was cross-sensitive to humidity only in certain areas of the heating pattern (e.g. $t > 10$ s) and for concentration levels above 15.6 ppm. This indicates a non-linear effect between humidity and CO on the sensor response. From this plot, we can see that the signals corresponding to the blanks are clearly distinguishable from the signals at 11 ppm, at least in the low temperature part of the heating cycle. Therefore, we can intuitively think that the LOD should be lower than 11 ppm. To get an initial estimate of the LOD and as a first step to understand how the PLS algorithm can model these sensor conductance patterns, PLS models were built using samples from the first experimental day. Considering the multiple sensor units, five to seven components were optimum based on the cross-validation results. A scatter plot of the scores (i.e. *score-plot*) of the first two LVs (\mathbf{t}_1 vs. \mathbf{t}_2) of one calibration model is shown in Fig. 4a. Since the first two LVs already captured 96% of the variance, the scores of the following LVs were not represented. The first thing we can notice in this plot is that there is a non-linear dependence between \mathbf{t}_1 and \mathbf{t}_2 , although the difference in variance between the two axes should be taken into

consideration. Second, the projection of the estimated regression vector $\hat{\mathbf{b}}$ (black solid line) into this subspace is confusing because of two reasons: (i) the direction of $\hat{\mathbf{b}}$ is not aligned to the analyte-capturing direction and (ii) the interference direction, which is the direction of the intra-class scattering, is not orthogonal to $\hat{\mathbf{b}}$. From this plot, we may (incorrectly) deduce that the PLS model was unable to reject the humidity interference and we cannot precisely assess the LOD.

An improved representation is shown in Fig. 4b, after the model was orthogonalized. In this plot, the x -axis captures the y -variation and the y -axis captures the y -orthogonal variation, which is related to the cross-sensitivity of the sensor to the interferences. The \mathbf{t}_2 scores were mean-centered at each concentration level to enhance the interpretability of the model. The ability of the model (and obviously of the sensor) to distinguish between two consecutive concentration levels is clearly identified in this plot. The orthogonal score-plot could be seen as a deformation of the sample space (c.f. Fig. 2) due to the combined effect of the sensor, the calibration model and the noise. The smaller the sensor cross-sensitivity to the noise the less scattered will be the points in the vertical direction (ξ_i). The intra-class scattering in the x -axis (ξ_a) measures the ability of the PLS model to reject the interference. The smaller the value of ξ_a the higher the selectivity of the model. If the scores of the same concentration level can be approximated by a straight line, the angle β between this line and the horizontal axis also indicates the selectivity of the model for that concentration. $\beta = 90^\circ$ would mean perfect rejection of the interference and $\beta = 0^\circ$ means no rejection at all. In this example, we can see that the vertical scattering (ξ_i) increases with concentration, which is coherent with the fact that the cross-sensitivity to humidity of the sensor conductance patterns was higher for 20 ppm than for 2 ppm (c.f. Figure 3b). We can visually check that the first non-zero concentration (2.2 ppm) is clearly separated from the blanks (0 ppm), so the LOD must be located between these two levels. From this picture, we can understand that the LOD is affected by the interferences because the first concentration level is tilted with an angle β . In an interference-free scenario ($\beta = 0^\circ$), the lesser degree of overlapping between the blanks and the low concentrations would produce a smaller LOD.

From this initial analysis we visually estimated the LOD in the region between 0 and 2.2 ppm. To estimate the LOD, it is recommended to use calibration samples with concentration close to the LOD during the model building step [46]. Therefore, a reduced range of 0-9 ppm was used for the rest of the analysis. The auto-scaled and baseline-corrected logarithmic sensor conductance patterns, in this reduced range, are shown in Fig. 5. The sensor response at low temperatures ($t > 5$ s) show more sensitivity to higher concentration levels (see how the distance between the red and orange traces is higher than the distance between the dark and light blue signals). On the other hand, the sensor response at high temperatures ($t < 5$ s) seem more sensitive to the lower concentrations. The variance of the signals at 2.2 ppm was higher than at the other concentration levels (heteroscedasticity), probably due to the uncertainty in the generation of the calibration concentrations (see Section 2.1). However, the greatest contribution to heteroscedasticity occurs along the response pattern, rather than along the samples. The measurement error during the temperature transitions is one order of magnitude higher than in other areas of the pattern. This can be clearly seen in the pooled row error covariance matrix (Fig. 6). Looking at the noise structure, we can see that the highest measurement noise was found in the temperature transitions ($t = 0$ and $t = 5$ s), probably because the sampling frequency of the acquisition system was not fast enough to accurately capture the rapid changes in sensor conductance during the temperature

changes. The measurements corresponding to high temperature of the sensing layer ($t < 5$ s) showed significant error compared to those at the beginning of the low temperatures ($t \in [6, 12]$ s), where the lowest measurement error was found. A moderate stable noise level was found at the end of the low temperatures ($t > 15$ s). We computed the correlation coefficient among variables of the response pattern and values above 0.8 were found (except in the temperature transitions).

In the presence of heteroscedastic and correlated measurement errors, MLPCR should produce models with lower number of components and lower prediction error than PLS. However, the empirical studies of Reis [38] confirm that this is not always the case: noise level, correlation level and heteroscedasticity play in role in determining the best linear regression method. Consequently, it seems that an empirical investigation in the particular case of temperature modulated MOX sensors is needed. A comparison of the cross-validation error between PLS and MLPCR is presented in Fig. 7. In both cases, the calibration models were built using 50 samples in the range 0-9 ppm (5 concentrations \times 10 replicates) collected in the first experimental day. For the same level of error (i.e. 0.4 ppm), PLS required less components than MLPCR (5 vs 10 LVs). This behavior –which was originally found between PCR and PLS [75]– is related to the fact that PLS latent variables consider the correlation with the Y target vector, providing parsimonious predicting spaces. This behavior has been confirmed for MLPCR by Reis [38]. The authors attribute the good results of PLS to the effective way of this method to find a low dimensional predictive space, onto which the regressors are projected prior to being used for predicting the response. These results do not agree with the MLPCR theory [37,72], which states that, in the case of non-iid noise, MLPCR should perform better than PLS. On the light of the analysis of Reis, we may assume that PLS is able to reject heteroscedastic noise, when this noise is not correlated with the Y vector. This can be seen by looking at the changes in the shape of the PLS regression vector as a function of the number of latent variables (Fig. 8). When only 1 LV is used, the regression vector equally weights the response pattern, regardless of the error of each variable. When two latent variables are used, suddenly the initial part of the low temperature section is heavily weighted while the rest is rejected. This part of the sensor response exhibits low measurement errors (see Fig. 6) and is less correlated with the rest of the pattern. PLS weights more this area by virtue of the higher correlation with Y, producing a more effective filtering of the noise in X.

Based on these results, we only considered PLS models for the rest of the analysis. New PLS models were built using all samples from the first experimental day in the range 0-9 ppm and the optimum number of LVs. After orthogonalization of the resulting models, x_D was estimated using the first orthogonal component and then compared against the multivariate LOD interval (See Table 1). Values of x_D ranging from 0.71 to 0.92 ppm, with an average value of 0.79 ppm, were found. The multivariate LOD ranges were centered at a mean value of 0.80 ppm and they were relatively narrow (mean width of 0.10 ppm). For the six sensors analyzed, x_D was contained inside $[x_{Dmin}, x_{Dmax}]$. From this table, we can conclude that both LOD methods provide equivalent results if the model complexity is well estimated. When the model complexity is underestimated or overestimated, x_D was either slightly higher than x_{Dmax} or slightly lower than x_{Dmin} , respectively (Fig. 9). Both LOD methods show a monotonically decreasing relationship with the number of latent variables, like a typical RMSEC vs LV curve. In the simplest model (1 LV), x_D (3.02 ppm) was 4.9% higher than x_{Dmax} (2.85 ppm). As the model complexity reaches its optimum value (5 LV), x_D (0.91 ppm) falls within the LOD range ([0.78, 0.93] ppm). x_D remained within $[x_{Dmin}, x_{Dmax}]$ beyond the optimum number of LVs, up to a point in which it exits the LOD range from the lower end (not shown in the Figure). The

width of the LOD range (i.e. $x_{D\max} - x_{D\min}$) increased as more latent variables were added to the model (from 0.03 to 0.15 ppm in the first ten LV), due to the term $h_{0\max}$ in the computation of $x_{D\max}$ (Eq. 30). $h_{0\max}$ represents the maximum distance of a given sample to the calibration center in the score space. Because the distance between two given points in a n -dimensional space increases with the dimensionality of the space, $h_{0\max}$ forces $x_{D\max}$ to separate from $x_{D\min}$ as the dimensionality of the PLS model increases.

Fig. 10 shows the orthogonal scoreplot of the calibration samples for different model complexities, together with the corresponding location of x_D , $x_{D\min}$ and $x_{D\max}$. In the left column (panels **a**, **c** and **e**), moving from top to bottom, we can appreciate the improvement in fitting of the data as the model complexity increases. The scores corresponding to the different concentrations become less scattered in the x-axis and less biased with respect the nominal concentration (black dashed lines), which results in a lower prediction error. Looking now at the right column (panels **b**, **d** and **f**), we can see that x_D , $x_{D\min}$ and $x_{D\max}$ shift towards the blanks as more latent variables are added to the model. When the model is too simple (panel **a**), the scores corresponding to different concentrations are partially overlapped, especially at low concentrations. In this case, neither x_D (2.4 ppm) nor the LOD range ([2.1, 2.2] ppm) satisfy the 5% false positive and false negative rates established by the IUPAC, because the t-student distributions of the blanks and the LOD overlap more than 5% (panel **b**). The underestimation of x_D , $x_{D\min}$ and $x_{D\max}$ is a side-effect of assuming homoscedasticity and using the pooled standard deviation $s_{y,x}$ in the computation of x_D (Eq. 22) and the pooled covariance matrix $\hat{\Sigma}_{pooled}$ in the formulas of $x_{D\min}$ and $x_{D\max}$ (Eqs. 29-30). We can see that the samples of 2.2 ppm show more variability in the x-axis than the rest of concentration levels. When the model complexity is adequate (panels **c** and **d**) or the models are overfitted (panels **e** and **f**), the PLS residuals are homoscedastic and the approximate distribution of the samples at the LOD level can satisfy the 5% overlapping with the blanks, for both LOD estimation methods.

The LOD is estimated using a set of calibration samples, but the sensors will be typically used to predict future samples not seen by the model during the calibration process. Here, a relevant question is whether the calibration LOD will satisfy the 5% false positives and false negatives in the prediction of new samples. To respond this question, we projected validation samples -taken days after calibration- into the calibration score plot (Fig. 11). The plot reveals that the calibration data did not represent all the variability contained in the validation data. The reasons for the increased variability could be related to the overfitting of the model to the calibration data, the cross-sensitivity to humidity and temperature -which varied slightly among different experimental days-, the open-loop control of the sensor temperature or the intrinsic instability of the sensor. The intersection between the blanks and the samples centered at the LOD contains 20 blanks (out of 100) and 20 samples at the LOD (out of 100). This represents an approximate number of false positives and false negatives of 20%, which clearly indicates that, for future operation of the sensor, the LOD was underestimated. To quantify the underestimation factor, we recomputed the LOD using both calibration and validation samples, as a function of the elapsed time between calibration and validation (Fig. 12). The mean LOD (considering multiple sensor units) smoothly increased from 0.79 ppm (calibration) to 1.1 ppm (two weeks after calibration), which represents a 43 % degradation. This highlights that the accuracy of the LOD value in MOX sensors is not as critical as in more precise and stable instruments, due to the expected variability of future data. Reporting a LOD with a safety margin would be a sensible choice in this case, so we could say that in our chemical measurement

process the sensors can detect concentrations above 2 ppm. In the presence of humidity as the main chemical interference, this sensor model can satisfy the LOD specifications of a carbon monoxide (CO) toxic alarm, where the permissible exposure limit is 35 ppm in 8 h, as established by the Occupational Safety and Health Administration (OSHA) [76]. On the other hand, the use for a smoke cessation monitor application, which requires a detection threshold of few ppm [77], needs to be evaluated considering all possible interferences in a CO concentration range closer to the LOD.

Sensor	PLS LV	$[x_{Dmin}, x_{Dmax}]$ (ppm)	x_D (ppm)	$\sqrt{\mathbf{b}^T \Sigma_x \mathbf{b}}$ (ppm)	$[h_{0min}, h_{0max}]$	$\frac{s_{y,x}}{\bar{A}}$ (ppm)	$\frac{K}{I}$
S1	7	[0.72, 0.81]	0.73	0.21	[0.06, 0.34]	0.21	0.997
S3	5	[0.67, 0.75]	0.73	0.20	[0.06, 0.31]	0.21	0.997
S4	5	[0.67, 0.73]	0.71	0.20	[0.06, 0.23]	0.21	0.997
S5	7	[0.78, 0.93]	0.91	0.23	[0.06, 0.47]	0.26	0.997
S6	7	[0.85, 0.98]	0.92	0.25	[0.06, 0.38]	0.27	0.995
S7	5	[0.78, 0.90]	0.86	0.23	[0.06, 0.38]	0.25	0.996

Table 1. Comparison between x_D , x_{Dmin} , x_{Dmax} and the main parameters related to their computation, for six FIS SB-500-12 sensor units. Sensor S2 was excluded from the analysis due to abnormal behaviour. The meaning of the columns is explained in the text. The term $2t_{1-\alpha,v}$ was 3.35, $h_0 = 0.06$, and $\sigma_{Ay} = 0.1$ ppm, for all sensor units.

3.2. Figaro TGS 3870-A04

The baseline-corrected logarithmic sensor conductance patterns of the Figaro TGS 3870-A04 (Fig. 13) reveal a high cross-sensitivity to humidity for all the CO concentrations that we tested (see panel **b** for 2.2 ppm). Changing the humidity content from 15% to 65% r.h. produces an effect in the sensor response at least comparable to a change in CO concentration from 0 to 20 ppm (c.f. panel **a**). PLS models with 6-7 components were built using the samples collected in the first experimental day, in the reduced concentration range 0-9 ppm. The models were then orthogonalized to facilitate their interpretation and to estimate the LOD. For example, interference-free signals reconstructed using the first component of the orthogonal model (Fig. 14) clarified which variables of the response pattern are more sensitive to CO. The most important parts of the signal for prediction purposes are the beginning of the high temperatures and the end of the low heater temperatures. This information could not be extracted from the original signals (c.f. Fig. 13a). Regarding LOD estimation, values of x_D ranging from 0.83 to 1.5 ppm, with an average value of 1.1 ppm, were found. The multivariate LOD ranges were centered at a mean value of 1.2 ppm and a mean width of 0.32 ppm. For the seven sensors units analyzed, x_D was contained inside the LOD range, similarly to the FIS devices. Fig. 15 shows the score plot of one orthogonalized PLS calibration model, together with the projection of the validation data. One interesting comment about the drift and its impact into the LOD reliability should be remarked. The 2.2 ppm validation samples (cyan open points) with the lowest humidity content (15%) drifted towards the blanks, probably due to the high cross-sensitivity

to humidity of this sensor model. Because the blanks and these samples get closer to each other in the x-axis, the model degrades its discrimination power between these two concentration levels. Consequently, the LOD (which was calculated in Day 1 when the scores of both samples were far away) might be no longer valid in Day 13 when the scores are much closer. The orthogonal scoreplot was a useful visualization tool to detect this unexpected situation. The average LOD (over seven sensor units) increased from 1.1 (calibration) to 1.6 ppm (two weeks after calibration), which represents a 46% degradation (Fig. 17).

4. Conclusions

MOX sensors have many challenges to be solved regarding their selectivity and stability. PLS is the de facto multivariate model to calibrate temperature-modulated MOX sensors in the presence of interferences. One research question was if MLPCR would outperform PLS in the analysis of MOX sensor signals with non-iid measurement errors. We found that, for the same level of error, PLS required less components than MLPCR. Unfortunately, the PLS models were still highly complex (5-7 LVs) to properly interpret the models and to estimate common figures of merit, such as the LOD. In the present study, we propose a method to estimate the LOD using the scores of the first component of an orthogonalized PLS model. To illustrate the proposed methodology, we used seven units of two commercial MOX sensors, with different degree of cross-sensitivity to humidity. The sensors tested in this article are just examples of many sensors with similar problems. Post-processing the PLS models with an orthogonalization step resulted in a reduction of the model complexity from 5-7 latent variables to two-component models. These simpler models revealed patterns of the data which were hidden in the original PLS representation and yielded average LOD values (estimated using the proposed approach) of 0.79 ppm (FIS) and 1.1 (Figaro). Through the orthogonal score plot, we showed that these LOD values were contained within the multivariate LOD range estimated by error-propagation, provided the model complexity was adequate. When the model complexity was underestimated, the proposed approach yielded slightly higher LOD values than the upper limit of the multivariate LOD range. The average LOD increased to 1.1 ppm (FIS) and 1.6 ppm (Figaro), when validation samples obtained two weeks after calibration were included in the analysis. This degradation of 43% (FIS) and 46% (Figaro) in the LOD value suggests that, if the interfering factors are not completely controlled, caution should be taken when using MOX sensors for applications requiring LOD values similar to the metrological LOD of the sensors. Adding a safety margin of 1 ppm to the original LOD estimates was recommended in our experiments to guarantee the LOD estimates in a two-week period. This means that the sensors tested in this experimental setup could be used to detect concentrations of CO of around 2 ppm in environments in which the main interfering source is the changes in humidity. This might need to be confirmed with more tests in a concentration range closer to the LOD. In the future work we will include other interfering gases in the sample composition and explore the performance of the sensors in longer timer periods (months).

Acknowledgements

This work was partially funded by the Spanish MINECO program, under grants TEC2014-59229-R (SIGVOL), PCIN-2013-195 (SENSIBLE) and BES-2015-071698 (SEVERO-OCHOA). The Signal and

Information Processing for Sensor Systems group is a consolidated Grup de Recerca de la Generalitat de Catalunya and has support from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (expedient 2014-SGR-1445). This work has received support from the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya and the European Social Fund (ESF). Additional financial support has been provided by the Institut de Bioenginyeria de Catalunya (IBEC). IBEC is a member of the CERCA Programme/Generalitat de Catalunya.

References

- [1] M. Utriainen, E. Kärpänä, H. Paakkanen, Combining miniaturized ion mobility spectrometer and metal oxide gas sensor for the fast detection of toxic chemical vapors, *Sensors Actuators B Chem.* 93 (2003) 17–24. doi:10.1016/S0925-4005(03)00337-X.
- [2] A. Modi, N. Koratkar, E. Lass, B. Wei, P.M. Ajayan, Miniaturized gas ionization sensors using carbon nanotubes., *Nature.* 424 (2003) 171–174.
- [3] H. Li, X. Mu, Y. Yang, A.J. Mason, Low power multimode electrochemical gas sensor array system for wearable health and safety monitoring, *IEEE Sens. J.* 14 (2014) 3391–3399.
- [4] A.J. Bandonkar, J. Wang, Non-invasive wearable electrochemical sensors: a review, *Trends Biotechnol.* 32 (2014) 363–371.
- [5] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R.P. Dick, Q. Lv, M. Hannigan, others, The next generation of low-cost personal air quality sensors for quantitative exposure monitoring, *Atmos. Meas. Tech.* 7 (2014) 3325.
- [6] S. Zampolli, I. Elmi, E. Cozzani, G.C. Cardinali, A. Scorzoni, M. Cicioni, S. Marco, F. Palacio, J.M. Gómez-Cama, I. Sayhan, others, Ultra-low-power components for an RFID Tag with physical and chemical sensors, *Microsyst. Technol.* 14 (2008) 581–588.
- [7] G. Adam, S. Lemaigre, X. Goux, P. Delfosse, A.-C. Romain, Upscaling of an electronic nose for completely stirred tank reactor stability monitoring from pilot-scale to real-scale agricultural co-digestion biogas plant, *Bioresour. Technol.* 178 (2015) 285–296.
- [8] A. Loutfi, S. Coradeschi, G.K. Mani, P. Shankar, J.B.B. Rayappan, Electronic noses for food quality: A review, *J. Food Eng.* 144 (2015) 103–111. doi:10.1016/j.jfoodeng.2014.07.019.

- [9] A. Perera, A. Pardo, D. Barrettino, A. Hierlermann, S. Marco, Evaluation of fish spoilage by means of a single metal oxide sensor under temperature modulation, *Sensors Actuators B Chem.* 146 (2010) 477–482.
- [10] P.K. Sekhar, E.L. Brosha, R. Mukundan, W. Li, M.A. Nelson, P. Palanisamy, F.H. Garzon, Application of commercial automotive sensor manufacturing methods for NO_x/NH₃ mixed potential sensors for on-board emissions control, *Sensors Actuators, B Chem.* 144 (2010) 112–119. doi:10.1016/j.snb.2009.10.045.
- [11] L. Boon-Brett, J. Bousek, P. Moretto, Reliability of commercially available hydrogen sensors for detection of hydrogen at critical concentrations: part II--selected sensor test results, *Int. J. Hydrogen Energy.* 34 (2009) 562–571.
- [12] Y.-S. Kim, I.-S. Hwang, S.-J. Kim, C.-Y. Lee, J.-H. Lee, CuO nanowire gas sensors for air quality control in automotive cabin, *Sensors Actuators B Chem.* 135 (2008) 298–303.
- [13] J. Nicolas, A.-C. Romain, Establishing the limit of detection and the resolution limits of odorous sources in the environment for an array of metal oxide gas sensors, *Sensors Actuators B Chem.* 99 (2004) 384–392. doi:10.1016/j.snb.2003.11.036.
- [14] C. Borrego, A.M. Costa, J. Ginja, M. Amorim, M. Coutinho, K. Karatzas, T. Sioumis, et al., Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise, *Atmos. Environ.* 147 (2016). doi:10.1016/j.atmosenv.2016.09.050.
- [15] M. Kuske, M. Padilla, A.C. Romain, J. Nicolas, R. Rubio, S. Marco, Detection of diverse mould species growing on building materials by gas sensor arrays and pattern recognition, *Sensors Actuators, B Chem.* 119 (2006) 33–40. doi:10.1016/j.snb.2005.02.059.
- [16] G.F. Fine, L.M. Cavanagh, A. Afonja, R. Binions, Metal oxide semiconductor gas sensors in environmental monitoring, *Sensors.* 10 (2010) 5469–5502.
- [17] T. Hübert, L. Boon-Brett, V. Palmisano, M.A. Bader, Developments in gas sensor technology for hydrogen safety, *Int. J. Hydrogen Energy.* 39 (2014) 20474–20483.
- [18] R. Kwor, Carbon monoxide detectors, *Carbon Monoxide Toxicity*. CRC; Boca Raton, FL. (2000) 61–82.

- [19] A.A. Tomchenko, G.P. Harmer, B.T. Marquis, Detection of chemical warfare agents using nanostructured metal oxide sensors, *Sensors Actuators B Chem.* 108 (2005) 41–55.
- [20] T. Šundić, S. Marco, A. Perera, A. Pardo, S. Hahn, N. Bârsan, U. Weimar, Fuzzy inference system for sensor array calibration: prediction of CO and CH₄ levels in variable humidity conditions, *Chemom. Intell. Lab. Syst.* 64 (2002) 103–122.
- [21] K. Brudzewski, S. Osowski, W. Pawlowski, Metal oxide sensor arrays for detection of explosives at sub-parts-per million concentration levels by the differential electronic nose, *Sensors Actuators B Chem.* 161 (2012) 528–533.
- [22] M. Righettoni, A. Amann, S.E. Pratsinis, Breath analysis by nanostructured metal oxides as chemo-resistive gas sensors, *Mater. Today.* 18 (2015) 163–171.
- [23] A.D. Wilson, M. Baietto, Advances in electronic-nose technologies developed for biomedical applications, *Sensors.* 11 (2011) 1105–1176.
- [24] G. Korotcenkov, B.K. Cho, Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey), *Sensors Actuators B Chem.* 156 (2011) 527–538. doi:10.1016/j.snb.2011.02.024.
- [25] A.C. Romain, J. Nicolas, Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview, *Sensors Actuators B Chem.* 146 (2010) 502–506. doi:10.1016/j.snb.2009.12.027.
- [26] G. Korotcenkov, B.K. Cho, Engineering approaches for the improvement of conductometric gas sensor parameters: Part 1. Improvement of sensor sensitivity and selectivity (short survey), *Sensors Actuators B Chem.* 188 (2013) 709–728.
- [27] J.G. Monroy, J. González-Jiménez, J.L. Blanco, Overcoming the slow recovery of mox gas sensors through a system modeling approach, *Sensors.* 12 (2012) 13664–13680.
- [28] J. Fonollosa, S. Sheik, R. Huerta, S. Marco, Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring, *Sensors Actuators B Chem.* 215 (2015) 618–629.

- [29] F.I.S. Inc., FIS GAS SENSOR SB-500-12, (2017).
http://www.fisinc.co.jp/en/common/pdf/SB50012E_P.pdf. Online
(accessed 14/08/2017)
- [30] Figaro U.S.A. Inc., TGS 3870-A04 - for the detection of Carbon Monoxide,
(2013). www.figarosensor.com/products/3870pdf.pdf. Online (accessed
14/08/2017)
- [31] O. Tomic, T. Eklöv, K. Kvaal, J.-E. Haugen, Recalibration of a gas-sensor
array system related to sensor replacement, *Anal. Chim. Acta.* 512 (2004)
199–206.
- [32] P.K. Clifford, D.T. Tuma, CHARACTERISTICS OF SEMICONDUCTOR GAS
SENSORS I. STEADY STATE GAS RESPONSE, 3 (1983) 233–254.
- [33] A.P. Lee, B.J. Reedy, Temperature modulation in semiconductor gas
sensing, *Sensors Actuators B Chem.* 60 (1999) 35–42.
- [34] S. Wold, A. Ruhe, H. Wold, W.J. Dunn III, The collinearity problem in
linear regression. The partial least squares (PLS) approach to generalized
inverses, *SIAM J. Sci. Stat. Comput.* 5 (1984) 735–743.
- [35] I.S. Helland, Some theoretical aspects of partial least squares regression,
in: *Chemom. Intell. Lab. Syst.*, 2001: pp. 97–107. doi:10.1016/S0169-
7439(01)00154-X.
- [36] I.S. Helland, On the structure of partial least squares regression,
Commun. Stat. - Simul. Comput. 17 (1988) 581–607.
doi:10.1080/03610918808812681.
- [37] P.D. Wentzell, D.T. Andrews, B.R. Kowalski, Maximum likelihood
multivariate calibration, *Anal. Chem.* 69 (1997) 2299–2311.
- [38] M.S. Reis, P.M. Saraiva, A comparative study of linear regression methods
in noisy environments, *Journal of chemometrics* (2005) 526–536.
doi:10.1002/cem.897.
- [39] P.T. Hernández, A.J.T. Naik, E.J. Newton, S.M. V Hailes, I.P. Parkin,
Assessing the potential of metal oxide semiconducting gas sensors for
illicit drug detection markers, *J. Mater. Chem. A.* 2 (2014) 8952–8960.
- [40] A.C. Olivieri, Analytical figures of merit: from univariate to multiway
calibration, *Chem. Rev.* 114 (2014) 5358–5378.
- [41] A.C. Olivieri, A simple approach to uncertainty propagation in

- preprocessed multivariate calibration, *J. Chemom.* 16 (2002) 207–217.
doi:10.1002/cem.716.
- [42] K. Faber, B.R. Kowalski, Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares, *J. Chemom.* 11 (1997) 181–238.
doi:10.1002/(SICI)1099-128X(199705)11:3<181::AID-CEM459>3.0.CO;2-7.
- [43] F. Allegrini, P.D. Wentzell, A.C. Olivieri, Generalized error-dependent prediction uncertainty in multivariate calibration, *Anal. Chim. Acta.* 903 (2016) 51–60.
- [44] M. Høy, K. Steen, H. Martens, Review of partial least squares regression prediction error in Unscrambler, *Chemom. Intell. Lab. Syst.* 44 (1998) 123–133.
- [45] F. Allegrini, A.C. Olivieri, IUPAC-Consistent Approach to the Limit of Detection in Partial Least- Squares Calibration, (2014).
- [46] L. Currie, IUPAC, Commission on Analytical Nomenclature, Recommendations in evaluation of analytical methods including detection and quantification capabilities., *Pure Appl. Chem.* 67 (1995) 1699–1723.
- [47] A. Lorber, P.O. Box, K. Faber, B.R. Kowalski, Net Analyte Signal Calculation in Multivariate Calibration, 69 (1997) 1620–1626.
- [48] M.. Ortiz, L.. Sarabia, a Herrero, M.. Sánchez, M.. Sanz, M.. Rueda, D. Giménez, M.. Meléndez, Capability of detection of an analytical method evaluating false positive and false negative (ISO 11843) with partial least squares, *Chemom. Intell. Lab. Syst.* 69 (2003) 21–33. doi:10.1016/S0169-7439(03)00110-2.
- [49] A. Singh, Multivariate decision and detection limits, *Anal. Chim. Acta.* 277 (1993) 205–214.
- [50] M.F. Delaney, Multivariate detection limits for selected ion monitoring gas chromatography—mass spectrometry, *Chemom. Intell. Lab. Syst.* 3 (1988) 45–51.
- [51] A.C. Olivieri, N.M. Faber, J. Ferré, R. Boqué, J.H. Kalivas, H. Mark, Uncertainty estimation and figures of merit for multivariate calibration (IUPAC Technical Report), *Pure Appl. Chem.* 78 (2006) 633–661.
doi:10.1351/pac200678030633.
- [52] J. Ferré, N.K.M. Faber, Net analyte signal calculation for multivariate

- calibration, *Chemom. Intell. Lab. Syst.* 69 (2003) 123–136.
- [53] R. Bro, C.M. Andersen, Theory of net analyte signal vectors in inverse regression, *J. Chemom.* 17 (2003) 646–652.
- [54] L. Xu, V. Bulatov, V. V Gridin, I. Schechter, Absolute analysis of particulate materials by laser-induced breakdown spectroscopy, *Anal. Chem.* 69 (1997) 2103–2108.
- [55] H.C. Goicoechea, A.C. Olivieri, Enhanced synchronous spectrofluorometric determination of tetracycline in blood serum by chemometric analysis. Comparison of partial least-squares and hybrid linear analysis calibrations, *Anal. Chem.* 71 (1999) 4361–4368.
- [56] A.M. de la Peña, A. Espinosa-Mansilla, M.I.A. Valenzuela, H.C. Goicoechea, A.C. Olivieri, Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxycillin and clavulanic acid by stopped-flow kinetic analysis, *Anal. Chim. Acta.* 463 (2002) 75–88.
- [57] B. Efron, G. Gong, A leisurely look at the bootstrap, the jackknife, and cross-validation, *Am. Stat.* 37 (1983) 36–48.
- [58] S. Wold, J. Trygg, A. Berglund, H. Antti, Some recent developments in PLS modeling, *Chemom. Intell. Lab. Syst.* 58 (2001) 131–150.
- [59] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), *J. Chemom.* 16 (2002) 119–128.
- [60] J. Trygg, O2-PLS for qualitative and quantitative analysis in multivariate calibration, *J. Chemom.* 16 (2002) 283–293.
- [61] R. Ergon, PLS post processing by similarity transformation (PLS+ ST): a simple alternative to OPLS Theoretical properties and proofs, (n.d.).
- [62] J. Gabrielsson, H. Jonsson, C. Airiau, B. Schmidt, R. Escott, J. Trygg, OPLS methodology for analysis of pre-processing effects on spectroscopic data, *Chemom. Intell. Lab. Syst.* 84 (2006) 153–158.
- [63] S.J. Bruce, P. Jonsson, H. Antti, O. Cloarec, J. Trygg, S.L. Marklund, T. Moritz, Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: From extraction to data analysis, *Anal. Biochem.* 372 (2008) 237–249.

- [64] M. Bylesjö, D. Eriksson, A. Sjödin, S. Jansson, T. Moritz, J. Trygg, Orthogonal projections to latent structures as a strategy for microarray data normalization, *BMC Bioinformatics*. 8 (2007) 207.
- [65] N. Souihi, A. Lindgren, L. Eriksson, J. Trygg, OPLS in batch monitoring-- Opens up new opportunities, *Anal. Chim. Acta*. 857 (2015) 28–38.
- [66] R. Ergon, Compression into two-component PLS factorizations, *J. Chemom.* 17 (2003) 303–312.
- [67] T.A. Lestander, C. Rhén, Multivariate NIR spectroscopy models for moisture, ash and calorific content in biofuels using bi-orthogonal partial least squares regression, *Analyst*. 130 (2005) 1182–1189.
- [68] R. Ergon, Informative score-loading-contribution plots for multi-response process monitoring, *Chemom. Intell. Lab. Syst.* 95 (2009) 31–34.
- [69] J. Burgues, J.M. Jimenez-Soto, S. Marco, Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, *Anal. Chim. Acta*. (2018).
- [70] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.* 58 (2001) 109–130. doi:10.1016/S0169-7439(01)00155-1.
- [71] R. Ergon, PLS post-processing by similarity transformation (PLS+ ST): a simple alternative to OPLS, *J. Chemom.* 19 (2005) 1–4.
- [72] P.D. Wentzell, M.T. Lohnes, Maximum likelihood principal component analysis with correlated measurement errors: theoretical and practical considerations, *Chemom. Intell. Lab. Syst.* 45 (1999) 65–85.
- [73] K. Danzer, L.A. Currie, Guidelines for calibration in analytical chemistry. Part I. Fundamentals and single component calibration (IUPAC Recommendations 1998), *Pure Appl. Chem.* 70 (1998) 993–1014.
- [74] L.A. Currie, Detection: International update, and some emerging dilemmas involving calibration, the blank, and multiple detection decisions, *Chemom. Intell. Lab. Syst.* 37 (1997) 151–181.
- [75] I.E. Frank, J.H. Friedman, A Statistical of Some Chemometrics View Regression Tools, *Technometrics*. 35 (1993) 109–135. doi:10.2307/1269656.
- [76] O. Safety, H.A. (OSHA), Carbon Monoxide In Workplace Atmospheres

(Direct-Reading Monitor), (1993).

<https://www.osha.gov/dts/sltc/methods/inorganic/id209/id209.html>.

- [77] S.E. Meredith, A. Robinson, P. Erb, C.A. Spieler, N. Klugman, P. Dutta, J. Dallery, A mobile-phone-based breath carbon monoxide meter to detect cigarette smoking, *Nicotine Tob. Res.* 16 (2014) 766–773. doi:10.1093/ntr/ntt275.

Figure Captions

Figure 1. Block diagram of the LOD estimation/validation method described in this paper. The MOX sensor is exposed to dynamic gas mixtures of CO and Humidity. The sensor temperature is modulated using a square heating waveform (V_H). Measuring the sensor resistance (R_s) during a heating cycle produces patterns characteristic of the target gas and the interferences. A PLS model is built with the signals obtained in Day 1. This model is then orthogonalized, and the LOD is estimated from the scores of the first orthogonal component. The LOD estimate is validated by projecting signals measured in the following days into the orthogonal PLS model.

Figure 2. Combinations of CO concentration (nominal value) and relative humidity (measured using a Sensirion SHT75) that were presented to the sensors in each experimental day.

Figure 3. Conductance of a temperature-modulated SB-500-12 sensor (unit S7) during one heating cycle. Data was extracted from the first week of measurements. (panel a) Several measurements corresponding to a constant humidity value of 52 ± 3 % r.h are overlapped. The color indicates the CO concentration. (panel b) Effect of humidity for several concentrations. (0, 11 and 20 ppm). The color indicates the humidity content of the sample.

Figure 4. Scores from (panel a) PLS and (panel b) orthogonalized PLS of a SB-500-12 sensor (unit S7) in the calibration set, demonstrating the better interpretability of the orthogonal scores. The projection of the estimated regression vector $\hat{\mathbf{b}}$ (Eq. 1) is depicted as a solid black line. The number in brackets indicate the percentage of variance captured by that latent variable. The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for \mathbf{T} in Eq. 3). The symbols in panel b are defined in the text.

Figure 5. Auto-scaled and baseline-corrected logarithmic sensor conductance patterns of a SB-500-12 sensor (unit S7). The solid line surrounded by a shaded area indicate the dispersion across replicates as the mean \pm MAD. $MAD = \frac{1}{n} \sum_{i=1}^n |x_i - mean(x)|$. The color indicates the CO concentration: 0 ppm (dark blue), 2.2 ppm (light blue), 4.4 ppm (green), 6.7 ppm (orange) and 8.9 ppm (red). The black dashed line represents the heater voltage.

Figure 6. Equal row error covariance matrix of a SB-500-12 sensor (unit S7), obtained by pooling the row error covariance matrices of each sample (Eq. 12). The row error covariance matrix of each sample was calculated from replicates of the auto-scaled and baseline-corrected logarithmic sensor signals (Eq. 11). The color indicates the variance/covariance (also indicated by the z-axis). The variance was saturated at 0.15 for visualization purposes (only the temperature transitions exceeded this value).

Figure 7. Comparison between PLS and MLPCR in cross-validation error, for six units of the SB-500-12. The dispersion of the RMSECV (ppm) among sensor units is represented as the mean \pm MAD. $MAD = \frac{1}{n} \sum_{i=1}^n |x_i - mean(x)|$

Figure 8. Evolution of the regression vector for different number of latent variables, for a SB-500-12 sensor (unit S7).

Figure 9. Limit of detection as a function of the number of latent variables, for a SB-500-12 sensor (unit S7). The optimum number of LV is indicated with an ellipse.

Figure 10. Calibration scoreplot of the first two orthogonal PLS components of sensor SB-500-12 (unit S7), using 2 LVs (panels a,b), 4 LVs (panels c,d) and 10 LVs (panels e,f). The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for \mathbf{T} in Eq. 3). Panels b, d and f show the blanks and the first concentration level, together with \mathbf{x}_D (vertical solid red line) and $[\mathbf{x}_{Dmin}, \mathbf{x}_{Dmax}]$ (shaded red area). The red crosses simulate the distribution of the scores at \mathbf{x}_D , by replication of the score distribution corresponding to the closest concentration level (i.e. 2.2 ppm in panel b, and 0.0 ppm in panels d and f). The t-

student curves below the blanks and x_D fit the distribution of the scores of the first orthogonal PLS component (t_1^*). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable.

Figure 11. Projection of data from all experimental days into the calibration score-plot of a SB-500-12 sensor (unit S7), using 4 LV. Only the blanks and the first concentration level are shown for visual clarity. The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for \mathbf{T} in Eq. 3). The solid circles represent the calibration data (Day 1) whereas the open circles represent the projection of the validation samples (Days 2-13). The red crosses simulate the distribution of the scores at x_D , by replication of the scores corresponding to the blanks. The t-student curves below the blanks and x_D fit the distribution of the scores of the first orthogonal PLS component (t_1^*). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable.

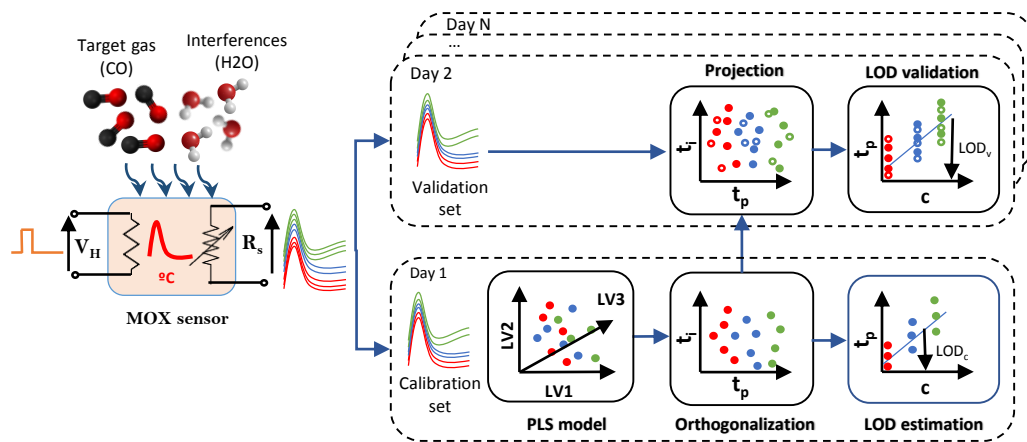
Figure 12. Boxplot of the limit of detection (x_D) of six SB-500-12 sensors, as a function of the days elapsed since calibration. In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.

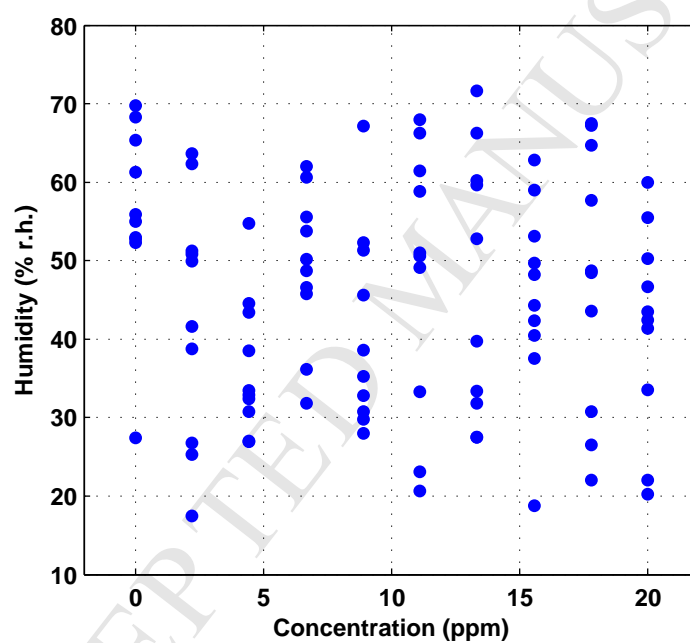
Figure 13. Conductance of a temperature-modulated TGS 3870-A04 sensor (unit S2) during one heating cycle. Data was extracted from the first week of measurements. (a) Several measurements corresponding to a constant humidity value of $52 \pm 3\%$ r.h are overlapped. The color indicates the CO concentration. (b) Effect of humidity at 2.2 ppm of CO. The color indicates the humidity content of the sample.

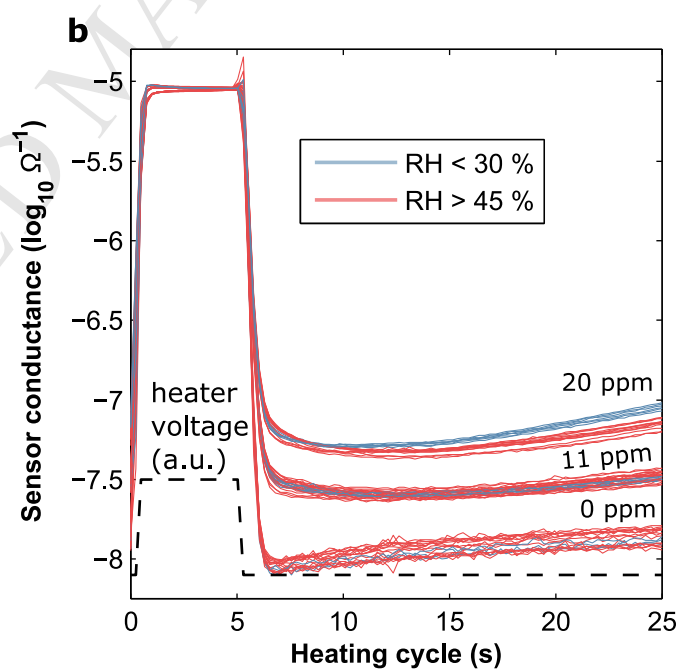
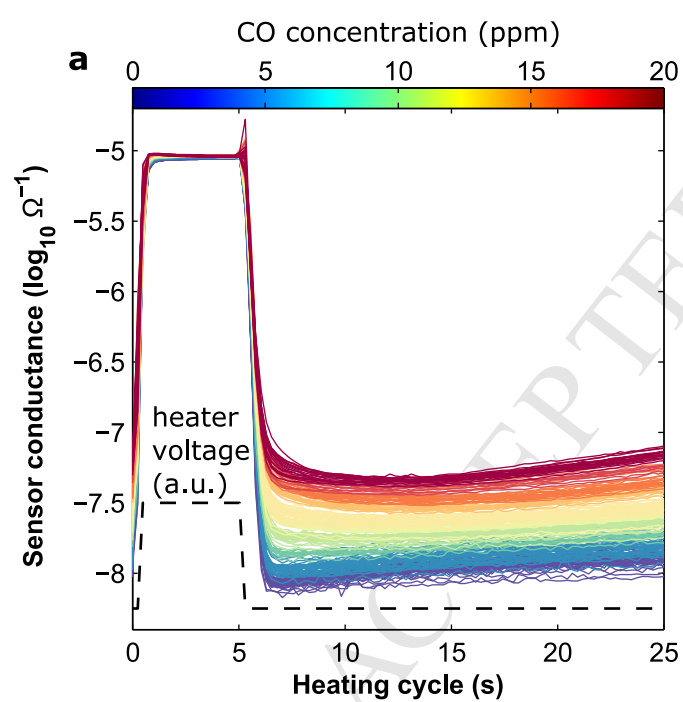
Figure 14. Reconstructed patterns of a TGS 3870-A04 sensor (unit S2), using the first orthogonal PLS component (Eq. 1). The color code represents the CO concentration (ppm).

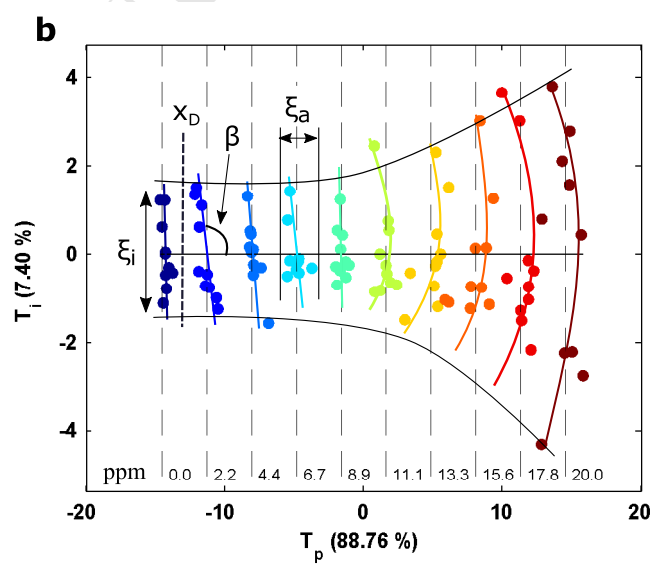
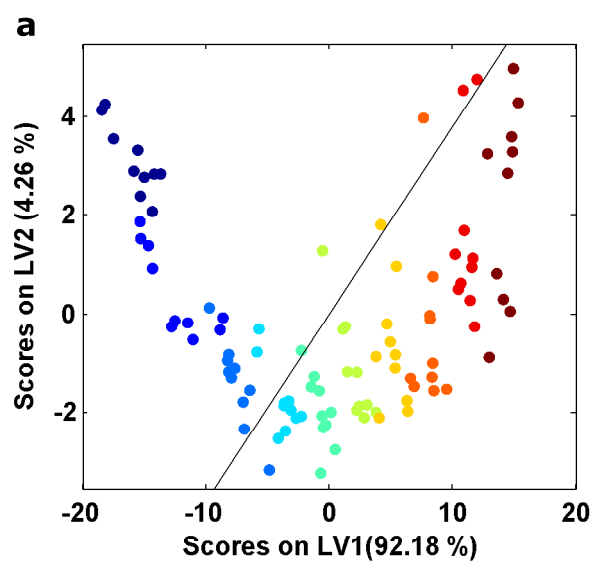
Figure 15. Projection of data from all experimental days into the calibration score-plot of a TGS 3870-A04 sensor (unit S2), using 6 LV. The solid circles represent the calibration data (Day 1) whereas the open circles represent the projection of the validation samples (Days 2-13). The number in brackets in the axis label indicates the percentage of variance captured by that latent variable. The color code represents the nominal CO concentration (ppm), also labeled at the top of vertical black dashed lines which represent the concentration in the score space (solving for \mathbf{T} in Eq. 3).

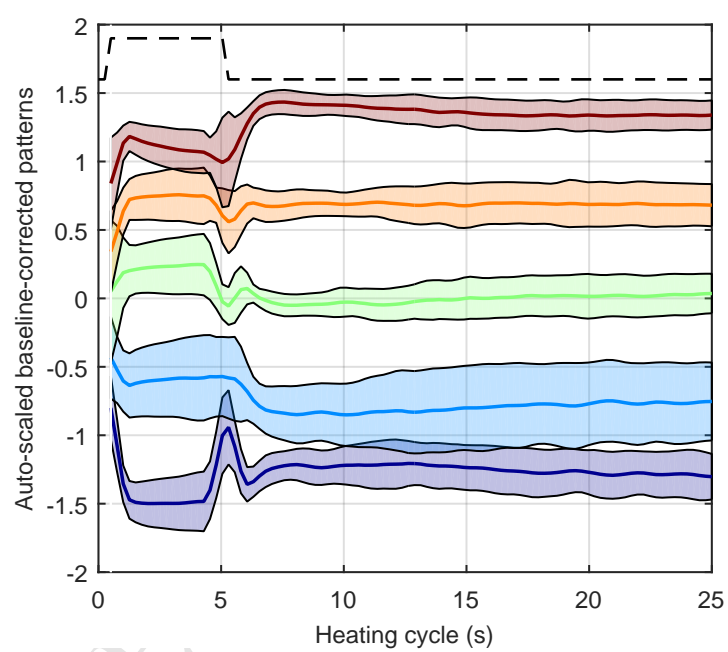
Figure 16. Boxplot of the limit of detection (x_D) of seven TGS 3870-A04 sensors, as a function of the days elapsed since calibration. In each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.

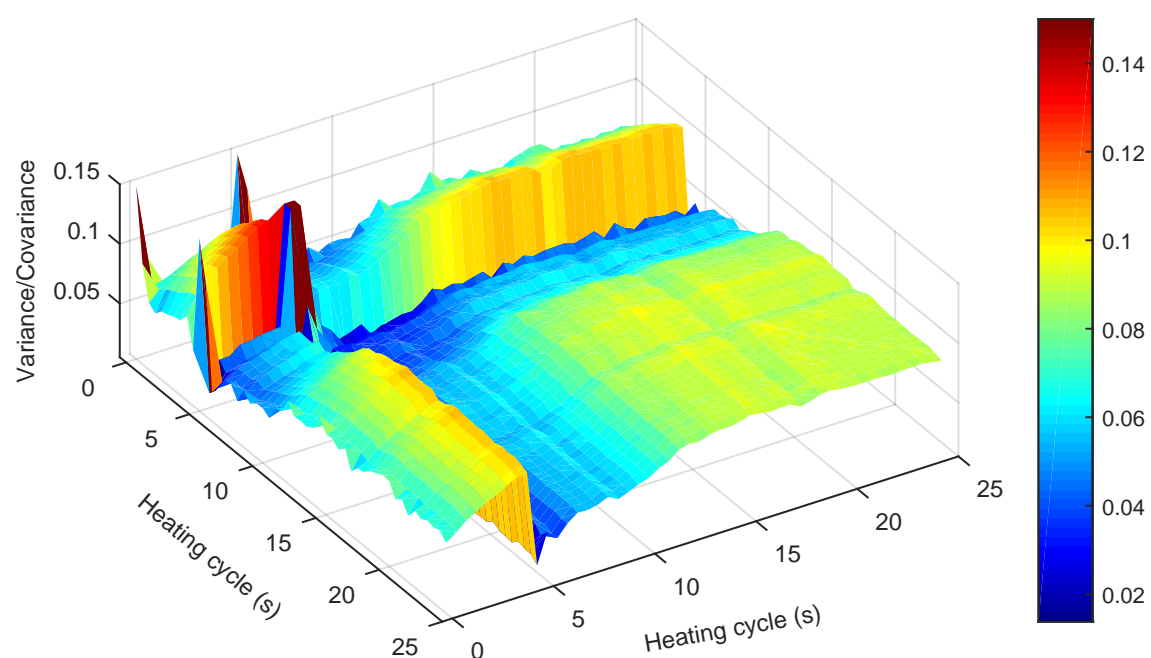


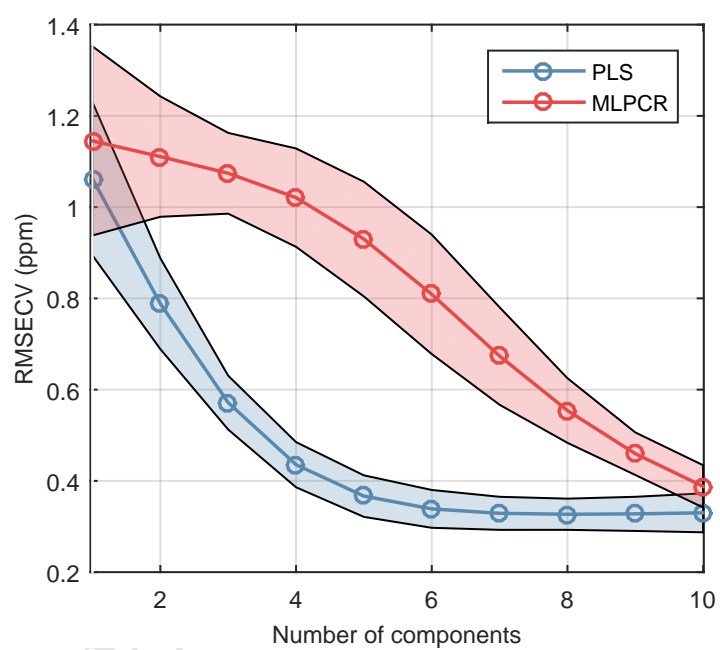


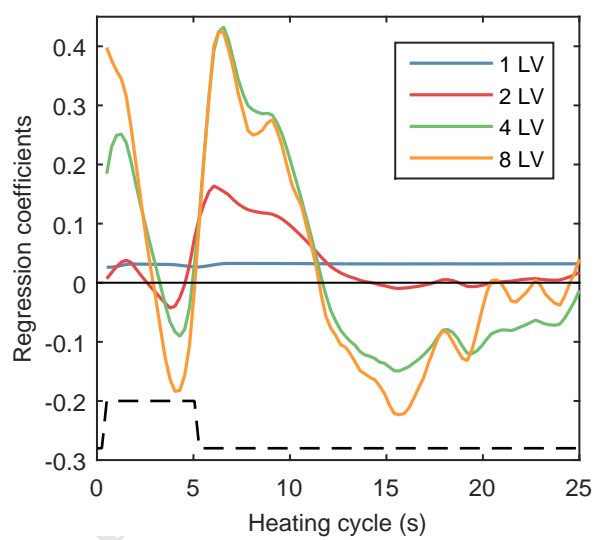


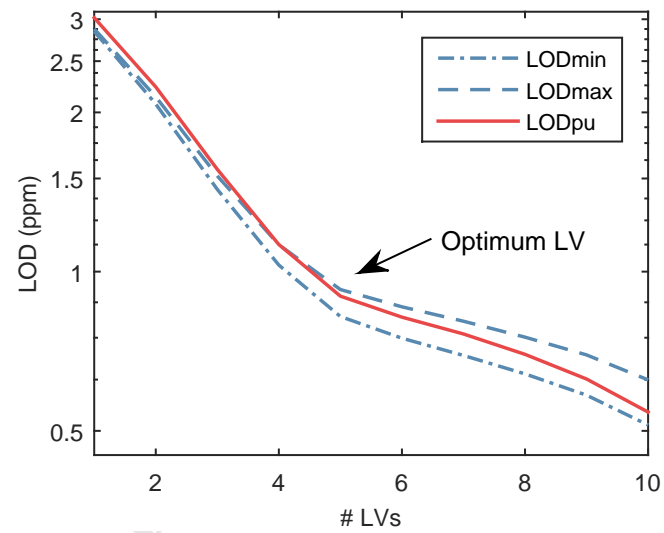


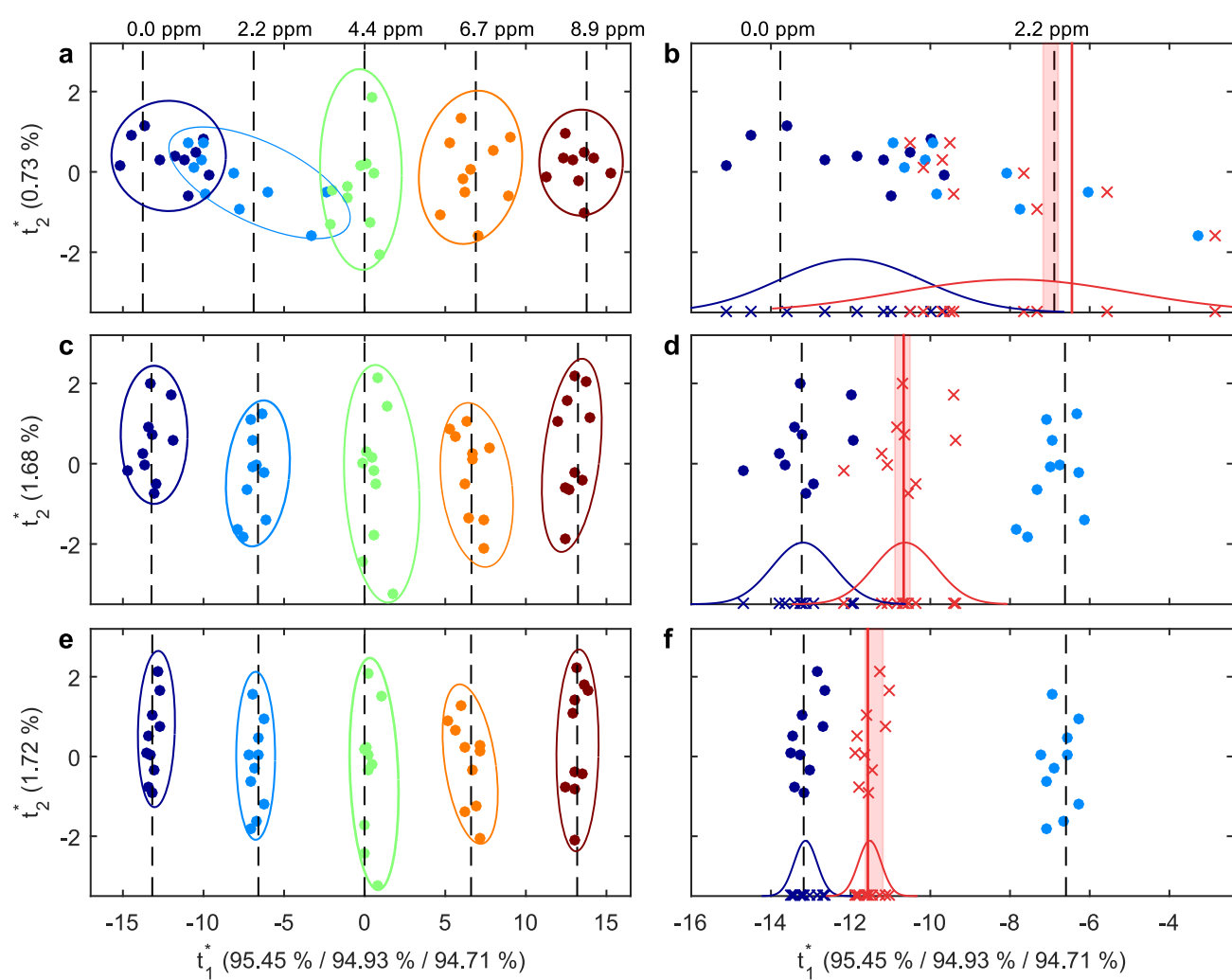


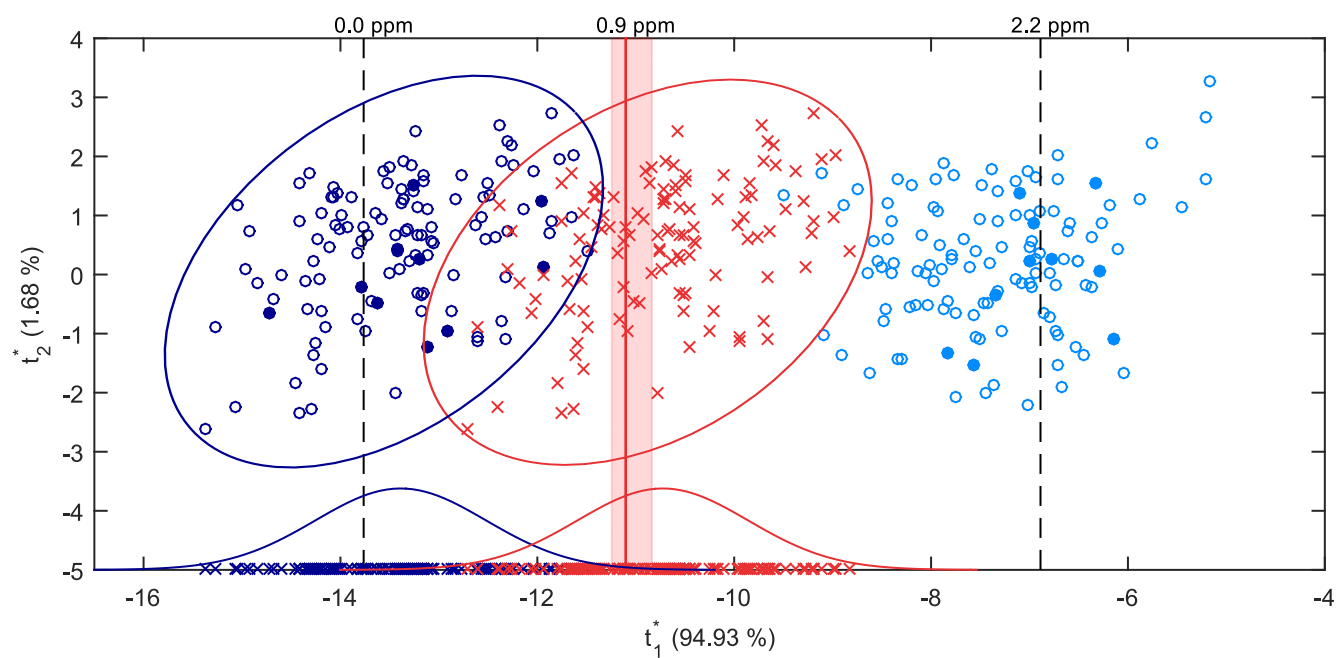


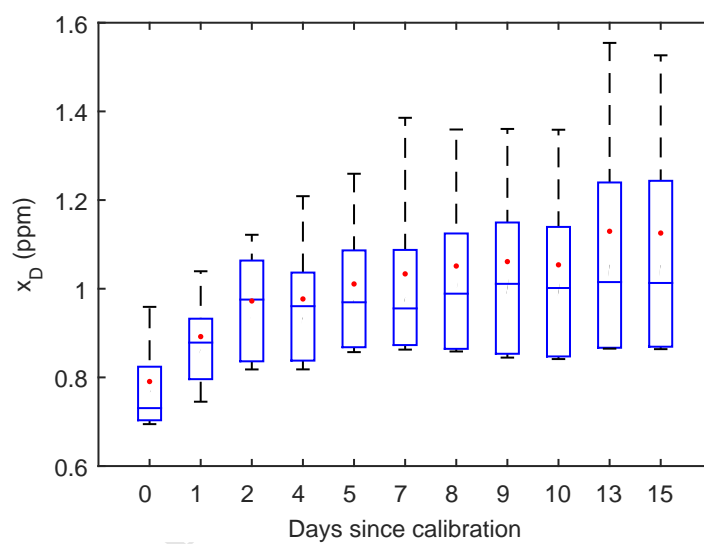


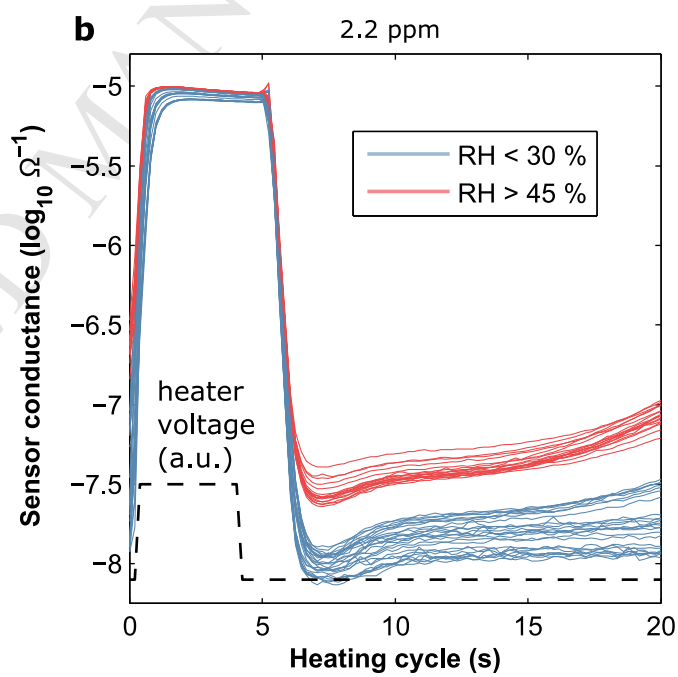
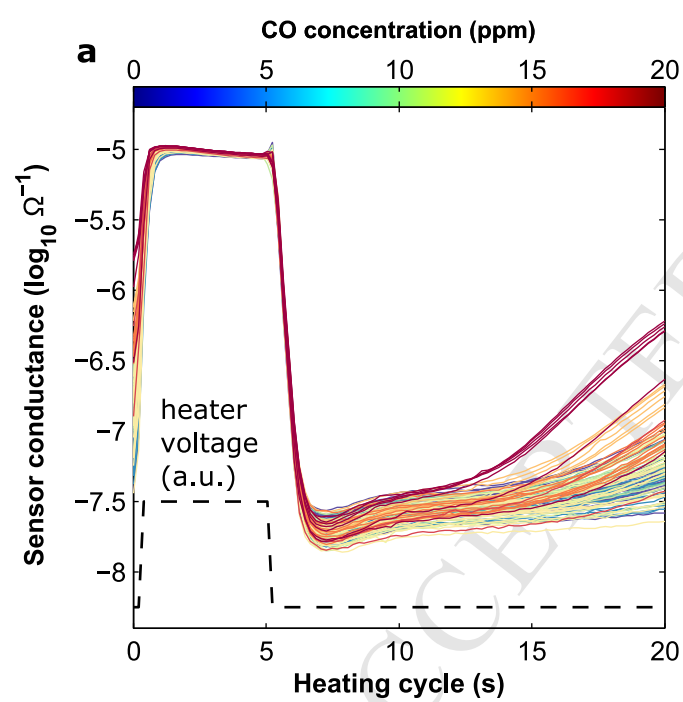


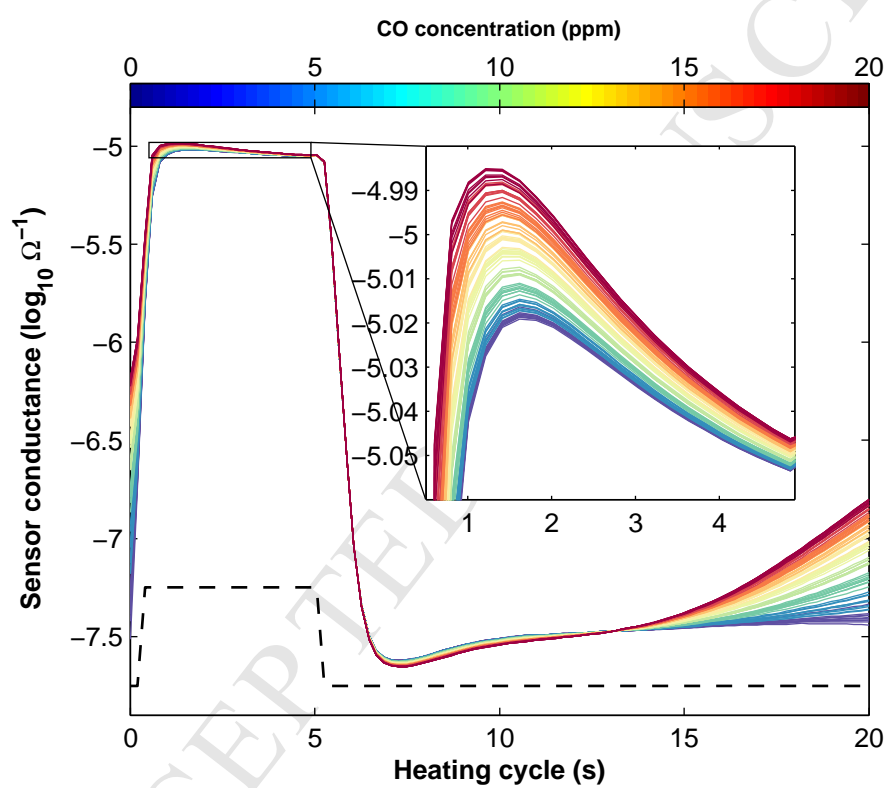


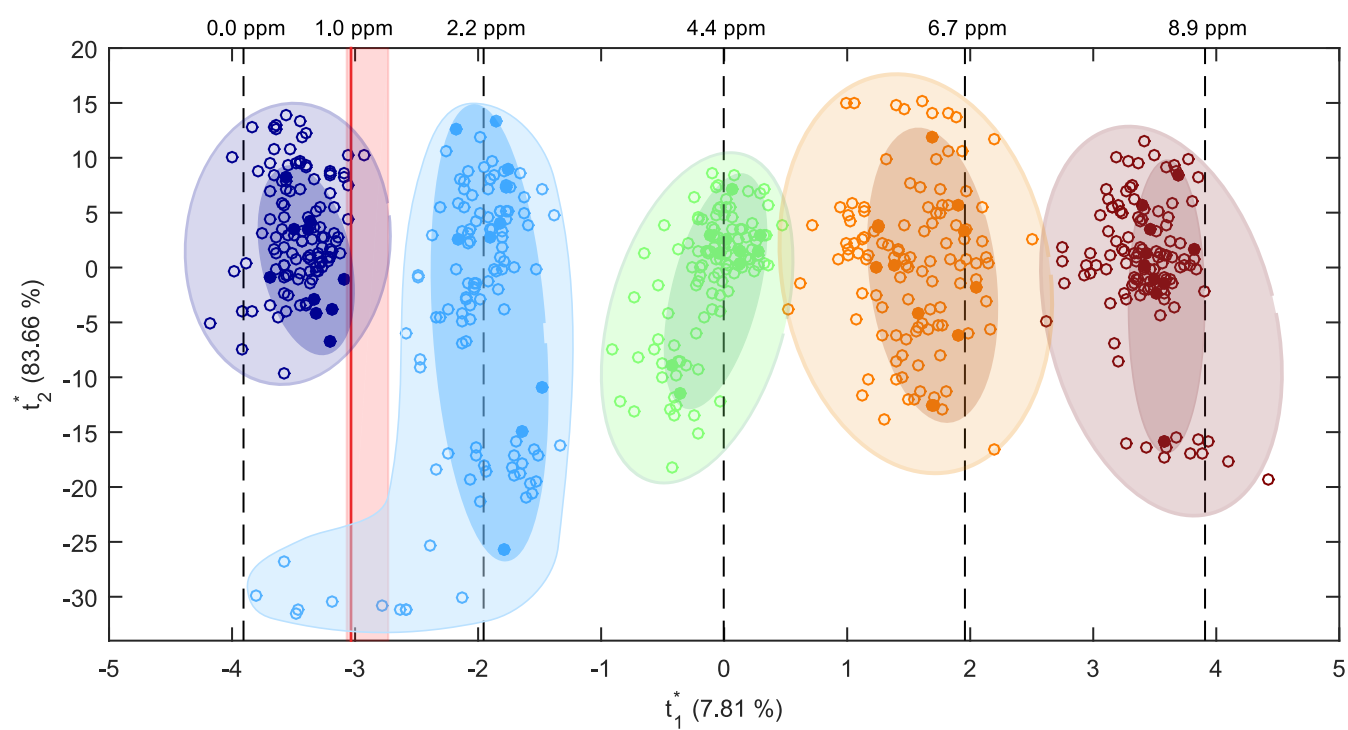


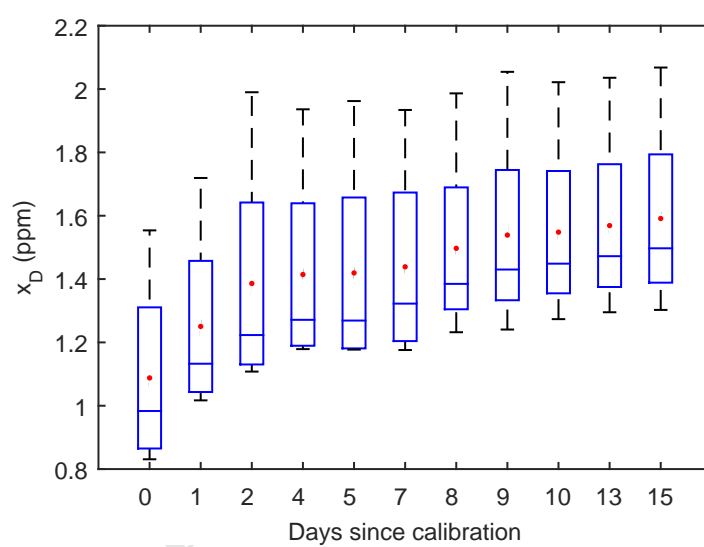












Highlights:

- A methodology to estimate the LOD in orthogonalized PLS models is proposed.
- It reveals the role that interferences (chemical noise) play in the model structure.
- The methodology allows the estimation of the LOD in temperature modulated metal oxide sensors.
- As an illustration, we estimated the LOD of carbon monoxide in the presence of uncontrolled humidity.

ACCEPTED MANUSCRIPT