

A unified protocol for sensor calibration and verification in applications of WWTPs and sewer systems monitoring

G. Xin¹ & J.-L. Bertrand-Krajewski²

¹*Aquateam – Norwegian Water Technology Centre, Norway*

²*INSA-Lyon, France*

Abstract

PREPARED, Enabling Change, is an EU-funded Collaborative Project under the Seventh Framework Programme (FP7). The project aims to gather urban utilities in Europe and worldwide to develop an advanced strategy in meeting the upcoming challenges for water supply and wastewater treatment brought by climate change. This work aims to increase technological capacity and improve performance of traditional water supply and wastewater treatment systems via better use of sensors. For newly developed multivariate sensors, calibration procedures are quite different from manufacturer to manufacturer. Manual operation during calibration and verification is commonly practised in this field due to complexity of the calibration procedure and occurrence of unpredictable events. In addition, many of the sensor manufacturers have not developed detailed instructions for sensor calibration. Therefore, a calibration and verification protocol is needed to guide and unify practice in this area where diversified practice is commonly observed. In this work, a unified protocol for sensor calibration and verification was developed and applied. Two sensor products, an optical sensor and an electronic nose sensor, are discussed as examples for calibration procedures, regression methods and verification methods of calibration in the application of wastewater treatment plant monitoring.

Keywords: PREPARED, multivariate sensors, calibration protocol, verification protocol, local calibration, global calibration, UV-vis spectrophotometer, electronic nose, Partial Least Squares (PLS) regression, Principle Component Analysis (PCA).



1 Introduction

On-line measurements of wastewater parameters in wastewater treatment plants (WWTPs) and in sewer systems can provide crucial information for plant operation and environment management. Some examples of continuous monitoring using ultraviolet visible (UV-vis) absorption and electronic nose sensors have demonstrated the feasibility of these techniques in the WWTP application; however, general failures to apply them in WWTPs are mainly because they need to have constant contact with wastewater or polluted air, resulting in instrument errors, frequent cleaning and recalibration [1].

For newly developed multivariate sensors, calibration procedures are quite different from manufacturer to manufacturer. Frequently, there are neither published calibration protocols nor available adequate standards or certified materials. Many of the sensor manufacturers have not developed detailed instructions for sensor calibration. Although automatic calibration is typically implemented for those sensors, manual operation during calibration and verification is commonly practiced in this field due to complexity of calibration procedure and occurrence of unpredictable events. Therefore, a calibration and verification protocol is needed to guide and unify practice in this area where diversified practice is commonly observed.

PREPARED, Enabling Change, is a EU-funded Collaborative Project under the Seventh Framework Programme (FP7). The project aims to gather urban utilities in Europe and worldwide to develop an advanced strategy in meeting the upcoming challenges for water supply and sanitation brought by climate change. In Work Area 3 of the project, the project team contributes to increased technological capacity and improved performance of traditional water supply and sanitation systems via better use of sensors and models. The objective of this study was to define a unified protocol for sensor calibration and verification. Two types of sensors, UV-vis spectrophotometers and electronic noses are discussed as examples in this article.

2 Calibration and regression methods

All sensors shall be calibrated periodically. Calibration allows a) verifying that the sensor works; b) verifying that the sensor has no offset, sensitivity and linearity errors (This is expected for an ideal sensor. There are no perfect sensors, however, only real sensors affected by imperfections. In this case, calibration allows estimating, quantifying and correcting these errors, and estimating the resulting uncertainties in measurements.); c) adjusting the sensor in order to reduce the observed errors below a level specified as the maximum acceptable level of error by the user (according to use of the measurement results e.g. for monitoring, real time control, etc.) and d) determining a calibration function to correct the residual errors of the sensor and to estimate uncertainties.

For a given sensor, the principle of calibration consists to observe the sensor outputs y (i.e. measured values, preferably at the end of the measurement chain as recorded in data loggers or SCADA systems) delivered when the sensor is submitted to known input values x , and to establish the corresponding calibration function $y = f(x)$. For simple sensors, x and y are scalars (e.g. water level sensors, turbidimeters). For more complex sensors, x and y may be vectors, matrix, etc. (e.g. UV-visible spectrophotometers, electronic nose). The principle does not change, but the determination of the calibration function f may be more complicated and requires sophisticated mathematical tools.

In practice, independently from the numerical and statistical methods to be applied, a calibration is basically carried out as follows: a) choice of N_x standards or certified materials to evenly cover the sensor range of measurement; b) for each x_i standard value, repeated readings of outputs values y_{ik} , with $k = 1$ to N_y . The repeated measurements y_{ik} are necessary to estimate properly uncertainties in sensors outputs from the observed standard deviation of the y_{ik} values. If only one single value y_i is measured, uncertainties in sensor measurements cannot be evaluated; c) determination of calibration function and related variables (uncertainties, etc.). For verification, the protocol is similar except for the 3rd step. If all measured values y_{ik} are within a specified interval around the previous calibration function, then the verification is accepted and the sensor is used with the previous calibration function. If some measured values y_{ik} are outside the specified interval, then the verification is rejected. A new calibration function shall be established. The specified interval is defined by the user according to his/her needs.

In order to establish multilinear or polynomial calibration functions, regression methods are used. Conventionally, two regression methods can be used: ordinary least square (OLS) regression method that is the most frequently used method and Williamson least square (WLS) regression method that may be applied in cases where uncertainties in both X and Y values have to be accounted for. But in cases of multivariate calibration functions, ordinary and Williamson regressions are no longer applicable. Other methods have to be used, like e.g. PLS (Partial Least Squares) regression methods.

PLS generalizes and fuses the principal component analysis (PCA) and the multiple regression methods [2]. It is especially useful in cases where the number of variables is comparable to or greater than the number of observations and/or where there are other factors leading to correlations between variables [3]. The aim of PLS is to predict a variable Y from a variable matrix X and to describe their common structure. The interest of PLS is that it detects the components of the matrix X that are also pertinent to explain Y . PLS decomposes both X and Y as a product of a common set of orthogonal factors and a set of specific loadings. With the notations of Abdi [2], the independent variables are decomposed according to eq. 1, where T is the score matrix and P the loading matrix. The columns of T are the latent vectors. When the number of columns is equal to the rank h of the matrix X , the latent vectors perform an exact decomposition of X . If

all the latent vectors are used, this regression is equivalent to a principal component regression. The matrix P is composed of the direction vectors of the principal axes.

$$X = TP^t \quad (1)$$

$$\hat{Y} = TBC^t \quad (2)$$

Likewise, the dependent variables matrix is estimated according to eq. 2, where B is a diagonal matrix with the regression weights as diagonal elements and C the matrix with the weights estimated for Y .

The objective is to obtain, for each latent vector, a pair of vectors $t = Xw$ and $u = Yc$, with a maximisation of $t'u$. When the first latent vector is found, its effect is removed from both X and Y and the procedure is re-iterated until the desired number of latent vectors is reached [2].

In PLS, the step in which latent variables are determined is a central point in the estimation of the model structure, because the estimation of parameters strongly depends on the obtained latent vectors. There are many methods, with various levels of efficiency, to choose the number of latent vectors. One of these methods calculates the explained variance percentage for X and Y simultaneously, for each number of latent vectors, starting with one latent vector and finishing with the highest possible number, which is equal to the number of independent variables measured. The number of latent vectors is chosen when adding an additional vector increases the percentage of explained variance for X and Y by only a small amount. Cross validation is another method, used to identify, among a given number of candidates, the parameters of the model minimizing the error of prediction. The basic idea consists to separate the initial data set in two sub-sets. The prediction model is then constructed with the first sub-set (calibration) and the second one is used to test the quality of the predictions (verification) [4].

Artificial neural network (ANN) is also commonly used for data analysis for electronic noses. ANN methods are very powerful and are inspired by the way the mammalian brain processes information. The non-linear character of these methods makes them interesting especially for the electronic nose technology.

3 Calibration and verification procedures

3.1 A unified calibration and verification protocol

A unified calibration protocol and a unified verification protocol that are evolved from a protocol developed by the Arizona Department of Health Services [5] are presented in Tables 1 and 2.

Table 1: The procedure of a unified calibration protocol.

Step	Description
1	The details of calibration procedures, including calculations, integrations, acceptance criteria, and associated statistics shall be included or referenced in the operating procedure.
2	An appropriate calibration model shall be selected for the calibration. This is applicably <i>a priori</i> for well-known laboratory instruments. For some field sensors, the best model can be chosen <i>a posteriori</i> based on calibration results and according to variance analysis criteria.
3	The minimum number of references (standards or surrogate tools when standard or certified values are not available) selected to establish calibration shall be more than the power of the calibration model, e.g. for a quadratic calibration model, the minimum shall be 5. As the most appropriate model is not always known <i>a priori</i> , the selected standards shall cover the full range of measurement of the sensor.
4	A blank or background sample shall be included in the standards.
5	The selected algorithm or regression technique shall be described mathematically, and shall provide equations, coefficients, or other parameters necessary to characterise the calibration function uniquely.
6	Acceptability criteria shall be established for calibration purposes. A simple criterion, for laboratory linear instruments, is the correlation coefficient of the regression model. Another criterion is the Fisher exact test, which helps to decide, among various calibration functions, which one performs better. The Fisher exact test approach is better in some cases, e.g. in the case when turbidity values have to be converted into equivalent total suspended solids (TSS) concentrations.
7	For both univariate and multivariate standards, eliminating standard responses (outliers) from calibration is allowed, only under the following conditions: a) the minimum number of standard points are maintained; b) the retained concentration standards shall cover the desired concentration range; c) the retained highest concentration standard defines the upper limit of the concentration range without having to dilute; d) an eliminated standard from calibration shall be adequately documented for cause, statistical evaluation and corrective action.
8	Once samples have been quantified using a finalised calibration function, calibration personnel shall not change the calibration model or the calibration function without performing another calibration. This is especially important for on-line continuous measurements.
9	Laboratories shall perform another calibration under the following conditions: a) after instruments have undergone non-routine maintenance; b) when repeated use or other conditions change their expected behaviour; or c) when the calibration cannot be verified.
10	Calibration personnel shall retain all the raw data necessary to reconstruct instrument calibration.

Table 2: The procedure of a unified calibration verification protocol.

Step	Description
1	The established calibration shall be verified with a reference method(s) prior to analysis of any samples. Samples for the reference method(s) shall be made or taken from the same source as the standards used for calibration unless the verification method allows use of a second source.
2	Standard operating procedures shall document the process for calibration verification, calculations and any additional statistics.
3	The number of calibration verification standards (or surrogate tools) to be analysed shall be based on the selected calibration model, algorithm or regression methods, and the number of standards used during calibration. The minimum number of the calibration verification standards shall be: a) when an instrument is tuned with one calibration standard, using a universally accepted scientific law or scale, then at least one calibration verification standard shall be analyzed at any concentration; b) when two calibration standards are used, at least one calibration verification standard shall be analyzed at a concentration in between the two standard concentrations; c) at least one verification standard shall be analysed when using average of responses or a linear regression analysis. The verification standard concentration may vary within the calibration range; d) at least two verification standards shall be analysed when using a quadratic regression. At least one verification standard concentration should be near the point of inflection; e) at least three verification standards shall be analysed when using a cubic regression. The concentration of two of the verification standards should be near the points of inflection.
4	Unless otherwise required by regulation, method, or program, calibration personnel shall establish acceptance criteria for the calibration verification. For on-line sensors, different values may be chosen, according to both the sensor uncertainties and the maximum level of error. For example, the acceptance may be obtained if the concentration of the standard is measured within +/- three times the standard uncertainty of the sensor in the range of the standard concentration.
5	When the calibration verification is outside the acceptance criteria, the laboratory may re-analyse the calibration verification that failed. If the results of the re-analysed calibration verification are outside acceptance criteria again, the laboratory shall take corrective action. If two consecutive calibration verifications are within criteria after the corrective action, the laboratory may proceed with the analysis. If the two consecutive continuing calibration verifications are outside the criteria, the laboratory shall perform another calibration.

3.2 An example for optical sensor calibration and verification

The submersible Spectro::lyser spectrophotometer from S::CAN (www.s-can.at) is a probe that utilises the UV-visible range (220-720 nm) for simultaneous measurement of multiple parameters. The probe is a 2-beam (one measuring beam and one reference beam), 256 pixel, UV-Vis spectrophotometer, with a Xenon lamp as a light source. All of the controller electronics are included in the probe, along with a data logger and a water level meter. The measurement principle is as follows: substances contained in the medium to be measured reduce the intensity of a light beam when it goes through this medium. The light intensity is measured by a detector over a range of wavelengths. Each molecule of a dissolved substance absorbs radiation at a certain and known wavelength. The concentration of the substance determines the strength of the sample absorption. A single measurement typically takes 45 seconds and the measurement results are recorded and displayed in real-time.

The global calibration is a software tool which allows the instrument to give relatively good values for the parameters of concern in one type of water medium. For each type of water medium, for example wastewater, the global calibration provides an average of calibration based on hundreds data points from multiple WWTPs. No samples need to be analysed for implementation of the global calibration since it is purely software controlled.

Due to different compositions of matrix of concern at a specific site, a local calibration may be required to improve trueness, precision and long term stability. For local calibrations, samples with known concentrations have to be used, and the calibration software (ana::lyte) is used to enhance the accuracy of the probe.

3.2.1 Global calibration of Spectro::lyser spectrophotometer

The method and procedure used in the global calibration is presented by Langergraber *et al.* [6]. Fig. 1 shows a schematic step-wise calibration procedure used for Spectro::lyser calibration in a wastewater application. As the first step, the spectral plausibility is validated to exclude wrong spectra. The most important wavelengths are identified automatically, usually 5 to 10 wavelengths selected. Then calibration is done with turbidity compensated spectra (at infrared 880 nm). The turbidity compensation has two tasks: measurement of turbidity/suspended solids and baseline compensation for measurement of dissolved substances. A recursive procedure includes detection and elimination of outliers, model building, multivariate calibration with PLS, and multiple cross validation. To perform PLS, the spectra and the reference measurements have to be normalised.

A leave-one-out cross validation (LOOCV) strategy is used for the calibration verification [7]. After the model is built using the training set and PLS, a test set is used to validate the robustness of the model. The model can be improved by integrating more principal components into the model. Up to a certain point, the model becomes “over sensitised” by integrating one more principal component, which is when the best model is obtained.

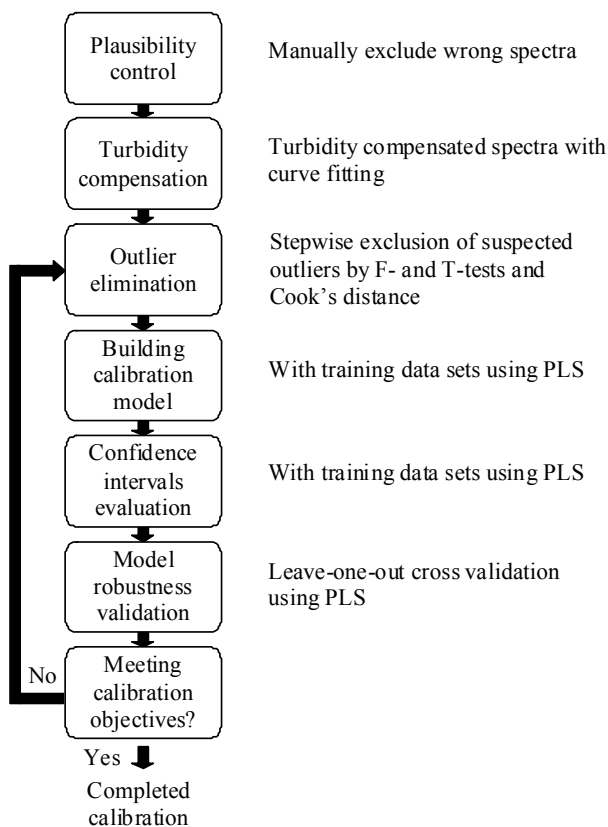


Figure 1: A step-wise procedure for Spectro::lyser sensor global calibration [6].

3.2.2 Local calibration of Spectro::lyser spectrophotometer

A local calibration is a correction of the global calibration. In case a specific PLS is necessary, the user should either contact the manufacturer or do his/her own PLS calibration. An example is given in Torres and Bertrand-Krajewski [8] for urban stormwater, where various calibration methods are compared. The local calibration is based on grab samples analysed for the parameters of concern and can be performed without demounting the sensor (in this case, it is very important that the grab sample has the same characteristics as the water measured by the sensor, e.g. by taking the grab sample at the same location and time as the sensor measurement). Before performing a local calibration, the following checks have to be made: a) the correct global calibration result has been downloaded to the probe; b) a reference analytical method is available and proven for the matrix of concern; c) the probe has been well maintained and is ready for local calibration; and d) the ana::lyte software has been correctly configured and the measurement plausibility has been checked.

In a local calibration procedure [9], a reading of the matrix of concern is executed and results are stored in the calibration database, along with a sample ID and the sampling time. Simultaneously, a grab sample from the same sampling location shall be taken for the reference measurement. Once the reference results are available, the results are linked with the data stored in the calibration database based on the sample IDs and the sampling times. As soon as all the data have been input in the database, a local calibration can be carried out. The local calibration can be carried out either automatically or manually. First, a parameter (e.g. nitrate) and a calibration type (linear or non-linear) are set. Then the local calibration is carried out according to the same principle used in the global calibration (Fig. 1). The result of the local calibration is a new recovery curve (Fig. 2). If there is any data set that falls out of the 95% confidence interval (the yellow lines in Fig. 2), the data set is labelled as an outlier and excluded in the next round of calibration. For the non-linear calibration type, the local calibration can be done only in the manual mode. When the non-linear option is selected, the desired order (power) of a calibration model can be entered. The number of available samples must be taken into account when making the selection: the number of calibration points must always exceed the order (power) of the calibration model. When a satisfied local calibration is obtained, it is saved and activated in the probe.

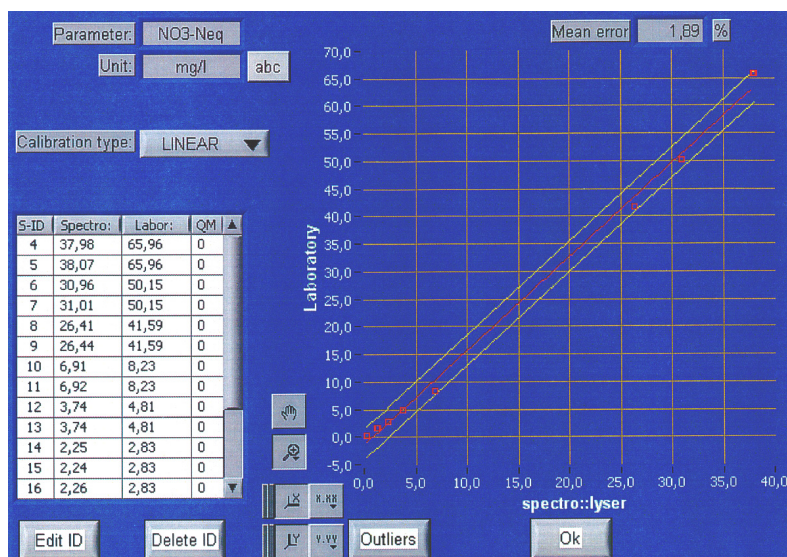


Figure 2: Screen display of ana:lyte for local calibration curve obtained for nitrate in milliQ water using a 10 mm path length [10].

3.3 An example for electronic nose calibration and verification

The EOS Ambiente (EOS507D) manufactured by SACMI (www.sacmi.it) is an electronic nose system based on gas sensors made of metal oxide semiconductors

(MOS) [11]. The system contains 6 sensors with different metal oxide coatings, specifically selected to interact with environmental odorous substances, such as sulphur, ammonia, aromatic compounds. The output of the electronic nose system is reported as UE (EOS unit) that can be correlated to the standard measurement unit for odour concentration (Odour Unit, OU, European norm EN 13725 [12]). Classification of data clusters is done by PCA (the training process) and then the signals in each class (typically associated with a specific chemical) are correlated to a corresponding olfactometric reference using the linear regression function (the recognition process).

A periodical calibration is required since MOS sensor signals drift with time. The automatic calibration is a multiple-step process using clean air generated inside the system and an internal gas reference. Calibration should also be carried out if the humidity (*DPWork*, humidity at the dew point) changes. Only local calibration shall be done for the sensor.

A schematic calibration and verification procedure for EOS507D is shown in Fig. 3. Prior to recognizing different odours or substances, the EOS system has to be trained with the specific odours or substances. In substance training, the system automatically changes concentrations of a specific substance to help recognition of its odour. At the end of the training process, the system stores a

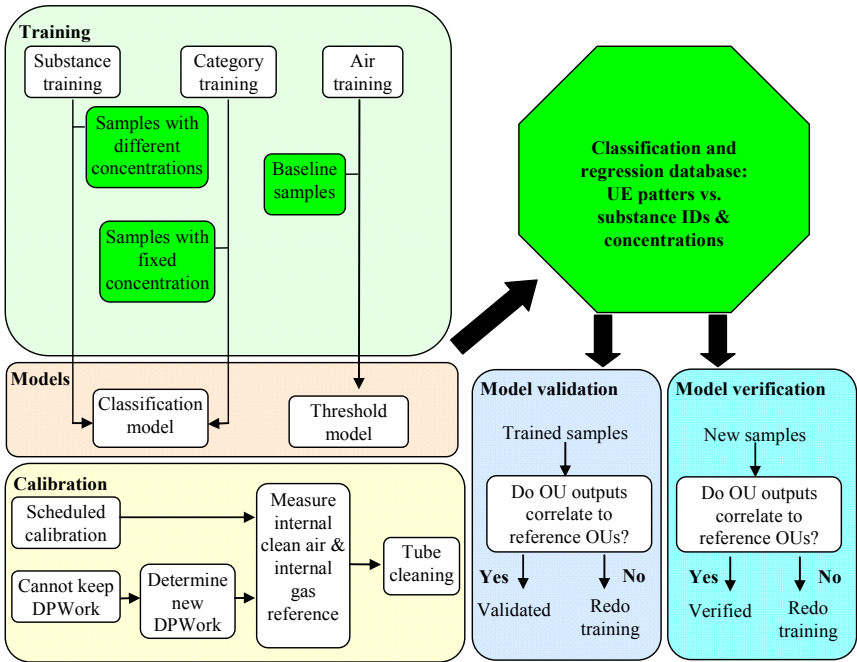


Figure 3: Schematic diagram for the EOS507D sensor calibration and verification procedure.

set of UE pattern correlated with the substance ID and its concentrations. This stored information is named *classification model*. Samples with fixed concentrations are used in category training. Its output is also stored in the *classification model*. For applications in specific environments where maximum odour limitations apply, the system is then trained by a threshold training process (air training). The stored information is named *threshold model*.

Once training is completed, the system is set in recognition mode. In this mode, the system measures the new substances submitted or the environment air continuously and generates output as recognized substance IDs and concentrations. If OU data are used in the training process, the outputs are in the unit of OUs. By submitting to the system the same samples used during training or new samples of the same substances, one can validate the models and/or verify the robustness of the models.

Three electronic noses from SACMI were used for continuous monitoring of odours emission from a WWTP in North Italy in a research project [13]. The results of the study show the critical importance of creation of a suitable training data set for the models, which can maximize the capacity of the electronic noses to the continuous monitoring, and also the opportunity of applying successfully calibrated electronic noses for continuous monitoring of odours at WTPs.

4 Discussion

The indisputable advantage of an on-line sensor can capture variations of matrix of concern at high frequency. However, uncertainty grows rapidly if systematic errors (biases), due to e.g. sensor drift or improper calibration, are allowed to occur. Even relatively small systematic errors may have dramatic effects on the final results from the sensor [14]. Therefore, one cannot overlook the critical importance of calibration and verification of data obtained from a continuously monitoring sensor.

Systematic errors in sensors may be very difficult to detect and remove. Sensor calibration is a way to identify, quantify and remove systematic errors. However, sensor calibration only qualifies the sensor itself, and frequently, non-instrumental errors create larger errors than the instrument errors [15]. Non-instrumental errors may include: *in situ* conditions of use, operator errors, different matrix, ignored offsets, wrong programming and settings, interferences and influential factors like temperature, pressure, humidity, etc. This is why the location and the conditions of use of the sensors should be accounted for as much as possible during the calibration and calibration verification procedures.

References

- [1] Bonastre, R., Ors, J. V., Capella, M. J. and Peris, M, In-line chemical analysis of wastewater: present and future trends. *Trends in Analytical Chemistry*, **24**(2), pp. 128-137, 2005.



- [2] Abdi, H., Partial least squares (PLS) regression. *Encyclopedia of social sciences research methods*. ed. M. Lweis-Beck, A. Bryman and T. Futing: Thousand Oaks, CA, pp. 792-795, 2003.
- [3] Partial Least Squares Regression (PLSR), 2005. Virtual Computational Chemistry Laboratory (VCCL). <http://146.107.217.178/lab/pls/>
- [4] Aji, S., Tavolaro, S., Lantz, F. and Faraj, A., Contribution of bootstrap techniques to PLS regression: application to the prediction model of gas-oil quality Control. *Oil and Gas Science and Technology*, **58(5)**, pp. 599-608, 2003.
- [5] Calibration protocol, version dated on 18 January 2007. The State Assessors Calibration Group of the Arizona Department of Health Services (AZDHS), Phoenix, AZ. www.azdhs.gov/lab/
- [6] Langergraber, G., Fleischmann, N. and Hofstaedter, F, A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in Wastewater. *Water Science and Technology*, 47(12), pp. 63-71, 2003.
- [7] Rudemo, M., Empirical choice of histograms and Kernel density estimators. *Scandinavian Journal of Statistics*, **9(2)**, pp. 65-78, 1982.
- [8] Torres, A., Bertrand-Krajewski, J. -L., Partial Least Squares local calibration of a UV-visible spectrometer used for in situ measurements of COD and TSS concentrations in urban drainage systems. *Water Science and Technology*, **57(4)**, pp. 581-588, 2008.
- [9] S::CAN. Manual of ana::lyte version 5.8. S::CAN Messtechnik GmbH, Vienna, Austria, pp. 1-85, 2009.
- [10] Kiwa, A short evaluation of the S::CAN spectro::lyser. Report No.: BTO 2005.002, Rijswijk, Netherlands, pp. 1-41, 2005.
- [11] SACMI. EOS507D software user's manual. SACMI, Imola, Italy, pp. 1-31, 2010.
- [12] EN 13725, Air quality - determination of odour concentration by dynamic olfactometry. CEN, Geneva, Switzerland, 2003.
- [13] Cappelli, L., Sironi, S., Centola, P., Del Rosso, R. and Grande, M. I., Electronic noses for the continuous monitoring of odours from a wastewater treatment plant at specific receptors: focus on training methods. *Sensors and Actuators B*, **131(1)**, pp. 53-62, 2008.
- [14] Fletcher, T. D. and Deletic A, Statistical observations of a stormwater monitoring programme; lessons for the estimation of pollutant loads. *Proc. of Novatech*, Lyon, France, pp. 1575-1582, 2007.
- [15] Bertrand-Krajewski, J.-L. and Muste, M., Understanding and managing uncertainty (Chapter 6). *Data requirements for Integrated Urban Water management*, ed. T. Fletcher and A. Deletic, Taylor and Francis, Urban Water series - UNESCO IHP, London, UK, pp. 65-90, 2007.

