Calibration transfer between UV-Vis spectrophotometers in drainage systems

Alexandre da Silva Pinheiro

Instituto Superior Técnico, Universidade de Lisboa

Lisbon, Portugal

alexandre.pinheiro@tecnico.ulisboa.pt

2015

Abstract - *The adequate monitoring of water quality is essential to support a proper management of urban drainage systems and to ensure a good physical, chemical and ecological status of water bodies. The use of the absorbance spectrum of radiation in the ultraviolet-visible range (UV-Vis) has been suggested as a viable indicator for estimating quality parameters such as chemical oxygen demand (COD) and total suspended solids (TSS). When a sample is submitted to an estimation model in a different equipment or under different environmental factors than those used to construct the initial model, this may lead to the occurrence of situations of poor effluent characterization. The differences between the spectra obtained in the two conditions can contribute to render the model invalid in the new system. The preservation of a model should be done with the application of calibration transfer techniques between equipments.*

This Master's Thesis aims to compare calibration transfer techniques between UV-Vis spectrophotometers, available in the literature, to ensure the suitability of the characterization of water quality in wastewater drainage systems through the collection of information using different equipments, namely bench and field (submersible) instruments. In order to characterize the water quality, spectrophotometric measurements in the UV-Vis range and reference laboratory procedures were used. PCA models (principal component analysis) were used to identify quality changes in the water matrix and calibration models were developed using PLS (partial least squares regression) for the estimation of COD and TSS in wastewater samples from UV-Vis spectra acquired in a bench scanning spectrophotometer.

Among the three calibration transfer techniques analyzed (SBC – slope and bias correction, SWS – single wavelength standardization and MSR – mean sample residual spectrum correction) the SBC technique was shown to be the most appropriate for the available samples. The SWS method is a good alternative with the particularity of being more independent of the PLS models used, while the MSR proved to be inadequate as calibration transfer technique in the present study.

Keywords – urban drainage; UV-Vis spectrophotometry; calibration transfer; principal component analysis; partial least squares regression

I.   INTRODUCTION

On the agenda of the European Union, in the context of the EU Water Framework Directive [1], are highlighted two major challenges: the quantity and quality of water. Currently, there is a growing imbalance between the quantity and quality of the available water and its necessities to a wide range of uses of this limited resource. The adoption of efficient measures for water management is relevant, and the treated wastewater emerge as alternative water resource, viable and relevant [2]. Municipal drainage corresponds mostly to unit, mixed and pseudo-separated operations systems that transport both domestic wastewater and rainwater.

Thus, in rainy weather is frequent the occurrence of water discharges, with significant pollutant loads in the receptor's water bodies. The management of drainage systems and rainwater inflow in the Wastewater treatment plant (WWTP) should take into account the minimization of discharged flows without any treatment, with minimum energy, reagents and production of sludge [3]. From the discharge of wastewater and the lower dilution capacity of the receiving environment, emerges a high potential of water resources quality deterioration. The protection of the receiving environment is an important component in EU policy, reflected in the legislative field by the establishment of various policies.

The use of continuous monitoring systems, based on the application of spectrophotometric techniques, in urban drainage systems, is an important step towards increasing the knowledge of the operating conditions and also to serve as an early warning for the possibility of anomalous situations. The spectral information allows the estimation water quality parameters such as chemical oxygen demand (COD) and total suspended

solids (TSS), with the support of statistical analysis techniques, once known the expected relationship between the spectrum format and the analytical parameter value obtained in the laboratory [4].

In wastewater drainage systems, the equipment is subjected to a very harsh environment and extreme hydraulic operation variability. The complexity of these systems can contribute to the occurrence of equipment damage and the data failure. Additionally, the dynamics of these systems requires that the equipment can shift from one section to another in the same drainage system. The use of an already calibrated model that relates the spectra obtained with the respective quality parameters is dependent on the conditions under which the spectra were acquired. Changing the conditions of new spectra acquisition can lead to the inadequacy of the model, contributing to a poor effluent characterization. The preservation of an already calibrated model, when changes occur in the context, the equipment should be subject to calibration transfer techniques.

II. OBJECTIVES

The present thesis intends to assess the feasibility of calibration transfer procedures that allows the combination of spectra acquired by different equipments, including submersible and bench spectrophotometers, and integrate them in the same model to estimate water quality parameters (COD and TSS). In this context, it is interesting to identify the calibration transfer techniques for the UV-Vis spectrophotometers in the literature and apply to a practical case – wastewater drainage system. The UV-Vis spectrophotometry has shown great potential in this field and, with this thesis, is intended to contribute to the enlargement the context of their application to the wastewater drainage systems.

III. UV-VIS SPECTROPHOTOMETRY IN WASTEWATER DRAINAGE SYSTEMS

In urban drainage, the wastewater quality monitoring has gained great relevance near the management entities. This necessity is mainly the result of the more recent requirements relating to compliance with legal discharge requirements. Given the requirements of treatment, the necessity of treatment plants improvements and the return of large investments, is necessary to have detailed system information. The implementation of monitoring programs in sewage collectors is not easy, since they must be taken into account the operational constraints relating to the monitoring site. These constraints [5], may occur due to the hydraulic, physical and environmental characteristics of the monitoring site. The quality parameters that should be monitored in urban drainage systems are related, not only, to the operational objectives, but especially with the need to

safeguarding of public health and the preservation of the receiving environment.

The characterization of the effluent quality transported in a drainage system is not simple, since there is great matrix variability. The composition may change daily or seasonally, depending on the basin characteristics [6] and the existence of storms flows. Due to the requirements of law, managing entities should implement monitoring programs to ensure the collection of representative and reliable data of the water quality variation. The main requirement needed for the operation of a drainage system is to know the composition of the effluent disposed at the level of organic matter and nutrients. As provided by law, the assessment of the contamination risk of a receiving environment with organic and solid load is usually measured by the levels of BOD_5, COD and TSS.

Typically, the effluent quality monitoring is obtained by periodic sampling campaigns at predefined locations. The sampling can be done manually or by using automatic collectors. These samples are sent to certified laboratories and analyzed according to standard procedures internationally recognized. In the last decade, has been developed surveillance and alert systems in real time as tools for forecasting and allowing to anticipate the occurrence of pollution accidents and assist the decision making to minimize their effects. These systems have benefited greatly from the generalization of monitoring networks for real-time, which has been supported by the continuous improvement of measuring equipment's. Among other devices, the availability of portable and submersible solutions with recourse to the acquisition of absorbance spectrum in the UV-Vis range has allowed the implementation of online monitoring campaigns in collectors. The use of this type of equipment is very useful since the shape of the spectrum in the UV-Vis range changes according to possible variations in the wastewater quality, as the affluence of rain water and industrial sources [7].

Calibration transfer of UV-Vis spectrophotometers

The spectrum is composed not only by the information relating to the sample, but also by information of the equipment used. This information, if present in all spectra obtained by the same equipment, doesn't affect the results of the statistical model, which incorporates the whole information in the algorithm [8]. The use of spectra acquired by different devices may affect the initial calibration model accuracy. However, in theory, it is possible to adapt the model to new samples. It is a lengthy procedure and, in some cases, it may not be feasible to rebuilt the entire model, since initial samples may be already destroyed, be chemically unstable or correspond to measurements online [9]. There are several circumstances that can introduce modifications in the new spectra that weren't taken into

2

account in the calibration procedure. According to Feudale [10] there are essentially three situations that may make the model invalid:

- Changes in the samples physic-chemical composition.
- Changes or equipment replacements.
- Changes in the environmental conditions to which the equipment is submitted.

In some cases, the response to the new samples is not significantly affected by the new measurement conditions which makes the existing model applicable without the need for corrections and may be adjusted gradually to the new conditions. When this doesn't happen there is the necessity to make a calibration transfer between spectrophotometers using numerical methods. There are several methods of calibration transfer that can be grouped in two ways, depending on the type of adjustment made [10]:

- Adjustment of the quality parameters estimated by the model;
- Adjustment of the spectra obtained from another device, to be similar to those obtained by the original equipment.

To determine the suitability of each method, several parameters can be applied, which highlight the root mean squared error (RMSE) and the slope (a) and the bias (b) from the regression line. The RMSE measures the deviations variance and is expressed in the same units of the greatness under study, expressed by equations (1). The relative root mean squared error (RMSErel) allows determining the magnitude of the error, in percent, and is obtained by the equation (2).

$$RMSE = \sqrt{\frac{\Sigma(O_i - P_i)^2}{n}} \qquad (1)$$

$$RMSE_{rel} = \frac{RMSE}{\bar{O}} x\, 100 \qquad (2)$$

Where:

n: number of elements in the sample
O_i: observed value for the element i
$\bar{O}$: average values determined in laboratory
P_i: estimated value for the element i

In a good calibration transfer model the regression line shall present a slope and bias the closest to 1 and 0, respectively. To RMSErel, in the context of monitoring in drainage systems, classification limits have been proposed [8], in which an acceptable model needs to submit a classifications between *Satisfactory* (10-20%) and *Very Good* (<5%).

IV. CASE STUDY

The monitoring site was selected taking into account the existence of preliminary studies on the same site, the management company availability, their geographical proximity, the characteristics of the contribution basin and sink.

The wastewater sampling campaigns were carried out in the Frielas subsystem that is operated by the SIMTEJO (Saneamento Integrado dos Municípios do Tejo e Trancão, S.A.), who collects and treat wastewater from the municipalities of Amadora, Lisboa, Loures, Mafra, Vila Franca de Xira and Sintra. The WWTP located in Frielas receives effluents from domestic and industrial origins and relevant contributions of rainwater. The monitoring site belongs to Rio da Costa interceptor and is located upstream of the wastewater treatment plant. The collector has a diameter of 1500 mm and receives flow from a basin with 40 km².

Two experimental campaigns were performed on 30 October 2013 and 18 March 2014. Each campaign lasted about 3 hours and 12 samples were taken from each campaign. The UV-Vis spectra obtained by bench and diode-array submersible spectrophotometer (ex situ) and the COD and TSS parameter determined by the standard techniques allowed the support to implementation and evaluation of the transfer calibration procedures in wastewater drainage systems.

V. METHODOLOGY

The methodologies described in this chapter have been proposed to meet the objectives of the dissertation, even though it was necessary to define some procedures suitable to the characteristics of the case study. A combination of two experimental campaigns of wastewater sampling has created the necessity of developing a procedure that allowed the combination of dataset from multiple campaigns. The choice of calibration transfer methods to be applied was made with the support of a preliminary analysis of acquired data. This analysis was performed on the diffence between the absorbance values given by the two equipments for each wavelength ($\Delta_{abs\ rel}$). Also, the existence of data from a campaign in the same place [8] allowed developing a methodology for validation of the models, later titled by online external validation.

Standard COD and TSS laboratory analyses were implemented on collect samples. COD was determined according to the procedure specified by ISO 15705:2002 using test kits (COD Cell Test ref. 14690WTW, Weilheim, Germany), the digestion step was implemented in a digital dry bath (Accublock D1200, Labnet, Woodbridge, New Jersey, USA) and the measurements were performed on a Spectro-Flex 6600 photometer (WTW, Weilheim, Germany). TSS was determined according to the Portuguese standard (NP EN 872 2000) using GF/C glass fiber filters and a HB43-S Moisture Analyser (Mettler Toledo, Greifensee, Switzerland).

UV-Vis spectra of the collected samples were acquired in a laboratory bench scanning spectrophotometer (Specord 200, Analytik Jena, Jena, Germany) between 190 and 800 nm, with 1 nm resolution, using a quartz cell with 10 mm of path length. Also were acquired spectra with an UV-Vis diode-array submersible probe (Spectro::lyser, S::can, Vienna, Austria) between 200 and 750 nm, with 2.5 nm resolution and 5 mm of optical path length. These spectra were acquired off line in the laboratory.

The principal component analysis (PCA), due to its ability to implement to spectral data, was used for matrix analysis and series trends verification. Spectroscopic data may be constituted by large and complex matrices and the PCA allows synthesizing such information. A PCA model enables the spectra representation in one space of reduced dimension, where each spectrum is expressed as scores values relating to the first principal components (PC). The spectra analysis through the PCA allowed:

- Identify outliers and clusters in the spectra data;
- Observe original and diluted samples groups;
- Analyze the differences between spectra obtained by a bench and a submersible spectrophotometer;
- Analyze the differences between spectra obtained in the same samples but with different ages, by a bench spectrophotometer.

The pre-processing procedures available were selected and applied the mean centering to the spectra data. The PCA model was chosen as a primary support in the identification of outliers and clusters and possible rejection thereof. Under this procedure, a spectrum is considered outlier if it is out of range at 95% confidence in the score plot and/or the Hotelling T^2 vs $Q_{Residuals}$. The acceptance or rejection of a spectrum classified as an outlier must be preceded by an evaluation of the causes that lead to demarcate of the remaining spectra.

The spectra analyzed in PCA models (Matrix X) associated with the analytical parameters values (Matrix Y) are the basis for the partial least squares (PLS) models construction. Calibration models were developed, with the spectra from equipment A, for each parameter (COD and TSS).

From the available data were formed two distinct groups: the cross-validation set (corresponding to 2/3 of the data) and the external validation set (corresponding to 1/3 of the data). The identification of the data, included in each group, was defined in order to maintain the same ratio between the original samples and the various dilutions.

From the available pre-processing techniques, the mean centering was elected and a full cross-validation procedure (contiguous blocks, with 6 datasets) to evaluate the adequacy of the model for the training set. IPLS models are implemented for the two parameters in question in order to select the range(s) of wavelengths that contribute(s) to more robust PLS models. To determine which PLS or IPLS model is more robust, both RMSEP and R^2 parameters from the two models are compared and the best is chosen.

PCA, PLS and IPLS models were developed with the support of Matlab R2007a (The Matworks Inc., EUA) with PLStoolbox 3.0 supplement (Eigenvector Research Inc., USA).

In all experimental studies, the data can be related to several campaigns which allowed obtaining multiple sets of data (spectral data and values for water quality parameters). In order to investigate the possibility of aggregating data from multiple campaigns on the same model was defined the procedure summarized in Figure 1.
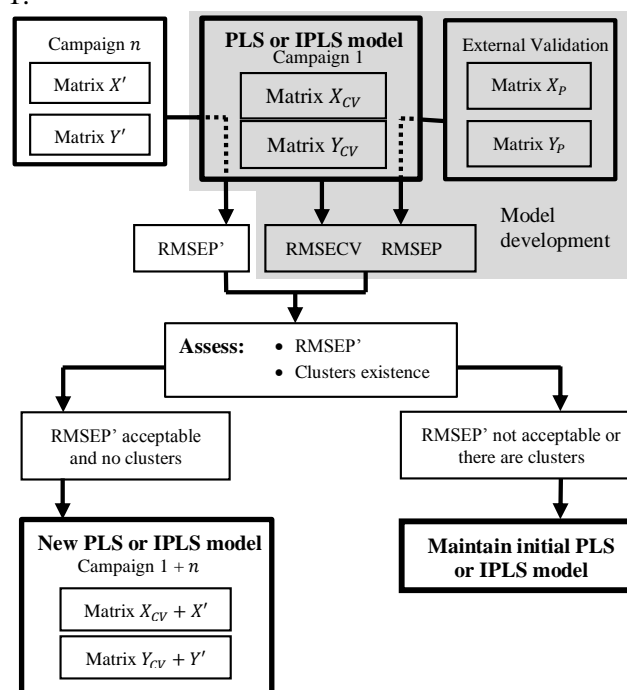

Fig. 1 - Procedure to determine the possible introduction of several matrices of data (different campaigns) in the same model for COD and TSS parameters

PLS and IPLS models were implemented for each analytical parameter (COD and TSS) using UV-Vis spectra acquired by the bench spectrophotometer and the respective COD or TSS values. These models were developed with only two-thirds of the first campaign data (X_{CV} and Y_{CV} matrices, cross-validation set). For each parameter, the developed PLS and IPLS models were compared and the robustness of each one was determined based on RMSECV. Then the values of the parameters are estimated for the remaining one-third of the data (X_P and Y_P matrices, external validation set) and determined its RMSEP.

Subsequently, both spectra data and COD/TSS values, of the second campaign (X' and Y'), were entered into the model and the RMSEP was determined for both models. If the spectra of each campaign do not differentiate into cluster and RMSEP associated to further campaigns are the same order of magnitude of

4

the first campaign, it is developed a set PLS or IPLS model for each parameter.

In order to know if it's possible to use the spectral data of samples analyzed at different times (after collection) and use that data in a single model, a procedure has been defined based on the same assumptions of the procedure expose in Figure 1.

Calibration transfer techniques

The calibration transfer was performed between a bench spectrophotometer (equipment A) used in the laboratory and a submersible diode-array spectrophotometer (equipment B), also used in the laboratory, but in offline mode. The bench spectrophotometer was considered the master equipment and the submersible spectrophotometer as secondary equipment. The calibration transfer techniques were applied to the spectra data from the wavelength range between 200 and 740 nm, with 5 nm resolution.

Initially, from the various calibration transfer techniques referenced, and given the fact that not all were recommended, the most appropriate for analysis were chosen. A preliminary analysis was made on the difference between the absorbance values given, by the two equipment's, for each wavelength ($\Delta_{abs\ rel}$). Thus, three calibration transfer techniques were selected to assess, specifically:

- *Slope and bias correction* (SBC);
- *Single wavelength standardization* (SWS);
- *Mean sample residual spectrum correction* (MSR);

In the following procedures, the COD and TSS parameters are called "parameter" or "Par" and the spectral data acquired by the equipment A and equipment B by "Spectra$_A$" and "Spectra$_B$", respectively.

The *slope and bias correction* (SBC) consist on direct comparison of parameter estimates for Spectra$_A$ and Spectra$_B$, for the same sample. This method includes the following steps:

- Development of a IPLS model for Spectra$_A$ and estimate the quality parameter (Par$_A$);
- Insert Spectra$_B$ in the model and make a preliminary parameter estimation (Par$_{B\ Pre}$);
- Determine the regression line between Par$_A$ (y-axis) and Par$_B$ (x-axis). The slope (a) and bias (b) are determined by: $Par_{B\ Pre} = a * Par_A + b$;
- Correction of the estimated parameters values for Spectra$_B$, using the equation: $Par_{B\ Corr} = (Par_B - b)/a$.

The *Single wavelength standardization* (SWS) performs a standardization of the spectral response by multiplication or division and covers the following steps:

- Determine an average absorbance value for each wavelength λ_i for Spectra$_A$. The resulting matrix consisting of all average absorbance values is called as standard spectrum (St$_A$);
- Determine the standard spectrum (St$_B$) for Spectra$_B$, repeating the previous step;
- Calculate the corrective factor (f_i) for each wavelength λ_i using the equation: $f_i = St_{A,i}/St_{B,i}$;
- Implement the corrective factor to each wavelength of Spectra$_B$ according to equation: $Spectra_{B,Cent\ i} = Spectra_{B,i} * f_i$;
- Insert Spectra$_{B\ Cent}$ in the previous developed IPLS model and obtain the parameters estimates (Par$_B$).

The *mean sample residual spectrum correction* (MSR) is based on the adding or subtracting the difference between the standard spectra A and B (St$_A$ and St$_B$) to Spectra$_B$. The procedure includes the following steps:

- Determine standard spectra (St$_A$ and St$_B$) to Spectra$_A$ and Spectra$_B$, as explained in the previous procedure;
- Determine the corrective factor (S$_i$), for each wavelength λ_i, using he equation: $S_i = St_{A\ i} - St_{B\ i}$ (calculated from the difference);
- For each wavelength λ_i of the Spectra$_B$, the absorbance value is corrected applying the corrective factor, getting the Spectra$_{B\ Corr}$ matrix, by the following equation: $Spectra_{B,Corr\ i} = Spectra_{B\ i} + S_i$;
- Insert Spectra$_{B\ Corr}$ in the previous developed IPLS model and obtain the estimated parameters (Par$_B$).

The calibration transfer techniques previously described were developed using the cross-validation set (corresponding to 2/3 of spectra). In the case of Spectra$_A$ are the same spectra used to develop the IPLS models. For this set was made a preliminary assessment, where the Par$_A$, obtained by inserting Spectra$_A$ in the IPLS models, were compared with Par$_B$ obtained through Spectra$_B$ after implement the SBC, SWS and MSR methodologies. The development of a linear regression between Par$_A$ and Par$_B$ allowed the evaluation of the suitability of each technique. The evaluation of the precision of the estimated parameters compared to the analytical values determined in laboratory was made using RMSE and RMSE$_{rel}$ for three sets of Spectra$_B$: the cross-validation set (2/3 of Spectra$_B$ acquired in laboratory), the external validation set (1/3 of Spectra$_B$ acquired in laboratory) and the online external validation set (Spectra$_B$ acquired online, *in situ*).

The online external validation counted with data from a campaign carried out in 2010 at the same location. From this campaign resulted water quality data, measured in laboratory, and spectrophotometric data, obtained by a bench spectrophotometer and a diode-array spectrophotometer installed in the collector

(online). These data was subjected to an exploratory analysis of the methodologies outlined in this work, with very satisfactory results [8].

VI. RESULTS AND DISCUSSION

Preliminary treatment of monitoring data

The two campaigns conducted have allowed the collection of 24 samples (12 for each campaign). For each sample from the first campaign (F1) was generated two more subsamples, by dilution 1:2 and 1:4 with tap water, totaling 72 samples. For each sample from the second campaign (F2) were generated more three subsamples, by dilutions 1:2, 1:4 and 1:6 with tap water, totaling 96 samples.

Standard COD and TSS laboratory analyses were performed on the original collected samples after collection and the dilution factor was used to determine the subsamples values. For the cross-validation set were obtained COD and TSS values between the ranges 27-697 mg O_2/L and 9-525 mg/L, respectively. For the online external validation set were obtained COD and TSS values ranged from 408-531 mg O_2/L and 175-225 mg/L, respectively

The first campaign samples was subjected to three analysis (spectrophotometric only) with different time spacing's. One analysis after collection and another two after 10 and 20 days.

For each sample and subsample of wastewater was determined the UV-Vis spectrum with two different equipments: a bench and a diode-array submersible spectrophotometer (Fig. 2).
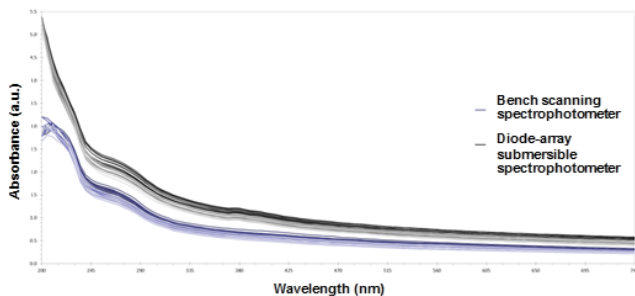

Fig. 2 – Example of UV-Vis spectra obtained on a bench spectrophotometer (equipment A) and a diode-array submersible spectrophotometer (equipment B) - results from raw spectra of the first campaign (F1)
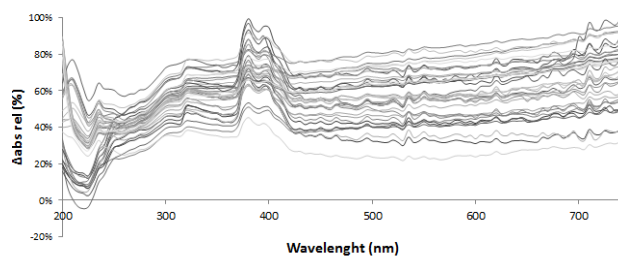

Fig. 3 - Differential analysis of spectra acquired on the same samples using a bench scanning spectrophotometer (equipment A) and a diode-array submersible spectrophotometer (equipment B), dataset from one of the sampling campaigns

The spectra obtained from equipments A and B were evidently different (Fig. 2 and 3) and was not detected absorbance profile shifts in the wavelength axis (Figure 3). In addition, differences are visible across the spectral range, however do not occur uniformly. The calibration transfer methodologies clearly had to involve the whole available spectral range. So, the SBC, SWS and MSR methods were selected for testing.
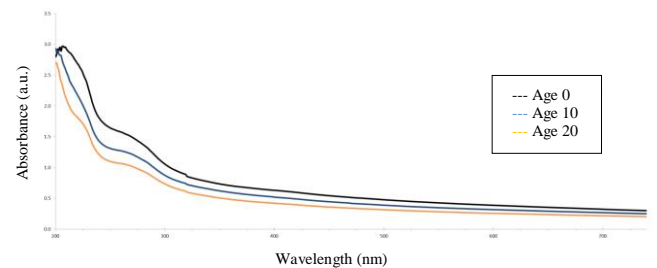
The Fig. 4 represents the spectra of the same sample analyzed in three different times after sampling (after 0, 10 and 20 days). It's possible to verify that have the same pattern over the wavelength range, but their absorbance values are different. There is a tendency that the absorbance decreases with the samples aging.

Matrix analysis of UV-Vis spectra

In Table 1 are marked the number of spectra used for the construction of each model, and the number of outliers for each dataset analyzed. For each model developed, it took only two principal components (PC) to obtain a high percentage of variance captured. Given the dataset size, the number of outliers is relatively low.

Table 1 - PCA models for matrix analysis after outliers removal

Model id	Matrix	n	outliers	CP	Captured variance (%)
1	LAB 1.1	72	0	2	99.7
2	LAB 1.2	72	0	2	99.8
3	LAB 1.3	72	3	2	99.9
4	LAB 2.1	96	9	2	99.9
5	SCAN 1.1	72	1	2	99.9
6	SCAN 2.1	96	4	2	99.9
7	LAB 1.1 + SCAN 1.1	72		2	99.7
8	LAB 2.1 + SCAN 2.1	90		2	99.8
9	LAB 1.1 + LAB 2.1	80		2	99.8
10	SCAN 1.1 + SCAN 2.1	82		2	99.9
11	LAB 1.1 + LAB 1.2 + LAB 1.3	107		2	99.8

Legend:
id: model identification; *n*: number of spectrums included in the matrix; *CP*: number of principal components;

Analytical parameters estimation based on UV-Vis spectra

PLS and IPLS models were implemented with the spectral data obtained by a bench spectrophotometer for the first campaign samples after the removal of outliers.

A total of 84 samples were used for model development (24 samples collected in two experimental campaigns more 60 samples obtained by dilution with tap water to obtain more samples by simulating dilution

phenomena) and 12 samples for online external validation.

Par	Matrix	Model	Proc	n	VL	λ (nm)	RMSE (mg/L)	RMSE$_{rel}$ (%)	R^2
CQO	LAB 1.1 + LAB 2.1 + CQO 1.1 + CQO 2.1	PLS	VC	50	2	200 - 740	18.2	8.8	0.988
		IPLS	VC	50	2	270 - 340	14.3	6.9	0.993
		IPLS	PE	27			19.9	9.5	0.989
SST	LAB 1.1 + LAB 2.1 + CQO 1.1 + CQO 2.1	PLS	VC	50	2	200 - 740	38.4	36.9	0.832
		IPLS	VC	50	2	540 - 560; 585 - 695	22.4	21.5	0.924
		IPLS	PE	26			27.8	33.6	0.843

Legend:
Par: parameter; **Proc**: procedure; **n**: number of spectrums included in the matrix; **λ**: optimal wavelength ranges (nm); **VL**: number of latent variables; **VC**: cross-validation; **PE**: external validation

By examining Table 2, it was concluded that for the estimation of COD and TSS is preferable to implement IPLS models rather than PLS. Thus, it was possible to reduce errors of 8.8% to 6.9% and from 36.9% to 24.5%, respectively. As in previous models, the latent variables requires were only two. Once more, it can be seen that the error associated with the estimation of the SST parameter are significantly higher than the COD parameter. This difference can relate to the phenomenon of attenuation of the light used to estimate the two parameters [11].

For the SST parameter, the fact that the IPLS models selected wavelengths in the visible region (where the light scattering phenomenon is more important) may

affect the absorbance values obtained, since this phenomenon is sensitive to variations in composition, size and morphology of suspended solids.

Calibration transfer between UV-Vis spectrophotometers

The bench spectrophotometer (equipment A) was considered as the main equipment and a submersible spectrophotometer (equipment B) as the secondary equipment. The calibration transfer procedures between spectrophotometers were applied to the spectra data from the wavelength range between 200 and 740 nm, with 5 nm increments.

In order to estimate the COD and TSS parameters on the residual water samples, the IPLS developed models were used for the spectral data of the first and second campaigns. These models were based in the spectral data obtained by the bench spectrophotometer.

Figures 4 to 6 represent the regression lines, before and after the application of the three calibration transfer methods. The regression line of the pre-transfer calibration shows the dispersion between the results of the equipment A (ParA), through IPLS models, and the obtained results relating to equipment B in the IPLS models. The regression line on the post-transfer calibration shows the dispersion between the estimated results for ParA and ParB, after the transfer calibration by SBC, SWS and MSR methods.
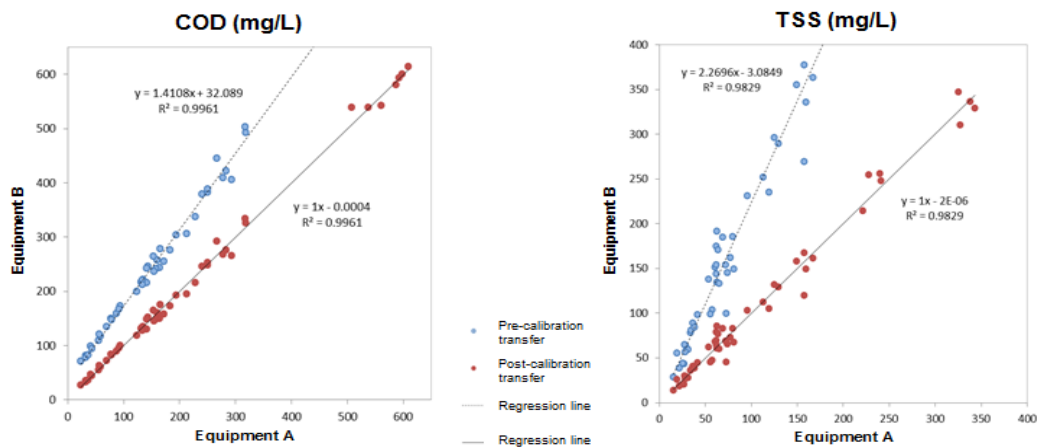

Fig. 4 - Transfer calibration by slope and bias correction (SBC) for COD and TSS parameters.
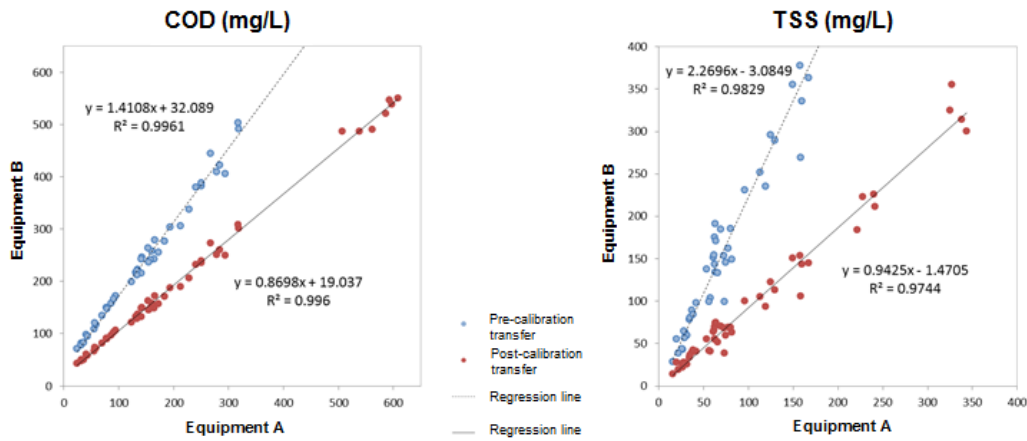
Fig. 5 - Transfer calibration by single wavelength standardization (SWS) for COD and TSS parameters.
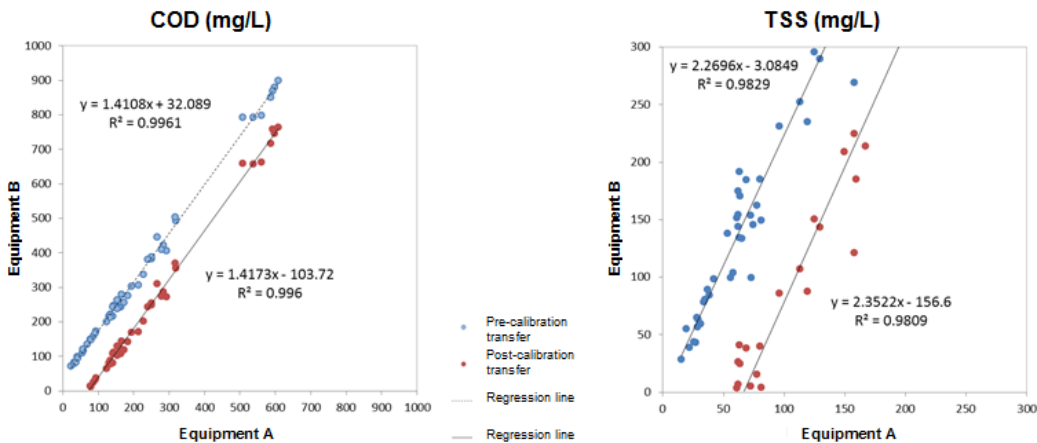

Fig. 6 - Transfer calibration by mean sample residual spectrum correction (MSR) for COD and TSS parameters.

On table 3, are shown the RMSE values obtained for the COD and TSS estimation parameters, using the spectral information from equipment B, before and after the application of the different transfer calibration methods on the IPLS model, developed with spectra from equipment A.

The RMSE errors were calculated in steps: cross-validation, external forecasting and external online prediction. Through the analysis Table 14 one can see that the RMSErel errors are considerably higher than when it is not applied any transfer-calibration procedure (above 50% and 165% for the estimated COD and TSS parameters, respectively). Any of the transfer-calibration procedures allows a substantial reduction in errors. In the case of COD parameter, it was possible to reduce errors up to 8-12% by SBC method. For the SST parameter the reduction could go up to 10-32% by the SWS method and up to 18-32% by SBC method. As in IPLS model, the prediction errors for the SST parameter were higher than for the COD parameter.

For the COD parameter, the best results were obtained by the SBC method. Errors (RMSErel) associated with the data set used in the cross-validation and the external forecast are very similar to those obtained in IPLS model (Table 12), not adding

significant errors in the estimates. The second best method was the SWS; however it significantly increased the forecast errors. In the case of TSS parameter, the SBC and SWS methods provide the best results. Errors (RMSErel) associated with the data set used in the cross-validation and the external forecast are very similar to those obtained in IPLS model (Table 12) not adding significant errors in the estimates.

From the results of direct standardization application of the spectral response by standard spectra convergence (MSR), was found that this method is not appropriate for the case study. The correction factor calculated using equation 3.6, allows negative absorbance results, which when correcting the spectral matrix from equipment B, leads to the estimation of negative values for COD and TSS.

The use of SWS and MSR methods allows an adjustment of the spectra obtained in equipment B, in order to resemble the spectra acquired by equipment A. To do this, it was used corrective factors, calculated for all spectrum of wavelengths (200-740 nm) which ensured the independence from IPLS model. If only was used the wavelength for regions identified in IPLS models, RMSE errors could possibly be lower (in relation to the SBC method). However, this procedure

could compromise the quality of the estimations in future updates of IPLS models, which can lead to new ranges of optimized wavelengths.

The online external forecast exhibited minimum error values in the ranges 12-14% for COD and 10-18% for TSS. The SBC method provided, again, the best results for the available data.

Table 3 - Calibration transfer results

Parameter	Calibration transfer technique	Cross-validation n = 50		External validation n = 26		On-line External validation n =11 (COD) n = 10 (TSS)	
		$RMSE$	$RMSE_{rel}$	$RMSE$	$RMSE_{rel}$	$RMSE$	$RMSE_{rel}$
		(mg/L)	(%)	(mg/L)	(%)	(mg/L)	(%)
COD	without	141.0	67.7	135.0	64.3	265.0	54.3
	SBC	17.5	8.4	20.0	9.5	58.7	12.0
	SWS	28.8	13.9	37.9	17.9	68.4	14.0
	MSR	75.6	36.4	69.8	33.3	135.0	27.6
TSS	without	178.0	170.0	155.0	188.0	368.0	169.0
	SBC	25.9	24.8	26.1	31.6	39.1	18.0
	SWS	26.1	25.0	25.9	31.3	22.9	10.5
	MSR	121.0	116.0	111.0	134.0	215.0	98.6

Legend:

n: number of spectrums included in the matrix

VII. CONCLUSION

Adequate monitoring of the flow characteristics in the collectors and pollutant loads discharged into a receiving environment is important for the proper management of urban drainage systems and to ensure the proper physical, chemical and ecological status of water bodies. The spectrum of radiation absorbance, in the ultraviolet-visible range, has been recently used in urban drainage, to estimate quality parameters such as COD and TSS.

The application of PCA analyzes holds the potential of monitoring with UV-Vis spectra, since it allows the identification of anomalies or changes in effluent quality matrix, such as inputs of flows with different characteristics. The comparative analysis made to spectra obtained from the bench spectrophotometer and submersible spectrophotometer has shown differences between the two sets and the need to conduct an adequate calibration transfer, in order to include the spectra in the same estimates model. It was also found that when a wastewater sample is not analyzed immediately after collection, follows an aging effect, where the physicochemical properties of the sample are changed. The identification of changes in the water matrix through the PCA analysis is an essential step prior to the implementation of estimates models.

The application of PLS and IPLS models to estimate COD and TSS in wastewater collectors, based on spectrophotometry, showed satisfactory results. It was found that the implementation of IPLS models for both parameters yielded the best results, since this type of model selects the spectral ranges that best relate to the analytical parameter. It was confirmed that it was not suitable to include samples of different ages in the same spectral model, due to significant estimation errors.

In situations where it is necessary to use spectra acquired by different equipment, or in the same equipment by changes in the spectra acquisition conditions, the calibration transfer techniques should be implemented. In this work, estimation models were developed for the parameters COD and TSS, based on the spectral information from a bench spectrophotometer. Were applied and evaluated, to the case study, three calibration transfer techniques between this equipment and a diode-array submersible spectrophotometer to purchase online spectra, namely SBC, SWS and MSR procedures.

Through an analysis of the values of quality parameters estimated by the spectra of both equipments we found that it would not be appropriate to include the spectra from the equipment B in the estimation models developed for spectra from equipment A, because in the absence of calibration transfer techniques, the regression lines showed a and b values unsatisfactory.

For the case study, the calibration transfer techniques that offered better results was the SBC, that directly standardized quality parameters in estimate. Although the method SWS has allowed $RMSE_{rel}$ values very similar, the parameters for SBC regression line were significantly better. However, given the SBC method being parameterized specifically for the available samples, the quality of the estimates may be affected if the model is changed, since the corrective factor depends on the model. However, this method was the most suitable of the three methods analyzed and submitted the same error that obtained in IPLS models for COD and TSS. Given the robustness limitations of the method SBC upon changes in IPLS model, the SWS may be an alternative method, since it is independent of IPLS model. This allows the model improve with new data from others campaigns. In this case study, it is not advisable to use the MSR method between the two spectrophotometers.

In the online external validation, the SBC and SWS methods offered good results. The reduced error associated to the prediction of this data set is a good indication to the possibility of using these calibration transfer techniques in spectra acquired in situ.

In summary, spectrophotometry through the acquisition of spectra in situ or in the laboratory, combined with PLS and IPLS models to estimate TSS and COD can contribute to the proper management of urban drainage systems. Through a simple calibration transfer technique (as SBC or SWS) between a bench spectrophotometer and a submersible

spectrophotometer, in situ, it is possible to keep improving a estimates model for quality parameters such as COD and TSS.

In the future can be interesting to explore in more detail the use of spectra acquired sometime after collection of the samples under the same estimates model. If it is necessary to obtain a spectrum of a sample in a laboratory, but it is not possible to make their immediate determination can be important to develop a procedure for modeling the spectral changes associated with aging samples.

It may also be interesting to explore the potential of the relationship between the spectrum in the UV-Vis range with others quality parameters and associate them with calibration transfer procedures used in this work.

REFERENCES

[1] Directive 2000/60/EC of the European Parliament and the Council of 23 October 2000 establishing a framework for Community action in the field of water policy.

[2] Viegas C., Sousa A., Póvoa P., Martins J., Rosa J. (2011) – Treated wastewater use in Portugal: challenges and opportunities. In *WATERuse Barcelona 2011*. 26-28 Sept. p. 39-43.

[3] Rosa J., Mesquita E., Silva C. (2014) – *Qualidade e tratamento de água e águas residuais. Desafios e Oportunidades*. Água & Resíduos. Vol. 84, p. 12-15.

[4] Aguado D., Rosen C. (2008) – Multivariate statistical monitoring of continuous wastewater treatment plants. Journal of Engineering Applications of Artificial Intelligence. Vol. 21, p. 1080-1091.

[5] Godley A. (2002) – Flow measurement in partially filled closed conduits. Flow Measurement and Instrumentation – Elsewier. Vol. 13. p. 197-201.

[6] Henze M., Harremoes P., La Cour Jansen J., Arvin E. (2002) – Wastewater treatment. Biological and chemical processes. Springer. p. 132.

[7] Pouet M. F., Baurès E., Vaillant S., Thomas O. (2004) – Hidden isosbestic point(s) in UV spectra. Applied Spectroscopy. Vol. 58, p. 486-490.

[8] Brito R. S. (2012) – Monitorização de Variáveis Hidráulicas e de Qualidade da Água em Drenagem Urbana. PhD thesis. Lisboa. Instituto Superior Técnico.

[9] Fan W., Liang Y., Yuan D., Wang J. (2008) – Calibration model transfer for near-infrared spectra based on canonical correlation analysis. Analytica Chimica Acta. Vol. 623, p. 22-29.

[10] Feudale R., Woody N., Huwei T., Myles A., Brown S., Ferré J. (2002) – Transfer of univariate calibration models. A review. Chemometrics and Intelligent Laboratory. Vol. 64, p. 181-192.

[11] Brito, R. S., Pinheiro, A., Pinheiro, H., Ferreira, F., Matos, J. S., Lourenço, N. D. (2015) – Calibration transfer between a bench scanning and a submersible diode array spectrophotometer for in situ wastewater quality monitoring in sewer systems. Submitted for publication.