

FUEL PARK WASTEWATER MONITORING WITH UV-VIS SPECTRA AND PARTIAL LEAST SQUARES MODELS

Nídia D. Lourenço¹, José C. Menezes¹, Helena M. Pinheiro¹, David Diniz²

¹IBB – Institute for Biotechnology and Bioengineering, Centre for Biological and Chemical Engineering, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais 1049-001 Lisboa, Portugal

²CLC, Companhia Logística de Combustíveis, S.A., E.N. 366 km 18, 2050 Aveiras de Cima, Portugal
helena.pinheiro@ist.utl.pt

In the context of the high application potentials for on-line and in-line measurements in wastewater quality monitoring, UV-Vis spectrophotometry has received recent attention. In the present work the development of Partial Least Squares (PLS) calibration models for the fast determination of parameters legally required at the discharge of a fuel park wastewater treatment plant was attempted using UV-Vis spectra and the corresponding standard analytical values. A PLS model was developed and validated for Chemical Oxygen Demand (COD) determination with 3.8 mg O₂/l RMSEP (root mean squared error of prediction). The models developed for 5-day Biochemical Oxygen Demand (BOD₅) and Total Suspended Solids (TSS) were validated (3.5 mg O₂/l and 2.9 mg/l RMSEP, respectively), but presented some predictive limitations mainly attributed to the narrow interval of concentration values present in the original dataset.

Key words: UV-Vis spectrophotometry; wastewater monitoring; Partial Least Squares (PLS); fuel park wastewater

МОНИТОРИНГ НА ОТПАДНИТЕ ВОДИ ОД СКЛАДИШТЕ ЗА ГОРИВА СО UV-VIS СПЕКТРИ И ПАРЦИЈАЛНИ РЕГРЕСИОНИ МОДЕЛИ

UV-Vis спектрометријата како метод со висок потенцијал на применливост во on-line и in-line мониторинг на квалитетот на отпадните води во последните години добива значително внимание. Во овој труд е направен обид со користење на UV-Vis спектрите и соодветните стандардни аналитички вредности да се развијат калибрациони модели за брзо определување на потребните законски параметри при испуштање отпадна вода од складиштата за горива во постројка за третман на отпадна вода. Моделот PLS беше развиен и валидиран за хемиска потрошувачка на кислород (COD) определен со 3,8 mg O₂/l RMSEP (корен на средната квадратна грешка на предвидување). Развиените модели за 5-дневна биолошка потрошувачка на кислород (BOD₅) и вкупните суспендирани цврсти материји (TSS) беа потврдени (3,5 mg O₂/l и 2,9 mg/l RMSEP, соодветно), но покажаа некои предвидливи ограничувања главно поради тесниот интервал на концентрациски вредности присутни во оригиналниот сет на податоци.

Клучни зборови: UV-Vis спектрофотометрија; мониторинг на отпадни води; парцијални регресиони модели; отпадни води од складишта за горива

INTRODUCTION

Regulatory authorities in Europe are increasingly becoming more demanding towards wastewater treatment activities, and the compliance assessment of the legal discharge requirements has been a powerful driving force for wastewater qual-

ity monitoring [1]. On the other hand, the available wastewater quality monitoring techniques based on the characterization of aggregate parameters (chemical oxygen demand, COD, and biochemical oxygen demand, BOD) still possess several drawbacks when used for the control of wastewater treatment systems [2]. In addition to the sampling

and sample conditioning problems, the standard analytical methods currently used do not allow a real-time knowledge of the phenomena influencing wastewater treatment plant operation. In fact, the reproducibility of some of these analytical techniques is uncertain, and most of the analyses are time-consuming, use high-cost and/or toxic reagents and produce wastes that require further treatment [3].

Given the high development and application potentials for on-line (automated sample extraction and delivery to the analyzer) and in-line (in situ analysis using a probe inside the process) measurements in wastewater quality monitoring [4], UV-Visible (UV-Vis) spectrophotometry can be very useful in this area, contributing to the correct operation of the treatment systems. UV-Vis spectrophotometry is fast and simple method that has been used for wastewater quality evaluation and organic matrix composition identification [5, 6], since most organic compounds and a few soluble minerals (such as nitrates) absorb in the UV-Vis region. The reported quantitative environmental applications of UV-Vis spectrophotometry include the estimation of organic matter and nitrate in wastewaters [7], and determination of polycyclic aromatic hydrocarbons (PAH) in soils [8]. In the present work UV-Vis spectra of samples collected in a fuel park Wastewater Treatment Plant (WWTP), after biological treatment, were acquired and used for the attempted development of Partial Least Squares (PLS) calibration models for four environmental monitoring parameters, namely, Chemical Oxygen Demand (COD), 5-day Biochemical Oxygen Demand (BOD₅), Total Suspended Solids (TSS) and Oil & Grease (OG).

EXPERIMENTAL

WWTP description and sampling strategy

Companhia Logística de Combustíveis (CLC) is a logistic fuel company responsible for the operation of the multi-product pipeline between the Sines oil refinery and the Aveiras fuel park, Portugal. CLC's activities involve the transport, storage and expedition of gaseous and liquid fuels, their clients being the company shareholders. The company received a safety, quality and environmental certification in 2004, and is responsible for the distribution of around 50% of all the fuels con-

sumed in Portugal. Wastewaters generated at the Aveiras fuel park include domestic effluents, liquid effluents with high hydrocarbon content due to occasional spills and aqueous effluents with detergents and lubricants, mainly from the gas cylinder washing zone. The latter two effluents constitute the industrial sewage and are treated together with the domestic sewage in the CLC's WWTP. The plant can be divided in three main areas, *i.e.*, the wastewater collection sumps, the tilted plate interceptor and the biological reactor. This reactor is an activated sludge Sequencing Batch Reactor (SBR) running in 24-hour cycles consisting of four main steps: sewage feed, aeration, biomass settling and treated effluent discharge. The samples used in the present work were collected weekly at the discharge line of the WWTP biological step and analyzed at the Instituto Superior Técnico Laboratory of Analyses (LAIST, certified laboratory).

Spectra acquisition

UV-Vis spectra of the collected samples were acquired in a UV-Vis Specord 200 spectrophotometer (Analytik Jena, Germany; Aspect Plus software, Zeiss, Germany) between 190 and 500 nm in a quartz cell of 10 mm path length, raw and after appropriate dilution in order to overcome signal saturation in the UV region.

PLS model development

The PLS regression method is the most widely used method for multivariate calibration and corresponds to a guided decomposition model where the dependent variables, Y, intervene directly in the decomposition of the independent variables, X. The purpose of this method is to determine a small number of latent factors that can predict Y (analytical data) using the data in X (spectra) as efficiently as possible. The UV-Vis spectral data obtained from CLC WWTP samples and the correspondent analytical reference data were exported to Grams[®]32/AI software extended with the PLSplus/IQ module (Thermo Galactic, USA) for PLS calibration development.

One essential requirement in establishing a PLS model is choosing the appropriate wavelengths since part of the information gathered in the full spectrum is redundant, and the signals measured at some wavelengths may be non-linear, represent noise, or contain useless information for

the intended purpose [9]. The UV-Vis region used in the developed PLS models was chosen by evaluation of the coefficient of determination (R^2) for each wavelength of the data set and selection of the region with higher values for this coefficient.

Due to the limited number of samples available, a full cross validation (leave-one-out) procedure was adopted to evaluate the predictive ability of the PLS models for the training set. Cross-validation was performed by running several sub-model validations where one object (sample) was left out of the data set in each run. The average results of each run were calculated and the quality of the results was evaluated by the root mean squared error of cross-validation, RMSECV, according to

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{\setminus i})^2}{n-1}} \quad (1)$$

where n is the number of samples in the training set, y_i the measured value of sample i and $\hat{y}_{\setminus i}$ the predicted value of sample i when the model is constructed without sample i . The optimal number of factors to retain in the PLS models was established by the described cross-validation procedure applied to the training set.

The validation of the final PLS models was attempted with the test set and the difference between the predicted and the measured values for the wastewater parameters was expressed by the root mean squared error of prediction, RMSEP, calculated according to

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

where n is the number of samples in the test set, y_i the measured value of sample i and \hat{y}_i the predicted value of sample i . The RMSEP was evaluated as a function of the number of latent factors kept in the PLS model.

RESULTS AND DISCUSSION

The development of PLS models was attempted using UV-Vis spectra of CLC treated wastewater samples and the corresponding analytical data obtained by reference standard methods (LAIST) for four parameters legally required

at discharge, *i.e.*, COD, BOD₅, TSS and OG. However, it was not possible to establish a PLS calibration model for OG. This could be attributed to the sampling difficulties typical of two-phase systems and to the narrow interval of values measured in the CLC treated wastewater samples (0.4 – 1.4 mg/l), which, furthermore, are close to the standard analytical technique detection limit (0.2 mg/l). PLS calibration models were developed for COD and BOD₅ using spectra of diluted samples and for TSS using spectra of raw samples. The 2nd derivative was applied (according to the Savitzky-Golay (SG) convolution derivation) for COD and BOD₅ calibrations since this procedure decreases spectral noise and reduces light scattering effects due to sample particles, which are mathematically filtered away [10]. All spectra were mean-centered prior to PLS calibration model development.

A PLS model using 5 latent factors with 0.92 R^2 was achieved for COD (Fig. 1a). The obtained RMSECV was 3.3 mg O₂/l and it should be noted that for COD measurements in the range of 5 to 50 mg O₂/l the standard analytical method mean squared error (MSE) is 13 mg O₂/l (APHA, 1995). For BOD₅ a PLS model using 3 latent factors with 0.88 R^2 and 3.1 mg O₂/l RMSECV was developed (Fig. 1b). The standard analytical method MSE for this parameter is around 11 mg O₂/l (detection limit of 2 mg O₂/l) (APHA, 1995). The PLS model achieved for TSS used 6 latent factors and had 0.84 R^2 and 1.2 mg O₂/l RMSECV (Fig. 1c). The detailed results of the obtained PLS models are summarized in Table 1.

The validation of the proposed PLS calibration models was attempted with a set of external samples (test set) and the results for COD, BOD₅ and TSS are presented in Fig. 2. The RMSEP was 3.8 mg O₂/l for COD, 3.5 mg O₂/l for BOD₅ and 2.9 mg/l TSS.

Table 1

Summary of the PLS calibration models developed for three parameters legally required at CLC WWTP discharge using UV-Vis spectra of CLC samples

Parameter	COD	BOD ₅	TSS
UV-Vis region	274–309	274–295	224–500
Factors	5	3	6
R^2	0.92	0.88	0.84
RMSECV	3.3 mg O ₂ /l	3.1 mg O ₂ /l	1.2 mg/l

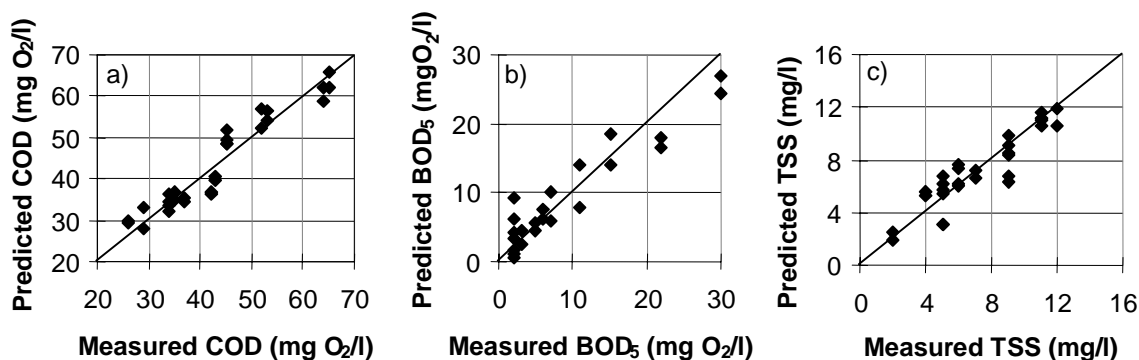


Fig. 1. Measured against predicted values for the PLS calibration models developed for COD (a), BOD₅ (b) and TSS (c) using UV-Vis spectra of samples from the CLC WWTP

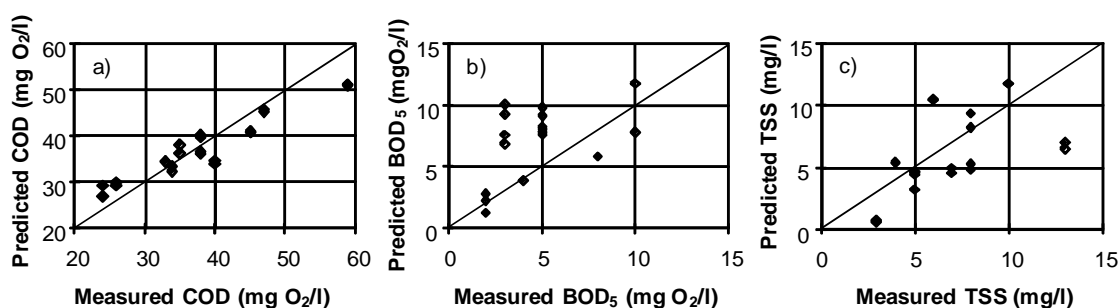


Fig. 2. Validation of the PLS calibration model for COD (a), BOD₅ (b) and TSS (c) using UV-Vis spectra of external samples from CLC WWTP. The samples were diluted for COD and BOD₅ validation and were raw for TSS validation

The results of the COD PLS model validation presented in Fig. 2a are highly satisfactory. A PLS calibration model for the parameter COD based on UV-Vis spectral analyses of WWTP samples with 9 factors, 0.91 R^2 and 23.5 mg O₂/l RMSEP was already reported [7]. However, the COD range covered by that PLS model was 50–500 mg O₂/l and is not comparable to the low range of values encountered in CLC WWTP samples (26–65 mg O₂/l).

Concerning COD multivariate calibration models based on UV-Vis spectra, it should be mentioned that significant deviations may occur if the matrix composition changes. In fact, whereas COD analysis determines almost 100% of organic carbon [11], the spectrophotometric method includes only the part of the organic carbon matrix that absorbs light in the UV-Vis region and there might be some important fractions of organic carbon present that do not show absorption in the UV-Vis spectra, such as short chain fatty acids, sugars and starch.

Although BOD₅ is a commonly used parameter for wastewater quality characterization, several

problems are inherent to its standard analytical method, since most of the variables present are difficult to control during the analysis, including seed quality, presence and concentration of natural biological inhibitors, the time taken to complete the analysis, and test method accuracy. The results presented in Fig. 2b for the BOD₅ PLS model validation showed higher predicted values than the standard analytical method but with a low RMSEP (3.5 mg O₂/l) value when compared to the standard analysis MSE (11 mg O₂/l). However, it should be noted that this method was developed and validated with a limited number of samples of a narrow interval of BOD₅ values very close to the reference method detection limit. For that reason the presented results can only suggest that the development of PLS calibration models for BOD₅ based on the UV-Vis spectra of wastewater samples is a promising technique for real-time estimation of BOD₅. This was already attempted by a near infrared (NIR) transmission PLS calibration model with 28.6 mg O₂/l RMSECV and 28.7 mg O₂/l RMSEP for BOD₅ values ranging from 35 to 275 mg O₂/l [12].

Similarly to the BOD₅ calibration model, the calibration model for TSS was developed with a limited number of samples of a narrow interval of values. The validation results presented in Fig. 2c show that the predicted TSS values can be higher or lower than the concentration values determined by the standard method, suggesting that UV-Vis spectra may contain more information related to the suspended organic matter than the simple standard TSS analysis. In fact, different wastewater samples with the same concentration of organic carbon could be distinguished by UV-Vis analysis, as well as different suspensions with the same TSS concentration [6]. Furthermore, a proposal of a typology of organic matter fractions taking into account the organic matter origin, its stability and the part corresponding to inorganic particles was established based on the UV response of different size range fractions of organic matter from different origins [6].

CONCLUSIONS

The preliminary results of the present work showed that UV-Vis spectra of aqueous wastewater samples contain information that can be extracted and used for WWTP monitoring. In the particular case of the treated wastewater samples from CLC, the use of spectrophotometric techniques in the UV-Vis region and their chemometric treatment (through PLS models) revealed a high application potential as a fast, simple and cost-effective (no sample preparation, no reagents consumption) method for COD determination, and also some potential for the parameters BOD₅ and TSS, in spite of the models being developed in a narrow interval of values close to the standard analytical methods' detection limits. The use of PLS UV-Vis calibration models for the fast determination of COD or BOD₅ can constitute a highly advantageous alternative to the analyzers installed at CLC (total organic carbon and OG) for the on-line or in-line control of the discharged treated wastewater quality.

Acknowledgement. N.D. Lourenço acknowledges the financial support from Fundação para a Ciência e a Tecnologia (FCT, Portugal) through a post-doctoral research grant (SFRH / BPD / 31497 / 2006, Portugal).

REFERENCES

- [1] B. N. Jacobsen, Compliance assessment for the 91/271/EEC directive and end-user's view on potentials for further on-line monitoring and control at wastewater treatment plants. *Talanta* 50 (4), 717–723 (1999).
- [2] M.-F. Pouet, O. Thomas, B. N. Jacobsen, A. Lynggaard-Jensen, P. Quevauviller, Conclusions of the workshop on methodologies for wastewater quality monitoring. *Talanta* 50 (4), 759–762 (1999).
- [3] A. Bonastre, R. Ors, J. V. Capella, M. J. Fabra, M. Peris, In-line chemical analysis of wastewater: present and future trends. *Trends Anal. Chem.* 24 (2), 128–137 (2005).
- [4] R. E. Clement, P. W. Yang, Environmental analysis. *Anal. Chem.* 73, 2761–2790 (2001).
- [5] O. Thomas, H. El Khorassani, E. Touraud, H. Bitar, TOC versus UV spectrophotometry for wastewater quality monitoring. *Talanta* 50, 743–749 (1999).
- [6] S. Vaillant, M. F. Pouet, O. Thomas, Basic handling of UV spectra for urban water quality monitoring. *Urban Wat.* 4, 273–281 (2002).
- [7] G. Langergraber, N. Fleischmann, F. Hofstädter, A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. *Water Sci. Technol.* 47 (2), 63–71 (2003).
- [8] E. Touraud, M. Crône, O. Thomas, Rapid diagnosis of polycyclic aromatic hydrocarbons (PAH) in contaminated soils with the use of ultraviolet detection. *Field Anal. Chem. Technol.* 2, 221–229 (1998).
- [9] M. L. Luis, J. M. G. Fraga, A. I. Jiménez, F. Jiménez, O. Hernández, J. J. Arias, Application of PLS regression to fluorimetric data for the determination of furosemide and triamterene in pharmaceutical preparations and triamterene in urine. *Talanta* 62, 307–316 (2004).
- [10] M. Karlsson, B. Karlberg, R. J. O. Olsson, Determination of nitrate in municipal waste water by UV spectroscopy. *Anal. Chim. Acta* 312, 107–113 (1995).
- [11] APHA, American Public Health Association, *Standard Methods for the Examination of Water and Wastewater*, 19th ed., A. D. Eaton, L. S. Clesceri, A. E. Greenberg (Eds.), Washington D.C., 1995.
- [12] A. B. Stephens, P. N. Walker, Near infrared spectroscopy as a tool for real-time determination of BOD₅. Paper N°. 006089, ASAE Meeting Presentation, Miami, USA, 1–15 (2000).