ORIGINAL PAPER

# Quantitative monitoring of an activated sludge reactor using on-line UV-visible and near-infrared spectroscopy

**Mafalda C. Sarraguça · Ana Paulo ·**
**Madalena M. Alves · Ana M. A. Dias · João A. Lopes ·**
**Eugénio C. Ferreira**

**Abstract** The performance of an activated sludge reactor can be significantly enhanced through use of continuous and real-time process-state monitoring, which avoids the need to sample for off-line analysis and to use chemicals. Despite the complexity associated with wastewater treatment systems, spectroscopic methods coupled with chemometric tools have been shown to be powerful tools for bioprocess monitoring and control. Once implemented and optimized, these methods are fast, nondestructive, user friendly, and most importantly, they can be implemented in situ, permitting rapid inference of the process state at any moment. In this work, UV-visible and NIR spectroscopy were used to monitor an activated sludge reactor using in situ immersion probes connected to the respective analyzers by optical fibers. During the monitoring period, disturbances to the biological system were induced to test the ability of each spectroscopic method to detect the changes in the system. Calibration models based on partial least squares (PLS) regression were developed for three key process parameters, namely chemical oxygen demand (COD), nitrate concentration (N-NO$_3^-$), and total suspended solids (TSS). For NIR, the best results were achieved for TSS, with a relative error of 14.1% and a correlation coefficient of 0.91.

The UV-visible technique gave similar results for the three parameters: an error of ~25% and correlation coefficients of ~0.82 for COD and TSS and 0.87 for N-NO$_3^-$. The results obtained demonstrate that both techniques are suitable for consideration as alternative methods for monitoring and controlling wastewater treatment processes, presenting clear advantages when compared with the reference methods for wastewater treatment process qualification.

**Keywords** Activated sludge process · UV-visible · Near-infrared · Chemometrics · On-line monitoring · Wastewater treatment

M. C. Sarraguça · J. A. Lopes
REQUIMTE, Serviço de Química-Física, Faculdade de Farmácia,
Universidade do Porto,
Rua Aníbal Cunha, 164,
4099-030 Porto, Portugal

A. Paulo · M. M. Alves · A. M. A. Dias · E. C. Ferreira (✉)
IBB—Institute for Biotechnology and Bioengineering,
Centre of Biological Engineering, Universidade do Minho,
Campus de Gualtar,
4710-057 Braga, Portugal
e-mail: ecferreira@deb.uminho.pt

## Introduction

The use of microorganisms to treat effluents by removing contaminants from wastewater is a common and environmentally friendly practice, since almost all wastewaters contain biodegradable constituents that can be removed biologically [1]. Activated sludge systems are among the most widely used secondary biological treatment processes. These systems involve inoculation with floc-forming bacteria, which oxidize the organic matter, stabilizing the wastewater under aerobic conditions. A settling tank is used to separate the biomass consituents according to their settling abilities.

Recent trends in environmental protection include a move towards increasingly stringent demands on water and wastewater treatment efficiency. The European directive 91-271 EEC contains a new set of recommendations for efficient wastewater treatment [2]. The main focus of this directive is the development of monitoring tools that will allow greater knowledge of water treatment processes, with the ultimate aim being to improve their efficiencies [3, 4].

Additionally, the sensitivities of these biological systems to sudden changes in feedstock composition or nutrient removal means that such tools should be able to monitor the characteristics of such systems in real time. This enhances the operator's ability to react rapidly to events and thus avoid pollutant discharges or biological system damage [2, 5].

Traditionally, such systems are monitored by manual sampling with the off-line analysis of a number of key process parameters, such as biological oxygen demand (BOD), chemical oxygen demand (COD), total organic carbon (TOC), total suspended solids (TSS), and nitrate (N-NO$_3^-$). Such analyses are time consuming (e.g., the standard BOD$_5$ test takes five days to generate a result, and the COD test requires 2–4 h), expensive (e.g., a series of techniques are needed) and hazardous to the environment (toxic chemicals such as solvents and reactants are required). Moreover, this approach provides only a snapshot of the system, making it unsuitable for on-line monitoring, where rapid feedback is necessary [6]. Therefore, simpler, faster, and less expensive analytical tools are needed for on-line monitoring and control.

Spectroscopic techniques such as UV-visible, near-infrared (NIR), and fluorescence possess all of the characteristics needed for effective on-line monitoring, given that they are noninvasive, nondestructive, versatile, and flexible measuring systems [7]. The main objective of such techniques is to establish robust and predictive relations between a given monitored parameter and one or more variables, which in this case are the measured absorbances at one or more wavelengths. Furthermore, the development of miniaturized systems and progress in telecommunications technology that has made inexpensive high-quality polymer- or silica-based optical fibers available have made in situ measurements a feasible proposition [7, 8]. UV-visible spectroscopy has been used for a long time as an alternative method of monitoring and controlling wastewater systems. Initially, selected wavelengths that correlated with the process parameters were used. The absorbance at 254 nm correlated with COD [9] and TOC [10] parameters in municipal wastewaters. Due to the sensitivity of such measurements to turbidity, a second wavelength was also used to correct for scattering effects [8]. COD and TOC were thus determined using a combination of the absorbance at 254 nm and that at 350 nm (for turbidity correction) [11]. More recently, the absorbance at 254 nm was used in synchronous fluorescence spectroscopy to estimate the dissolved chemical oxygen demand (DCOD), COD, ammonia and the turbidity in municipal wastewaters for fingerprinting purposes [8]. The main drawback associated with the use of a single wavelength, even corrected for turbidity, is that frequent calibration is required to guarantee good results [12].

Moreover, this univariate approach is based on the fact that the pollution present in the effluent has a defined peak of maximum absorbance that always occurs at the same wavelength. However, this value can vary significantly depending on the matrix composition [6]. However, fast computational processing tools have become available in the last few decades in association with the development of more robust and precise spectrophotometers, allowing a shift towards a multiple-wavelength approach [6, 13, 14]. The multi-wavelength approach can give better results than single wavelength procedures, especially when monitoring effluents that are characterized by a constant variation in composition [13, 14]. However, when this approach is used, it is necessary to reduce the number of wavelengths that are used in the correlation in order to reduce the amount of superfluous information, and thus rapidly and clearly identify relevant patterns among the data and identify the status of the system at any time. Different mathematical procedures have been reported in the literature to be efficient methods for reducing data in the multivariate approach. Deconvolution methods, for instance, have been used to determine DCOD COD, TOC, BOD, TSS, and nitrate [12, 15, 16]. Escalas et al. [17] used a modified UV deconvolution method to estimate the dissolved organic carbon (DOC) in wastewaters. Chemometric methods such as artificial neural networks (ANN) and partial least squares (PLS) have also been used to determine COD, nitrate, and TSS [6, 18–21]. Despite the advantages mentioned above, some disadvantages of the method have also been reported, such as its sensitivity to turbidity [8], its inability to detect saturated bonds [7, 22], and fouling of the probe tip [3] when performing on-line measurements.

Similar to the UV-visible technique, NIR is increasingly being applied to qualitatively and quantitatively monitor processes and to process diagnostics [23]. The food and pharmaceutical industries were the first to use NIR in a systematic approach [24]. The application of NIR to wastewater treatment processes is less common, and only a few works have been reported in this field. Stephens et al. [25] tested a NIR-visible method of evaluating BOD$_5$ and Sousa et al. [26] used NIR to determine COD, both in wastewater processes. In situ NIR spectroscopy was also used to monitor a lab-scale activated sludge system, and promising results for the application of NIR to biological processes were obtained [5]. More recently, a NIR trans-flectance probe was used to monitor an on-line sequential batch reactor for the aerobic treatment of dairy residues, with PLS regression used to calibrate parameters such as total solids (TS), TSS, and COD [27].

Some of the advantages of using NIR are shared with UV-visible spectroscopy, while others are intrinsic to NIR, such as the ability to use it in highly scattering and strongly absorbing media [5], as well as the possibility of

simultaneous determining chemical and physical properties [28, 29]. One of the main disadvantages of using NIR spectroscopy in aqueous systems is the high absorbance of water, which leads to a broad peak that can limit the detection of other compounds that are present in smaller quantities [30].

In this work, UV-visible and NIR spectrometers connected to immersion probes were used to estimate key parameters that characterize an activated sludge reactor: COD, TSS, and nitrate ($N-NO_3^-$). Partial least squares (PLS) was used as the regression model, with the root mean square error of cross-validation (RMSECV) used as an indicator of the model's accuracy.

## Materials and methods

### Activated sludge system

A complete-mixed lab-scale activated sludge reactor was inoculated with both heterotrophic microorganisms and nitrifying bacteria recovered from a municipal wastewater treatment plant. The system consisted of a tank with a total volume of 25 L and 17 L of suspended biomass followed by a 2.5 L cylindrical settler. It was fed with a synthetic influent based on peptone and meat extract as carbon sources, as prepared according to Marquéz et al. [31]. The pH of the system was controlled with a pH meter and a control pump (Model BL 7916–BL 7917, Hanna Instruments, Woonsocket, RI, USA). Complete mixing inside the reactor was guaranteed by supplying a continuous inflow of air bubbles through an air diffuser placed at the bottom of the reactor. An oxygen probe (TriOmatic 690, WTW, Weilheim, Germany) was used to measure the amount of dissolved oxygen. The concentration of dissolved oxygen was maintained at above 7 mg $O_2$ $L^{-1}$. Sludge recirculation from the settler to the reactor was guaranteed by an air pump.

### In situ process monitoring

In situ monitoring of the lab-scale reactor was achieved using two immersion probes that measured in the UV-visible and NIR spectroscopic ranges. The UV-visible spectra were acquired using a portable dispersive UV-visible spectrometer (model USB4000, Ocean Optics, Dunedin, FL, USA) with a 3648-element linear CCD array detector that provided measurements in the wavelength range between 230 and 700 nm and a spectral resolution of ~0.3 full width at half-maximum (FWHM). A DH-2000 (Ocean Optics) deuterium tungsten halogen bulb was used as the light source. This light source combines the continuous spectra of the deuterium and tungsten halogen light

sources, producing powerful and stable output from 215 to 2000 nm. The sampling accessory was an immersion probe with an optical path length of 1 cm, which was connected to the light source and the spectrometer by two optical fibers (model TP300-UV-Visible, Ocean Optics).

The NIR spectra were acquired using a portable dispersive spectrometer (model NIR-512, Ocean Optics). It featured a temperature-regulated 512-element indium–gallium–arsenide (InGaAs) array detector that was effective in the 900–1700 nm wavelength range and had a spectral resolution of 3.0 FWHM. The detector temperature was kept constant at -4.0 °C. Spectra were acquired using a transflectance probe (model T300RT, Ocean Optics) with an optical path length of 1 cm. The probe was connected through optical fibers (QP400-2-VISNIR, Ocean Optics) to a tungsten halogen lamp filled with krypton gas and possessing a spectral range of between 350 nm and 2200 nm (SL1 light source, Stellarnet, Tampa, FL, USA) as well as to the NIR spectrometer. The operating mode was the same in both techniques: ten scans were made and then averaged, the boxcar width was 5, and dark correction was used when acquiring spectra. The integration time was adjusted until the peaks at 550–600 nm for UV-visible and 1100–1200 nm for NIR were close to 60,000 intensity units. OIBase32 software (Ocean Optics) was used for spectrometer configuration, control, and data acquisition.

For in situ monitoring, the probes were immersed in the settler at the same time, acquiring spectra simultaneously. Spectra were acquired every monitoring day (2–3 times a week) for no more than 45 min to avoid fouling the optical path. Probe tips and sample windows were rinsed with distilled water and cleaned with smooth paper. The NIR probe sample window was dried before spectral acquisition was initiated. The UV-visible probe was immersed in tap water and the NIR probe was positioned in contact with the air in a stable position when acquiring the reference spectra.

### Off-line process monitoring

The system was regularly monitored for parameters as COD, TSS, and nitrate concentration. The samples were collected from the settler COD and nitrates were quantified after centrifugation and filtration. COD determination was based on a colorimetric method, in closed reflux, according to the method 5220D from the Standard Methods for the Examination of Water and Wastewater [32]. The samples from the activated sludge process were analyzed immediately after being collected. The TSS was determined according to the method 2540D from the Standard Methods for the Examination of Water and Wastewater [32]. The nitrate ($N-NO_3^-$) was determined by HPLC (JASCO, Tokyo,Japan) with automatic injection, using a UV detector (210 nm). The column used was a Varian (Palo Alto, CA,

USA) Metacarb 87H operating at a temperature of 60 °C. The eluent was a solution of sulfuric acid (0.005 mol L$^{-1}$) with a flow rate of 0.70 mL min$^{-1}$ and a pressure of between 70–80 kg cm$^{-2}$. The software for the HPLC (Varian Star Workstation) was used to integrate the peaks.

The off-line UV-visible technique is used in most routine analysis, since almost all laboratories have the equipment needed for this. Therefore, it is necessary to compare the performances of the in situ and off-line UV-visible techniques in order to assess the real impact of using immersion probes to monitor the systems of interest here. For this reason, unfiltered samples taken from the settler were analyzed in a UV-visible spectrometer (model V560, JASCO) using a quartz cell with a path length of 1 cm. The results were then compared with those obtained with the in situ UV-visible technique.

All measurements were performed in triplicate, and the average values obtained were used to achieve the desired correlations.

Theory

All calculations were performed using MATLAB (version 6.5, Mathworks, Inc., Natick, MA, USA).

The models used for COD, TSS, and nitrate prediction, based on the collected UV-visible and NIR spectra, were developed using the PLS1 algorithm [33]. An internal full cross-validation (leave-one-out) was performed to optimize the model, with the number of model components (latent variables) chosen by the lowest root mean square error of cross-validation, RMSECV.

$$RMSECV = \left[ \left( \mathbf{Y}_C - \widehat{\mathbf{Y}}_C \right)^t \left( \mathbf{Y}_C - \widehat{\mathbf{Y}}_C \right) N_C^{-1} \right]^{0.5} \qquad (1)$$

In Eq. 1, $\widehat{\mathbf{Y}}_C$ and $\mathbf{Y}_C$ are the estimated PLS cross-validation and the measured reference value for the $i$th sample, respectively. $N_C$ is the number of calibration samples. Further details on PLS can be found elsewhere [33]. Spectral preprocessing methods were applied to the raw spectra to remove undesirable spectral variations, like baseline drift, light-scattering effects, and temperature variations. The methods used were Savitzky–Golay (SG), multiplicative scatter correction (MSC), and standard normal variate (SNV) [34]. Mean centering (MNCN) was always applied after spectral preprocessing. Different combinations of these methods were tested using different parameters such as the filter window and derivative order for the SG method. An algorithm for wavelength selection was used in order to optimize the PLS model. This algorithm, known as bootstrapping, is a statistical method that generates new datasets by sampling with replacement from the original data set. Different models are obtained with these datasets. The statistical significance of each
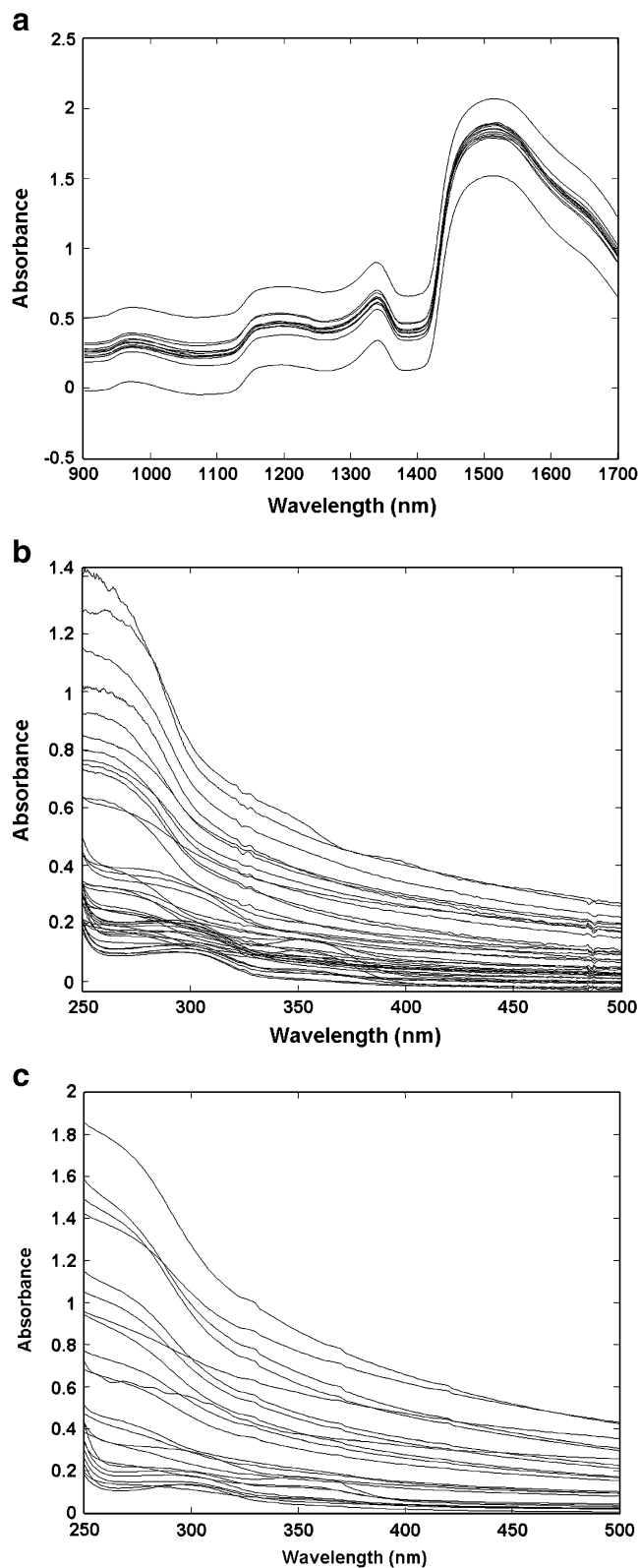


Fig. 1 Raw spectra obtained in situ with the NIR analyzer (**a**), in situ with the UV-visible analyzer (**b**), and off-line with the UV-visible analyzer (**c**). The spectra were collected between 900 and 1400 nm for NIR and between 250 and 500 nm for UV-visible

regression coefficient is assessed by determining the confidence interval at the 95% level by iterating the standard deviation. If the interval includes the zero value then the corresponding wavenumber is discarded [35]. The statistically significant wavenumbers were then used in a PLS1 model optimized as described above.

## Results and discussion

Typical spectra acquired during the monitoring period, for both the NIR and UV-visible ranges, are presented in Fig. 1. The NIR spectra depicted in Fig. 1a show a broad band between 1400 and 1700 nm corresponding to the vibration of the first overtone of the O–H bond of water. This band, which is typical of the NIR spectrum of an aqueous solution, makes the technique difficult to use for some other species, since it masks any other bands present in this spectral range, such as that for the first overtone of the C–H stretching vibration. Figures 1b and c represent typical spectra measured in the UV-visible range and obtained with the on-line immersion probe and the off-line method using quartz cuvettes, respectively. As expected, the spectra observed for both methods are similar, with the majority of the information present in the spectral region between 250 and 400 nm. The number of samples used to establish the correlations as well as the concentration ranges for each of the studied parameters and the average experimental errors associated with the reference methods are presented in Table 1. The determination of the nitrate concentration has the lowest experimental error (1.72%), while COD and TSS were measured with average errors of 4.79% and 4.27%, respectively.

As explained in the previous section, variable selection was performed using a bootstrapping technique, which was employed to select the wavelength ranges that would best enhance the results of the PLS correlation models by eliminating superfluous and correlated information. The spectral ranges, the preprocessing techniques used in each case, and the results obtained with the PLS calibration are compiled in Table 2. The number of PLS components was chosen based on the minimum RMSECV. The dependence of the RMSECV on the number of latent variables can be seen in Figure 2. For all methods and parameters, the minimum RMSECV value is well defined, thus providing a good indication of the appropriate number of components for each model.

All of the models were improved through the use of the variable selection technique, with the exception of the model for TSS determination using the spectra measured off-line. However, a poor correlation was obtained between the experimental and predicted COD data when using the NIR range. One of the main reasons for this is that, as mentioned above, NIR spectroscopy is very sensitive to the presence of water, as bands from water mask the majority of the bands corresponding to the first overtone of C–H stretching vibrations, which results in a loss of sensitivity to the presence of organic matter. Also, information from the middle and far infrared regions is lost because the detector used in this work only detects up to 1700 nm. Nevertheless, the narrow range of the COD experimental data measured here (between 17.24 and 99.53 mg $O_2$ $L^{-1}$) is believed to be the major contributor to this poor correlation, as a wide calibration range is one of the crucial factors needed to generate a robust calibration model [27, 36].

NIR spectroscopy does not detect the presence of inorganic compounds, and so the correlation obtained for nitrate is only an indirect measure, as now explained. In order to achieve improved nitrification process efficiency in an activated sludge system, the reproductive rate of the nitrifying bacteria must be greater than their removal rate from sludge wasting [37]. To increase the number of bacteria in the reactor, the biomass content was increased by limiting the sludge purge frequency. When the nitrification process was occurring at a satisfactory rate, operational problems related to an increase in the TSS in the effluent were identified. These were possibly related to a low food-to-microorganism ratio [1]. Hence, it is probably the variation in TSS in the effluent that is effectively detected by the NIR spectra. Therefore, the value of 0.61 for the

**Table 1** Dataset used to develop the PLS models for COD, N-$NO_3^-$ and TSS

| Technique | NIR | | | In situ UV-visible | | | Off-line UV-visible | | | Reference methods |
|---|---|---|---|---|---|---|---|---|---|---|
| | Number of samples | Concentration[a] (mgL$^{-1}$) | | Number of samples | Concentration[a] (mgL$^{-1}$) | | Number of samples | Concentration[a] (mgL$^{-1}$) | | Average error % |
| Parameter | | Min | Max | | Min | Max | | Min | Max | |
| COD | 14 | 17.2 | 99.5 | 39 | 0.0 | 163.5 | 24 | 17.2 | 163.5 | 4.79 |
| N-$NO_3^-$ | 14 | 85.0 | 164.4 | 30 | 0.0 | 164.4 | 19 | 0.0 | 164.4 | 1.72 |
| TSS | 13 | 41.0 | 165.0 | 30 | 16.7 | 184.2 | 21 | 16.7 | 184.2 | 4.27 |

[a] The COD concentration is expressed in mg $O_2$ $L^{-1}$

**Table 2** Spectroscopy-based COD, N-NO₃⁻ and TSS PLS modeling results

| Technique | NIR | | | In situ UV-visible | | | Off-line UV-visible | | |
|---|---|---|---|---|---|---|---|---|---|
| | COD | N-NO$_3^-$ | TSS | COD | N-NO$_3^-$ | TSS | COD | N-NO$_3^-$ | TSS |
| Range (nm) | 900–1400 | 900–1400 | 900–1400 | 250–500 | 250–380 | 250–380 | 250–500 | 250–380 | 250–380 |
| Preprocessing | SNV and MNCN | SNV and MNCN | SNV and MNCN | SNV and MNCN | SG[c] and MNCN | SG[c] and MNCN | SG[c] and MNCN | SG[c] and MNCN | SG[d] and MNCN |
| Number of wavelengths | 136 | 26 | 142 | 145 | 206 | 10 | 25 | 9 | [b] |
| PLS latent variables | 8 | 2 | 4 | 6 | 3 | 4 | 9 | 3 | 9 |
| RMSECV (mg L$^{-1}$)[a] | 27.8 | 33.7 | 12.5 | 15.4 | 19.0 | 19.8 | 18.2 | 23.8 | 24.1 |
| Relative standard deviation (%) | 70.3 | 24.0 | 19.7 | 19.9 | 6.8 | 13.9 | 45.5 | 11.1 | 81.1 |
| Relative error (%) | 52.4 | 37.8 | 14.1 | 23.1 | 26.6 | 28.9 | 24.2 | 35.3 | 29.0 |
| Correlation coefficient | 0.23 | 0.61 | 0.91 | 0.82 | 0.87 | 0.82 | 0.77 | 0.79 | 0.81 |

[a] For COD the units are mg O$_2$ L$^{-1}$.

[b] Results are not shown because variable selection did not improve the results in this case.

[c] Fifteen-point Savitzky–Golay filter with a second-order polynomial fit and first derivative.

[d] Five-point Savitzky–Golay filter with a second-order polynomial fit and second derivative

correlation coefficient only reflects a tendency; it does not reflect the true correlation between the NIR spectra and the nitrate concentration.

The best results obtained in the NIR range were those for the TSS, with a relative error (RMSECV divided by the nominal value of the parameter) of 14.1%, corresponding to a correlation coefficient of 0.91. This result indicates the high sensitivity of NIR spectroscopy to physical changes in the system.

The UV-visible in situ technique yielded better results, with similar relative errors for the three parameters of 23.1%, 26.6%, and 28.9%, for the COD, N-NO₃⁻, and TSS, respectively. The off-line UV-visible technique gave similar

results to the in situ technique for COD and TSS. However, the relative error for nitrate prediction was almost 10% higher for the off-line technique.

The relative standard deviation (RSD) was calculated using the same bootstrap technique as used for variable selection [35]. The value was calculated by finding the average of the results of 500 different models obtained by bootstrapping. The RSD for the NIR model is comparable with the RSDs for the other two techniques, indicating that, even with a lower number of calibration samples, the NIR results have the same level of precision as the results obtained using the UV-visible techniques. Upon comparing the RSDs of the three parameters for the NIR technique, is



**Fig. 2** Root mean square error of cross-validation (RMSECV) as a function of the number of PLS latent variables
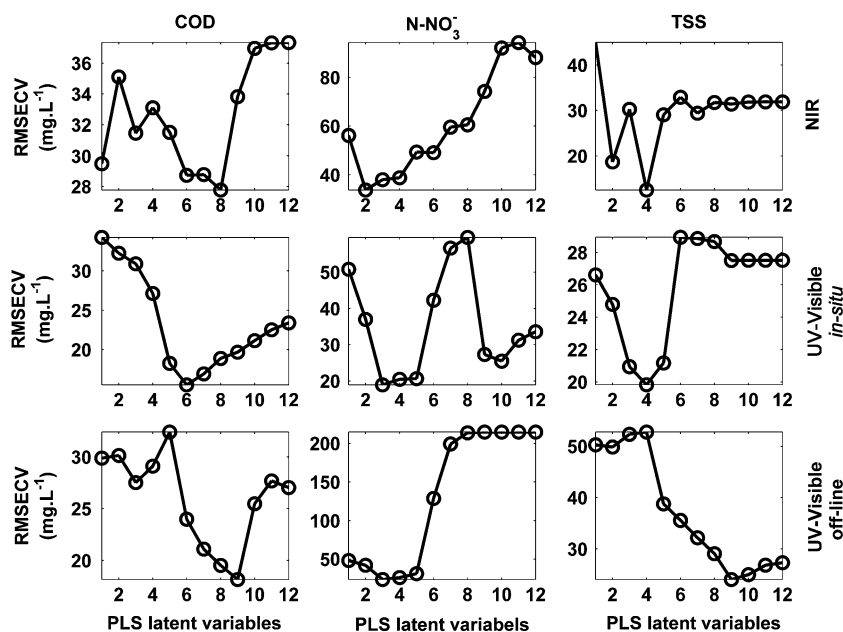
becomes clear that the main problem with the COD result is not the number of calibration samples, since the other two parameters have more or less the same number of calibration samples, but the lower sensitivity of this technique in aqueous environments and the small concentration range used for this parameter. Figures 3, 4, 5 show plots of reference versus predicted values for the parameters studied, and for each of the spectral ranges and/or techniques used (in situ and off-line). Clear, linear relations are obtained in all cases. Even though, considering actual current legislation, the obtained results still exhibit significant and limiting errors in relation to quantitative measurements, they still represent promising input data for on-line artificial intelligence monitoring and control systems, such as fuzzy logic or neural network-based control systems [38]. A comparison of UV-visible off-line and in situ measurements allows us to conclude that the in situ technique is clearly advantageous, since, aside from its better predictive capabilities, the method avoids the need for sampling and pretreatment, which would be needed to improve the results obtained with the data acquired off-line.

## Conclusions

This work assessed the abilities of in situ spectroscopic techniques to monitor an activated sludge reactor by estimating the key parameters of COD, N-NO$_3^-$, and TSS. Characteristic spectra were collected in both the NIR and UV-visible spectroscopic ranges through the use of immersion probes on the settler. PLS was used as the regression technique to correlate the acquired spectra with the monitored parameters, and a variable selection methods was employed to remove superfluous data.
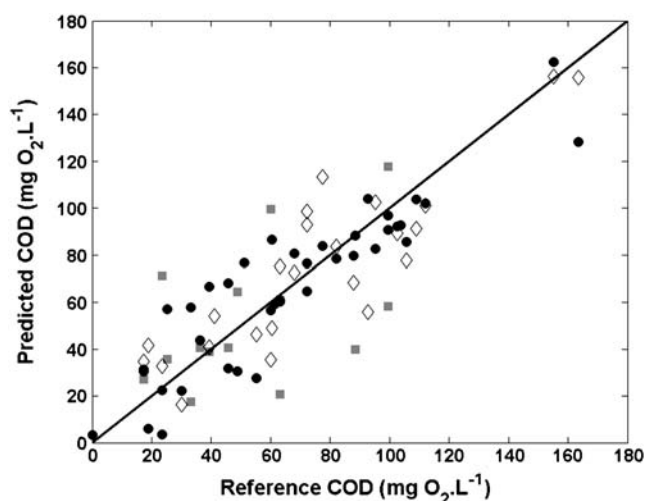


Fig. 3 Reference versus predicted COD for NIR (*gray squares*), in situ UV-visible (*black circles*), and off-line UV-visible (*white diamonds*)
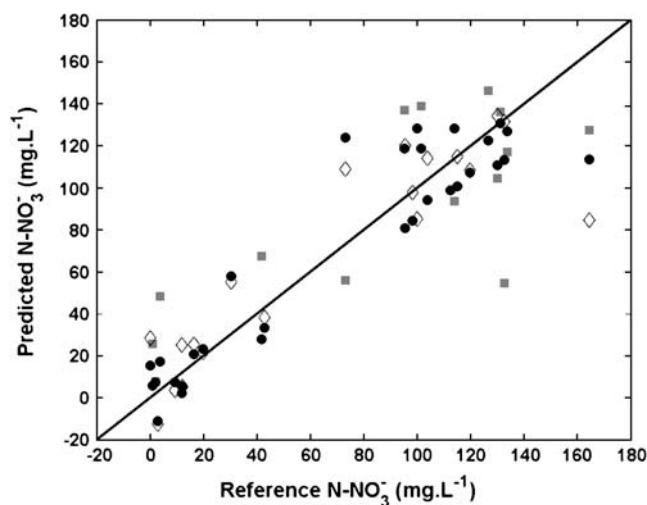


Fig. 4 Reference versus predicted N-NO$_3^-$ for NIR (*gray squares*), in situ UV-visible (*black circles*), and off-line UV-visible (*white diamonds*)

The NIR modeling results were not as accurate as the results obtained with UV-Vis, particularly when modeling the COD. The presence of water has a large effect on NIR spectra, masking spectral features that could be important for estimating the COD. Despite this issue, the results presented here show that both spectroscopic ranges can be used in practice to predict key process parameters (with different accuracies) on-line/in situ, thus avoiding the need for time-consuming off-line analyses, which are currently usually used to infer the status of a biological system at any given moment. A comparison between the off-line and in situ UV-visible methodologies showed that the in situ technique yielded the most accurate results. Moreover, this technique enables the the biological system to be continuously monitored, removing the need for sampling and off-line analysis. Future work will focus on enlarging the
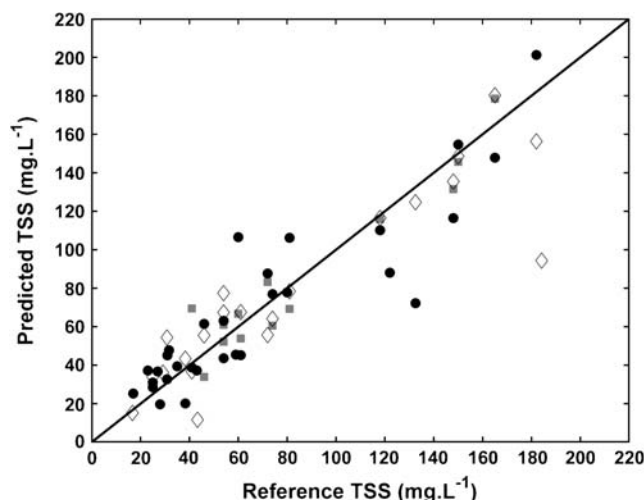


Fig. 5 Reference versus predicted TSS for NIR (*gray squares*), in situ UV-visible (*black circles*), and off-line UV-visible (*white diamonds*)

number and range of samples measured, in order to test the transferability and robustness of the correlations presented in this work for predicting COD, N-NO$_3^-$ and TSS values under different operating conditions.

# References

1. Metcalf and Eddy, Inc., Tchobanoglous G, Burton FL, Stensel HD (2003) Wastewater engineering: treatment and reuse, 4th edn. McGraw-Hill, New York
2. Maribas A, Silva MDL, Laurent N, Loison B, Battaglia P, Pons MN (2008) Monitoring of rain events with a submersible UV/VIS spectrophotometer. Water Sci Technol 57:1587–1593
3. Bourgeois W, Burgess JE, Stuetz RM (2001) On-line monitoring of wastewater quality: a review. J Chem Technol Biotechnol 76:337–348
4. Vaillant S, Pouet MF, Thomas O (1999) Methodology for the characterisation of heterogeneous fractions in wastewater. Talanta 50:729–736
5. Dias AMA, Moita I, Pascoa R, Alves MM, Lopes JA, Ferreira EC (2008) Activated sludge process monitoring through in situ near-infrared spectral analysis. Water Sci Technol 57:1643–1650
6. Fogelman S, Zhao HJ, Blumenstein M (2006) A rapid analytical method for predicting the oxygen demand of wastewater. Anal Bioanal Chem 386:1773–1779
7. Pons MN, Le Bonte S, Potier O (2004) Spectral analysis and fingerprinting for biomedia characterisation. J Biotechnol 113:211–230
8. Wu J, Pons MN, Potier O (2006) Wastewater fingerprinting by UV-visible and synchronous fluorescence spectroscopy. Water Sci Technol 53:449–456
9. Mrkva M (1975) Automatic UV-control system for relative evaluation of organic water-pollution. Water Res 9:587–589
10. Dobbs RA, Dean RB, Wise RH (1972) Use of ultraviolet absorbance for monitoring total organic carbon content of water and wastewater. Water Res 6:1173–1180
11. Matsche N, Stumwohrer K (1996) UV absorption as control-parameter for biological treatment plants. Water Sci Technol 33:211–218
12. Thomas O, Theraulaz F, Domeizel M, Massiani C (1993) UV sectral deconvolution—a valuable tool for waste-water quality determination. Environ Technol 14:1187–1192
13. Langergraber G, Fleischmann N, Hofstaedter F, Weingartner A (2004) Monitoring of a paper mill wastewater treatment plant using UV/VIS spectroscopy. Water Sci Technol 49:9–14
14. Rieger L, Langergraber G, Siegrist H (2006) Uncertainties of spectral in situ measurements in wastewater using different calibration approaches. Water Sci Technol 53:187–197
15. El Khorassani H, Trebuchon P, Bitar H, Thomas O (1999) A simple UV spectrophotometric procedure for the survey of industrial sewage system. Water Sci Technol 39:77–82
16. Thomas O, Theraulaz F, Agnel C, Suryani S (1996) Advanced UV examination of wastewater. Environ Technol 17:251–261
17. Escalas A, Droguet M, Guadayol JM, Caixach J (2003) Estimating DOC regime in a wastewater treatment plant by UV deconvolution. Water Res 37:2627–2635
18. Karlsson M, Karlberg B, Olsson RJO (1995) Determination of nitrate in municipal waste-water by UV spectroscopy. Anal Chim Acta 312:107–113
19. Langergraber G, Fleischmann N, Hofstadter F (2003) A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. Water Sci Technol 47:63–71
20. Langergraber G, Gupta JK, Pressl A, Hofstaedter F, Lettl W, Weingartner A, Fleischmann N (2004) On-line monitoring for control of a pilot-scale sequencing batch reactor using a submersible UV/VIS spectrometer. Water Sci Technol 50:73–80
21. Rieger L, Langergraber G, Thomann M, Fleischmann N, Siegrist H (2004) Spectral in-situ analysis of NO$_2$, NO$_3$, COD, DOC and TSS in the effluent of a WWTP. Water Sci Technol 50:143–152
22. Vaillant S, Pouet MF, Thomas O (2002) Basic handling of UV spectra for urban water quality monitoring. Urban Water 4:273–281
23. Roggo Y, Chalus P, Maurer L, Lema-Martinez C, Edmond A, Jent N (2007) A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J Pharm Biomed Anal 44:683–700
24. Reich G (2005) Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. Adv Drug Deliv Rev 57:1109–1143
25. Stephens AB, Walker PN (2002) Near-infrared spectroscopy as a tool for real-time determination of BOD5 for single-source samples. Trans ASAE 45:451–458
26. Sousa AC, Lucio MMLM, Bezerra OF, Marcone GPS, Pereira AFC, Dantas EO, Fragoso WD, Araujo MCU, Galvao RKH (2007) A method for determination of COD in a domestic wastewater treatment plant by using near-infrared reflectance spectrometry of seston. Anal Chim Acta 588:231–236
27. Pascoa RNM, Lopes JA, Lima JLFC (2008) In situ near infrared monitoring of activated dairy sludge wastewater treatment processes. J Near Infrared Spectrosc 16:409–419
28. Blanco M, Villarroya I (2002) NIR spectroscopy: a rapid-response analytical tool. Trends Anal Chem 21:240–250
29. Buning-Pfaue H (2003) Analysis of water in food by near infrared spectroscopy. Food Chem 82:107–115
30. Chen D, Hu B, Shao XG, Su QD (2004) Removal of major interference sources in aqueous near-infrared spectroscopy techniques. Anal Bioanal Chem 379:143–148
31. Marquez MC, Costa C, Jul M (2004) Kinetics and effects of hydraulic residence time and biomass concentration on removal of organic pollutants in a continuous unsteady state activated sludge process. Chem Biochem Eng Q 18:423–428
32. APHA, AWWA, WPC (1989) Standard methods for the examination of water and wastewater. American Public Health Association/American Water Works Association/Water Pollution Control, Washington, DC
33. Geladi P, Kowalski BR (1986) Partial least-squares regression—a tutorial. Anal Chim Acta 185:1–17
34. Miller CE (2000) Chemometrics for on-line spectroscopy applications—theory and practice. J Chemometrics 14:513–528
35. Wehrens R, VanderLinden WE (1997) Bootstrapping principal component regression models. J Chemometrics 11:157–171
36. Sarraguça MC, Lopes JA (2009) Quality control of pharmaceuticals with NIR: from lab to process line. Vib Spectrosc 49:204–210
37. Gerardi MH (2002) Nitrification and denitrification in the activated sludge process. Wiley, New York
38. Dias AMA, Ferreira EC (2009) Computational intelligence techniques for supervision and diagnosis of biological wastewater treatment systems. In: do Carmo Nicoletti M, Jain LC (eds) Computational intelligence techniques for bioprocess modelling, supervision and control. Springer, Berlin, pp 127–162