Quantification of pharmaceutical compounds in wastewater samples by near infrared spectroscopy (NIR)

C. Quintelas[a,*], D.P. Mesquita[a], E.C. Ferreira[a], A.L. Amaral[a,b]

[a] CEB - Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal
[b] Instituto Politécnico de Coimbra, ISEC, DEQB, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal

ARTICLE INFO

ABSTRACT

The quantification of pollutants, as pharmaceuticals, in wastewater is an issue of special concern. Usually, typical methods to quantify these products are time and reagent consuming. This paper describes the development and validation of a Fourier transform near-infrared (FT-NIR) spectroscopy methodology for the quantification of pharmaceuticals in wastewaters. For this purpose, 276 samples obtained from an activated sludge wastewater treatment process were analysed in the range of $200 \, cm^{-1}$ to $14,000 \, cm^{-1}$, and further treated by chemometric techniques to develop and validate the quantification models. The obtained results were found adequate for the prediction of ibuprofen, sulfamethoxazole, 17β-estradiol and carbamazepine with coefficients of determination (R^2) around 0.95 and residual prediction deviation (RPD) values above four, for the overall (training and validation) data points. These results are very promising and confirm that this technology can be seen as an alternative for the quantification of pharmaceuticals in wastewater.

1. Introduction

Emerging contaminants, as pharmaceuticals, are compounds of special concern due to the widespread usage and growing presence in aqueous systems. The increased environmental presence of these compounds is mainly due to the exponential growth of human population and the development of industrial, agricultural, and health care activities required to suport their well-being. Acording to Gomes et al. [1], many of these new substances are being marketed with little information about their ecotoxity to non-target species and their risks to the environment. These lack of information results of several factors: i) the required information often depends on the quantities produced and placed in the market; ii) the ecotoxicological evaluation is sometimes performed after their entrance in the market, and iii) the lack of fast and reliable methods to quantify these compounds in the environment.

Several studies have shown the relevance and impact of the presence of these compounds in water bodies. Houtman [2] studied the emerging contaminants in surface waters, and their relevance for the production of drinking water in Europe, and found that compounds as the natural estrogenic hormones 17β-estradiol, estriol and estrone, the synthetic estrogenic hormone 17α-ethynylestradiol, and pharmaceuticals as bezafibrate, ibuprofen, sulfamethoxazole, metoprolol, carbamazepin, as well as compounds used as X-ray contrast media, are the most significative compounds found in Europe's surface waters. A similar panorama was found by Castiglioni et al. [3], which analysed the presence of 80 emerging contaminants and realise that almost all emerging contaminants were ubiquitous in untreated wastewater, at concentrations up to the µg/L range, with the most abundant classes being pharmaceuticals and anthropogenic markers. Also Han Tran et al. [4] analysed the ocurrence of 60 emerging compounds in Asia, Europe, and North America and verified that the concentrations of the most important emerging compounds in raw influents in the Asian region tend to be higher than those in European and North American countries. Furthermore, these authors state that these discrepancies could be attributed to the differences in usage patterns in each region, climate conditions, population size and/or density, analytical methods and particularly, sampling strategies.

The traditional techniques used for the determination of pharmaceuticals and other emerging compounds include titrimetric techniques, chromatografic techniques, high-performance liquid chromatography (HPLC), gas chromatography (GC), and electrochemical and electrophoretic methods [5]. These techniques are time or reagent consuming and faster, and labour and environmental friendly methods are welcome. In this context, this study presents a very simple, non-destructive, inexpensive and green strategy applied to the determination of ibuprofen (IBU), carbamazepine (CRB), β-estradiol (E₂), ethinylestradiol (EE₂) and sulfamethoxazole (SMX) concentrations using FT-NIR spectroscopy, in aqueous solutions. This technique presents, as most

---

* Corresponding author.
  E-mail address: cquintelas@deb.uminho.pt (C. Quintelas).

interesting advantages, the absence of reagents use, its nondestructive character allowing to reuse the sample after measurement [6], speed and the possibility of online monitoring. The main difficulty of this technique is the complexity of the spectra due to their nature, and for this reason FT-NIR is not used as a direct analysis technique. Recent developments in chemometrics allowed to overcome these difficulties and the combination between FT-NIR and chemometrics is now considered as a promising technology able to quantify a wide range of organic compounds.

2. Materials and methods

2.1. Sample preparation

Experiments with different initial ibuprofen (IBU), carbamazepine (CRB), β-estradiol (E_2), ethinylestradiol (EE_2) and sulfamethoxazole (SMX) concentrations (within the range of mg L^{-1}) were conducted in Erlenmeyer flasks, with 300 mL of working volume in batch mode, operated at room temperature during, approximately, five days. The agitation was kept constant at 150 rpm. The flasks were inoculated with activated sludge from a domestic wastewater treatment plant with the use of an initial mixed liquor suspended solids (MLSS) concentration of 3 g L^{-1}. A synthetic medium was fed to the system in the beginning of each experiment and contained (per liter): 2.55 g $C_2H_3O_2Na\cdot3H_2O$, 0.34 g $C_3H_5NaO_2$, 0.59 g NH_4Cl, 0.95 g $MgSO_4\cdot7H_2O$, 0.44 g $CaCl_2\cdot2H_2O$, 0.03 g EDTA, and 3.16 mL of a trace metals solution. The trace metals solution is described in Smolders et al. [7]. Samples were taken at different time intervals, varied from 0 h to 53 h, centrifuged (13400 rpm for 10 min) and the aqueous phase was stored at 4 °C. Prior to analysis, the liquid samples were thawed and homogenized by vortexing. The pharmaceuticals concentrations were determined using a UHPLC system.

2.2. UHPLC analysis

The chromatographic analysis was performed using a Shimadzu Corporation apparatus (Tokyo, Japan) consisting of a UHPLC equipment (Nexera) with one multi-channel pump (LC-30 CE), an autosampler (SIL-30AC), an oven (CTO-20AC), a diode array detector (M-20A) and a system controller (CBM-20A) with built-in software (LabSolutions). For the analysis of all the five compounds a Kinetex 5 u EVO C18 column (150 × 4.6 mm i.d.) supplied by Phenomenex, Inc. (CA, USA) was used. For IBU, the mobile phase, with a flow rate of 1.5 mL min^{-1}, consisted of sodium phosphate (20 mM; pH 2.4) (pump A) and acetonitrile (pump B). The starting mobile phase composition was 80% A, decreasing to 30% A in 10 min and remaining in this percentage for 5 min. The sample was monitored by the diode array detector from 190 to 400 nm, and the chromatograms were extracted at 225 nm. The column oven was set to 40 °C and the volume of injection was 10 μL. For SMX, the mobile phase, with a flow rate of 0.3 mL min^{-1}, consisted of water (pump A) and acetonitrile (pump B) and the isocratic mode, with 40% A and 60% B, was used. The sample was monitored by the diode array detector from 190 to 400 nm, and the chromatograms were extracted at 267 nm. The column oven was set to 25 °C and the volume of injection was 20 μL. For CRB, The mobile phase, with a flow rate of 1.0 mL min^{-1}, consisted of water (pump A) and acetonitrile (pump B) and the isocratic mode, with 70% A and 30% B, was used. The sample was monitored by the diode array detector from 190 to 400 nm, and the chromatograms were extracted at 220 nm. The column oven was set to 25 °C and the volume of injection was 10 μL. For E_2 and EE_2, the mobile phase, with a flow rate of 0.8 mL min^{-1}, consisted of water (pump A) and acetonitrile (pump B) and the isocratic mode, with 55% A and 45% B, was used. The sample was monitored by the diode array detector from 190 to 400 nm, and the chromatograms were extracted at 220 nm. The column oven was set to 25 °C and the volume of injection was 20 μL. The standard errors for the

HPLC measurements were 0.029 mg L^{-1}, 0.016 mg L^{-1}, 0.122 mg L^{-1}, 0.043 mg L^{-1}, and 0.077 mg L^{-1}, respectively for IBU, SXF, CARB, EE_2 and E_2 and the values of R^2 for the calibration curves were around 1 for all the compounds.

2.3. Near infrared scanning

Near infrared (NIR) spectra were recorded on a Fourier-transform near infrared spectrometer (FTLA 2000, ABB, Thermo Electron Corporation) equipped with an indium–gallium–arsenide (InGaAs) detector, from 14,000 to 200 cm^{-1}, in transmitance mode using a flow cell with a 0.7 mm pathlength. For each sample, 32 scans were made with a spectral resolution of 8 cm^{-1} and then averaged. Samples were temperature equilibrated at 23 °C (approximately 3 min) in the instrument before scanning. The integration time was adjusted until the peaks at 1100–1200 nm for NIR were close to 60,000 intensity units. Grams/ AI software (Thermo Electron Corporation) was used for spectrometer configuration, control, and data acquisition. Distilled water was used as background. A typical NIR spectrum is presented as Supplementary material (Fig. S1).

2.4. Kolmogorov–Smirnov test and boxplot analysis

The Y dataset employed in the chemometric analyses consisted of the IBU, SMX, E_2, EE_2 and CRB concentrations, monitored throughout the time length of the different experiments in this work, whilst the X dataset consisted of the collected FT-NIR spectra (ranging from 14,000 to 200 cm^{-1}). The following chemometric techniques were employed to this data in a sequential manner, as follows: i) Kolmogorov–Smirnov test and boxplot analysis to infer normal distributions and identify Y dataset outliers; ii) principal component analysis (PCA) to identify samples interrelationships (clusters) and X dataset outliers; and iii) partial least squares (PLS) regression to derive the models for each studied compound.

The first test employed was the Kolmogorov–Smirnov test to infer the null hypothesis that a given dataset is normally distributed. This test was first employed for the Y datasets, and then for selected X datasets (suspected of presenting X outliers). Next, a boxplot analysis was performed, given that a normally distributed data was observed, returning a box graph with the median as the central mark, the 25th and 75th percentiles as the edges, and the most extreme data points (not considering the outliers) as the whiskers. The whisker length was defined as a factor (1.5) of the interquartile distance between the 25th and 75th percentiles, covering 99.3% of the normally distributed data. As a result, outliers are plotted individually, outside the box, and can be identified by visual inspection.

2.5. Principal components analysis

A principal components analysis (PCA) was further performed to the X dataset (wavelengths values), in order to both allow for the recognition of possible interrelationships (clusters) between samples as for the identification of possible outliers. For the performed PCA study, as well as for the PLS, the complete datasets, upon the exclusion of the found E_2 outlier, were employed resulting in a total of 36, 60, 59, 60 and 60 samples for the IBU, SMX, E_2, EE_2 and CRB datasets, respectively. Subsequently, 2/3 of these samples were used for modeling (calibration) purposes and 1/3 for validation in the PLS analysis.

2.6. Partial least squares regression

Next a PLS analysis was employed in order to predict the studied pharmaceuticals concentrations from the wavelength dataset. A standard normal variate (SNV) methodology was employed to remove undesirable X data matrix variations, alongside a cross-validation (CV) technique to test its predictive significance.

Two different methodologies were employed regarding the performed PLS analysis, the first [M1] employing the raw dataset, and the second [M2] employing on an iterative method. The second methodology consisted of the following sequential steps: i) determination of each wavelength weights for the entire wavelength range in an initial PLS analysis, ii) arrangement of the wavelength values according to weight similarity and, iii) final PLS analysis with the averaged wavelength values [8].

Furthermore, a total of 500 possible random samples combinations, for each dataset, were screened to select the most unbiased calibration and validation datasets. For all PLS analyses, the maximum number of PLS components allowed was set at half of the calibration data.

All of the above analyses were performed in Matlab 7.11 (The Mathworks, Inc. Natick, USA). Further details regarding the PCA and PLS techniques can be found in Einax et al. [9].

3. Results and discussion

3.1. Analytical data

The minimum, average and maximum values for the ibuprofen (IBU), sulfamethoxazole (SMX), β-estradiol (E_2), ethinylestradiol (EE_2) and carbamazepine (CRB) concentrations in the experiments are presented in Table 1. A total of 6 experiments were employed for the IBU study, whereas 10 experiments were used for the SMX, E_2, EE_2 and CRB studies. Each of these samples was then divided into two groups, the calibration (modeling) group, with 2/3 of the samples, and the validation, with the remaining 1/3 of the samples. The number of samples used for each studied pharmaceutical is presented in Table 2.

3.2. Kolmogorov–Smirnov test and boxplot analysis results

Prior to the boxplot analysis, a Kolmogorov–Smirnov test was first employed in order to determine if the Y datasets (IBU, SMX, E_2, EE_2 and CRB concentrations) were normally distributed. The obtained results, for all the studied datasets, allowed confirming that all Y datasets were not normally distributed. For that reason, the boxplot analysis was not performed for the Y dataset (although it was performed further on within the PCA analysis), and no Y dataset outliers' exclusion was performed, at this time.

3.3. PCA results

A principal components analysis (PCA) was further performed to the X dataset (wavelengths values), as depicted in Fig. 1. The first, second and third principal components (or latent variables – LV) explained a total of 45.2%, 68.7%, 77.1%, 81.0% and 83.8% respectively for the IBU, SMX, E_2, EE_2 and CRB X dataset variation in the performed PCA analyses.

The PCA analysis revealed an outlier (Fig. 1c in the corner box analysis), for the E_2 wavelength dataset, later removed for the subsequent analysis (thus leading to a final 59 samples dataset). To that purpose, a boxplot analysis was also performed and presented in Fig. 2. The whisker length for this analysis was of 1.5, that is, 1.5 times the

Table 1
Minimum (Min.), maximum (Max.) and average (Avg.) values in mg L^{-1} in the IBU (ibuprofen), SMX (sulfamethoxazole), E_2 (17β-estradiol), EE_2 (17α-ethynylestradiol) and CRB (carbamazepine) concentrations (Y) dataset.

Samples	No.	Min.	Max.	Avg.
IBU	36	0.05	1.99	0.24
SMX	60	0.31	7.60	2.96
E_2	59	0	2.75	1.22
EE_2	60	0.44	6.93	3.15
CRB	60	0.56	8.12	4.02

Table 2
Calibration and validation datasets size for the each studied pharmaceutical.

Samples	Calibration	Validation
IBU	24	12
E_2	40	19
EE_2	40	20
SMX	40	20
CARB	40	20

interquartile distance between the 25th and 75th percentiles, for the identifications of samples falling outside a 99.3% coverage of a normally distributed data. Taking into consideration the obtained results, it could be seen that one data point regarding the E_2 X dataset fell outside this limit, regarding the 3rd LV. A Kolmogorov–Smirnov test was then performed allowing confirming that the 3rd LV data was normally distributed and, hence, the data point that fell outside the 99.3% coverage of a normally distributed data was removed from subsequent analyses. No further outliers could be determined for the IBU, SMX, EE_2 and CRB datasets.

It could also be found for the IBU, SMX, E_2, EE_2 and CRB (Fig. 1) that a number of possible clusters seem to emerge. Analyzing these clusters in detail, it was possible to infer some dependency of the overall spectra over time, but not strictly with respect to the studied compounds concentration, and rather related with the fact that, being the biological wastewater treatment a dynamic and multifaceted system, a significant number of changes in the wastewater composition occur over time. It should be kept in mind that, being the synthetic medium constant for all the experiments, changes not relating to the pharmaceutical concentrations could be attributed to small alterations in the activated sludge inoculum and/or to the biota composition and/or activity evolution over time. Furthermore, it could be seen that both the number of possible different clusters and the time span that they cover diverge from experiment to experiment and, therefore, it was considered not to be of practical use (for real wastewaters monitoring) to divide the dataset into such clusters in the subsequent analysis regarding the pharmaceutical compounds concentration prediction.

3.4. PLS results

A partial least squares (PLS) regression was performed to the IBU, SMX, EE_2 and CRB data obtained by the UHPLC analysis as Y dataset and the FT-NIR spectra (14,000–200 cm^{-1}) as the X dataset. In order to determine the best model results for the studied pharmaceutical compounds, the obtained equations, coefficients of determination (R^2), root mean square error (RMSE) (in relative percentage of the samples range) and residual predictive deviation (RPD) values were obtained and are presented in Table 3. According to Fearn [10], the RPD parameter, defined as the ratio between the standard deviation (SD) of a population and the standard error of cross validation (SECV) for a prediction, when larger than 3 is considered fair and recommended for screening purposes. Both the RMSE and RPD values for the overall (tr + val) and the validation (val) data points are presented.

In order to select the most unbiased calibration and validation datasets, a screening of 500 possible random combinations, for these datasets, was employed. In accordance, the best overall (calibration + validation) results were chosen, reflecting the most unbiased datasets combination. This procedure ensured that both calibration and validation datasets were distributed along the full range of the pharmaceuticals' concentrations.

For the SMX prediction, the iterative [M2] methodology provided the best results, whereas regarding the IBU, E_2, EE_2 and CRB predictions, the best results were obtained by the raw dataset [M1] methodology. For the overall samples of the SMX prediction, an R^2 value of 0.948, an RMSE value of 4.91% (7.90% for the validation samples) and an RPD value of 4.41 (2.74 for the validation samples) were obtained.
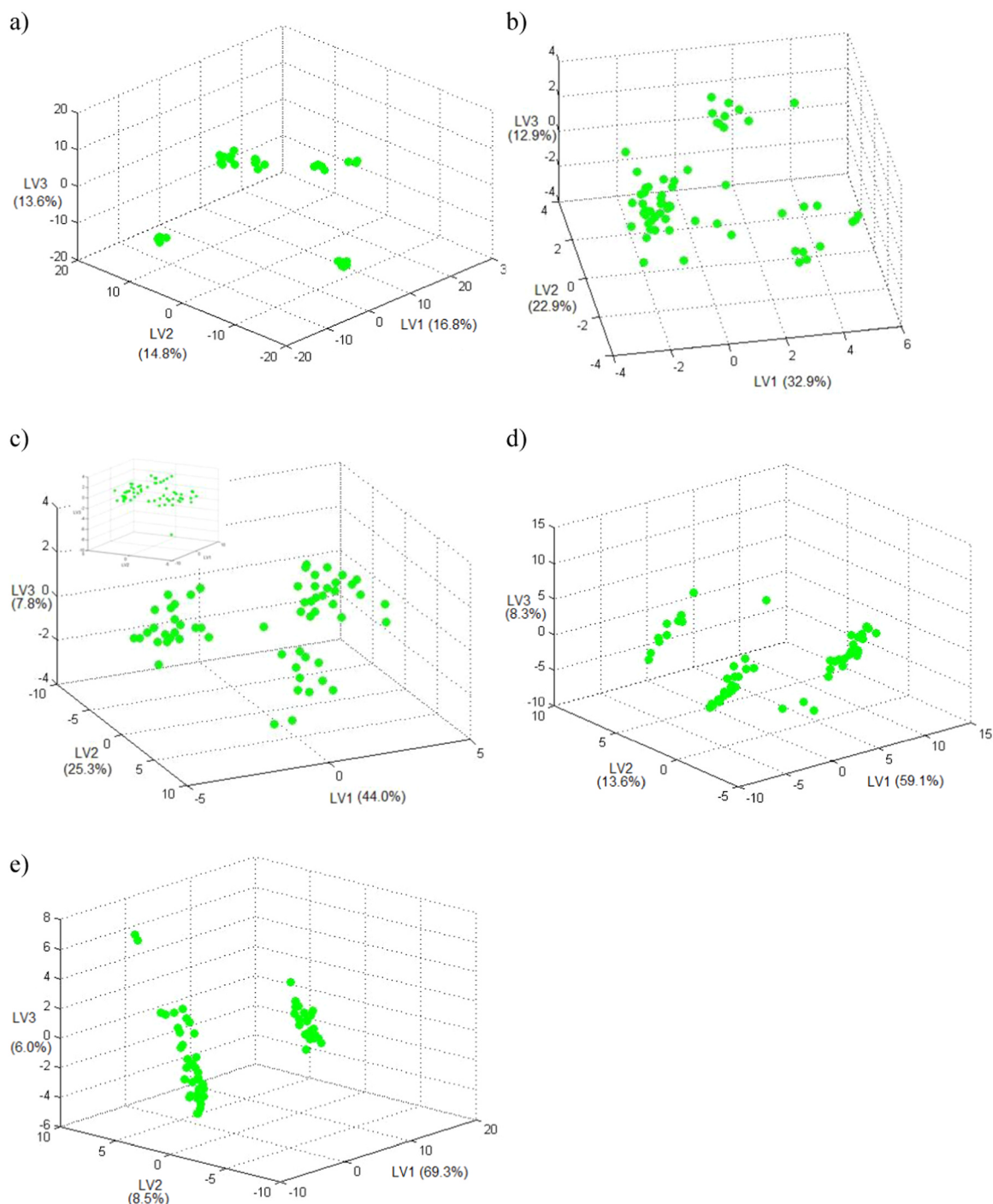
Fig. 1. PCA analysis of the X dataset, showing the first (LV1), second (LV2) and third (LV3) latent variables. a) IBU, b) SMX, c) E_2, d) EE_2, and e) CRB.

Considering the overall data RPD value of 4.41, and even though falling short of the value 3 regarding the validation data, these values confirm the potential of the developed iterative model [M2] for the SMX prediction. In fact, taking into account that the RMSE did not surpass 8% of the studied concentration range, regarding both the overall and the validation samples, this methodology seems to be quite promising.

Regarding the E_2 and EE_2 estrogens prediction, the overall samples presented R^2 values of 0.951 (E_2) and 0.858 (EE_2), RMSE values of 6.16% (E_2) and 10.12% (EE_2) and RPD values of 4.69 (E_2) and 2.83 (EE_2). These results, alongside the validation data RMSE (10.72% for E_2 and 17.50% for EE_2) and RPD (2.70 for E_2 and 1.64 for EE_2) values were worse than for the SMX prediction. Indeed, for the EE_2 prediction the

obtained results are still somewhat far from the desirable RMSE values below 10% of the studied range and RPD values above the value 3, thus further development on this methodology should be envisaged.

On the other hand, the best prediction results of the studied pharmaceuticals were obtained for the CRB compound. In fact, the obtained values for the overall samples, were as follows: R^2 value of 0.963, RMSE value of 5.10% (8.79% for the validation samples) and RPD value of 5.44 (3.16 for the validation samples). Indeed, both the overall and the validation data RPD values (5.44 and 3.16, respectively) were higher than the value 3, whereas the RMSE values remained below the 10% value of the studied range. Thus, the raw dataset [M1] methodology seems to be quite promising regarding the CRB prediction.
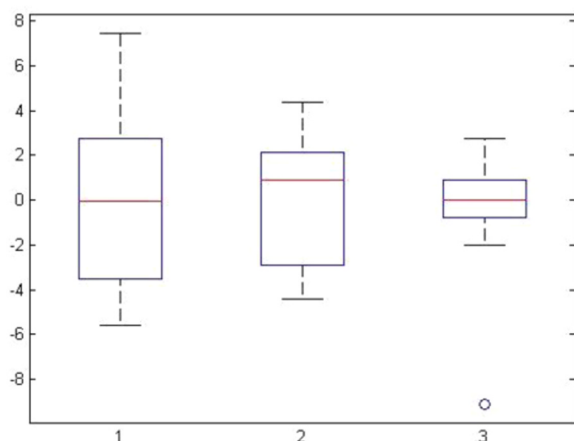
Fig. 2. Boxplot analysis of the three first LV's (representing 75.0%) of the E_2 X dataset for a 1.5 whisker length.

With respect to the IBU compound prediction, by the raw dataset [M1] methodology, the overall samples presented an R^2 value of 0.943, an RMSE value of 5.47% (9.17% for the validation samples) and an RPD value of 4.261 (2.54 for the validation samples). Again, the obtained RPD values (though falling short of the value 3 for the validation data) and RMSE values (below 10% of the studied concentration range in both cases) confirm the potential of the raw dataset model [M1] for the IBU prediction. It should be noted, though, that a quite different result was obtained for the iterative model [M2] for the IBU prediction, presenting a much worsen prediction ability with respect to the other studied compounds.

It should also be noticed that, in all cases, the number of PLS components or latent vectors (LVs), was larger for the raw dataset model [M1] with respect to the iterative model [M2]. This is in accordance to the fact that the raw model dataset is larger than the iterative model dataset (original variables transformed prior to the final PLS analysis into fewer averaged wavelength values) fed to the PLS analysis. Nonetheless, the number of LVs found for the optimal models (ranging from 6 for the IBU model to 17 for the CRB model), represented a massive decrease from the initial 3578 different wavelengths, while comprising the rule of being less than half of the calibration data points for each studied compound.

The studied prediction model results, regarding both the raw dataset [M1] and the iterative model [M2] methodologies, for all the studied pharmaceutic compounds concentrations, are next presented in Fig. 3. In all cases there seems to be a clear overfitting of the calibration data, regarding the [M1] methodology, explained by the large amount of the X dataset original variables (3578 different wavelengths). That being the case, the wavelengths selected by the model LVs concerning the calibration dataset may not be the most adequate with respect to

the validation dataset. This overfitting seems to be less profound for the iterative [M2] methodology given that it employs solely a fraction of the X dataset variables (original variables transformed prior to the final PLS analysis into fewer averaged wavelength values). Indeed, regarding the difference between the overall and the validation RMSE values, smaller differences were obtained, in most cases, for this methodology.

The use of FT-NIR by the pharmaceutical industry is well known. The main applications include quality control, raw materials identification, water content, active drug identification and in the control of individual steps in the production process ([6,11–14]) and are main focused in the evaluation and determination in solid dosage forms (tablets). More recently, some works describe the quantification and identification of active drugs in other pharmaceutical matrices as acetaminophen determination in low-dose pharmaceutical syrup [15], non-invasive quantification of 5 fluorouracil and gemcitabine in aqueous matrices by direct measurement through glass vials [16], determination of amoxicillin in suspension formulations [17] and quantification of active ingredients (metronidazole, erythromycin, salicylic acid and urea) in creams and ointments [18].

FT-NIR techniques are also frequently used to monitor wastewater treatment processes. The parameters usually evaluated by this technique include polyhydroxyalkanoates, extracellular polymeric substances, biochemical oxygen demand, chemical oxygen demand, total suspended solids, total organic and inorganic carbon, bicarbonate alkalinity, biochemical methane potential, volatile fatty acids, nitrate, ammonia [19] and volatile organic pollutants as tetrachloroethylene and dichlorobenzene ([20,21]) and fenitrothion [22]. However, the quantification of pharmaceuticals in wastewaters by FT-NIR is a field to explore and, to the authors' knowledge, no relevant works exist in the literature describing the use of this technique at this level. On the other hand, the growing increase of pharmaceuticals presence in wastewaters claims for rapid measurements, such as the approach described in the present report. In fact, the technology presented here presents several advantages over the traditional analytical procedures and is very promising for near real time monitoring of wastewater treatment processes.

4. Conclusions

A FT-NIR transmission spectroscopy methodology was developed for the determination of IBU, CRB, E_2, EE_2 and SMX concentrations in wastewaters. To achieve this aim, a chemometric approach was used employing first a Kolmogorov–Smirnov test to check the normality of the data, a boxplot analysis for outliers identification and a PCA analysis aiming to identify samples interrelationships and define the final datasets. Next, a PLS analysis was performed in order to obtain a prediction model suitable for the pharmaceuticals quantification purposes. This procedure resulted in relatively high coefficients of determination (R^2), and low root mean square errors (RMSE), for the prediction ability of almost all the studied compounds. Furthermore, the residual

Table 3
Equation (Eq.), coefficient of determination (R^2), root mean square error (RMSE) (%), number of samples (nS), number of PLS components or latent vectors (LVs) and residual predictive deviation (RPD) of the PLS analysis for the studied model results.

		Eq. (tr + val)	R^2 (tr + val)	RMSE (tr + val)	RMSE (val)	RPD (tr + val)	RPD (val)	nS	LVs
IBU	**M1**	**y = 1.022x**	**0.943**	**5.47**	**9.17**	**4.26**	2.54	36	**6**
	M2	y = 0.714x	0.644	13.71	21.07	1.70	1.11	36	1
SMX	M1	y = 0.998x	0.923	5.89	10.20	3.67	2.12	60	15
	M2	**y = 1.001x**	**0.948**	**4.91**	**7.90**	**4.41**	2.74	60	**11**
E_2	**M1**	**y = 0.987x**	**0.951**	**6.16**	**10.72**	**4.69**	2.70	59	**12**
	M2	y = 0.994x	0.927	7.62	12.64	3.74	2.28	59	10
EE_2	**M1**	**y = 0.981x**	**0.858**	**10.12**	**17.50**	**2.83**	1.64	60	**16**
	M2	y = 1.018x	0.840	10.77	17.54	2.66	1.63	60	10
CRB	**M1**	**y = 1.000x**	**0.963**	**5.10**	**8.79**	**5.44**	3.16	60	**17**
	M2	y = 1.002x	0.950	6.03	8.93	4.60	3.11	60	9

* All the *p-values* were below 0.0001.
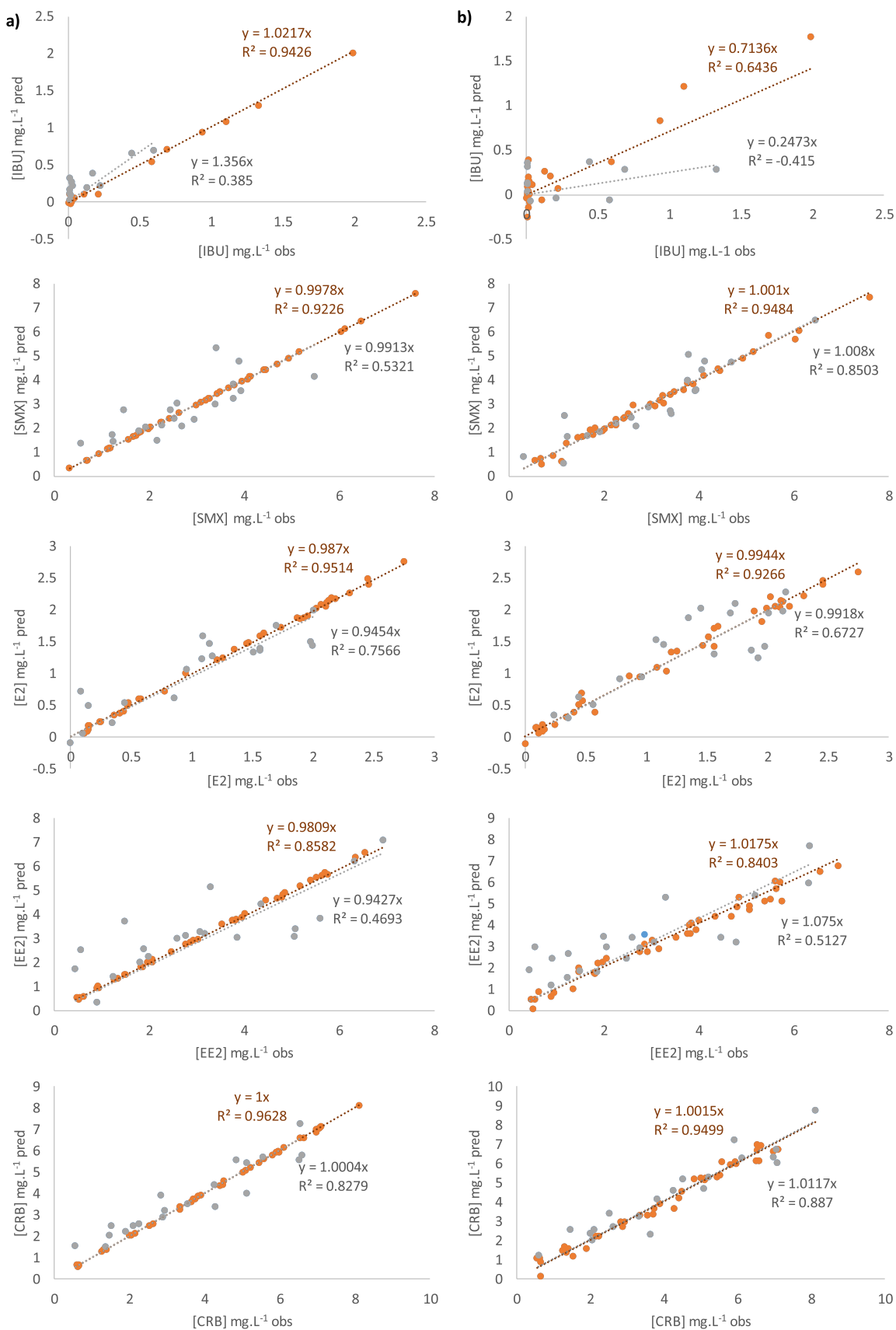
Fig. 3. Model results for the IBU, SMX, E$_2$, EE$_2$ and CRB predictions. a) M1 model, b) M2 model. The orange circles represent the calibration data and the gray circles represent the validation data (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

prediction deviation (RPD) for the overall (training and validation) samples was above three (except for EE_2) configuring its adequacy towards these compounds monitoring. The proposed method is not only a simple procedure, non-expensive and fast technique, but also it has relatively high sensitivity for the studied pharmaceutical compounds. Therefore, this methodology can be considered as promising towards the replacement of HPLC and GC analysis for routine pharmaceuticals quantification in wastewater, decreasing costs in reagents and analysis time.

Acknowledgments

Conflict of interest

There is no conflict of interest.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.talanta.2018.10.076.

References

[1] A.R. Gomes, C. Justino, T. Rocha-Santos, A.C. Freitas, A.C. Duarte, R. Pereira, Review of the ecotoxicological effects of emerging contaminants to soil biota, J. Environ. Sci. Health Part A Toxic/Hazard. Subst. Environ. Eng. 52 (2017) 1093–4529, https://doi.org/10.1080/10934529.2017.1328946/10.1080/10934529.2017.1328946 (J. J. Environ. Sci. Heal. Part A).

[2] C.J. Houtman, Emerging contaminants in surface waters and their relevance for the production of drinking water in Europe, J. Integr. Environ. Sci. (2010), https://doi.org/10.1080/1943815X.2010.511648.

[3] S. Castiglioni, E. Davoli, F. Riva, M. Palmiotto, P. Camporini, A. Manenti, E. Zuccato, Mass balance of emerging contaminants in the water cycle of a highly urbanized and industrialized area of Italy, Water Res. 131 (2018) 287–298, https://doi.org/10.1016/j.watres.2017.12.047.

[4] N. Han Tran, M. Reinhard, K. Yew-Hoong Gin, Occurrence and fate of emerging contaminants in municipal wastewater treatment plants from different geographical regions-a review, Water Res. 133 (2018) 182–207, https://doi.org/10.1016/j.watres.2017.12.029.

[5] M.R. Siddiqui, Z.A. AlOthman, N. Rahman, Analytical techniques in pharmaceutical analysis: a review, Arab. J. Chem. 10 (2017) S1409–S1421, https://doi.org/10.1016/j.arabjc.2013.04.016.

[6] J. Luypaert, D.L. Massart, Y. Vander Heyden, Near-infrared spectroscopy applications in pharmaceutical analysis, Talanta (2007), https://doi.org/10.1016/j.talanta.2006.12.023.

[7] G.J. Smolders, J. van de Meij, M.C.M. van Loosdrecht, J.J. Heijnen, Model of the anaerobic metabolism of the biological phosphorus removal process: stoichiometry and pH influence, Biotechnol. Bioeng. 43 (1994) 461–470.

[8] Z. Genisheva, C. Quintelas, D.P. Mesquita, E.C. Ferreira, J.M. Oliveira, A.L. Amaral, New PLS analysis approach to wine volatile compounds characterization by near infrared spectroscopy (NIR), Food Chem. 246 (2018) 172–178, https://doi.org/10.1016/j.foodchem.2017.11.015.

[9] J.W. Einax, H.W. Zwanziger, S. Geiss, Chemometrics in Environmental Analysis, Wiley-VCH, Weinheim, Germany, 1997.

[10] T. Fearn, Assessing calibrations: SEP, RPD, RER and R2, NIR News 13 (2002) 12–14.

[11] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies, J. Pharm. Biomed. Anal. 44 (2007) 683–700, https://doi.org/10.1016/j.jpba.2007.03.023.

[12] S.S. Rosa, P.A. Barata, J.M. Martins, J.C. Menezes, Development and validation of a method for active drug identification and content determination of ranitidine in pharmaceutical products using near-infrared reflectance spectroscopy: a parametric release approach, Talanta 75 (2008) 725–733, https://doi.org/10.1016/j.talanta.2007.12.008.

[13] L. Alvarenga, D. Ferreira, D. Altekruse, J.C. Menezes, D. Lochmann, Tablet identification using near-infrared spectroscopy (NIRS) for pharmaceutical quality control, J. Pharm. Biomed. Anal. 48 (2008) 62–69, https://doi.org/10.1016/j.jpba.2008.05.007.

[14] M. Blanco, R. Cueva-Mestanza, A. Peguero, Controlling individual steps in the production process of paracetamol tablets by use of NIR spectroscopy, J. Pharm. Biomed. Anal. (2010), https://doi.org/10.1016/j.jpba.2009.09.038.

[15] E. Ziémons, J. Mantanus, P. Lebrun, E. Rozet, B. Evrard, P. Hubert, Acetaminophen determination in low-dose pharmaceutical syrup by NIR spectroscopy, J. Pharm. Biomed. Anal. 53 (2010) 510–516, https://doi.org/10.1016/j.jpba.2010.06.003.

[16] L.M.M. Lê, E. Caudron, A. Baillet-Guffroy, L. Eveleigh, Non-invasive quantification of 5 fluorouracil and gemcitabine in aqueous matrix by direct measurement through glass vials using near-infrared spectroscopy, Talanta 119 (2014) 361–366, https://doi.org/10.1016/j.talanta.2013.10.060.

[17] M.A.M. Silva, M.H. Ferreira, J.W.B. Braga, M.M. Sena, Development and analytical validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy, Talanta 89 (2012) 342–351, https://doi.org/10.1016/j.talanta.2011.12.039.

[18] L.B. Schlegel, M. Schubert-Zsilavecz, M. Abdel-Tawab, Quantification of active ingredients in semi-solid pharmaceutical formulations by near infrared spectroscopy, J. Pharm. Biomed. Anal. 142 (2017) 178–189, https://doi.org/10.1016/j.jpba.2017.04.048.

[19] D.P. Mesquita, C. Quintelas, A.L. Amaral, E.C. Ferreira, Monitoring biological wastewater treatment processes: recent advances in spectroscopy applications, Rev. Environ. Sci. Biotechnol. 16 (2017) 395–424, https://doi.org/10.1007/s11157-017-9439-9.

[20] K. Michel, B. Bureau, C. Boussard-Plédel, T. Jouan, J.L. Adam, K. Staubmann, T. Baumann, Monitoring of pollutant in waste water by infrared spectroscopy using chalcogenide glass optical fibers, Sens. Actuators B Chem. 101 (2004) 252–259, https://doi.org/10.1016/j.snb.2004.03.014.

[21] R. Lu, B. Mizaikoff, W.-W. Li, C. Qian, A. Katzir, Y. Raichlin, G.-P. Sheng, H.-Q. Yu, Determination of chlorinated hydrocarbons in water using highly sensitive mid-infrared sensor technology, Sci. Rep. 3 (1–6) (2013) 2525, https://doi.org/10.1038/srep02525.

[22] C. Gu, Q. Tang, B. Xiang, J. Xu, Determination of fenitrothion in water by near infrared spectroscopy and chemometric analysis, Anal. Lett. 48 (2015) 1481–1493, https://doi.org/10.1080/00032719.2014.986680.