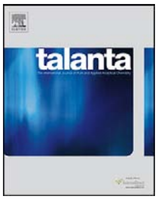
Talanta 82 (2010) 575–581

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta



Monitoring and classification of wastewater quality using supervised pattern recognition techniques and deterministic resolution of molecular absorption spectra based on multiwavelength UV spectra deconvolution

Christos M. Tsoumanis, Dimosthenis L. Giokas, Athanasios G. Vlessidis*

Laboratory of Analytical Chemistry, Department of Chemistry, University of Ioannina, 45110 Ioannina, Greece

ARTICLE INFO

Article history:
Received 18 December 2009
Received in revised form 5 May 2010
Accepted 10 May 2010
Available online 19 May 2010

Keywords:
Classification
In situ monitoring
Pattern recognition techniques
UV spectra deconvolution
Wastewater

ABSTRACT

A field flow approach for the in situ monitoring of wastewater quality is developed and assessed in this work, based on a combination of methods employing deconvolution of molecular absorption spectra and in situ/on-line analysis of wastewater effluent of various origin. The approach involves in situ immersion probes to monitor basic physicochemical parameters followed by UV spectrum deconvolution in order to provide a rapid estimate of organic matter, suspended solids and nitrate and on-line analysis of phosphates in a fully automated setup. The collected data are then treated with a series of supervised pattern recognition techniques in order to classify wastewater effluent according to their origin in three major categories namely municipal, industrial and hospital. The results suggest that the method affords a good approximation of realistic concentrations, as determined by reference methods, while it affords a good classification among various wastewater effluents of different origin. In that manner, the method enables a rapid inference of treated wastewater quality and a robust assessment of treatment process state, especially with regards to violations of effluent quality parameters.

1. Introduction

Wastewater quality monitoring is an essential pre-requisite in order to ensure compliance with environmental legislation regarding treated effluent discharge in natural waters and control the process efficiency of wastewater treatment facilities. For many years, the most popular methods for wastewater quality assessment have been based on a succession of procedures involving manual sampling, storage and laboratory analysis, which provide only a snapshot of the effluent quality and do not permit a continues feedback on wastewater quality characteristics.

To satisfy the need for incessant monitoring, many automated procedures have been proposed which enable the on-line and in situ determination of most wastewater quality key parameters [1–4]. Among these, spectrophotometric methods are the most popular due to their versatility and flexibility, enabling them to be adapted to various procedures with minimal maintenance and reagent consumption. In most occasions cited inside the literature, indirect spectrophotometric measurements are performed deploying the introduction of reagents for the development of an analyte specific colour, thus obtain quantitative information on the presence or absence of the target analyte(s) based on the absorbance measurement in a given wavelength [5,6]. In a simpler approach, a univariate correlation between selected wavelengths and quality parameters has been used, based on the assumption that wastewater effluent have a defined peak of maximum absorbance that always occurs at the same wavelength [7,8]. This assumption practically relies on the fact that, in most occasions, effluent composition does not vary significantly and that many organic compounds contains chromophores, usually an aromatic ring or conjugated double bonds, that exhibit specific maxima in the UV region (e.g. detergents show a shoulder at 225 nm, phenolics between 260 and 320 nm, etc.) while several inorganic compounds also absorb light in certain wavelength regions (e.g. nitrate around 205–210 nm) [9,10]. However, this approach requires continues calibration to alleviate errors induced by variable wastewater composition [9].

An alternative approach to wastewater quality monitoring is the multiwavelenght exploitation of the UV spectra by interpretating the UV profile of the sample in order to identify relevant patterns among the data either after appropriate separation of sample matrix [11] or by mathematically resolving UV spectra to its basic components [12–15]. This approach alleviates the drawbacks of single-wavelength procedures providing a more representative description of sample matrix composition especially in effluents of variable composition minimizing interferences (diffusion phenomena, matrix composition, etc.) [16]. Statistical resolution techniques have been brought into this effort, in order to separate identical

---

* Corresponding author. Tel.: +30 2651008781: fax: +30 2651008401.
E-mail address: avlessid@cc.uoi.gr (A.G. Vlessidis).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.05.009

spectra from spectra of low purity, thus enable the determination of specific compounds or classify them into categories depending on their source [17,18].

Despite the progress made, most of the articles cited inside the literature regarding the determination of sample constituents using mathematical resolution of sample spectra, deal with batch pre-treatment or manual handling procedures [19–21] a situation which does not enable their application in the real-time monitoring of wastewater quality. Moreover, the majority of articles describing the resolution of UV spectra as a tool to characterize the quality of wastewater have been focused on municipal wastewater and their subsequent discharge in water bodies [7–10,15–18] or to various types of industrial wastewater [22–26]. On the other hand, studies describing the potential of UV spectrum for characterization and quality assessment of non-municipal wastewater of organic origin are scant. To this end, the aim of this work is twofold. The first is to evolve a simple and automated measurement system that can combine UV spectra deconvolution with established analytical methods as a complementary approach to the on-line/in situ monitoring of conventional wastewater parameters. The second is to combine these data in appropriate chemometric algorithms that can be used to classify samples of organic origin based on their source to three major categories, namely municipal, industrial and hospital. The last feature could be of interest in regular monitoring assessment of wastewater quality in order to assess potential violations of effluent quality or even identify illegal discharges in open channels, used to transfer effluent from various facilities.

2. Experimental

2.1. Reagents

Stock standard solutions of K_2HPO_4 (100 mg L^{-1} as P–PO_4^{3-}), KNO_3 (100 mg L^{-1} as NO_3^-), $(NH_4)_6Mo_7O_{24}$–$4H_2O$ (0.02 mol L^{-1}) and potassium antimony (III) oxide tartrate hemihydrate (0.03 mol L^{-1}) were prepared on a weekly basis by dissolving the required amounts in distilled water. Fresh solutions of ascorbic acid (1.5%, w/v) were prepared daily in distilled water. Ferrous ammonium sulphate and potassium dichromate were obtained from Mallinckrodt Baker (Phillipsburg, USA). Membrane filters (0.45 μm) for the determination of total suspended solids were purchased from Whatman (Piscataway, NJ, USA).

2.2. Instrumentation

A laboratory scale configuration, simulating the real-time operational monitoring of wastewater effluent was devised. The overall setup of the manifold employed for the measurements is schematically depicted in Fig. 1. A fixed volume dilution conduit was used for each wastewater effluent type to avoid absorbance values above 2.0, for the reasons discussed below in more detail. On-line UV spectrum acquisition over the range of 200–700 nm was performed with a Shimadzu SPD-M6A UV/diode-array detector operating under the Class M10A Software (version 1.20). A Secomam S-1000 flow spectrophotometer employed with a silica flow cell of 80 μL and 10 mm path length (Hellma) was used for recording UV spectra at a fixed wavelength for the analytical detection of orthophosphate using the manifold setup described in our previous work [27]. Batch UV–vis absorption measurements for the verification analysis of COD and nitrate were performed with matched quartz cells of 1 cm path length in a Jenway 6405 UV/Vis spectrophotometer (Essex, UK). A pH meter, WTW-552 model glass-electrode was employed for the pH adjustment of the solutions. Conductivity and temperature measurements were conducted with a Crison (Barcelona, Spain) conductivity meter (model CM 524). Turbidity was recorded with a TB502-IR Online Turbidity Analyzer (Global Water Instrumentation, Inc., USA) and ammonium by NH4D-sc ammonium sensor, available from HACH.

Clogging of the filter (MF Millipore, mixed cellulose ester filter, 0.45 μm) employed to remove suspended solids prior to single-wavelength UV analysis for orthophosphate was empirically evidenced by decreased flow rate with time. However, the low content of the samples in suspended solids and the necessary dilution significantly increases the life-time of the filter. Maintenance operations of pH, conductivity, turbidity and ammonia analyzers were performed according to the instructions of the manufacturers.

2.3. Samples

Spot effluent wastewater samples, selected to represent samples of organic origin, were collected during the period from December 2008 to July 2009 from various wastewater treatment facilities in the wider region of Ioannina city (NW Greece). All treatment facilities are designed at least for secondary treatment with various modifications. Municipal wastewaters were collected from the effluent stream of the municipal wastewater treatment plant (WWTP) of Ioannina city (designed for full nitrification/denitrification to treat 35,000 m^3 day^{-1} of municipal wastewater with an average load of 280 mg L^{-1} BOD and 350 mg L^{-1} SS) and from the outlet of a draining ditch receiving the treated effluent from the WWTP of Ioannina city and a diary product industry as well as wastewater from local settlements not connected to the sewage network. Industrial wastewater effluent were collected from two local industries with wastewater of organic origin; the first, producing various milk-based products (designed to treat 2500 m^3 day^{-1} of wastewater with an average load of 2000 mg L^{-1} BOD and 700 mg L^{-1} SS) and the second a poultry slaughter-house (designed to treat 300 m^3 day^{-1} of wastewater with an average load of 1500 mg L^{-1} BOD and 1100 mg L^{-1} SS). Both industrial treatment plants are designed to treat raw wastewater in a two-stage process involving high rate aeration followed by extended aeration. Dietary effluent is directed to the influent stream of the municipal wastewater treatment plant for further treatment while slaughter-house effluent is directed for underground discharge. Hospital wastewater effluent were taken from the effluent streams of the wastewater treatment facilities of the largest regional hospitals (designed to treat approx. 400 m^3 d^{-1} of wastewater with an average load of 340–600 mg L^{-1} BOD and 300–400 mg L^{-1} SS). Both effluent streams are directed to the inlet of the municipal WWTP for additional treatment.

All samples were immediately analyzed to avoid alteration of the sample's matrix initial composition. A portion of the samples was used for batch analysis while the remaining sample was placed in a beaker, which was used as a pilot-scale tank for on-line sampling. Approximately, 150 samples were collected during the 8-month sampling period form each wastewater treatment plant and their UV spectrum was monitored in the 200–700 nm region to determine the working absorbance profile of the wastewater samples. The obtained spectra showed that absorbance in the visible range did not provide any significant information with regards to sample discrimination. To this end, a working range of 200–400 nm was decided.

2.4. Samples for cross validation

A series of samples from the effluent streams of all treatment facilities examined within the confines of the present study were taken 5 months after the initial sampling survey and re-assessed for the examined quality parameters. The results were used to cross-

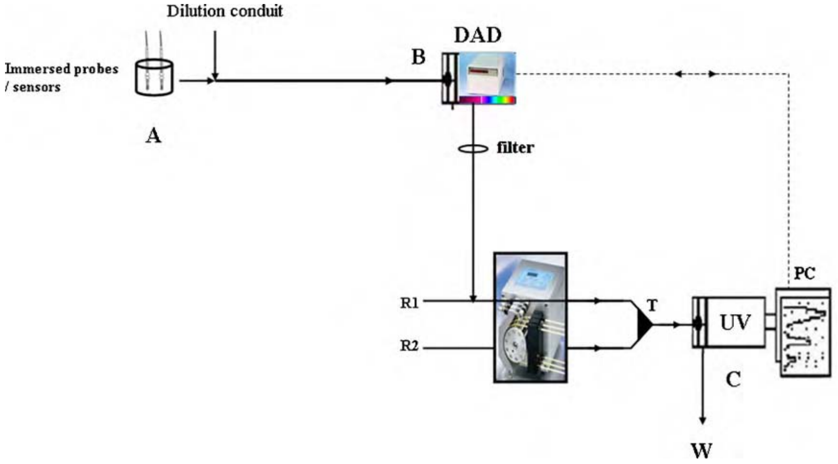


Fig. 1. Manifold setup employed for the measurements. Detection steps: A (pH, conductivity, turbidity, temperature and ammonium); B (COD, TSS, NO_3) and C (PO_4^{3-}); R1, molybdenum complex solution (incl. antimony tartrate); R2, ascorbic acid/glycerol solution; W, wastes.

validate the efficiency of the discriminant model developed for wastewater effluent classification.

2.5. Analytical procedures

The analytical determination of wastewater quality parameters was performed using standard procedures [28,29] on filtered samples. Briefly, COD was determined in the filtered samples titrimetrically with FAS (ferrous ammonium sulfate) after oxidation with the dichromate reflux method (RSD (%) = 13.9, limit of detection/LOD = 7.1 mg L^{-1}). Orthophosphate was determined spectrophotometrically using a field flow system based on the molybdenum blue method (RSD (%) = 4.3/LOD = 0.12 mg L^{-1}) while nitrate was determined spectrophotometrically after reduction with Cd by the Griess assay (RSD (%) = 6.5/LOD = 0.1 mg L^{-1}). Suspended solids were determined after filtration [29]. All results are expressed in mg L^{-1} unless otherwise stated.

2.6. Multiwavelength UV spectra devonvolution (UVSD)

The method behind UV spectra deconvolution is based on a procedure of matrix algebra, relying on the assumption that an unknown UV spectrum can be re-constituted by a linear combination of a small number (p) of well-defined and representative reference spectra (REF_i), using a deterministic or semi-deterministic method [15,31]. The coefficient contribution (α_i) of each reference spectrum in the linear combination can be calculated through the resolution of a matricial system, by either the Gaussian or the Rank approaches [12,13,32] assuming that the contribution coefficients of the reference spectra are constant along the wavelength range. In other words, the deconvolution method calculates the best fit between measured and restituted spectra, with a least-squared method and the associated quadratic error for a procedure check [10], using the following relation, established for each wavelength:

$$S_s(\lambda) = \sum_{i=1}^{n} \alpha_i \times \mathrm{REF}_i(\lambda) \pm r$$

where $S_s(\lambda)$ and $\mathrm{REF}_i(\lambda)$ are the absorbance of the sample and the ith REF at each wavelength λ, respectively, α_i is the contribution coefficient of the ith reference spectra, n is the number of reference spectra and r is the error of restitution of each wavelength. The total error estimation is computed from the equation:

$$\mathrm{Error}_{\mathrm{Total}} = \sum_{i=1}^{n} \mathrm{ES}(\lambda_i)^2$$

where $\mathrm{ES}(\lambda)$ is the error at wavelength λ, defined as the difference between measured and restituted absorbencies.

All additive parameters can then be estimated using with the same linear combination by replacing the reference spectra by the parameter value [25,33], through the relation:

$$P_s(\lambda) = \sum_{i=1}^{n} \alpha_i \times \mathrm{PAR}_i(\lambda) \pm r_n$$

where $P_s(\lambda)$ and $\mathrm{PAR}_i(\lambda)$ are respectively the values of the parameter associated with the unknown spectrum $S_s(\lambda)$ and with the reference spectra $\mathrm{REF}_1, \ldots, \mathrm{REF}_n$, and r_n is the error on the computation of the parameter value. Quality parameters for spectrum restitution are the quadratic error (standard error of the estimate) and the multiple regression coefficients (determination coefficient $-R^2$). Calculations were performed in commercial spreadsheet software.

2.7. Pattern recognition techniques

Preliminary assessment of the experimental data was conducted with the aid of Kolmogorov–Smirnov test in order to evaluate departures from normality. Principal Components Analysis (PCA) was used to afford data reduction with minimal loss of information. Redundant variables that did not contribute to the explanation of the observed variance were extracted by estimating the extraction communalities which are estimates of the variance in each variable accounted for by the factors (or components) in the factor solution. According to this method, variables with low extraction communalities should be considered for removal leaving those variables that hold a high value in the initial data matrix [31,32].

Discriminant analysis (DA) was used to classify and discriminate wastewater effluent based on the composition of its basic components determined by UVSD and conventional methods. Three groups representing municipal, industrial and hospital wastewater and the discriminant function for each group was calculated as:

$$f(G_i) = k_i + \sum_{j=1}^{n} \omega_{ij} p_{ij}$$

where i is the number of groups (G), k_i is the constant inherent to each group, n is the number of parameters used to classify a set of data into a given group, ω_j is the weight coefficient, assigned by DA to a given selected parameter (p_j).

Statistical analysis was performed with the aid of SPSS 14.0 automated modules.

3. Results and discussion

3.1. Acquisition of spectra

The deconvolution method by Thomas et al. [15,33] was used, taking a base of five-reference spectra for municipal wastewater; suspended matter, colloidal matter, dissolved matter, detergents and nitrate. These spectra have been shown to provide a representative description of municipal wastewater effluent [9,14,15,30,33]. Since all samples were effluent wastewater with at least secondary treatment, the same base of reference spectra as in the original reference was used [34], except for organic matter because the UV absorbance of each wastewater sample is influenced by organic matrix composition [17], therefore a separate set of reference spectra for dissolved organic matter were used for each wastewater in order to correctly restitute the UV signal (local calibration). These reference spectra were selected from a data set of approx. 60 samples to correspond to the specific wastewater class and were the difference between the raw absorbance (after appropriate dilution) and the respective spectra acquired after sample filtration through 0.025 μm filter. The spectra were recorded in the range of 210–350 nm to eliminate the absorbance due to the presence of nitrates. Differences in the absorbance pattern of each reference spectra were validated by registering the signal at equal parameter levels (determined by conventional methods), thus giving a qualitative indication of the dissolved matter composition in each sample. Calculations and selection of reference spectra were performed in spreadsheet software using the procedure described previously [12,13,35].

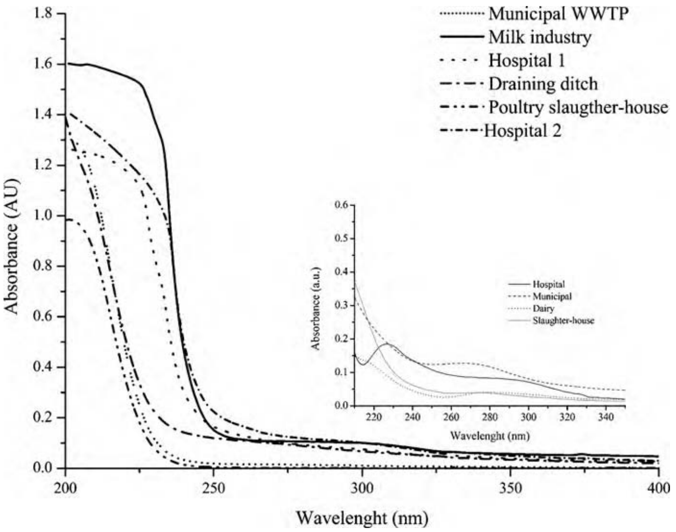


Fig. 2. UV spectra of different wastewater effluent after dilution. Inset graph: typical reference spectra for organic matter.

The UV spectra of Fig. 2 represent real samples measured after appropriate dilution to ensure absorbance below 2.0 (a.u.) thus avoid spectra saturation [33,36]. The dilution factor was separately determined in each occasion after a preliminary measurement of the sample absorbance. As we can observe, all samples exhibit a peak at 205–210 nm which is attributed to nitrate [9,17,37]. In most samples a second characteristic peak is observed at approx. 230–240 and 280–300 nm which is most probably related to the use of detergents or even compounds containing hydroacids and esters [10]. The residual absorption after 300 nm is most probably related to the presence of suspended solids [17,34,36].

3.2. Deconvolution basis calibration

The efficiency of the calibration step in estimating the different parameters was evaluated through least squares linear regression between measured (by conventional methods) and estimated (by UVSD) values. The calibration of the deconvolution bases was

Table 1
Concentration of COD, nitrate and total suspended solids determined by standard methods and comparison with UVSD modeling results for various wastewater effluent samples.

Parameter	Sample	Estimated values (mg L^{-1})[a]	Experimentally measured values (mg L^{-1})[a]	Recovery (%)	Standard error of prediction (%)[b]	Relative error (%)[c]	Determination coefficient (R^2)
NO_3^-	All	12–452	8.8–447.5	104.8 ± 15.8	3.9		0.9747
	WWTP	25.0	24.0	104.2	19.2	4.2	0.9188
	Ditch	17.9	14.8	120.9	13.4	20.9	0.9471
	Milk industry	255.0	259.8	98.2	10.2	−1.8	0.9236
	Slaughter-house	243.0	260.7	93.2	12.3	−6.8	0.9887
	Hospital 1	237.0	227.0	104.4	6.5	4.4	0.9008
	Hospital 2	293.0	266.9	109.8	3.6	9.8	0.9188
TSS	All	11–17	8.5–16	96.9 ± 9.5	5.2		0.8978
	WWTP	15.1	13.7	110.2	8.0	10.2	0.8941
	Ditch	14.2	12.1	117.4	7.1	17.4	0.9278
	Milk industry	16.7	14.0	119.3	13.2	19.3	0.8680
	Slaughter-house	16.8	14.0	120.0	6.1	20.0	0.9032
	Hospital 1	11.2	12.0	93.3	8.4	−6.7	0.9112
	Hospital 2	11.8	14.0	84.3	11.2	−15.7	0.924
COD	All	19.4–153	15–156	105.6 ± 10.3	7.7		0.9481
	WWTP	39.8	38.0	104.7	10.3	4.7	0.9622
	Ditch	24.1	24.3	99.2	8.5	−0.8	0.916
	Milk industry	57.3	50.0	114.6	5.3	14.6	0.8895
	Slaughter-house	88.1	80.0	110.1	4.8	10.1	0.9695
	Hospital 1	51.0	42.0	121.4	6.5	21.4	0.9589
	Hospital 2	64.9	71.0	91.4	7.7	−8.6	0.9481

[a] Mean of 25 samples.

[b] Based on calculated recoveries.

[c] Based on estimated and measured values.

performed using 35 samples from each wastewater effluent, chosen randomly to represent the overall sampling period with a relatively wide distribution of parameter values which permits better calibration of the deconvolution basis [38]. The output of the linear regression between the conventionally measured concentrations of the examined parameters and those estimated by UVSD revealed that the method affords good prediction of nitrate yielding coefficients of determination between 0.9561 and 0.9785, a fair estimation of TSS with determination coefficients in the range of 0.8756–0.9150 and a satisfactory prediction of COD values with coefficients of determination between 0.8986 and 0.9528. The efficiency of the UVSD method to predict these variables was decreasing from municipal to industrial and finally to hospital wastewater, especially for COD values. These findings came as no surprise because COD does not represent all dissolved organic matter is the samples but an estimate of oxidizable matter present in the medium. Furthermore, the nature of suspended solids and the degree of treatment in each process varies, which justifies the variation of solids effluent composition in each wastewater medium.

3.3. *Method validation*

Twenty-five effluent wastewater samples from each treatment facility were used to formally validate the model. The validation results given in Table 1 show that the determination of nitrate, suspended solids and COD content of various wastewater samples can be estimated satisfactory by UV deconvolution. The least square linear regression analysis of the data (Fig. S1-Supplementary material) revealed that the predicted values are in good agreement with the measured values, yielding satisfactory determination coefficients and quadratic errors below the limit value of 0.2, which is higher than those used in previous works [30,33] in order to enable the deconvolution procedure to be applied to all samples using a single computational approach. The detection limits, defined as the lowest concentration determined with the UVSD approach in all samples, were 5 mg L^{-1} for NO_3^-, 3.5 mg L^{-1} for TSS and 9 mg L^{-1} for COD.

3.4. *Wastewater classification*

In order to enable wastewater classification, supervised pattern recognition techniques were employed using a set of common wastewater quality variables, namely pH, temperature, conductivity, turbidity, ammonium, phosphates, COD, total suspended solids and nitrate. All variables selected for that purpose can be obtained online via continues monitoring with the aid of the automated manifold setup propose in this work.

To pursue wastewater classification based on the above parameters, a data set consisting of 600 numerical values representing the aforementioned 9 variables retrieved from 25 wastewater samples of various origin, were assayed. Since sample size was greater than 50, departures from normality were assessed with the aid of Kolmogorov–Smirnov test suggesting that the raw numerical data are not normally distributed. Therefore, the logarithmic transformation of the raw data was employed [32].

Before proceeding to data discrimination with DA, it was decided to test if any variables are redundant thus they can be removed form the subsequent analysis. The extraction communal-

Table 2
Standard discriminant analysis (DA) of effluent wastewater quality parameters.

Statistical parameters	Standard DA mode	
	Function 1	Function 2
(a) Classification functions and statistical parameters		
Chi-square	79.00	13.71
Eigenvalue	4.45	0.43
Wilk's l	0.128	0.700
Canonical *R*	0.9037	0.5473
% Explained variance	91.2	8.8
P values	0.00	0.017
Degrees of freedom (*df*)	12	5

Groups	Function at group centroids	
	Function 1	Function 2
Municipal	2.407	0.138
Hospital	−1.057	−1.117
Industrial	−2.047	0.543

Experimental parameters	Classification functions		
	Municipal	Hospital	Industrial
NO_3^-	54.10	55.69	51.21
TSS	152.27	152.78	153.87
COD	−7.62	−5.37	0.74
pH	2966.86	2906.89	2931.06
PO_4^{3-}	−51.19	−44.69	−40.49
Conductivity	−42.94	−39.16	−39.05
Constant	−1375.90	−1342.97	−1370.68

	Classification matrix			
		Groups assigned by DA		
	% correct cases assigned by DA	Municipal	Hospital	Industrial
(b) Classification matrix				
Municipal	100.0	18	0	0
Hospital	100.0	0	10	0
Industrial	87.5	1	1	14
Total	95.5	19	11	14

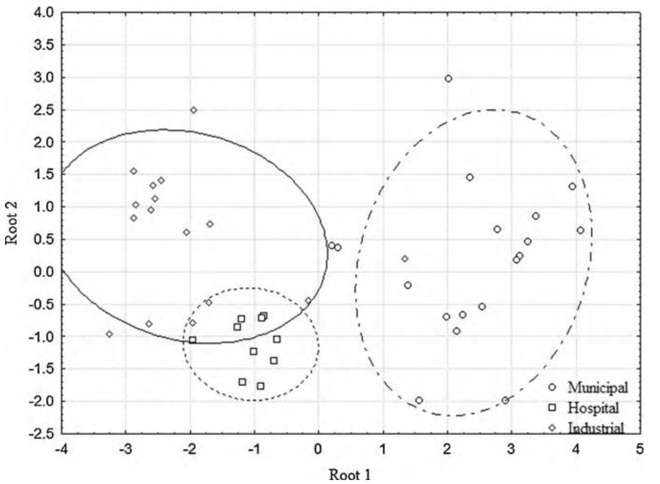


Fig. 3. Canonical score graph of the discriminant functions for the general classification of wastewater effluent.

ities advice towards the removal of ammonium, temperature and turbitity since these variables do not contribute to the interpretation of the variance of the data matrix. Therefore, these three variables were latent and were not further considered in the DA modules in order to reduce uncertainty. Discriminant analysis was therefore performed using six out of the original nine variables inserted in the model.

Standard, forward and backward DA was run in the logarithmically transformed data in order to examine which mode affords the best discrimination efficiency. The best results were obtained with standard DA assigning 95.5% of the cases correctly (Table 2). The 2D canonical score graph of Fig. 3 and the results of Table 2 show that standard DA with the reduced variables enables a good discrimination among wastewater effluent samples classifying municipal, hospital and industrial wastewater in well-defined groups. However, the classification of industrial effluent were the poorest among the wastewater samples reflecting the high variability of the wastewater quality (wide scattering in the scatter plot of Fig. 3) due to differences in the influent wastewater as determined by the variation in the production line. Even in that case though, more than 87.5% of the cases were correctly assigned.

The good discrimination of the wastewater effluent based on three general categories encouraged us to investigate the potential of differentiating among all samples in a single standard DA module. The results were also satisfactory classifying 79.5% of the cases correctly. According to the canonical score graphs of Fig. 4 root 1 discriminates between municipal, hospital and dairy wastewater while root 2 between the two types of industrial wastewater. Finally root 3, aids the discrimination of hospital with slaughterhouse wastewater which are not discriminated by the first root. Therefore, standard DA using these parameters can be further deployed to provide more detailed information on the kind of wastewater even within similar wastewater samples. However, these results should be interpreted with caution since the samples assessed in this work represent effluent quality from at least secondary treatment. Therefore, some over-lapping is reasonable since effluent must meet certain criteria depending on the subsequent discharge (draining ditch, underground discharge, etc.).

3.5. Cross validation

Samples collected 5 months after the initial model calibration were used to validate its applicability in the classification of wastewater quality. All municipal wastewater samples were correctly classified while hospital wastewater was correctly assigned by DA at 80%, with two samples being wrongly assigned to municipal. On the other hand, 4 out of 10 industrial wastewater samples were misclassified. Considering the fact that spot samples were employed for both calibration and validation data and the different periods of sampling (reflecting changes in the quality of wastewater) the overall performance is deemed satisfactory for municipal and hospital wastewater but mediocre for industrial wastewater. However, classification techniques are by definition data driven, therefore appropriate training sets are necessary to select variables that lead to a meaningful feature detection. In this context, calibration on a regular basis should be conducted.

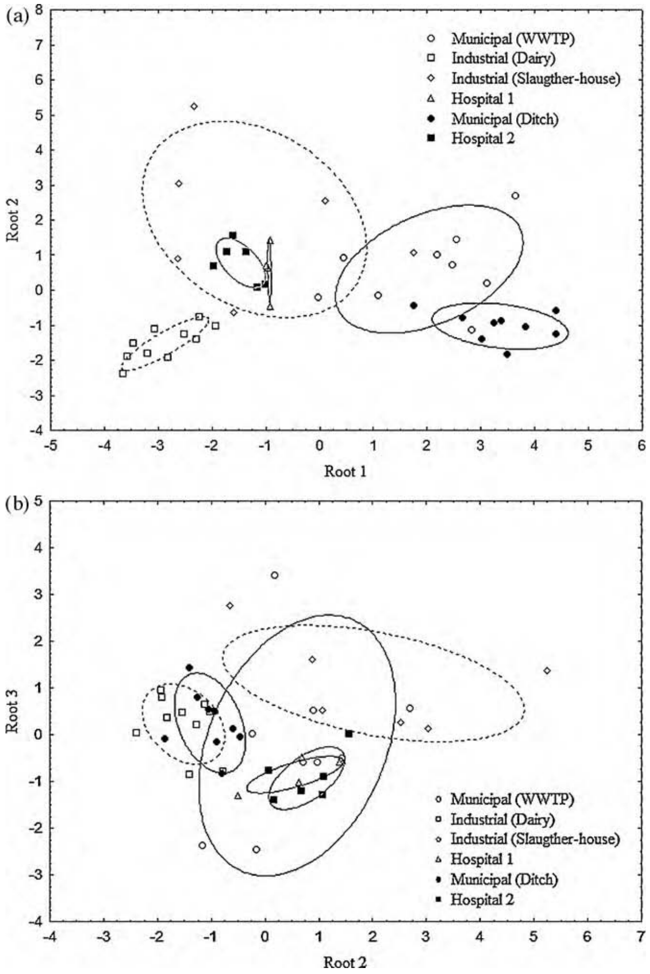


Fig. 4. Canonical score graph of the discriminant functions for the classification of wastewater effluent according to their origin.

4. Conclusions

In this work an in situ setup for wastewater quality monitoring was devised and assessed in various wastewater effluent of municipal, industrial and hospital origin. The method initially employs a set of sensing probes for the determination of physicochemical parameters and ammonium, followed by UV spectra deconvolution with a diode array detector for the online determination of COD, suspended solids and nitrate and a simple UV-based flow injection manifold for the assessment of phosphates. Thus, the need for off-line analysis is alleviated enabling the control of most effluent quality parameters in real time with minimum experimental effort and reagents consumption.

Method application in various wastewater samples suggests that UV spectra deconvolution is a powerful tool that can provide a rapid assessment of wastewater quality. Combined with pattern recognition techniques it can also be used to provide a clear discrimination among municipal, industrial and hospital wastewater affording a fair classification even among wastewater samples of the same origin. The last feature is important in the protection of water quality since it can be a valuable tool in the identification of illegal discharges in open channels used to transfer treated effluent. At the same time the method can be deployed to the real-time control and monitoring of various treatment processes, especially with regards to violations in effluent quality parameters and treatment process failure.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.talanta.2010.05.009.

References

[1] N.J. Blundell, P.J. Worsfold, H. Casey, S. Smith, Environ. Int. 21 (1995) 205–209.
[2] M. Miró, J.M. Estela, V. Cerdà, Talanta 60 (2003) 867–886.
[3] M. Miró, J.M. Estela, V. Cerdà, Talanta 62 (2004) 1–15.
[4] M. Miró, J.M. Estela, V. Cerdà, Talanta 63 (2004) 201–223.
[5] C.Z. Katsaounos, D.L. Giokas, A.G. Vlessidis, E.K. Paleologos, M.I. Karayannis, Sci. Total Environ. 305 (2003) 157–167.
[6] L. Monser, S. Sadok, G.M. Greenway, I. Shah, R.F. Uglow, Talanta 57 (2002) 511–518.
[7] N. Matsche, K. Stumwohrer, Water Sci. Technol. 33 (1996) 211–218.
[8] J. Wu, M.N. Pons, O. Potier, Water Sci. Technol. 53 (2006) 449–456.
[9] H. Khorassani, G. Theraulaz, O. Thomas, Acta Hydrochim. Hydrobiol. 26 (1998) 296–299.
[10] A.G. Vlessidis, M.E. Kotti, D.L. Giokas, N.P. Evmiridis, Water Environ. J. 21 (2007) 157–164.
[11] O. Thomas, N. Mazas, S. Gallot, B. Clement, Int. J. Environ. Anal. Chem. 52 (1993) 179–188.
[12] S. Gallot, O. Thomas, Int. J. Environ. Anal. Chem. 52 (1993) 149–158.
[13] S. Gallot, O. Thomas, Fresenius J. Anal. Chem. 346 (1993) 976–983.
[14] S. Suryani, F. Theraulaz, O. Thomas, Trends Anal. Chem. 14 (1995) 457–463.
[15] O. Thomas, F. Theraulaz, C. Agnel, S. Suryani, Environ. Technol. 17 (1996) 251–261.
[16] B. Roig, O. Thomas, Manage. Environ. Qual. Int. J. 14 (2003) 398–404.
[17] L. Rieger, G. Langergraber, M. Thomann, N. Fleischmann, H. Siegrist, Water Sci. Technol. 50 (2004) 143–152.
[18] G. Langergraber, N. Fleischmann, F. Hofstadter, Water Sci. Technol. 47 (2003) 63–71.
[19] M.C. Sarraguça, A. Paulo, M.M. Alves, A.M.A. Dias, J.A. Lopes, E.C. Ferreira, Anal. Bioanal. Chem. 395 (2009) 1159–1166.
[20] A. Vargas, G. Buitron, Water Sci. Technol. 53 (2006) 457–463.
[21] L.R. Chevalier, C.N. Irwin, J.N. Craddock, Adv. Environ. Res. 6 (2002) 369–375.
[22] H. El Khorassani, P. Trebuchon, H. Bitar, O. Thomas, Water Sci. Technol. 39 (1999) 77–82.
[23] F. Pouly, E. Touraud, J.-F. Buisson, O. Thomas, Talanta 50 (1999) 737–742.
[24] B. Roig, C. Gonzalez, O. Thomas, Talanta 50 (1999) 751–758.
[25] E. Dupuit, M.F. Pouet, O. Thomas, J. Bourgois, Environ. Model. Soft 22 (2007) 1153–1163.
[26] N.D. Lourenço, C.L. Chaves, J.M. Novais, J.C. Menezes, H.M. Pinheiro, D. Diniz, Chemosphere 65 (2006) 786–791.
[27] I.A. Tsoulfanidis, G.Z. Tsogas, D.L. Giokas, A.G. Vlessidis, Microchim. Acta 160 (2008) 461–469.
[28] M.E. Kotti, A.G. Vlessidis, N.C. Thanasoulias, N.P. Evmiridis, Water Res. Manage. 19 (2005) 77–94.
[29] APHA, AWWA and WEF, Standard methods for examinations of water and wastewater (1987), Washington DC, USA.
[30] O. Thomas, F. Theraulaz, M. Domeizel, C. Massiani, Environ. Technol. 14 (1993) 1187–1192.
[31] C.Z. Katsaounos, D.L. Giokas, I.D. Leonardos, M.I. Karayannis, Water Res. 41 (2007) 406–418.
[32] C.Z. Katsaounos, D.L. Giokas, A.G. Vlessidis, M.I. Karayannis, Desalination 213 (2007) 311–333.
[33] A. Escalas, M. Droguet, J.M. Guadayol, J. Caixach, Water Res. 37 (2003) 2627–2635.
[34] European Commission, BCR Information Chemical Analysis-Wastewater Quality Monitor, Project MAT1 CT93-0008, Final Report (EUR 16914 EN).
[35] J.E. Billo, Excel for Chemists: A comprehensive guide, 2nd ed., Wiley, 2001.
[36] O. Thomas, H.E.l. Khorassani, E. Touraud, H. Bitar, Talanta 50 (1999) 743–749.
[37] P.H. Nam, B. Alejandra, H. Frederic, B. Didier, S. Olivier, P. Andre, Talanta 76 (2008) 936–940.
[38] M. Hassouna, F. Theraulaz, C. Massiani, Talanta 71 (2007) 861–867.