Talanta 82 (2010) 1254–1260

Contents lists available at ScienceDirect

Talanta

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

Partial least squares X-ray fluorescence determination of trace elements in sediments from the estuary of Nerbioi-Ibaizabal River

Javier Moros[a], Ainara Gredilla[b], Silvia Fdez-Ortiz de Vallejuelo[b], Alberto de Diego[b], Juan Manuel Madariaga[b], Salvador Garrigues[c], Miguel de la Guardia[c,*]

[a] Department of Analytical Chemistry, Faculty of Sciences, University of Málaga, E29071 Málaga, Spain
[b] Department of Analytical Chemistry, University of the Basque Country, P.O. Box 644, 48080 Bilbao, Basque Country, Spain
[c] Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, 46100 Burjassot, Valencia, Spain

ARTICLE INFO

Article history:
Received 12 February 2010
Received in revised form 3 June 2010
Accepted 28 June 2010
Available online 6 July 2010

Keywords:
Partial least squares
X-ray fluorescence
Estuarine sediments
Metals
Trace elements

ABSTRACT

The feasibility of partial least squares (PLS) regression modeling of X-ray fluorescence (XRF) spectra of estuarine sediments has been evaluated as a tool for rapid trace element content monitoring. Multivariate PLS calibration models were developed to predict the concentration of Al, As, Cd, Co, Cr, Cu, Fe, Mg, Mn, Ni, Pb, Sn, V and Zn in sediments collected from different locations across the estuary of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country). The study was carried out on a set of 116 sediment samples, previously lyophilized and sieved with a particle size lower than 63 μm. Sample reference data were obtained by inductively coupled plasma mass spectrometry. 34 samples were selected for building PLS models through a hierarchical cluster analysis. The remaining 82 samples were used as a test set to validate the models. Results obtained in the present study involved relative root mean square errors of prediction varying from 21%, for the determination of Pb at hundreds $\mu g\,g^{-1}$ level, up to 87%, for Ni determination at little tens $\mu g\,g^{-1}$ level. An average prediction error of ±37% for the 14 elements under study was obtained, being in all cases mean differences between predicted and reference results of the same order than the standard deviation of three replicates from a same sample. Residual predictive deviation values obtained ranged from 1.1 to 3.9.

1. Introduction

Estuaries suffer from a series of contaminant inputs of a variety of industrial and domestic sources including sewage out falls, agricultural and road run-off, atmospheric deposition and pollution from shipping [1]. Their sediments show a great capacity to integrate and accumulate trace elements from low concentrations in the overlying water column [2]. Trace elements are serious pollutants due to their toxicity, persistence, and non-degradability in the environment [3] and, because of that, assessment of sediment contamination from metals must be carefully controlled to prevent their adverse effects on aquatic organisms, including those living in sediments. Unfortunately, there is not a single method able to quantify all the parameters related to sediment pollution, such as heavy metals and organic micropollutants concentration, in order to obtain a diagnostic on their impact on biota. Thus, the use of several complementary methods rather than a single one is becoming a standard practice to obtain independent information of the degree and possible impacts of sediment contamination.

The integrated application of different sediment assessment methods can be proposed in a diagnosis and/or remediation project if the whole cost of the analysis is reasonable. In this sense, less expensive analytical protocols are required to obtain valuable information. However, in recent studies [4] soil and sediment pollution characteristics have been predicted through the modelization of IR data in both, middle [5] and near [6] domains and mineral phases have been identified by Raman spectroscopy [7], differentiating those of natural origin from those of anthropogenic sources. Raman spectroscopy has been also used to evaluate the interaction between calcium, magnesium and arsenic in estuarine media [8].

In spite of the interesting results obtained for trace element prediction in sediments by using vibrational spectroscopic techniques it is clear that vibrational spectrum of mineral elements in sediment samples is quite poor and prediction capabilities of partial least squares (PLS) models are based in small changes in the signals due to the interaction between the organic matter and the pollutants [5,6].

Destructive methods as inductively coupled plasma/optical emission spectroscopy (ICP–OES), and especially, inductively cou-

* Corresponding author. Tel.: +34 96 354 4838; fax: +34 96 354 4845.
E-mail address: miguel.delaguardia@uv.es (M. de la Guardia).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.06.047

pled plasma/mass spectrometry (ICP–MS) are widely used for multi-elemental analysis of sediment samples [9,10] but are costly and too much time is required to perform the analysis. On the other hand, from a green analytical point of view, X-ray fluorescence (XRF) provides a reagent-free environmentally friendly and non-destructive technique, which permits to determine the mineral content of samples without any complex sample pre-treatment, thus minimizing laboratory space, resources and safety implications, also avoiding the use of chemical reagents. Furthermore, XRF presents the possibility of making in situ analysis and could be considered a green technique [11]. The secondary XR fluorescence spectrum identifies the elements present in the sample and its intensity is proportional to their concentration. The disadvantage of the XRF technique is that the detection limits (typically mg kg^{-1}) are higher than those obtained by using ICP–OES or ICP–MS for dissolved trace elements and that the quantitative processing of the spectral data using external calibration is time consuming, as compared with other instrumental techniques in order to avoid matrix sample interferences, and requires skilled operators. So, XRF is normally used for semi-quantitative and quantitative multi-elemental analysis of mineral phases in pigments, rocks, sediments, soils and contaminated lands [12,13].

A large variety of multivariate methods are used in combination with XRF results. Orthogonal multivariate regression models are those most employed for quantitative analysis, as well as principal components regression (PCR) and partial least square (PLS) [14]. PLS combines spectral and analytical information of a number of samples to build predictive models by establishing a direct correlation between spectra of samples and properties of interest (generally concentration values) of the species contained in them. All these implementations to guarantee the accuracy in quantifying trace element concentrations have opened the possibility to use XRF as a rapid and low cost technique to obtain basic information of total concentration of metals in the assessment of potentially polluted soils and sediments [15].

The purpose of the present paper has been the development of a new analytical methodology, based on the combination of energy-dispersive X-ray fluorescence (μ-EDXRF) measurements and chemometric data treatment to quantify the amount of several elements in a unique step through the evaluation of selective regions from the whole spectrum. Developed method has been tested for the determination of minor and trace elements in sediment samples, corresponding to eight sampling locations and four sampling campaigns per year (during 2005–2008), of the estuary of the Nerbioi-Ibaizabal River (Bay of Biscay, Basque Country). Reference values of trace elements concentration corresponding to those samples used for both building and evaluating PLS models were obtained by ICP/MS after acid leaching of the sediment.

2. Experimental

2.1. Sampling, sample preparation and analysis

A total amount of 116 sediment samples were collected from eight sampling locations in the estuary of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Basque Country, north of Spain), located on the continental shelf of the Cantabrian coastline in the south-eastern part of the Bay of Biscay [16] (see Fig. 1)

Sediments were sampled approximately every three months (15 sampling campaigns from January 2005 to October 2008), using plastic sampling materials and latex gloves to avoid sample contamination with metals. Approximately 500 g of surface sediment samples (~2 cm depth) were collected at low tide, from eight well identified stations strategically distributed along the estuary. Samples were set inside cleaned plastic bags and transported to the laboratory in portable cooler boxes at 4 °C to reduce the effects of microbiological activity.

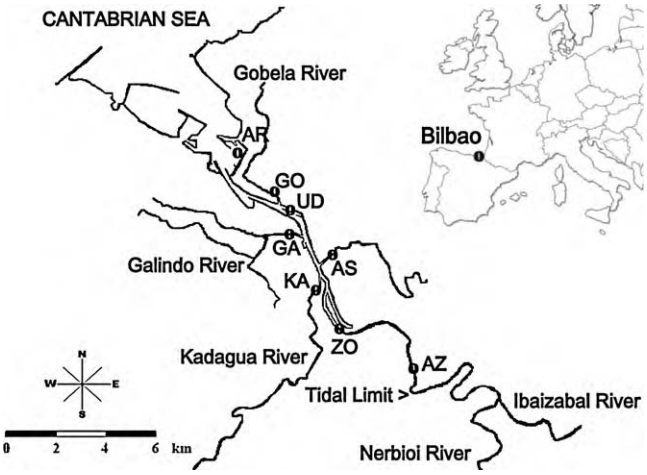

Fig. 1. Nerbioi-Ibaizabal estuary. *Note*: Sampling sites are indicated with star marks: Alde Zaharra (AZ), Zorroza (ZO), Kadagua (KA), Asua (AS), Galindo (GA), Udondo (UD), Gobela (GO) and Arriluze (AR).

In the laboratory, sediment samples were frozen at −20 °C and then lyophilized at 150 mTorr and −52 °C in a Cryodos apparatus (Telstar, Terrassa, Spain) for 48 h. Finally, samples were sieved at 63 μm and kept in the refrigerator at 4 °C until analysis.

XRF were obtained using the portable ArtTAX μ-EDXRF equipment by Röntec (nowadays Bruker AXS) equipped with a molybdenum anode working at a maximum voltage of 50 kV and a maximum current of 0.7 mA. The X-rays are collimated by a tantalum collimator with a diameter of 1.5 mm. The equipment also includes a CCD camera which allows to obtain an image of the sample analyzed (see inset of Fig. 2) and a motor-driven *XYZ* positioning unit to focus on different parts of the samples controlled by the computer. 200 mg of each homogenized sediment sample compacted under a total pressure of 10 tons, thus providing sample pellets with 10 mm diameter and 1 mm width. Independent triplicate spectra were obtained for each pellet. All the measurements were performed with an exposition time of 1000s at a voltage of 50 kV and a current of 0.6 mA. For the identification and quantification of elements with a low atomic number a helium flow (99.999%,

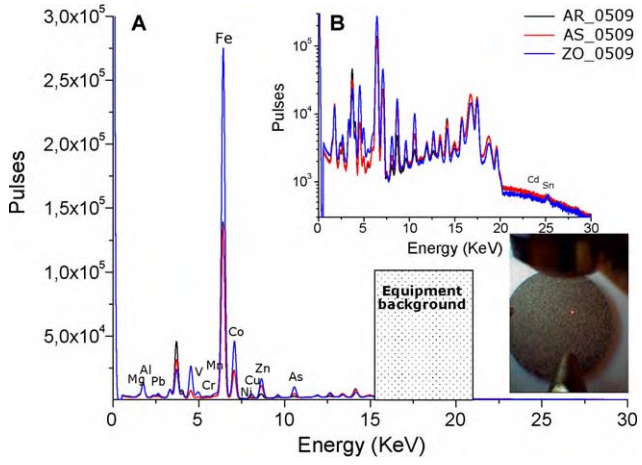

Fig. 2. Micro-XRF spectra acquired from the surface of samples corresponding to different sampling points (ARriluze, GObela and UDondo) obtained in different sampling campaigns: 2006 March, 2007 January and 2008 January, respectively: (A) lineal and (B) logarithmic scale for intensity values. Inset: photography showing sample pellet measuring point to obtain XRF spectra.

Air liquid, Spain) was used. All spectra processing and manipulation were carried out by using ARTAX 4.9.13.2 program from Bruker AXS.

The extractable concentration of Al, As, Cd, Co, Cr, Cu, Fe, Mg, Mn, Ni, Pb, Sn, V and Zn in the sediment samples were determined by ICP–MS after acid (HNO_3 and HCl mixtures) digestion accelerated with focused ultrasound energy as described elsewhere [17]. During the extraction step samples were sonicated with a HD 2070 Sonopuls Ultrasonic Homogenizer (Bandelin, Germany) equipped with a GM 2070 generator (70 W, 20 kHz), an UW 2070 ultrasonic converter, a SH 70 GQ horn and GS 6 glass probe (6 mm). The analysis of the extracts was carried out by inductively coupled plasma/mass spectrometry (ICP–MS) using an Elan 9000 ICP–MS spectrometer from PerkinElmer (Ontario, Canada), equipped with a Ryton cross-flow nebulizer, a Scott-type double pass spray chamber and standard nickel cones. Preparation of calibrations and sample analysis was done inside a clean room (class 100). The nitric acid (69%, Tracepur) and hydrochloric acid (36%, Tracepur) used in the extraction step were purchased from Merck (Darmstadt, Germany). Milli-Q (Millipore, Billerica, MA, USA) quality water with a conductivity lower than 0.05 $\mu S\,cm^{-1}$ was used to prepare sample and standard solutions. Multi-elemental standard solutions of the analytes were prepared by weight in 1% HNO_3 from individual commercial stock solutions (Specpure from Alfa Aesar, War Hill, MA, USA) of Al, As, Cd, Co, Cr, Cu, Fe, Mg, Mn, Ni, Pb, Sn, V and Zn (1000 $mg\,L^{-1}$ in 5% HNO_3). Be^{9}, Sc^{45}, In^{115} and Bi^{209}, also from Specpure stock standard solutions (1000 $mg\,L^{-1}$ in 5% HNO_3), were added to blank, standard and sample solutions as internal standard to yield a 10 $\mu g\,L^{-1}$ concentration level. 99.999% argon (Praxair, Spain) was used as plasmogen and carrier gas in the ICP–MS. The certified reference material NIST-SRM 1646a (estuarine sediment from the National Institute of Standards and Technology, Gaithersburg, USA) was used to validate the accuracy of the ICP–MS analytical procedure.

These determined extractable concentration values were used as reference data for building the partial least squares (PLS) multivariate models as well as for the subsequent evaluation of their predictive capabilities.

2.2. Chemometric data treatment

Hierarchical cluster analyses (HCA) of XRF spectra as well as PLS models for mineral elements determination were developed by using OPUS program Version 4.1 from Bruker Gmbh (Bremen, Germany).

HCA is an exploratory data tool used for describing a system of organized observations through their classification into groups of samples sharing properties in common (see Ref. [18] for additional information on HCA). In this study HCA was used to evaluate the spectra differences between samples considered in order to select an appropriate subset of samples for building the PLS prediction models [19]. Previous to HCA a vector normalization spectral pre-treatment was applied. For PLS modeling XRF spectra were mean-centered and after that different spectral pre-treatments were applied to obtain calibration models with a high predictive capability for each element considered in this study.

The root mean square error of calibration (RMSEC), the root mean square error of prediction (RMSEP), the coefficient of multiple correlation (r) and the mean of the differences between the predicted spectroscopic PLS values and ICP–MS reference data (d_{x-y}), together with the pooled standard deviation for validation samples (s_{trip}) and the residual predictive deviation (RPD), ratio of validation dataset standard deviation to RMSEP (SDR) were employed as quality indicators of the different models assayed [20].

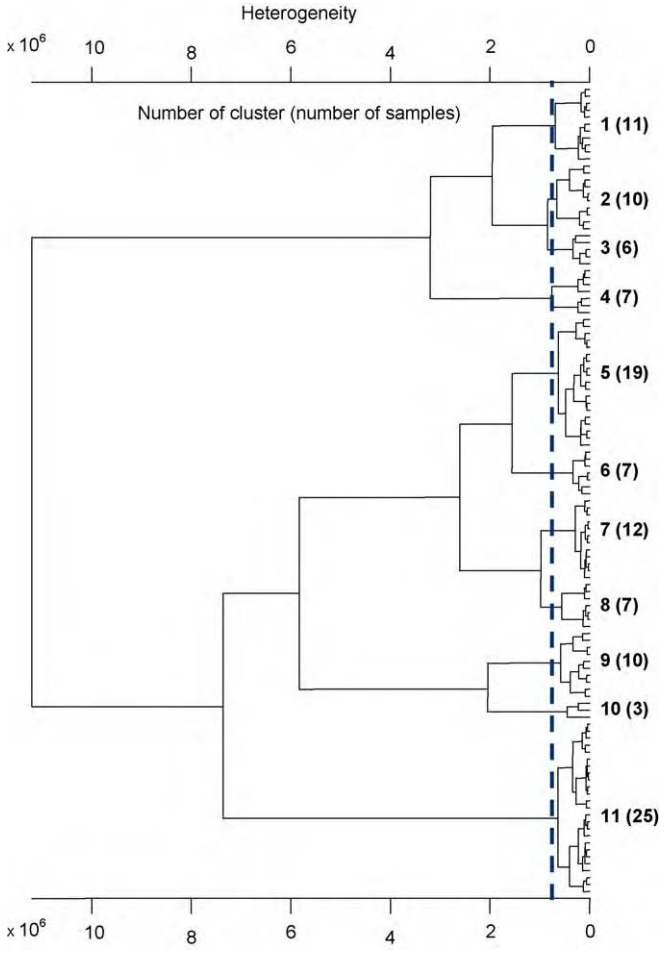

Fig. 3. Dendrographic classification of sediment samples from their XRF spectra. The Euclidean distance on the vector normalized spectra and the Ward linkage methods were used. For details about cluster group composition and sample ordering see data reported in Table 1.

3. Results and discussion

3.1. Reference data from trace elements

Metal concentrations in sediment samples used in the present work were obtained by the reference ICP–MS procedure [17] and correspond to the mean of four sample replicates, being precision values (reproducibility between days) comprised between 1.1% for Pb and 4.7% for Mn. Element contents in sediment samples considered in this study ranged between 900 and 15,500 $\mu g\,g^{-1}$ for Al, from 0.6 to 220 $\mu g\,g^{-1}$ for As, from non detected to 18 $\mu g\,g^{-1}$ for Cd, between 0.3 and 32 $\mu g\,g^{-1}$ Co, from 5 to 175 $\mu g\,g^{-1}$ for Cr, between 15 and 575 $\mu g\,g^{-1}$ for Cu, from 5200 to 36,720 $\mu g\,g^{-1}$ for Fe, between 330 and 8340 $\mu g\,g^{-1}$ for Mg, from 65 to 935 $\mu g\,g^{-1}$ for Mn, between 3.3 and 320 $\mu g\,g^{-1}$ for Ni, from 21 to 445 $\mu g\,g^{-1}$ for Pb, from non detected to 68 and 37 $\mu g\,g^{-1}$ for Sn and V, respectively and between 40 and 2060 $\mu g\,g^{-1}$ for Zn.

3.2. X-ray fluorescence spectra of sediment samples

Fig. 2 shows the characteristic XRF spectra of three samples taken from different points of the estuary during different sampling campaigns. As it can be seen the spectra contain characteristic peaks of the analytes of interest and non-explicit information, also suitable to be modeled as a function of the concentration of trace elements in samples of the same type previously characterized by a specific and well validated methodology. Additionally, the micro-

Table 1
Characteristics of the Nerbioi-Ibaizabal estuarine sediment samples classified into clusters after dendrographic treatment of XRF data.

Element	Cluster number (number of samples in each cluster)										
	1 (11)	2 (10)	3 (5)	4 (7)	5 (19)	6 (7)	7 (12)	8 (7)	9 (10)	10 (3)	11 (25)
Al	4870 ± 1254	4799 ± 461	5680 ± 2897	5061 ± 1946	3808 ± 2326	1816 ± 1590	5550 ± 3936	4226 ± 3075	3877 ± 2140	4107 ± 1770	4434 ± 1790
As	85 ± 47	87 ± 70	145 ± 67	61 ± 58	15 ± 14	11 ± 11	20 ± 15	3.5 ± 1.8	79 ± 39	45 ± 16	17 ± 14
Cd	5 ± 3	7 ± 3	7 ± 4	4 ± 3	2 ± 3	0.9 ± 0.8	2.0 ± 1.9	0.01 ± 0.03	4 ± 2	3.6 ± 0.6	2 ± 4
Co	7 ± 2	6.2 ± 1.0	7.7 ± 1.1	10 ± 4	7 ± 3	9.3 ± 1.0	5 ±3	3.4 ± 1.8	9 ± 2	19 ± 11	6 ± 2
Cr	44 ± 12	54 ± 15	62 ± 16	65 ± 12	42 ± 28	29 ± 10	33 ± 10	8 ± 3	41 ± 7	89 ± 75	36 ± 19
Cu	278 ± 158	301 ± 126	325 ± 95	194 ± 62	59 ± 38	49 ± 22	68 ± 31	39 ± 53	227 ± 145	187 ± 88	81 ± 56
Fe	18,167 ± 4694	22,434 ± 7423	24,599 ± 5615	22,334 ± 6745	10,776 ± 2627	8379 ± 1601	10,310 ± 2752	6626 ± 1074	17,954 ± 6627	23,916 ± 6825	10,562 ± 1928
Mg	3085 ± 881	3128 ± 660	2203 ± 1226	2250 ± 661	2841 ± 1795	5739 ± 1841	2681 ± 2550	1362 ± 673	4522 ± 1866	4645 ± 2573	2360 ± 1465
Mn	348 ± 199	500 ± 312	328 ± 231	473 ± 151	277 ± 125	227 ± 89	339 ± 241	249 ± 102	318 ± 151	574 ± 148	241 ± 146
Ni	22 ± 8	29 ± 11	21 ± 6	69 ± 39	26 ± 15	34 ± 20	18 ± 4	6 ± 2	28 ± 17	164 ± 136	21 ± 7
Pb	235 ± 72	240 ± 61	307 ± 72	259 ± 72	93 ± 75	92 ± 82	120 ± 79	44 ± 19	223 ± 78	321 ± 89	119 ± 102
Sn	26 ± 7	31 ± 10	38 ± 12	32 ± 10	10 ± 5	8 ± 4	9 ± 6	3 ± 3	25 ± 8	40 ± 24	12 ± 6
V	22 ± 8	19 ± 12	29 ± 7	13 ± 9	6 ± 6	6 ± 4	5 ± 6	2 ± 3	20 ± 4	11.0 ± 1.7	5 ± 4
Zn	829 ± 353	979 ± 309	1069 ± 290	872 ± 248	346 ± 213	280 ± 139	583 ± 426	190 ± 244	765 ± 178	1009 ± 195	500 ± 457
	Identification of samples corresponding to each cluster										
	AR_05-05	GA_05-12	GO_05-05	GA_05-01	AR_05-09	AR_08-10	AR_05-12	KA_05-01	AR_08-01	GA_08-01	AR_06-10
	GO_07-10	GA_06-07	UD_05-05	UD_05-01	AR_07-07	AS_08-10	AR_06-03	KA_05-05	GO_08-07	GA_08-04	KA_06-10
	GO_06-10	GA_06-03	GO_07-07	GA_07-04	AR_08-04	AZ_08-07	AR_06-07	KA_07-10	GO 08-10	GA_08-10	KA_07-01
	GO_07-01	UD_06-07	UD_07-07	GA_07-01	AR_08-07	AZ_08-10	AS_05-12	KA_05-12	UD_08-10		AR_07-01
	GO_07-04	UD_05-12	UD_05-09	GA_05-05	AZ_08-01	ZO_08-10	AS_06-03	KA_06-03	GA_08-07		KA_07-04
	GA_06-10	UD_06-03		GA_07-10	AS_05-09	KA_08-10	AZ_05-12	KA_06-07	GO_08-01		AS_06-10
	UD_07-04	GO_05-12		GA_07-07	ZO_05-09	ZO_08-07	AZ_06-03	KA_07-07	UD_08-04		AS_07-01
	GO_05-01	GO_06-03			ZO_08-01		AS_06-07		UD_08-07		AS_07-04
	UD_06-10	GO_06-07			ZO_08-04		ZO_05-12		GO_05-09		ZO_06-10
	UD_07-01	GO_08-04			AZ_05-09		ZO_06-03		UD_08-01		AZ_06-07
	UD_07-10				AZ_08-04		ZO_06-07				ZO_07-04
					KA_08-01		AS_07-07				AS_07-10
					KA_08-04						ZO_05-05
					KA_08-07						AR_07-04
					AS_08-01						AZ_07-01
					AZ_07-10						ZO_07-01
					AS_08-04						AZ_06-10
					AS_08-07						AR_07-10
					AZ_07-04						AZ_05-01
											ZO_07-10
											AS_05-01
											AS_05-05
											AZ_07-07
											ZO_07-07
											ZO_05-01

Notes: Samples are identified with two letters (corresponding to the sampling point) and four numbers (the first two numbers for the year and the last ones corresponding to the month of sampling) and are ordered according their dendographic classification (from up to down).
All trace element concentration values correspond to the averaged value from the concentration of all samples contained in the cluster, expressed in $\mu g\,g^{-1}$ units. ±s refers to the standard deviation of the average value.

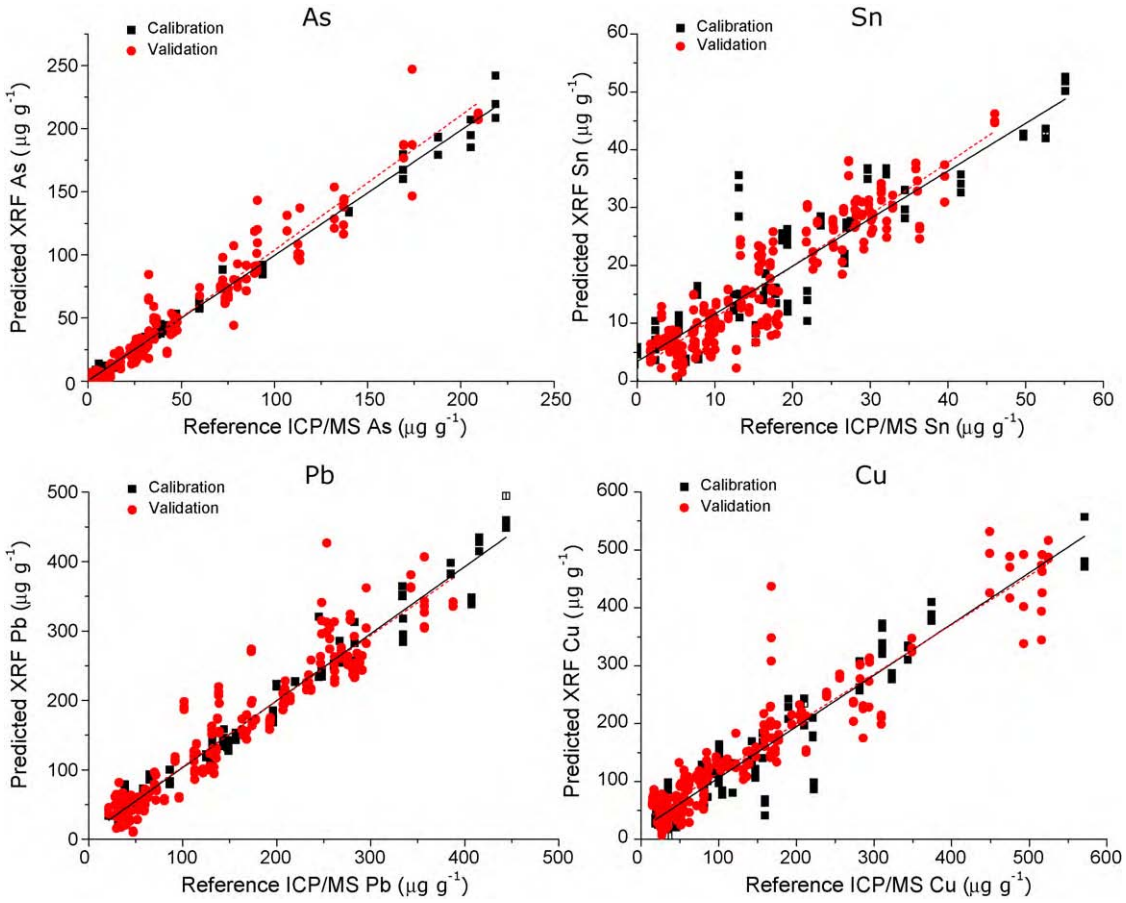

Fig. 4. Regression between PLS–XRF predicted data and reference concentration values obtained by ICP–MS for the determination of As, Cu, Pb and Sn in Nerbioi-Ibaizabal estuary sediment samples. Concentration values are expressed in $\mu g\,g^{-1}$. (■) Calibration samples and (●) validation samples.

X-ray spectra obtained from the surface of the sediment pellets include the Compton and Rayleigh scattering lines of Mo and Zr, the source and a component of the detector and the collimator which are employed to establish the equipment background.

All collected spectra were classified using the Euclidean distance with Ward linkage upon considering the voltage range between 0.5 and 28 keV on the raw spectra, being considered the spectral regions between 28 and 50 keV as useless and uninformative in order to determine the trace element concentration in the Nerbioi-Ibaizabal estuary.

Fig. 3 depicts dendrographic classification obtained from sediment sample spectra and as seen, 11 clusters are well defined for a cut-off value of 0.7×10^6. Samples grouped in each cluster joint together the averaged value for the concentration of each element, companied by its standard deviation, for overall sediment samples contained in each group are detailed in Table 1.

Table 2
Descriptive statistics for calibration and validation datasets used for PLS–XRF based prediction of trace elements in estuarine sediment samples and figures of merit of reference method based on ICP–MS measurements [17].

Element	Mean concentration value (concentration range)		Reference ICP–MS method	
	Calibration setNumber of samples: 34	Validation setNumber of samples: 82	Repeatability (%)	LD ($\mu g\,g^{-1}$)
Al	4903 ± 3236 (1095–15,487)	4169 ± 1817 (902–10,506)	2.9	18
As	50 ± 61 (1.0–219)	40 ± 45 (0.6–209)	3.7	0.9
Cd	4 ± 4 (0.0–17.4)	3 ± 3 (0.0–13.7)	2.1	1.0
Co	8 ± 3 (1.93–16.0)	7 ± 4 (2.36–31.6)	4.0	0.2
Cr	42 ± 22 (5.0–81.8)	42 ± 25 (6.0–175)	2.5	0.5
Cu	154 ± 126 (18.6–571)	137 ± 131 (14.8–525)	1.7	23
Fe	15,483 ± 8146 (5227–34,721)	13,731 ± 6395 (5996–34,245)	4.2	12
Mg	3425 ± 2121 (330–8117)	2799 ± 1696 (699–8335)	3.3	6.0
Mn	372 ± 220 (129–934)	301 ± 182 (66–936)	4.7	0.9
Ni	27 ± 24 (3.4–120)	30 ± 37 (6.2–320)	2.6	4.0
Pb	180 ± 126 (21–444)	153 ± 101 (22–387)	1.1	1.0
Sn	19 ± 15 (0.0–55.1)	17 ± 12 (1.7–67.7)	2.2	1.9
V	13 ± 11 (0.0–37.3)	10 ± 9 (0.0–36.8)	3.5	0.3
Zn	679 ± 479 (42–2060)	577 ± 375 (41–1739)	2.0	4.0

Note: All values are expressed in $\mu g\,g^{-1}$ units. Averaged ± its corresponding standard deviation concentration values of each trace metal calculated from the overall samples included in each set. Values indicated between brackets correspond with minimum and maximum concentration levels found for samples considered for each set. Repeatability values were obtained from four replicates. LD data were calculated from the standard deviation of eight measurements of a blank.

Table 3
Modeling parameters and prediction capabilities of PLS–XRF models developed for the determination of trace elements in estuarine sediment samples from the Nerbioi-Ibaizabal estuary.

Element	Spectral range (keV)	Spectral pre-processing	F	r	RMSEC ($\mu g\,g^{-1}$)	RMSEP ($\mu g\,g^{-1}$)	RRMSEP (%)	$d_{(x-y)}$ ($\mu g\,g^{-1}$)	s_{trip} ($\mu g\,g^{-1}$)	RPD ($\mu g\,g^{-1}$)	LD ($\mu g\,g^{-1}$)
Al	28.0–25.3/ 22.5–19.8/ 3.3–0.5	None	2	0.61	2601	1708	41	472	223	1.1	70
As	28.0–19.8/ 17.0–8.7	Vector normalization	10	0.97	6	11	29	−0.19	9	3.9	4
Cd	25.3–22.5/ 14.3–8.7/ 3.3–0.5	Vector normalization	4	0.84	1.9	1.7	62	0.3	0.3	1.8	0.4
Co	28.0–22.5/ 19.8–17.0/ 11.5–6.0	Second derivative	7	0.72	2.4	2.8	40	−0.16	0.5	1.4	0.5
Cr	6.8–6.6/ 6.4–6.2/ 5.8–5.2	Straight line subtraction	7	0.76	10	16	39	1.6	8	1.5	2
Cu	22.5–17.0/ 8.8–6.0	Straight line subtraction	9	0.95	42	44	32	10	23	3.0	14
Fe	9.0–8.6/ 8.2–7.8/ 7.4–7.0	Min–max normalization	8	0.79	4085	4240	31	1505	1365	1.5	5300[a]
Mg	25.3–19.8/ 14.3–0.5	First derivative	9	0.85	961	910	32	9	257	1.9	37
Mn	6.0–0.5	First derivative + straight line subtraction	8	0.86	80	94	31	13	29	1.9	6
Ni	8.8–6.0	None	9	0.75	7	26	87	−5	7	1.4	1.3
Pb	14.3–11.5	Straight line subtraction	7	0.95	21	31	21	1.7	16	3.2	7
Sn	25.3–22.5/ 19.8–17.0/ 14.3–8.7/ 3.3–0.5	Constant offset elimination	1	0.88	6	6	33	−0.13	1.5	2.1	2
V	25.3–19.8/ 14.3–8.7	Multiplicative scattering correction	4	0.92	5	4	35	0.12	1.1	2.6	5
Zn	28.0–22.5/ 17.0–8.7/ 3.3–0.5	None	7	0.78	127	236	41	3	35	1.6	13

Notes: F is the number of PLS factors employed to build the model, r is the correlation coefficient, RMSEC and RMSEP are the root mean square error of calibration and prediction, respectively. RRMSEP is the RMSEP divided by the mean value of trace element in the validation dataset. s_{trip} is the standard deviation of three replicates. $d_{(x-y)}$ is the mean difference between predicted and reference trace element content value. RPD is the residual predictive deviation. LD is the limit of detection established from the standard deviation of 10 measurements of a blank spectrum of the sample holder and air.

[a] Due the difference between spectra of blank and samples LD for Fe was established from the standard deviation of several measurements of a validation sample with the lower content of Fe.

A relative correlation between clusters structure and regions across the estuary can be established: (i) clusters 1–4 and clusters 9 and 10 contain sample spectra from Galindo (GA), Gobela (GO) and Udondo (UD) locations (that correspond to the most polluted sediment sites in the estuary [16]), being cluster 10 totally integrated by the spectra of GA sediments sampled in 2008 and cluster 4 by spectra of GA sediments taken in 2005 and 2007; (ii) clusters 5–8 and cluster 11 contain sample spectra from sediments collected in Arriluze (AR), Asua (AS), Alde Zaharra (AZ), Kadagua (KA) and Zorroza (ZO) locations (the less polluted sites in the estuary), being cluster 8 integrated by the main part of samples collected in KA from 2005 to 2007.

In fact, clustering seems to be mainly correlated with the concentration of elements exhibiting the most sensitive X-ray fluorescence signals. However there is not a simple correlation between trace element concentration and sample clustering, being thus the classification dependent on the matrix sediment additionally than on the trace element composition. So, in the present study HCA was used to evaluate the spectra differences between samples considered in order to select an appropriate subset of samples for building the PLS prediction models (see Refs. [6,19] for additional details on the use of HCA for properly selecting a representative calibration set for modeling and also for verifying the nature of new samples previously to be predicted). Table 2 summarizes the descriptive statistics of the calibration set established from 34 samples selected from the 11 clusters and validation independent set, composed by the 82 available samples not used for calibration. As it can be seen trace element concentration of validation samples are included in the range of the calibration set, thus avoiding the extrapolation of data.

Samples employed for calibration were randomly selected including samples of all cluster, taken from different sampling locations till to have the rounded value of the root square of the total number of samples included in each cluster.

3.3. Partial least square modelization of XRF data for trace metals prediction

Partial least square (PLS) multivariate models were built from the selected calibration set of XRF spectra of sediment samples. Several spectral regions in the voltage domain [21] as well as different spectra treatments [22] were tested. Models were compared in terms of RMSEP values of trace metal content for samples not used for calibration in order to evaluate the model prediction capabilities. However, as a consequence of the high volume of results generated, only the best ones have been summarized in this paper. Fitting features as well as prediction capabilities of models obtained for each trace element are summarized in Table 3.

Spectral region was optimized for each element under consideration by splitting the overall selected spectral range (0.5–28 keV) in 10 equally large regions and successively calculating combinations of them. In the case of Fe and Cr determination a refined selection of the spectral interval was made by splitting spectral ranges comprised between 5–9 and 5–7 keV, respectively. As it can be seen, from one to four combined spectral regions may be used to obtain the best prediction capability in each case.

Additionally, from none to different spectral treatments such as first (FD) or second (SD) order derivative, multiplicative scattering correction (MSC), constant offset elimination (COE), vector normalization (VN), straight line subtraction (SLS) and min–max normalization (MMN) as well as a combination between some of them (FD + SLS) were also required in some cases. Regarding the number of PLS factors they were selected to obtain the minimum RMSECV to have a good predictive capability. This parameter ranged from a minimum value of 1 (for Sn) to a maximum one of 10 (for As), being this parameter a descriptive index on the complexity of the relationship between XRF spectra and element concentration.

In order to evaluate the sensitivity of the proposed PLS–XRF method the limit of detection values for each element were established from 10 measurements of a blank spectra of the sample holder and air. LD obtained varied from 0.4 $\mu g\,g^{-1}$ for Cd to 70 $\mu g\,g^{-1}$ for Al and, in some cases, were of the same order than those provided by the reference method applied to sediment analysis (see Table 2).

As an example, Fig. 4 shows the comparison between the calibration and validation sets for arsenic (As) and tin (Sn), as elements with low concentration in the studied system, and copper (Cu) and lead (Pb), as elements with relative high concentrations in all the samples. As it can be appreciated, fitting for calibration and validation results are coincident within the confidence interval of the respective slopes, thus indicating that there are not significant differences between both, calibration and validation datasets.

The high fluctuation of concentration values is due to the seasonal trends observed in this estuary (minimum content in winter and maximum in summer) for the whole set of considered elements [6].

XRF prediction capabilities provided better results than those achieved in a previous study using vibrational spectroscopy [6]. XRF models provided, as a trend, the highest RPD values, RRMSEP values ranging between 21% and 41%, with few exceptions given for Cd (62%) and Ni (87%), and r values spanned from 0.72 to 0.97 (avoiding Al results) which are also better than those obtained by vibrational models. Nevertheless, optimal number of factors to be considered in the vibrational PLS models were, in general, lower than those determined for XRF–PLS models.

4. Conclusions

It has been demonstrated, from results obtained in the present study, that the concentration of several trace elements as As, Cd, Cr, Cu, Fe, Mg, Ni, Sn, V and Zn in estuarine sediments from the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country) could be properly modelized (exception given for Al, Co, and Ni exhibiting RPD values lower than 1.5) through the combined use of XRF spectra and PLS treatment. It results in a green analytical tool for environmental studies suitable to extract concentration data at trace levels from spectra of untreated samples just doing an appropriate calibration, with samples previously analyzed by a reference procedure, and considering some spectral pre-processing strategies.

The fact that the XRF spectra contain both, intrinsic and extrinsic informations on the mineral composition of samples is clearly the reason to obtain more accurate results by XRF than those previously found from vibrational spectrometry and evidences the capability of PLS based analytical procedures for processing simultaneously information regarding to different chemical compounds and obtained from different analytical signals.

Acknowledgements

Authors acknowledge the financial support of the BERRILUR 3 Strategic Research Project (Etortek Programme from the Basque Government, ref. IE09-242) as well as Ministerio de Educación y Ciencia (Projects AGL2007-64567 and CTQ2008-05719/BQU). Ainara Gredilla is also grateful to the UPV/EHU for her pre-doctoral fellowship.

References

[1] K.L. Spencer, Mar. Pollut. Bull. 44 (2002) 933.
[2] C. Christophoridis, D. Dedepsidis, K. Fytianos, J. Hazard. Mater. 168 (2009) 1082.
[3] K.S. Kumar, K.S. Sajwan, J.P. Richardson, K. Kannan, Mar. Pollut. Bull. 56 (2008) 136.
[4] J. Moros, M.J. Martínez-Sánchez, C. Pérez-Sirvent, S. Garrigues, M. de la Guardia, Talanta 78 (2009) 388.
[5] J. Moros, R.J. Cassella, M.C. Barciela-Alonso, A. Moreda-Piñeiro, P. Herbello-Hermelo, P. Bermejo-Barrera, S. Garrigues, M. de la Guardia, Vib. Spectrosc. 53 (2010) 204.
[6] J. Moros, S. Fdez-Ortiz de Vallejuelo, A. Gredilla, A. de Diego, J.M. Madariaga, S. Garrigues, M. de la Guardia, Environ. Sci. Technol. 43 (2009) 9314.
[7] U. Villanueva, J.C. Raposo, K. Castro, A. de Diego, G. Arana, J.M. Madariaga, J. Raman Spectrosc. 39 (2008) 1195.
[8] J.C. Raposo, O. Zuloaga, J. Sanz, U. Villanueva, P. Crea, N. Etxebarria, M.A. Olazabal, J.M. Madariaga, Mar. Chem. 99 (2006) 42.
[9] W.D. Lopes, R.E. Santelli, E.P. Oliveira, M.D.B. de Carvalho, M.A. Bezerra, Talanta 79 (2009) 1276.
[10] F.D. Gomes, J.M. Godoy, M.L.D.P. Godoy, Z.L. de Carvalho, R.T. Lopes, J.A. Sanchez-Cabeza, L.D. de Lacerda, J.C. Wasserman, Mar. Pollut. Bull. 59 (2009) 123.
[11] S. Armenta, S. Garrigues, M. de la Guardia, Trends Anal. Chem. 27 (2008) 497.
[12] K. Castro, S. Pessanha, N. Proietti, E. Princi, D. Capitani, M.L. Carvalho, J.M. Madariaga, Anal. Bioanal. Chem. 391 (2008) 433.
[13] M. Gurhan Yalcin, I. Narin, M. Soylak, Environ. Geol. 54 (2008) 1155.
[14] M.J. Adams, J.R. Allen, Analyst 123 (1998) 537.
[15] Q.T. Huang, L.Q. Zhou, Z. Shi, Z.Y. Li, Q. Gu, Spectrosc. Spect. Anal. 29 (2009) 1434.
[16] S. Fernández, U. Villanueva, A. de Diego, G. Arana, J.M. Madariaga, J. Mar. Syst. 72 (2008) 332.
[17] S. Fdez-Ortiz de Vallejuelo, A. Barrena, G. Arana, A. de Diego, J.M. Madariaga, Talanta 80 (2009) 434.
[18] B.S. Everitt, S. Landau, M. Leese, Cluster Analysis, Arnold, London, UK, 2001.
[19] F.A. Iñón, J.M. Garrigues, S. Garrigues, A. Molina, M. de la Guardia, Anal. Chim. Acta 489 (2003) 59.
[20] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics (Parts A and B), Elsevier Science B.V., Amsterdam, 1997.
[21] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413.
[22] J. Luypaert, S. Heuerding, Y. Vander Heyden, D.L. Massart, J. Pharmaceut. Biomed. Anal. 36 (2004) 495.