Trends in Analytical Chemistry 76 (2016) 30–39

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

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

Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review

Ainara Gredilla *, Silvia Fdez-Ortiz de Vallejuelo, Nerea Elejoste, Alberto de Diego, Juan Manuel Madariaga

Department of Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country, UPV/EHU, P.O. Box 644, 48080 Bilbao, Basque Country, Spain

ARTICLE INFO

Keywords:
Environmental
Chemometrics
Spectroscopy
Non-destructive analysis
Infrared spectroscopy
X-ray fluorescence
Raman Spectroscopy

ABSTRACT

The design of environmentally friendly analytical techniques and the modification of methods that are in use are important issues to be addressed in *Green Analytical Chemistry*. Potential applications of non-destructive spectroscopic techniques to achieve these principles are discussed in this review. Special emphasis is placed in environmental samples. Taking into account the current tendency to combine the analysis of environmental samples by non-destructive spectroscopic techniques and chemometrics, the most important applications concerning this task are also compiled. Although multiple objectives are pursued, regression techniques, frequently in combination with infrared spectroscopy, stand out, with around 300 references since 2009. More than half of these publications deal with the design of models to predict properties of soils. Nonetheless, the combination of regression techniques and non-destructive spectroscopy is continuously progressing, and they have been applied, for example, in the analysis of different environmental samples by remote or *in-situ* multispectral sensors with satisfactory results.

Contents

Abbreviations: AAS, Atomic Absorption Spectrometry; ANN, Artificial Neural Networks; ARIMA, AutoRegressive Integrated Moving Average Modelling; ATR, Attenuated Total Reflectance; BP-ANN, Back Propagation Artificial Neural Networks; BRT, Boosted Regression Tree; CA, Cluster Analysis; CARS-PLS, Competitive Adaptive Reweighted Sampling Partial Least Squares; DRIFT, Diffuse Reflectance Infrared Fourier Transform; EC-MLR, Equidistant Combination Multiple Linear Regression; ED-XRF, Energy Dispersive X-ray Fluorescence; EDXRFS, Energy Dispersive X-ray Fluorescence Scattering; FA, Factor Analysis; GA-PLS, Genetic Algorithm Partial Least Squares; GRNN, Generalized Regression Neural Networks; ICP-AES, Inductively Coupled Plasma Atomic Emission Spectrometry; ICP-MS, Inductively Coupled Plasma Mass spectrometry; IR, Infrared; k-NN, k-Nearest Neighbour; LDA, Linear Discriminant Analysis; LS-SVM, Least Squares Support Vector Machine; LWR, Locally Weighted Regression; MCR, Multivariate Curve Resolution; MCR-ALS, Multivariate Curve Resolution Alternating Least Squares; MIR, Middle Infrared; MLR, Multiple Linear Regression; MW-PLS, Moving Window Partial Least Squares; NIR, Near Infrared; PARAFAC, Parallel Factor Analysis; PCA, Principal Component Analysis; PCR, Principal Component Regression; PLS, Partial Least Squares; PLS-BPNN, Partial Least Squares Back Propagation Neural Networks; PLS-DA, Partial Least Squares Discriminant Analysis; SEM-EDX, Scanning Electron Microscopy Energy Dispersive X-Ray; SML, Stepwise Multiple Linear; SPA-MLR, Successive Projections Algorithm Multiple Linear Regression; SPA-LS-SVM, Successive Projections Algorithm Least Squares Support Vector Machine; SPA-PLS, Successive Projections Algorithm Partial Least Squares; STATIS, Structuration des Tableaux A Trois Indices de la Statistique; SVM, Support Vector Machine; UVE-GA-PLS, Uninformative Variable Elimination Genetic Algorithm Partial Least Squares; UVE-PLS, Uninformative Variable Elimination Partial Least Squares; VIS-NIR, Visible-Near Infrared; XRD, X-ray Diffraction; XRF, X-ray Fluorescence.

* Corresponding author. Tel.: +34 946015445; Fax: +34 94601 3500.

E-mail address: ainara.gredi@gmail.com (A. Gredilla).

http://dx.doi.org/10.1016/j.trac.2015.11.011
0165-9936/

1. Green chemistry concept

Currently, the traditional perception of Chemistry by the society is not positive. The Industrial Revolution brought about a drastic and unrestrained development of heavy industry. As the generated waste was directly discharged into the environment without prior treatment, numerous water bodies, soil plots and the atmosphere itself resulted seriously damaged. At that time, therefore, it was difficult to separate Chemistry from pollution. Nowadays we perceive Chemistry in a radically different way, and the stigma once attached to it is gradually disappearing. This new perception promotes Chemistry as a tool for the protection and improvement of the environment. In 1990, Paul Anastas and John Warner proposed the concept of *Green Chemistry* [1]. They defined *Green Chemistry* as a new philosophy based on eliminating, or at least reducing, the use or generation of hazardous substances in the design, synthesis and application of chemical products.

The *Green Chemistry* concept aims to unify the interests of the market, the society and the environment, nonetheless, chemical analysis is undoubtedly one of its principal areas of application. Particularly in this specific field, apart from reducing the use of hazardous agents, the decrease in energy consumption and generation of waste products is of special interest [2]. Thus, the development of new environmentally friendly analytical techniques and the modification of already existing methods are key points in the progress of the *Green Analytical Chemistry* concept.

2. Classical chemical analysis

The classical approach of an analytical process includes four main steps: *i)* sampling, *ii)* sample (pre)treatment, *iii)* analyte separation and finally *iv)* analyte detection. All of them could involve a considerable consumption of energy and produce a significant amount of hazardous substances.

Provided that the reagents used for cleaning purposes are carefully selected and handled, the sampling step should not constitute an important threat to the environment. Moreover, the use of passive sampling techniques instead of traditional approaches (active sampling) optimises the process in terms of time and money [3].

The pretreatment of solid, liquid and gaseous samples is usually a long and expensive process. Sometimes, it involves harmful steps for the environment. Pretreatment of a liquid, for example, often includes filtration, preconcentration and clean-up of the sample. The environmental effects derived from the pretreatment of solid samples are even stronger. In the classical approach, the analytes must be extracted into a liquid phase. The process usually takes a long time and there is a high risk to contaminate the sample and/or to degrade or transform the analyte [4].

The separation of analytes is mostly attempted by chromatographic and electrophoretic techniques. For detection, mass spectrometry and optical spectroscopy are the most popular approaches [5,6]. Separation and detection, however, are not nowadays critical stages concerning *Green Chemistry* criteria.

3. Non-destructive spectroscopy

Non-destructive spectroscopy has been studied in depth during the last years in order to overcome the drawbacks that, concerning the principles of *Green Analytical Chemistry*, result from the use of traditional analytical techniques [7]. The non-destructive spectroscopic techniques provide analytical information without damaging the sample and in most of the cases, in addition, sample treatment (extraction or digestion) is not required. In the view of *Green Chemistry,* this is indeed the most important advantage of non-destructive spectroscopy. Sampling can be also avoided, since *in situ* analyses are susceptible to be done using portable equipment [8]. Moreover, non-destructive spectroscopy potentially enables the simultaneous analysis of several analytes in a single run, with a consequent saving of reagents and time.

Vibrational techniques (Infrared (IR) and Raman spectroscopy) and X-ray fluorescence (XRF) are amongst the non-destructive spectroscopic techniques that have attracted more interest in the last years. They all rely on the interaction between electromagnetic radiation and matter, and the information obtained depends on the source of the energy used to irradiate the sample. Although the information obtained by IR could be classified as the poorer one, its applicability is by far the most important [9,10]. Numerous reviews such those of Moros et al. [7] and Armenta et al. [11] have compiled the most important works based on IR. Even though signals frequently overlap, Near Infrared (NIR, 780–2500 nm) is still the most used technique. Middle Infrared (MIR, from about 700 nm to 2500 nm) have also numerous applications, but the use of Far Infrared (FAR, 400 nm and 700 nm) is still limited.

Traditionally, dispersive instruments have been used to obtain infrared spectra, but as they were only able to measure one wavelength at each time, the use of Fourier Transform Infrared (FT-IR) spectrometers has been continuously increasing over time. A FT-IR spectrometer is able to collect all wavelengths simultaneously and, consequently, most of the IR instruments commercialised nowadays are FT-IR spectrometers.

Transmittance is the pioneering and most straightforward IR mode, however, applications based on the measurement of reflected light can be used for samples that there are difficult to analyse by the conventional transmittance methods. Reflectance methods can be divided into three principal categories: Attenuated Total Reflectance Spectroscopy (ATR), Specular Reflectance Spectroscopy and Diffuse Reflectance Spectroscopy (DRS). The information obtained from all of them is comparable if the differences emerging from the way the infrared light beam is absorbed by the sample are appropriately corrected. ATR spectroscopy is based on the total internal reflection phenomenon. Depending on the ATR cell, liquid, solid or gas samples can be analysed. Specular Reflectance Spectroscopy and DRS are based on external reflectance, so that the radiation reflected from a surface is measured in both cases. Specular reflection is defined as light reflected from a smooth surface at a definite angle, whereas DRS is based on the reflection produced when energy penetrates the surface of a sample and this tend to reflect light in all directions. Since DRS is usually coupled with Fourier transformation for a more appropriate interpretation of the signals, it is commonly known as DRIFT. Applications of DRIFT are by far more abundant than those of specular reflectance spectroscopy [7].

Raman Spectroscopy is considered a non-destructive technique because the power of its laser can be modulated according to the sample that is being analysed to avoid the thermodescomposition of the sample. Raman Spectroscopy provides us with revealing information, since the spectrum obtained is characteristic of each molecule of the sample area that is being analyzed. However, the application of this technique is still limited for quantitative analysis because the pretreatment of the spectra could be tedious.

XRF, unlike vibrational techniques, provides us with elemental information of the sample. It has been largely used in medical and environmental applications with satisfactory results [12]. In recent times, many attempts have been done to combine vibrational and elemental spectroscopic techniques. The combination of Scanning Electron Microscopy coupled to Energy Dispersive X-Ray Spectroscopy (SEM-EDX) and Raman spectrometry is a successful example of this combination [13].

Despite the multiple applications of the abovementioned spectroscopic techniques, this review will only focus in the analysis of environmental samples such as soils (including rocks and sediments), water, air particles, and biota (animals and plants mainly).

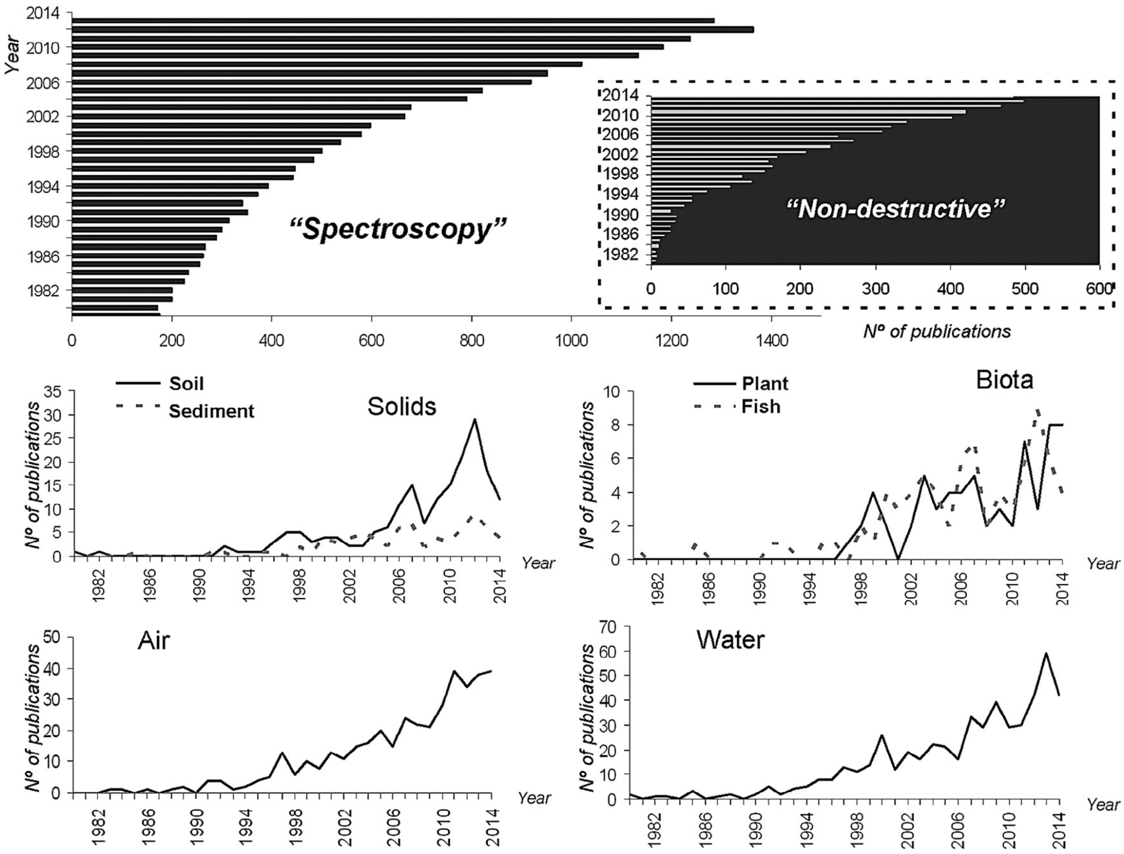

Fig. 1. Amount of publications (X axis) per year between 1980 and 2014 (Y axis) including "spectroscopy" (top and left) and the results obtained after refining the search with the concept "non-destructive" (top and right). The plots in the bottom correspond to the amount of publications (Y axis) per year from 1980 to 2014) (X axis) concerning non-destructive spectroscopy applied to the study of *i)* solid samples (sediment, soil, biota), *ii)* air or *iii)* water samples.

As shown in Fig. 1, the amount of publications devoted to the application of non-destructive spectroscopic techniques to the analysis of environmental samples has been continuously increasing over the period 1980-2014. Data in Fig. 1 was obtained after a bibliometric analysis carried out using the on-line version of SciFinder. This tool integrates the data bases included in CAplus and MEDLINE. 19.999 documents including books, reviews and articles were recovered with "spectroscopy" as search topic. The search was then refined including successively the concepts "non-destructive" (5666 documents) and "environmental" (455 documents). If the specific environmental matrixes usually studied are used to refine the "non-destructive" search, a clear increasing trend since 1980 is observed in all the cases. It should be clarified that the plot showing the evolution in the amount of works dealing with water samples is only a estimation, since it includes all the works containing the word "water", not necessarily as the studied matrix.

4. The role of chemometrics in the analysis of environmental samples by non-destructive spectroscopic techniques

The combination of chemometric and analytical techniques to produce methods according to the *Green Chemistry* concept has become an important challenge in the last few decades. A large variety of chemometric methods has been used up to date in the Analytical Chemistry field. We have tried to arrange this overview in a comprehensive way, but it becomes obvious that numerous overlaps may happen (see Fig. 2).

- *Time/spatial series analysis*: in environmental chemical analysis, and specifically in pollution monitoring, time and spatial series

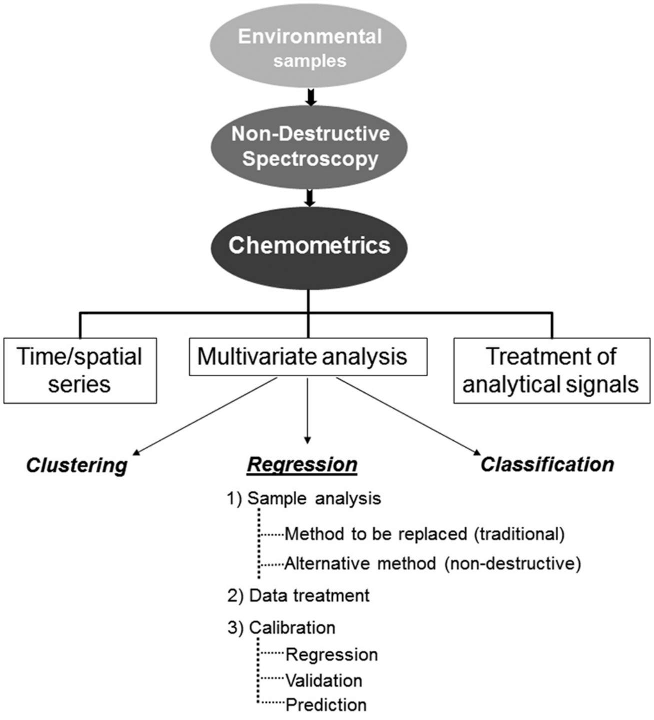

Fig. 2. Main chemometric techniques applied in combination with non-destructive spectroscopy in the analysis of environmental samples, and the main steps in a multivariate calibration approach.

are often investigated to detect systematic trends of analytes over time or space. When chronologically or strategically ordered data are analysed the main objectives consist on: *i)* forecasting, *ii)* process control, *iii)* process simulation and iv) development of new physical and biological models. Techniques used in this group include smoothing and filtering, regression, correlation and cross correlation, autoregression and AutoRegressive Integrated Moving Average modelling (ARIMA) [14], among others. The amount of publications included in this field and devoted to environmental issues is scarce. Most of them deal with monitoring studies, and due to the amount of data compiled, the chemometric techniques usually applied are, statistically speaking, more complex than those included in this category. That is, clustering or classification techniques have been more widely applied in this kind of studies because they provide the researcher with more useful information in a single statistical step, such as the main features of a given spatial or temporal trend. The few works focused in the abovementioned methods do not use directly spectral information. Logically, it is more difficult to investigate temporal and spatial variability of a given variable in a given area using spectral data rather than numeric discrete variables. Consequently, the works published in this field combining non-destructive spectroscopy and chemometric techniques for the analysis of time/spatial series use concentration values obtained from the integration of spectral bands or signals (depending on the method used) rather than the whole spectra. For example, XRF, which provides us with elemental information, is, in combination with correlation analysis, the most popular technique in environmental studies. Researches focused in aquatic systems stand out here. The evaluation of the ecological state of rivers and estuaries is an important issue due to their very high temporal and spatial variability. With this objective, sediments have been analysed by means of XRF [15,16]. Although the review of Simonescu [17] revealed that IR has numerous applications in environmental studies concerning pollution monitoring, most of them do not include chemometric techniques to study the possible existence of trends in spatial/temporal series. Similarly, Raman spectroscopy is not commonly used with this purpose.

- *Treatment of analytical signals:* A spectrum must be pretreated before further manipulation in order to obtain reliable information from it. If the objective is signal or band identification, a simple baseline correction may be enough. However, if the goal is quantification, different preprocessing methods should be thoroughly checked, and combined. NIR and MIR (Middle Infrared) spectroscopy are the most usual techniques in works concerned with environmental samples and spectral preprocessing. The noise present in IR spectra is behind this fact. Noise may have different origins, such as interference of chemical factors or errors in the experimental process, among others. Signal-to-noise ratio decreases when noise increases and, consequently, the abovementioned factors can distort the information extracted from the IR spectra. The pre-processing techniques most frequently used in NIR (and also in MIR) spectroscopy can be divided into two main groups: scatter correction methods and spectral derivatives. The scatter-corrective group includes Multiplicative Scatter Correction (MSC), Inverse MSC, Extended MSC (EMSC), Extended Inverse MSC, de-trending, Standard Normal Variate (SNV) and normalization. Among methods based on derivates two techniques stand out: Norris-Williams (NW) derivatives and Savitzky-Golay (SG) polynomial derivative filters. Both are normally applied to minimize baseline deviations caused by dispersion effects. Approximately the 90% of the IR spectra derived from environmental analysis is preprocessed. This is even more notable in the analysis of biological samples [18]. Pretreatment of spectral signals obtained from plant samples prior to composition assessment is, therefore, an important task [19,20]. The objective of most publications based on IR is not preprocessing itself; most usually, they have more ambitious objectives and signal treatment is only a previous step. Prediction of physico-chemical properties of soils, e.g., the evaluation of humic and fulvic acid contents [21,22], is a common objective of publications including IR preprocessing. The number of works dealing with preprocessing of Raman or XRF spectra is significantly lower. For instance, when Raman spectral preprocessing is carried out the main objective is usually the elimination of fluorescence. With this purpose background correction procedures, which minimize the effect of a varying background caused by fluorescence of the sample, filtering and scaling methods and techniques for dimension reduction have been applied for the spectral preprocessing of Raman spectra obtained after the analysis of solid environmental samples. In the case of XRF the preprocessing methods are usually based in derivation of the signal. Not only numerical derivatization is used, Savitzky-Golay or Norris algorithms are also applied. Afseth et al. 2006 [23] and Bocklitz et al. 2011 [24] investigated the effectiveness of different methods to preprocess Raman spectra obtained from biological samples. However, the main objective of works based on XRF or Raman spectroscopy usually is not the preprocessing itself; they all have more arduous objectives, such as regression for the prediction of the elemental composition of soils or soils quality assessment [25,26].
- *Multivariate Analysis*: As observed in Fig. 2, chemometric methods most often used in the analysis of multivariate environmental data can be divided into 3 main categories according to the objective pursued: *i)* clustering, *ii)* classification and *iii)* regression:
  - Clustering tries to identify clusters of objects (samples) or variables (experimentally measured parameters) with similar characteristics, which enables an assessment of the general structure of the dataset. Depending on how data is organised, two- or three- dimensional methods can be distinguished. That is, each sample can be graphically depicted in a two- or three-dimensional space. If the samples (objects) are organised in one direction and the variables in another one, a single 2D data table is obtained (two-way methods, Fig. 3). Data can be also organised according to a multi-set or multi-way arrangement. In both cases, a simultaneous analysis of several 2D data tables is performed (multi-set or multi-way methods, Fig. 3). Cluster Analysis (CA), Principal Component Analysis (PCA), Multivariate Curve Resolution (MCR) and Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) are the most popular two-dimensional techniques. Three dimensional approaches include Structuration des Tableaux A Trois Indices de la Statistique (STATIS), Parallel Factor Analysis (PARAFAC), PARAFAC 2 and Tucker3 [27].

    Although all these techniques are increasingly applied in environmental studies [14,28], their use in combination with pure spectral data is not very popular. Clustering has been more frequently used to deal with spectral data resulting from NIR and MIR measurements and, to a lesser extent, from XRF and Raman spectroscopy. When the spectral information is directly used, clustering techniques are usually applied prior to any other chemometric technique, such as regression analysis. For example, PCA has been performed before regression to detect spectral outliers [29] or to reduce the amount of spectral data to deal with [26,30]. This approach is important in NIR or VIS-NIR (Visible-Near Infrared) hyperspectral imaging, since PCA has also been used for reducing the spectral dimension of the images [31]. Clustering techniques are also useful to define correctly the calibration and validation sets before regression analysis [22,25] or to compare the results obtained by different spectroscopic techniques [32]. If concentration values calculated from IR, XRF or Raman spectra are also considered, the

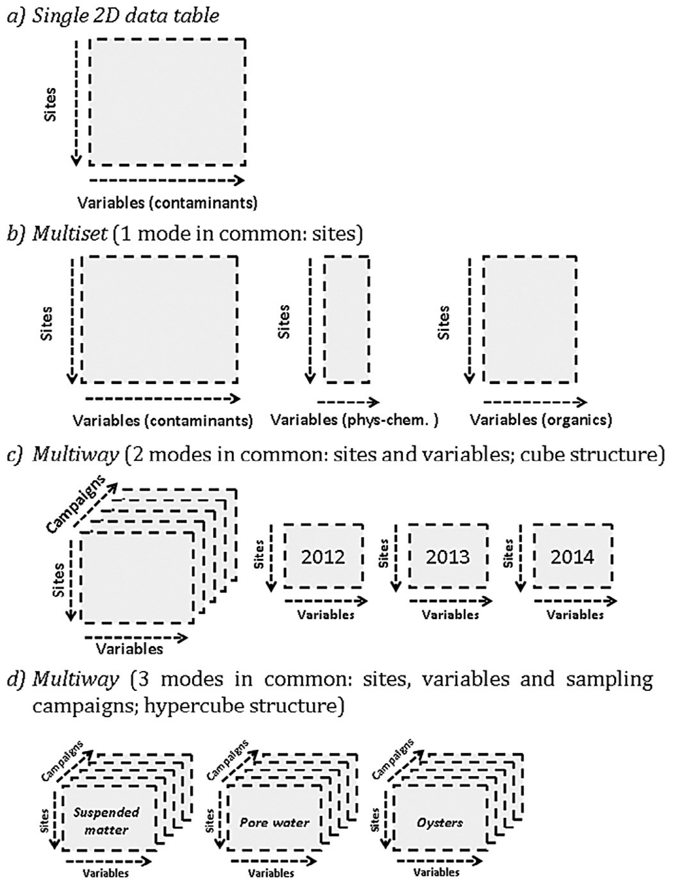

Fig. 3. Different ways to organise data: several compartments (suspended matter, pore waters and oysters) have been sampled in different sampling campaigns (2012, 2013 and 2014) at different sites, and several variables have been measured in the samples (concentrations of metals, organics and physico-chemical parameters).

use of clustering techniques significantly increases. In this case, the identification of contamination sources is a common objective. XRF is the technique most widely used with this purpose. Clustering is a good alternative for the analysis of airborne particles. For instance, XRF and ED-XRF (Energy Dispersive XRF) were combined with PCA to identify sources of aerosol particles both in rural areas and cities [33,34].

Clustering methods are divided into two groups: those based on distance or vicinity and those based on factor analysis. The first group includes CA in its different modalities, that is, single linkage method, k-means clustering, etc. The most popular technique based on Factor Analysis is PCA. Due to the complexity of the results obtained after the application of clustering techniques to spectral data, it is difficult to find research works using spectral wavelengths as variables. This is more evident in works using XRF [35] and Raman spectroscopy [36]. The combination of IR spectra and clustering techniques pursues different objectives. In soils, for example, IR has been used to characterise samples using PCA as classification tool or to explore the structural variability of soils [37,38]. In plants, green tools for mine phytoremediation strategies were identified by NIR and PCA [39] and discrimination of cultivars was attempted using olive leaves and the combination between FT-MIR and PCA [40]. CA and PCA combined with IR appear in the bibliography as a clustering tool for the study of different natural environments, such as seaside areas [41] and estuaries [42]. XRF and Raman spectra have also been occasionally used together with CA to define clustering models in river sediments and complex biological samples [43].

- The *classification techniques* use information about a characteristic that represents a difference among samples, so that a classification can be proposed. New samples can be ascribed to one or another group according to the proposed model. Such methods include Linear Discriminant Analysis (LDA) [44], k-Nearest Neighbour (k-NN) [45], Partial Least Squares Discriminant Analysis (PLS-DA) [46], and different variants of Neural Networks, such as Artificial Neural Networks (ANN), among others. Although they are good alternatives for the analysis of environmental samples, their use is not so popular, and in most cases the suitability of the result obtained by a classification technique is checked by the simultaneous use of a reference clustering technique, such as PCA [47].
- Finally, regression techniques are used in multivariate calibration to establish the relation between two matrices by means of a mathematical model. Each matrix contains different information about the same group of objects (samples). Most often, the first matrix contains information obtained by means of a method to be replaced, and the second matrix contains information obtained by an alternative method. In environmental analytical chemistry, there may be several reasons for substituting a method. The method may be costly, it may fail to comply with the concepts of *Green Chemistry* or the analytical process may be long and arduous [48].

The first step in the calibration process is to obtain information about the samples that will be used as calibrants (see Fig. 2). On the one hand, information obtained using the method to be substituted (usually the analyte concentrations) and, on the other hand, information provided by an alternative method (in spectroscopy, a spectrum).

The second step requires the pretreatment of the information obtained. Spectral data must be pretreated (baseline correction,. . .) before further manipulation. Sometimes, it is necessary to select the part of the spectrum that contains representative information of the sample, removing wavelengths with minimum or no information at all, that only add noise and distort the model rather than improving it [49].

The third step (calibration) involves regression, validation and prediction. First of all, samples must be divided into two subsets, the training (or calibration) set and the validation set. Each subset must be representative of the experimental conditions considered. To this end, different algorithms able to deal with the covariance between the analyte concentration data and the spectral data are normally used. The training data set is used to construct a mathematical model by regression analysis and this model is then validated through the validation or test set in order to check its quality [48]. Later on, in prediction the variables corresponding to unknown samples (the prediction set, completely independent from the training and validation sets) are predicted using the previously validated mathematical models.

The most popular multivariate regression techniques include Multilinear Regression (MLR), Principal Component Regression (PCR) and Partial Least Squares (PLS) [7]. There are some other techniques derived from PLS or MLR, such as, Genetic Algorithms Partial Least Squares (GA-PLS), Moving Window Partial Least Squares (MW-PLS), Uninformative Variable Elimination Partial Least Squares (UVE-PLS), Successive Projections Algorithm Multiple Linear Regression (SPA-MLR) and Equidistant combination Multiple Linear Regression (EC-MLR) [50]. All these techniques are based on PLS, or MLR, but they all include a previous step which investigates all the possible combinations among variables (spectral wavelengths in this case) to design the best calibration model. As it can be observed in Tables 1 and 2, two or more regression methods are often used

Table 1
Research works, published between 2009 and 2014, that include regression analysis of datasets with spectral data obtained from the analysis of solid environmental samples using non-destructive spectroscopic techniques (“Compare results” makes reference to the results obtained by means of different regression techniques). In this compilation biological samples are not included (see Table 2)

Matrix	Regression technique	Spectral range	Predicted variable	Reference
Soil	PLS, SVM	VIS-NIR	Soil organic Carbon/Compare results	[51]
Soil	PLS	NIR	Metal (Zn, Cu, Pb, Cr and Ni) content	[52]
Soil	PLS	MIR	As concentration	[29]
Soil	PLS, UVE-PLS, GA-PLS	VIS-NIR	Soil organic carbon and hot water-extractable carbon/Compare results	[53]
Soil	PLS	NIR, MIR	Total carbon and nitrogen, C/N ratio, substrate-induced respiration and denitrifying enzyme activity	[54]
Soil	PLS	MIR	Organic and inorganic carbon	[55]
Soil	PLS	VIS-NIR	Soil organic carbon	[56]
Soil	PLS, GRNN	MIR	Carbonate content/Compare results	[57]
Soil	PLS	NIR	Zn concentration	[58]
Soil	PLS, PLS-BPNN	VIS-NIR	Soil organic matter/Compare results	[59]
Soil	PLS, SML, SVM	VIS-NIR	Soil total nitrogen/Compare results	[60]
Soil	PLS	MIR, XRF	Metal concentration (Al, As, Ba, Ca, Ce, Co, Cr, Cs, Cu, Fe, Ga, Hf, K, La, Mg, Mn, Na, Nb, Ni, P, Pb, Rb, Sc, Si, Sr, Th, Ti, V, Y, Zn and Zr)	[61]
Soil	PLS	VIS-NIR	Cu contamination	[62]
Soil	PLS	NIR	Hot water extractable organic carbon and $K_2Cr_2O_7$ oxidizable organic carbon	[63]
Soil	PLS	NIR, MIR	Total organic carbon, total nitrogen, cation exchange capacity, electric conductivity and pH	[64]
Soil	PLS	VIS-NIR	Total nitrogen, total carbon, soil organic carbon and soil inorganic carbon	[65]
Soil	PLS	MIR, XRF	Carbonate content	[66]
Soil	PLS	VIS-NIR	Sand, silt and clay, organic carbon, total nitrogen, cation exchange capacity and calcium carbonate	[67]
Subsoil	PLS	MIR, XRD	Ratio of cation exchange capacity to clay	[68]
Soil	PLS, PCR	ED-XRF	Elemental concentration (V, Mn, Co, Ni, Cu, Zn, As, Sr, Ba, Pb and Pb)/Compare results	[69]
Soil	PLS	MIR	Soil properties (labile and stable inorganic and organic phosphorous forms and contents of organic carbon)	[70]
Soil	PLS	VIS-NIR	Investigate the effect of moisture content on the prediction accuracy of soil organic carbon and pH	[71]
Soil	PLS	MIR	Soil properties (pH, total carbon, total nitrogen, extractable phosphorus, exchangeable potassium, calcium, magnesium and texture)	[72]
Soil	PLS	NIR	Organic carbon, cationic exchange capacity and clay and sand contents	[30]
Soil	PLS, GA-PLS, SVM	VIS-NIR	Thermolabile organic carbon, inert organic carbon fraction and the sum of the total soil organic carbon/Compare results	[73]
Soil	CARS-PLS	VIS-NIR, MIR	Organic carbon, total nitrogen, microbial biomass-carbon, pH and hot water extractable carbon	[74]
Soil	PLS	NIR, MIR	Soil organic carbon and total nitrogen and C/N ratio	[75]
Soil	PLS	MIR	Soil organic carbon, soil inorganic carbon as carbonate, total carbon content, cation exchange capacity and pH	[76]
Soil	PLS, LS-SVM, LWR	MIR	Sand, silt and clay contents and total carbon and total nitrogen/Compare results	[77]
Soil	PLS, BRT	VIS-NIR	Total petroleum hydrocarbon/Compare results	[78]
Soil	PLS	VIS-NIR	Soil organic carbon	[79]
Soil	PLS	NIR	Soil texture, cation-exchange capacity, pH, total carbon and total nitrogen, organic carbon and potentially mineralizable nitrogen	[80]
Soil	PLS	MIR	Detect changes in $^{14}NO_3$-N and $^{15}NO_3$-N concentrations	[81]
Soil	PLS	NIR	Total organic carbon, total nitrogen content, organic carbon in humic substances, and concentrations of NH_4^+, Ca^{2+}, Mg^{2+}, K^+ and phosphorus	[82]
Soil	PLS	VIS-NIR	Soil organic and inorganic carbon	[83]
Soil	PLS	Raman	Phosphorus concentration	[84]
Soil	PLS, ANN	ED-XRF	Micronutrients (Fe, Cu, and Zn) and macronutrients (NO_3^-, SO_4^{2-}, and $H_2PO_4^-$)/ Compare results	[26]
Soil	PLS, PCR	XRF	Chromium speciation/Compare results	[85]
Soil	EC-MLR, MW-PLS, SPA-MLR	NIR	Organic matter/Compare results	[50]
Soil	PLS, ANN	MIR	Total carbon and total nitrogen/Compare results	[86]
Soil	PLS, ANN	ED-XRF	C, N, Na, Mg, P, Fe, Cu, Zn, NO_3^-, $H_2PO_4^-$, SO_4^{2-}/Compare results	[87]
Coal	PLS	NIR	Total moisture, inherent moisture, ash, volatile matter, fixed carbon and sulphur	[88]
Poultry manure	PLS, ANN	NIR	Ammonia nitrogen, total potassium, total nitrogen, total phosphorus/Compare results	[89]
Sediment	PLS	XRF	Al, As, Cd, Co, Cr, Cu, Fe, Mg, Mn, Ni, Pb, Sn, V and Zn	[25]
Sediment	PLS	NIR, MIR	Al, As, Cd, Co, Cr, Cu, Fe, Mg, Mn, Ni, Pb, Sn, V and Zn	[90]
Sediment	PLS	MIR	Polyphosphate content	[91]
Sediment	PLS	VIS-NIR, NIR, MIR	Organic and inorganic carbon, biogenic silica and total nitrogen	[92]
Sediment	PLS	MIR	As, Cd, Cr, Pb, Sb and Sn and acid soluble, reducible and oxidizable Cr fractions, elemental composition (carbon, nitrogen and hydrogen content), pH, redox potential and the amount of fulvic and humic acids	[22]

in the same work. PLS is the regression technique preferably used, with more than 300 publications in the last 5 years (without including biota). If works dealing with PLS models are sorted according to the environmental matrix studied, the 80% is concerned with the analysis of solid samples, and only 2 out of 10 investigate other solid sample than soil (e.g. sediment). Moreover, although it should not be used as a reference, when any other regression technique different from PLS is applied, PLS is usually employed to check the reliability

Table 2
Research works, published between 2009 and 2014, that include regression analysis of datasets with spectral data obtained from the analysis of biological and non-solid environmental samples using non-destructive spectroscopic techniques ("Compare results" makes reference to the results obtained by means of different regression techniques)

Matrix	Regression Technique	Spectral range	Predicted variable	Reference
Plant (Grass)	PLS	NIR	Total Nitrogen content	[93]
Plant (Flos Lonicerae)	PLS, GA-PLS, PCR	NIR	Bacteriostatic activity/Compare results	[94]
Plant (Eucalyptus urophylla)	PLS	NIR, XRD	The microfibril angle of crystalline cellulose	[95]
Plant (Bog)	PLS	NIR	δ^{13}C and δ^{15}N, total carbon and nitrogen	[96]
Plant (Ligustrum japonicum)	PLS	NIR hyperspectral imaging	Water content	[97]
Plant (Auricularia auricular)	PLS, MLR, LS-SVM, SPA-MLR, SPA-PLS, SPA-LS-SVM	NIR	Protein content/Compare results	[98]
Plant (Tobacco)	PLS	NIR	Three phenolic compounds: Chlorogenic acid, scopoletin and rutin/Compare results	[99]
Plant (Epipremnum aureum)	PLS	NIR hyperspectral imaging	Water content	[100]
Plant (Bamboo)	PLS, MLR, BP-ANN, LS-SVM	VIS-NIR	Antioxidant activity indices/Compare results	[101]
Plant (Acacia mearnsii)	PLS	NIR	Tannin content	[20]
Plant (Rosa laevigata)	UVE-PLS, UVE-GA-PLS	NIR	Polysaccharide content/Compare results	[102]
Bacteria (Lactobacillus kefir)	PLS	Raman	Quantification of metal ions	[19]
Fragrant mushroom	PLS, LS-SVM	MIR	Protein/Compare results	[103]
Water (wastewater)	PLS	MIR	Petroleum hydrocarbons	[104]
Water	PLS	NIR	Hydrocarbons	[105]
Water (tanned solid wastes)	PLS	NIR	Hydrocarbons	[106]
Water	PLS	MIR	Pyrethroid and organochlorine pesticides	[107]
Atmosphere (Fog and rain)	PLS	IR	Diethyl ether	[108]

of the results. In those cases, as in all the fields of analytical chemistry, a previously validated methodology should be used to check the applicability of the proposed method.

Similar works concerning water or air samples are scarce. The organic matter content in soils is clearly one of the variables more frequently estimated, but other properties such as pH or nitrogen content are also usually calculated. IR spectroscopy is the preferred technique in most of the approaches. The lack of specificity of IR spectral bands, with only vibrational information, increases the need of combining IR with multivariate calibration methods [11].

As can be seen in Table 1, the use of VIS-NIR, NIR and MIR for the analysis of solid samples have been similar. The objectives pursued, in addition, are similar in all the cases, that is, the prediction of physico-chemical properties of soils or the design of models to predict the elemental composition of soils. While some authors claim that the models with more predictive ability are those based on MIR [74], some others affirm that better results are obtained in the NIR region [54]. It can be concluded that the use of IR spectra is a good approach to predict soil properties whatever it is the region employed. The spectral range might be sample specific. Moreover, an appropriate pretreatment of data and a correct definition of the datasets for calibration, validation and prediction are crucial steps in the development of high quality models for the prediction of physico-chemical properties of solid samples. Despite the fact that the relation between NIR and/or MIR spectra of a soil or sediment and its elemental content is not evident, PLS models based on VIS-NIR, NIR and MIR measurements have been proposed to estimate the concentration of trace and major elements in solid samples with satisfactory results [22]. Analytical techniques to be replaced in these cases always include an acid extraction step followed by detection by Atomic Absorption Spectrometry (AAS), Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES) or Inductively Coupled Plasma Mass spectrometry (ICP-MS) [22,25,26,47,85].

If we focus on the different IR spectrometers that are currently on the market, few publications are based on transmission measurements [66]. This fact is not characteristic of environmental samples. Reflectance methods have replaced the traditional analysis using transmittance IR. A notable amount of publications do not specify which has been the type of reflectance spectrometry used. This is more common in works focused on non-solid environmental samples. Specular Reflectance Spectroscopy is not usual in the analysis of environmental samples.

DRIFT is undoubtedly the technique preferably chosen for the analysis of both solid [29,56,61,64,68,74,76,92] and non-solid environmental samples [95]. DRIFT and ATR are both affected by differences in particle size, but ATR allows a direct analysis of the sample using a zinc selenide (ZnSe) crystal or diamond, being not necessary to mix the sample before producing a pellet with a hydraulic press. Nonetheless, in ATR measurements the penetration depth of the IR beam is only of a few millimetres and decreases at shorter wavelengths; consequently, the lack of a good contact between the sample and the internal reflection element of the ATR may result in unsatisfactory results. For this reason, DRIFT is more used in environmental studies. The analysis of non-homogeneous samples (which is usually the case of environmental samples) is also preferably attempted by DRIFT.

In regression studies in which environmental samples are involved, ATR is more frequently combined with MIR [22]; however, DRFIT spectrometers have been also properly combined with VIS-NIR [53], NIR [64,90] and MIR [29,57,61] with satisfactory results.

It is not easy to find publications combining the analysis of environmental samples by Raman spectroscopy and regression techniques. Only a few examples devoted to the production of models to predict physico-chemical properties of the samples are worth mentioning. For instance, the relationship between sand soil phosphorus concentration and Raman spectra was explored by PLS (Partial Least Squares) [84]. XRF, on the other side, has been used to predict the concentration of metals and metalloids in samples. Consequently, it can be used as a prediction tool for the ecological assessment of environmental areas [25]. Moreover, energy dispersive X-ray fluorescence scattering (EDXRFS) has been successfully applied to measure macronutrients (NO_3^-, SO_4^{2-}, and $H_2PO_4^-$) in soils [26]. Portable XRF has also been used with similar objectives [109].

Due to its more homogeneous surface, leaf is the plant part which has been more often investigated in this field. Protein or water contents are the variables most often quantified in plants through regression models obtained after spectroscopic measurements. Models to predict nutritional values in fishes have also been proposed [110]. In the case of water, most of the works are concerned with contamination or pollution issues. Some important research works carried out between 2009 and 2014 regarding biota and non-solid samples, regression analysis and non-destructive spectroscopic techniques are summarised in Table 2. Clearly, XRF is not widely used to analyse environmental water samples because, due to the influence of the matrix effect, the sensitivity of XRF is lower than that obtained by ICP-MS or ICP-AES. In XRF analysis a preconcentration step is required, so that this methodology does not offer clear advantages compared to other traditional methods. Some authors claim that Raman spectroscopy is more appropriate than IR for the analysis of water, because water is a “weak Raman scattering but a strong absorber of infrared radiation” [111]. Nevertheless, in practice, works including regression techniques, IR and biological or water samples are more popular. Unlike in solids samples, where NIR and MIR are similarly used, NIR is the spectral range usually chosen to analyse non-solid environmental samples. This is probably because the analysis of non-solid samples by IR is still quite recent. The low cost of the materials, such as glass and quartz, that can be used as cell windows in NIR spectrometry, together with the fact that the signal of water is practically negligible in NIR, are behind the wide use of this IR region to analyse water or plants.

5. Conclusions

IR is for sure the non-destructive spectroscopic technique more frequently used nowadays in the analysis of environmental samples. This is because IR spectroscopy is a relatively cheap and easy to operate technique that provides us with low baseline noise spectra and non-specific bands. The information obtained from IR spectroscopy, however, is rather limited. Although the use of Raman spectroscopy is increasing, especially in the analysis of solid samples, its applications are still limited, mainly due to the high amount of organic matter usually present in environmental samples that typically complicates the acquisition of quality data and the identification of peaks.

The combination with multivariate chemometrics is clearly the reason that explains the rise in the application of non-destructive spectroscopic techniques to the analysis of solid environmental samples. Traditional statistical tools, such as those used to investigate time/spatial series, are not usually combined with non-destructive techniques or, at least, not alone. This is because clustering and classification techniques provide us with information which is not usually provided by correlation analysis or similar methods.

Certainly, soil is the natural compartment more frequently studied. Spectroscopic analysis of water or biota requires specific equipment and, sometimes, a tedious sample pretreatment. Moreover, when solvents are used interferences are prone to appear, making data treatment more difficult. Despite all these aspects, the non-destructive analysis of airborne particles and plants is slowly increasing. This is more obvious in works concerning clustering and IR. Special equipment is not necessarily required to analyse biological and water samples by IR, a few drops of water or biological tissue are enough for a direct analysis by ATR.

Acknowledgments

This work has been financially supported by the Ministry of Economy and Competitiveness of the Spanish Government through the research project DebaSed (Ref. CTM2014-55270-R) and the Basque Government through the Consolidated Research Group Project 2013–2018 (ref. IT-742-13).

References

[1] P.T. Anastas, J.C. Warner, Chemistry: Theory and Practice, Oxford University Press, New York, 1998.
[2] M. de la Guardia, S. Armenta, Green Analytical Chemistry, Elsevier, Great Britain, 2011.
[3] A. Kot, B. Zabiegała, J. Namieśnik, Passive sampling for long-term monitoring of organic pollutants in water, TrAC Trend. Anal. Chem. 19 (2000) 446–459.
[4] S. Garrigues, M. de la Guardia, Non-invasive analysis of solid samples, TrAC Trend. Anal. Chem. 43 (2013) 161–173.
[5] J.S. Becker, H.J. Dietze, Inorganic trace analysis by mass spectrometry, Spectrochim. Acta B 53 (1998) 1475–1506.
[6] R.E. Shepherd, Chromatographic and related electrophoretic methods in the separation of transition metal complexes or their ligands, Coordin. Chem. Rev. 247 (2003) 159–196.
[7] J. Moros, S. Garrigues, M. de la Guardia, Vibrational spectroscopy provides a green tool for multi-component analysis, TrAC Trend. Anal. Chem. 29 (2010) 578–591.
[8] J. Aramendia, L. Gomez-Nubla, I. Arrizabalaga, N. Prieto-Taboada, K. Castro, J.M. Madariaga, Multianalytical approach to study the dissolution process of weathering steel: the role of urban pollution, Corros. Sci. 76 (2013) 154–162.
[9] M. Blanco, A. Peguero, Analysis of pharmaceuticals by NIR spectroscopy without a reference method, TrAC Trend. Anal. Chem. 29 (2010) 1127–1136.
[10] S. Armenta, G. Quintás, S. Garrigues, M. de la Guardia, Mid-infrared and Raman spectrometry for quality control of pesticide formulations, TrAC Trend. Anal. Chem. 24 (2005) 772–781.
[11] S. Armenta, M. de la Guardia, Vibrational spectroscopy in soil and sediment analysis, TrAC Trend. Anal. Chem. 2 (2014) 43–52.
[12] F. Zambello, J. Enzweiler, Multi-element analysis of soils and sediments by wavelength-dispersive X-ray fluorescence spectrometry, J. Soil. Sed. 2 (2002) 29–36.
[13] L. Gomez-Nubla, J. Aramendia, A. Alonso-Olazabal, S. Fdez-Ortiz de Vallejuelo, K. Castro, L.A. Ortega, et al., Darwin impact glass study by Raman spectroscopy in combination with other spectroscopic techniques, J. Raman Spectrosc. (2015) doi:10.1002/jrs.4700.
[14] J.W. Einax, H.W. Zwanzinger, S. Geiss, Chemometrics in environmental analysis, Angewandte Chemie International Education, Weinheim (Germany), 1997.
[15] A. Keshav Krishna, K. Rama Mohan, N.N. Murthy, A multivariate statistical approach for monitoring of heavy metals in sediments: a case study from Wailpalli watershed Nalgonda District, Andhra Pradesh, India, Res. J. Environ. Earth Sci. 3 (2011) 103–113.
[16] T. Akyuz, N. Mukhamedshina, S. Akyuz, E. Sari, A.A. Mirsagatova, Toxic and trace element analysis of surface sediments from the Gulf of Saros by INAA and XRF methods, J. Radioanal. Nucl. Chem. 273 (2007) 747–751.
[17] C.M. Simonescu, Application of FTIR Spectroscopy in Environmental Studies, in: M.A. Farrukh (Editor), Advanced Aspects of Spectroscopy, Pakistan, 2012, pp. 49–84.
[18] Ã.S. Rinnan, F.V.D. Berg, S.R.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TrAC Trend. Anal. Chem. 28 (2009) 1201–1222.
[19] E. Gerbino, P. Mobili, E.E. Tymczyszyn, C. Frausto-Reyes, C. Araujo-Andrade, A. Gomez-Zavaglia, Use of Raman spectroscopy and chemometrics for the quantification of metal ions attached to Lactobacillus kefir, J. Appl. Microbiol. 112 (2012) 363–371.
[20] C.M. Menezes, A.B.D. Costa, R.R. Renner, L.F. Bastos, M.F. Ferrao, V.L. Dressler, Direct determination of tannin in Acacia mearnsii bark using near infrared spectroscopy, Anal. Methods 6 (2014) 8299–8305.
[21] B. Butkute, A. Slepetiene, Application of near infrared reflectance spectroscopy for the assessment of soil quality in a long-term pasture, Commun. Soil Sci. Plant Anal. 37 (2006) 2389–2409.
[22] J. Moros, R.J. Cassella, M.C. Barciela-Alonso, A. Moreda-Piñeiro, P. Herbello-Hermelo, P. Bermejo-Barrera, et al., Estuarine sediment quality assessment by Fourier-transform infrared spectroscopy, Vib. Spectrosc. 53 (2010) 204–213.
[23] N.K. Afseth, V.H. Segtnan, J.P. Wold, Raman spectra of biological samples: a study of preprocessing methods, Appl. Spectrosc. 60 (2006) 1358–1367.
[24] T. Bocklitz, A. Walter, K. Hartmann, P. Rösch, J. Popp, How to pre-process Raman spectra for reliable and stable models?, Anal. Chim. Acta 704 (2011) 47–56.
[25] J. Moros, A. Gredilla, S. Fdez-Ortiz de Vallejuelo, A. de Diego, J.M. Madariaga, S. Garrigues, et al., Partial least squares X-ray fluorescence determination of trace elements in sediments from the estuary of Nerbioi-Ibaizabal River, Talanta 82 (2010) 1254–1260.
[26] M.I. Kaniu, K.H. Angeyo, M.J. Mangala, A.K. Mwala, S.K. Bartilol, Feasibility for chemometric energy dispersive X-ray fluorescence and scattering (EDXRFS) spectroscopy method for rapid soil quality assessment, X-Ray Spectrom. 40 (2011) 432–440.
[27] E. Acar, B. Yener, Unsupervised multiway data analysis: a literature survey, IEEE Trans. Knowl. Data Eng. 21 (2009) 6–20.
[28] S. Mas, A. De Juan, R. Tauler, A.C. Olivieri, G.M. Escandar, Application of chemometric methods to environmental analysis of organic pollutants: a review, Talanta 80 (2010) 1052–1067.
[29] N.K. Niazi, B. Singh, B. Minasny, Mid-infrared spectroscopy and partial least-squares regression to estimate soil arsenic at a highly variable arsenic-contaminated site, Int. J. Environ. Sci. Technol. 12 (2015) 1965–1974.
[30] F. Goge, R. Joffre, C. Jolivet, I. Ross, L. Ranjard, Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database, Chemom. Intell. Lab. Syst. 110 (2012) 168–176.
[31] P.J. Williams, P. Geladi, T.J. Britz, M. Manley, Near-infrared (NIR) hyperspectral imaging and multivariate image analysis to study growth characteristics and

differences between species and strains of members of the genus Fusarium, Anal. Bioanal. Chem. 404 (2012) 1759–1769.

[32] I. Janotkova, L. Prokes, T. Vaculovic, M. Hola, J. Pinkas, I. Steffan, et al., Comparison of inductively coupled plasma optical emission spectrometry, energy dispersive X-ray fluorescence spectrometry and laser ablation inductively coupled plasma mass spectrometry in the elemental analysis of agricultural soils, J. Anal. At. Spectrom. 28 (2013) 1940–1948.

[33] I.J.K. Aboh, D. Henriksson, J. Laursen, M. Lundin, F.G. Ofosu, N. Pind, et al., Identification of aerosol particle sources in semi-rural area of Kwabenya, near Accra, Ghana, by EDXRF techniques, X-Ray Spectrom. 38 (2009) 348–353.

[34] D. Song, C. Yang, Geochemistry and source apportionment of atmospheric particulate matter in Jiaozuo city, Appl. Mech. Mater. 71–78 (2011) 2867–2872.

[35] M.B.B. Guerra, E.P. Limaneto, M.T. Antunes, E.R. Pereira-Filho, C.E.G.R. Schaefer, Post-fire study of the Brazilian Scientific Antarctic Station: toxic element contamination and potential mobility on the surrounding environment, Microchem. J. 110 (2013) 21–27.

[36] O.E. Rivera-Betancourt, R. Karls, B. Grosse-Siestrup, S. Helms, F. Quinn, R.A. Dluhy, Identification of mycobacteria based on spectroscopic analyses of mycolic acid profile, Analyst 138 (2013) 6774–6785.

[37] A.P. Fernandez-Getino, Z. Hernandez, A. Piedra Buena, G. Almendros, Exploratory analysis of the structural variability of forest soil humic acids based on multivariate processing of infrared spectral data, Eur. J. Soil Sci. 64 (2013) 66–79.

[38] N. Jing, X. Jiang, Q. Wang, Y. Tang, P. Zhang, Attenuated total reflectance/Fourier transform infrared (ATR/FTIR) mapping coupled with principal component analysis for the study of in vitro degradation of porous polylactide/hydroxyapatite composite material, Anal. Methods 6 (2014) 5590–5595.

[39] N.S. Mokgalaka-Matlala, T. Regnier, S. Combrinck, C.R. Kouekam, I.M. Weiersbye, Chemometrics and vibrational spectroscopy as green tools for mine phytoremediation strategies, Spectrochim. Acta A 100 (2013) 138–143.

[40] F. Aouidi, N. Dupuy, J. Artaud, S. Roussos, M. Msallem, I. Perraud-Gaime, et al., Discrimination of five Tunisian cultivars by Mid InfraRed spectroscopy combined with chemometric analyses of olive Olea europaea leaves, Food Chem. 131 (2012) 360–366.

[41] M. Mecozzi, E. Pietrantonio, Carbohydrates proteins and lipids in fulvic and humic acids of sediments and its relationships with mucilaginous aggregates in the Italian seas, Mar. Chem. 101 (2006) 27–39.

[42] J. Subburaj, T.R. Barathkumar, V.T. Prem, M. Hangaraj, J. Sivasubramanian, Spectroscopic analysis of five phylogenetically distant fungi (Division: Ascomycete) from Vellar estuary, southeast coast of India – a pilot study, J. Microbiol. Biotechnol. Food Sci. 2 (2013) 2355–2359.

[43] A.K. Kniggendorf, T.W. Gaul, M. Meinhardt-Wollweber, Hierarchical cluster analysis (HCA) of microorganisms: an assessment of algorithms for resonance Raman spectra, Appl. Spectrosc. 65 (2011) 165–173.

[44] A. Walter, S. Kuhri, M. Reinicke, T. Bocklitz, W. Schumacher, P. Rösch, et al., Raman spectroscopic detection of Nickel impact on single Streptomyces cells – possible bioindicators for heavy metal contamination, J. Raman Spectrosc. 43 (2012) 1058–1064.

[45] L.J. Ni, L.G. Zhang, J. Xie, J.Q. Luo, Pattern recognition of Chinese flue-cured tobaccos by an improved and simplified K-nearest neighbours classification algorithm on near infrared spectra, Anal. Chim. Acta 633 (2009) 43–50.

[46] S.C. Park, M. Kim, J. Park, H. Chung, H.Y. Kim, Wide area illumination Raman scheme for simple and nondestructive discrimination of seawater cultured pearls, J. Raman Spectrosc. 40 (2009) 2187–2192.

[47] M.H. Eltayeb, J. Injuk, W. Maenhaut, R.E. Van Grieken, Elemental Composition of Mineral Aerosol Generated from Sudan Sahara Sand, J. Atmos. Chem. 40 (2001) 247–273.

[48] K.M. Esbensen, Mulitivariate Data Analysis – In Practice, CAMO Process AS, Esbjerg (Denmark), 2001.

[49] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in Partial Least Squares Regression, Chemom. Intell. Lab. Syst. 118 (2012) 62–69.

[50] T. Pan, M. Li, J. Chen, Selection method of quasi-continuous wavelength combination with applications to the near-infrared spectroscopic analysis of soil organic matter, Appl. Spectrosc. 68 (2014) 263–271.

[51] Y. Gao, L. Cui, B. Lei, Y. Zhai, T. Shi, J. Wang, et al., Estimating soil organic carbon content with visible-near-infrared (Vis-NIR) spectroscopy, Appl. Spectrosc. 68 (2014) 712–722.

[52] M. Todorova, A.M. Mouazen, H. Lange, S. Atanassova, Potential of near-infrared spectroscopy for measurement of heavy metals in soil as affected by calibration set size, Water Air Soil Pollut. 225 (2014) 1–19.

[53] M. Vohland, C. Emmerling, Determination of total soil organic C and hot water-extractable C from VIS-NIR soil reflectance with partial least squares regression and spectral feature selection techniques, Eur. J. Soil Sci. 62 (2011) 598–606.

[54] M. Rabenarivo, L. Chapuis-Lardy, D. Brunet, J.L. Chotte, L. Rabeharisoa, B.G. Barthes, Comparing near and mid-infrared reflectance spectroscopy for determining properties of Malagasy soils, using global or LOCAL calibration, J. Near Infrared Spectrosc. 21 (2013) 495–509.

[55] J.A. Baldock, B. Hawke, J. Sanderman, L.M. Macdonald, Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra, Soil Res. 51 (2013) 577–595.

[56] M. Nocita, A. Stevens, G. Toth, P. Panagos, B. Van Wesemael, L. Montanarella, Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach, Soil Biol. Biochem. 68 (2014) 337–347.

[57] C. Du, Z. Ma, J. Zhou, K.W. Goyne, Application of mid-infrared photoacoustic spectroscopy in monitoring carbonate content in soils, Sens. Actuators B 188 (2013) 1167–1175.

[58] L. Yao, J. Chen, T. Pan, Rapid analysis of Zinc in soil with long-NIR spectroscopy, Appl. Mech. Mater. 365–366 (2013) 733–736.

[59] Y. Tian, J. Zhang, X. Yao, W. Cao, Y. Zhu, Laboratory assessment of three quantitative methods for estimating the organic matter content of soils in China based on visible/near-infrared reflectance spectra, Geoderma 202–203 (2013) 161–170.

[60] T. Shi, L. Cui, J. Wang, T. Fei, Y. Chen, G. Wu, Comparison of multivariate methods for estimating soil total nitrogen with visible/near-infrared spectroscopy, Plant Soil 366 (2013) 363–375.

[61] J.M. Soriano-Disla, L. Janik, M.J. McLaughlin, S. Forrester, J. Kirby, C. Reimann, The use of diffuse reflectance mid-infrared spectroscopy for the prediction of the concentration of chemical elements estimated by X-ray fluorescence in agricultural and grazing European soils, Appl. Geochem. 29 (2013) 135–143.

[62] Y. Liu, Y. Chen, Feasibility of estimating cu contamination in floodplain soils using VNIR spectroscopy-a case study in the le'an river floodplain, China, Soil Sediment Contam. 21 (2012) 951–969.

[63] R. Fan, X. Zhang, Y. Shen, X. Yang, A. Liang, Near-infrared spectroscopic assessment of hot water extractable and oxidizable organic carbon in cultivated and uncultivated Mollisols in China, Acta Agr. Scand. Sect. B-Soil PL 63 (2013) 37–45.

[64] L. Han, D. Li, W. Fang, Y. Wang, R. Gaussoin, Analysis of soil chemical properties of sand-based turfgrass rootzone using fourier transform-infrared spectroscopy, Commun. Soil Sci. Plan. 43 (2012) 2709–2721.

[65] H. Yang, B. Kuang, A.M. Mouazen, Quantitative analysis of soil nitrogen and carbon at a farm scale using visible and near infrared spectroscopy coupled with wavelength reduction, Eur. J. Soil Sci. 63 (2012) 410–420.

[66] V.J. Bruckman, K. Wriessnig, Improved soil carbonate determination by FT-IR and X-ray analysis, Environ. Chem. Lett. 11 (2013) 65–70.

[67] A.P. Leone, R.A. Viscarra-Rossel, P. Amenta, A. Buondonno, Prediction of soil properties with PLSR and vis-NIR spectroscopy: application to Mediterranean soils from southern Italy, Curr. Anal. Chem. 8 (2012) 283–299.

[68] P.M. Bloesch, Prediction of the CEC to clay ratio using mid-infrared spectroscopy, Soil Res. 50 (2012) 1–6.

[69] S. Akbulut, Validation of classical quantitative fundamental parameters method using multivariate calibration procedures for trace element analysis in ED-XRF, J. Anal. At. Spectrom. 29 (2014) 853–860.

[70] D. Rueckamp, C. Martius, L. Bornemann, D. Kurzatkowski, L.P. Naval, W. Amelung, Soil genesis and heterogeneity of phosphorus forms and carbon below mounds inhabited by primary and secondary termites, Geoderma 170 (2012) 239–250.

[71] Y. Tekin, Z. Tumsavas, A.M. Mouazen, Effect of moisture content on prediction of organic carbon and pH using visible and near-infrared spectroscopy, Soil Sci. Soc. Am. J. 76 (2012) 188–198.

[72] S.I. Muhati, K.D. Shepherd, C.K. Gachene, M.W. Mburu, R. Jones, G.O. Kironchi, et al., Diagnosis of soil nutrient constraints in small-scale groundnut (Arachis hyopaea L.) production systems of Western Kenya using infrared spectroscopy, J. Agric. Sci. Technol. (2011) 111–127.

[73] M. Vohland, J. Besold, J. Hill, H.C. Fruend, Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy, Geoderma 166 (2011) 198–205.

[74] M. Vohland, M. Ludwig, S. Thiele-Bruhn, B. Ludwig, Determination of soil properties with visible to near- and mid-infrared spectroscopy: effects of spectral variable selection, Geoderma 223-225 (2014) 88–96.

[75] H.T. Xie, X.M. Yang, C.F. Drury, J.Y. Yang, X.D. Zhang, Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for brookston clay loam soil in Southwestern Ontario, Canada, Can. J. Soil Sci. 91 (2011) 53–63.

[76] L.P. D'acqui, A. Pucci, L.J. Janik, Soil properties prediction of western Mediterranean islands with similar climatic environments by means of mid-infrared diffuse reflectance spectroscopy, Eur. J. Soil Sci. 61 (2010) 865–876.

[77] B. Igne, J.B. Reeves III, G. McCarty, W.D. Hively, E. Lund, C.R. Hurburgh, Evaluation of spectral pretreatments, partial least squares, least squares support vector machines and locally weighted regression for quantitative spectroscopic analysis of soils, J. Near Infrared Spectrosc. 18 (2010) 167–176.

[78] S. Chakraborty, D.C. Weindorf, C.L.S. Morgan, Y. Ge, J.M. Galbraith, B. Li, et al., Rapid identification of oil-contaminated soils using visible near-infrared diffuse reflectance spectroscopy, J. Environ. Qual. 39 (2010) 1378–1387.

[79] G.M. Vasques, S. Grunwald, W.G. Harris, Spectroscopic models of soil organic carbon in Florida, USA, J. Environ. Qual. 39 (2010) 923–934.

[80] C. Nduwamungu, N. Ziadi, G.F. Tremblay, L.E. Parent, Near-infrared reflectance spectroscopy prediction of soil properties: effects of sample cups and preparation, Soil Sci. Soc. Am. J. 73 (2009) 1896–1903.

[81] C. Du, R. Linker, A. Shaviv, J. Zhou, In situ evaluation of net nitrification rate in Terra rossa soil using a Fourier transform infrared attenuated total reflection 15N tracing technique, Appl. Spectrosc. 63 (2009) 1168–1173.

[82] A. Vergnoux, N. Dupuy, M. Guiliano, M. Vennetier, F. Theraulaz, P. Doumenq, Fire impact on forest soils evaluated using near-infrared spectroscopy and multivariate calibration, Talanta 80 (2009) 39–47.

[83] C.L.S. Morgan, T.H. Waiser, D.J. Brown, C.T. Hallmark, Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy, Geoderma 151 (2009) 249–256.

[84] L. Zheng, W.S. Lee, M. Li, A. Katti, C. Yang, H. Li, et al., Analysis of soil phosphorus concentration based on Raman spectroscopy, Proc. SPIE 8527 (2012) 852718/1–852718/8.
[85] J. Malherbe, F. Claverie, Toward chromium speciation in solids using wavelength dispersive X-ray fluorescence spectrometry Cr Kβ lines, Anal. Chim. Acta 773 (2013) 37–44.
[86] G. Aslan-Sungur, F. Evrendilek, N. Karakaya, K. Gungor, S. Kilic, Integrating ATR-FTIR and data-driven models to predict total soil carbon and nitrogen towards sustainable watershed management, Res. J. Chem. Environ. 17 (2013) 5–11.
[87] M.I. Kaniu, K.H. Angeyo, A.K. Mwala, F.K. Mwangi, Energy dispersive X-ray fluorescence and scattering assessment of soil quality via partial least squares and artificial neural networks analytical modeling approaches, Talanta 98 (2012) 236–240.
[88] Y. Wang, M. Yang, G. Wei, R. Hu, Z. Luo, G. Li, Improved PLS regression based on SVM classification for rapid analysis of coal properties by near-infrared reflectance spectroscopy, Sens. Actuators B Chem. 193 (2014) 723–729.
[89] L.J. Chen, L. Xing, L.J. Han, Quantitative determination of nutrient content in poultry manure by near infrared spectroscopy based on artificial neural networks, Poult. Sci. 88 (2009) 2496–2503.
[90] J. Moros, S. Fdez-Ortiz de Vallejuelo, A. Gredilla, A. de Diego, J.M. Madariaga, S. Garrigues, et al., Use of reflectance infrared spectroscopy for monitoring the metal content of the estuarine sediments of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country), Environ. Sci. Technol. 43 (2009) 9314–9320.
[91] A. Khoshmanesh, P.L.M. Cook, B.R. Wood, Quantitative determination of polyphosphate in sediments using Attenuated Total Reflectance-Fourier Transform Infrared (ATR-FTIR) spectroscopy and partial least squares regression, Analyst 137 (2012) 3704–3709.
[92] A. Hahn, P. Rosen, P. Kliem, C. Ohlendorf, B. Zolitschka, Comparative study of infrared techniques for fast biogeochemical sediment analyses, Geochem. Geophys. Geosyst. 12 (2011) 1–8.
[93] N. Shetty, A. Rinnan, R. Gislum, Selection of representative calibration sample sets for near-infrared reflectance spectroscopy to predict nitrogen concentration in grasses, Chemom. Intell. Lab. Syst. 111 (2012) 59–65.
[94] N. Xin, Q. Meng, Y. Li, Y. Hu, Near infrared spectral similarity combined with variable selection method in the quality control of Flos Lonicerae: a preliminary study, Chin. J. Chem. 29 (2012) 2533–2540.
[95] P.R.G. Hein, B. Clair, L. Brancheriau, G. Chaix, Predicting microfibril angle in Eucalyptus wood from different wood faces and surface qualities using near infrared spectra, J. Near Infrared Spectrosc. 18 (2010) 455–464.
[96] T. Kleinebecker, S.R. Schmidt, C. Fritz, A.J.P. Smolders, N. Hoelzel, Prediction of δ13C and δ15N in plant tissues with near-infrared reflectance spectroscopy, New Phytol. 184 (2009) 732–739.
[97] H. Kobori, S. Tsuchikawa, Prediction of water content in Ligustrum japonicum leaf using near infrared chemometric imaging, J. Near Infrared Spectrosc. 17 (2009) 151–157.
[98] F. Liu, Y. He, G. Sun, Determination of protein content of auricularia auricula using near infrared spectroscopy combined with linear and nonlinear calibrations, J. Agric. Food Chem. 57 (2009) 4520–4527.
[99] Z. Mao, R. Shan, J. Wang, W. Cai, X. Shao, Optimizing the models for rapid determination of chlorogenic acid, scopoletin and rutin in plant samples by near-infrared diffuse reflectance spectroscopy, Spectrochim. Acta A 128 (2014) 711–715.
[100] S. Higa, H. Kobori, S. Tsuchikawa, Mapping of leaf water content using near-infrared hyperspectral imaging, Appl. Spectrosc. 67 (2013) 1302–1307.
[101] D. Wu, J. Chen, B. Lu, L. Xiong, Y. He, Y. Zhang, Application of near infrared spectroscopy for the rapid determination of antioxidant activity of bamboo leaf extract, Food Chem. 135 (2012) 2147–2156.
[102] H. Yan, B.X. Han, Q.Y. Wu, M.Z. Jiang, Z.Z. Gui, Rapid detection of Rosa laevigata polysaccharide content by near-infrared spectroscopy, Spectrochim. Acta A 79 (2011) 179–184.
[103] P.C. Nie, Y. Yang, Y. He, Comparison of calibrations modeling for determination of protein of mushroom base on middle infrared spectroscopy, Key Eng. Mater. 460-461 (2011) 357–362.
[104] M.D. Ruiz, I.T. Bustamante, A. Dago, N. Hernandez, A.C. Nunez, D. Porro, A multivariate calibration approach for determination of petroleum hydrocarbons in water by means of IR spectroscopy, J. Chemom. 24 (2010) 444–447.
[105] A. Gammoun, S. Tahiri, A. Albizane, M. Azzi, J. Moros, S. Garrigues, et al., Separation of motor oils, oily wastes and hydrocarbons from contaminated water by sorption on chrome shavings, J. Hazard. Mater. 145 (2007) 148–153.
[106] A. Gammoun, J. Moros, S. Tahiri, S. Garrigues, M. de la Guardia, Partial least-squares near-infrared determination of hydrocarbons removed from polluted waters by using tanned solid wastes, Anal. Bioanal. Chem. 385 (2006) 766–770.
[107] A. Colume, J.L. Diewok, B. Lendl, Assessment of FTIR spectrometry for pesticide screening of aqueous samples, Int. J. Environ. Anal. Chem. 84 (2004) 835–844.
[108] L. Shao, C.W. Roske, P.R. Griffiths, Detection of chemical agents in the atmosphere by open-path FT-IR spectroscopy under conditions of background interference: II. Fog and rain, Anal. Bioanal. Chem. 397 (2010) 1521–1528.
[109] A. Sharma, D.C. Weindorf, T. Man, A.A.A. Aldabaa, S. Chakraborty, Characterizing soils via portable X-ray fluorescence spectrometer: 3. Soil reaction (pH), Geoderma 232–234 (2014) 141–147.
[110] K. Khodabux, M.S.S. L'omelette, S. Jhaumeer-Laulloo, P. Ramasami, P. Rondeau, Chemical and near-infrared determination of moisture, fat and protein in tuna fishes, Food Chem. 102 (2007) 669–675.
[111] S. Ewen, D. Geoffrey, Modern Raman Spectroscopy: A Practical Approach, Wiley, Chichester (England), 2005.