JAAS

Identification and classification of meteorites by a handheld LIBS instrument coupled with a fuzzy logic-based method

Journal:	*Journal of Analytical Atomic Spectrometry*
Manuscript ID	Draft
Article Type:	Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Senesi, Giorgio; CNR, Istituto di Nanotecnologia (NANOTEC) - PLasMI Lab Manzari, Paola; Istituto Nazionale di Astrofisica - Istituto di Astrofisica e Planetologia Spaziali (INAF-IAPS) Consiglio, Arianna; CNR, Istituto di Tecnologie Biomediche (ITB) De Pascale, Olga; CNR, Istituto di Nanotecnologia (NANOTEC) - PLasMI Lab

Consiglio Nazionale delle Ricerche

CNR**NANOTEC** - Istituto di Nanotecnologia

Partita IVA IT 02118311006 – C.F. 80054330586
PEC: protocollo.nanotec@pec.cnr.it

Bari, Italia
June 1, 2018

Dear Editor,

Enclosed please find the manuscript entitled "**Identification and classification of meteorites by a handheld LIBS instrument coupled with a fuzzy logic-based method**" submitted for possible publication as an original research paper in the Journal of Analytical Atomic Spectrometry. The on-going advances in laser, spectrometer and detector technologies have contributed significantly to the development and construction of compact and user-friendly portable handheld LIBS instruments capable of operating also outside the laboratory. The progress in LIBS technology has opened new opportunities to geologists by lowering the instrumentation cost that might be prohibitive for small and medium scale laboratories. In particular, handheld LIBS instruments can be used to rapidly identify geological samples in the field and select the most appropriate among them eligible for further analysis in conventional laboratories. The fuzzy logic-based algorithm applied in this work is based on pattern recognition and produces a set of rules used in the interpretation of results, featuring a very high accuracy. In particular, the handheld LIBS instrument appears very promising for performing in-situ measurements during field operations, object classification, identification of material surface alterations and/or decide previously any further laboratory analysis.

I look forward to hearing from you about the feasibility of this manuscript for publication.
Thank you for your consideration.

Sincerely yours,

Dr. Giorgio S. Senesi

Sede di Lecce
c/o Campus Ecotekne
Via Monteroni – 73100 Lecce
E-mail:
amministrazione.lecce@nanotec.cnr.it

Sede SecondariaBari
Via Amendola, 122/D
70126 Bari
☎ +39 (0)80 5929501
Fax N. +39 (0)80 5929520

Sede Secondaria Roma
c/o Dipartimento di Fisica N.E.
Università Sapienza - Piazzale
Aldo Moro, 5 00185 ROMA
☎ +39 (0)6 49913720
Fax N. +39 (0)6 49693308

Sede Secondaria Cosenza
Ponte P. Bucci, Cubo 31/C
87036 Rende (CS)
☎ +39 (0)984 496008
Fax N. +39 (0)984 494401

Consiglio Nazionale delle Ricerche

CNR**NANOTEC** - Istituto di Nanotecnologia

Partita IVA IT 02118311006 – C.F. 80054330586
PEC: protocollo.nanotec@pec.cnr.it

Corresponding author: giorgio.senesi@nanotec.cnr.it

Dr. Giorgio S. Senesi (researcher full time)

CNR - Istituto di Nanotecnologia (NANOTEC) – PlasMI Lab

Via Amendola, 122/D - 70126 Bari, Italia

Tel. +39.080.5929505

Fax +39.080.5929520

Journal Name

ARTICLE

Identification and classification of meteorites by a handheld LIBS instrument coupled with a fuzzy logic-based method

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Giorgio Saverio Senesi,*[a] Paola Manzari [b], Arianna Consiglio [c] and Olga De Pascale [a]

A handheld laser-induced breakdown spectroscopy (LIBS) instrument is proposed as a novel tool able to provide information on the nature of meteorites and discriminate among iron, stone, stony-iron meteorites and meteor-wrongs. Further, a novel fuzzy logic –based inference algorithm is applied to LIBS broadband spectra for the identification of meteorites and their classification according to the origin and nature. The identification of meteorites is a decision-making problem based on a compromise among human experience, visual evidence and analytical data, which fuzzy logic is proved able to solve. The final model is able to classify correctly 25 out of 26 samples and provides a set of IF-THEN rules that describe how some selected wavelengths are involved in the classification task.

1 Introduction

Meteorites are "objects" recovered from private and nationally organized search efforts mostly in hot (Africa, Australia) and cold (Antarctica) deserts.[1] Often they are found by nomadic inhabitants of north African desert countries and sold to collectors and researchers worldwide. Meteorites are classified in a variety of ways that highlight their specific features. They are divided in three macro-groups, i.e. stone, iron, and stony-iron meteorites, independently from any relationship with their proper parent-body and based merely on visual description. Differently, the scientific classification of meteorites is largely based on their mineralogical and petrographic characteristics and their whole-rock chemical and O-isotopic compositions. In the current classification, meteorites are divided into chondrites, primitive achondrites and achondrites,[1] their specific names are assigned by the Meteorite Nomenclature Committee (NomCom) and are included as classified meteorites in the Meteoritical Bulletin. In particular, the classification of iron meteorites is based on the bulk chemical analysis of the metal associated to structural and other (e.g., cosmic-ray exposure ages) data on the specific meteorite history.[2]

In the last decade Laser-Induced Breakdown Spectroscopy (LIBS) has been shown to be a very promising technique for meteorite analysis[3 and references therein] by using both bench-top[4-7] and portable[8,9] system and the ChemCam laboratory unit.[10-12] Recently, handheld LIBS instruments have become increasingly common for the rapid in-situ chemical identification and characterization of metallic objects of extraterrestrial origin with respect to human artifacts. In particular, this type of instruments allows to obtain a preliminary chemical identification and classification of meteorites, which can be used to identify and recognize mislabelled/unlabelled and dubious meteorite specimens in the field and in museums and private collections. Very recently, Senesi et al.[13] used for the first time a handheld LIBS instrument to identify the qualitative differences in the elemental composition among a certified iron meteorite named Agoudal, a suspected meteorite fragment and a pig iron product. Further, the quantification of the major elements Fe, Ni and Co in the iron meteorite and Fe, Mn, Si and Ti in the other two samples was attempted using a Calibration Free (CF)-LIBS method.[13]

The simplest LIBS predictors used for meteorite classification are the presence or absence of a spectral line in the spectrum and/or whether the ratio of spectral line intensities exceeds a specific threshold.[14] and references therein However, the success of such univariate approaches is dependent on a successful preliminary line assignment supported by adequate spectral resolution and limits of detection.[14] and references therein As LIBS offers the possibility to collect a large amount of data, so reducing the error of the measurements, these can be managed by multivariate analysis of LIBS spectra, so facilitating their interpretation. In particular, parametric and nonparametric multivariate chemometric methods can simultaneously analyse several features of the LIBS spectrum, thus resulting in a robust and efficient performance. As LIBS spectra are often complex and non-linear in consequence of several effects such as differential laser-material coupling, matrix effects and self-absorption of strong lines in the plasma

[a] CNR- Istituto di Nanotecnologia (NANOTEC) PLasMI Lab, Via Amendola 122/D, 70126 Bari, Italia. Email: giorgio.senesi@nanotec.cnr.it
[b] Istituto Nazionale di Astrofisica - Istituto di Astrofisica e Planetologia Spaziali (INAF-IAPS), via Fosso del Cavaliere 100, Roma, Italia.
[c] CNR - Istituto di Tecnologie Biomediche (ITB), Via Amendola 122/D, 70126 Bari, Italia.

† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

emission, chemometric and machine learning methods have been utilized widely for data dimension reduction, pattern recognition, cluster analysis, classification and data quantification.[15-20]

Currently the parametric methods most used in LIBS analysis are principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA), while the most used nonparametric ones are the artificial neural networks (ANNs). Each method has its advantages and limitations. For example, PCA can reduce effectively the dimensionality of LIBS spectra by computing a linear combination of the variables with the highest variance, but does not account for between group-to-group and within-group variability in classification problems.[21] Differently, PLS-DA can divide correctly the space of data, but works only with data linearly separable,[22] a condition that is not always met in LIBS spectra. ANNs produce accurate black box models and can also work with non-linear classifications, but do not allow any human interpretation of the computation of results.[23] Newer promising chemometric methods include random forest (RF) and support vector machines (SVM). However, still rare are systematic studies that compare the performance of different approaches on the same LIBS datasets.[14 and references therein,24]

The aim and novelty of this work are to explore the capability of a handheld LIBS instrument, in conjunction with either PCA or PLS-DA or a never tested fuzzy logic computational approach, to analyse, discriminate, classify and validate the LIBS spectra acquired from various classified and unclassified meteorite samples of different origin and nature. In particular, the approach based on fuzzy sets and genetic algorithms can produce an inference system that allows the interpretation and evaluation of broadband LIBS spectra of meteorites (and meteor-wrongs), and the efficient selection of the most significant features in order to obtain a class discrimination by reducing the complexity of the classification task.

2 Materials

2.1. Classified Meteorites

The seven classified meteorites investigated in this work are described below.

Two fragments (one fragment was weathered) of the Agoudal iron meteorite, which was found nearby the village with the same name in the High Atlas Mountains in Morocco.[25,26] This meteorite consisted of abundant coarse-grained kamacite (a-Fe, Ni), with an Fe:Ni ratio between 90:10 to 95:5, schreibersite (Fe, Ni)3P, rhabdite (tetragonal schreibersite) and troilite ((Fe1-x)S).[25,27] Its elemental composition measured by ICP-MS was: 5.5 %wt Ni, 4.1 mg g-1 Co, 58 μg/g Ga, Ir < 0.04 μg/g and Au ~ 1 μg/g. Based on these data the meteorite was assigned to the IIAB magmatic iron group.[25]

The Dronino ungrouped iron meteorite, which was discovered nearby the village of Dronino in the Ryazan oblast in Russia. It consisted of fine-grained ataxite containing mainly kamacite (7.0 ± 0.5 %wt Ni and 0.75 %wt Co) and also taenite (26.3 ± 0.5 %wt Ni and 0.35%wt Co) with sulfide inclusions (~ 10% v/v).[28] In this meteorite P and phosphides were undetectable by EMP, and its elemental composition (INAA data) was: Ni 98.1 mg/g, Co 5.54 mg/g, Cr 37 μg/g, Cu 32 μg/g, Ga <0.3 μg/g, As 3.52 μg/g, W 0.38 μg/g, Ir 1.68 μg/g and Au 0.284 μg/g. The meteorite resulted close to IVA iron meteorites in Ni, Ir and P amounts, but the low Au and Ga contents distinguished the Dronino elemental pattern from that of all known iron meteorite groups.

The Morasko iron meteorite, which was found in 1914 in an area north of Poznañ (central-west Poland),[29,30] and references therein was classified as an IAB MG octahedrite. INAA analyses yielded: Co 4.58 mg/g, Ni 65.6 mg/g, Cu 156 μg/g, Ga 97.4 μg/g, As 10.5 μg/g, Ru 10.9 μg/g, W 1.65 μg/g, Ir 1.15 μg/g, Pt 9.8 μg/g and Au 1.448 μg/g.[30] The predominant iron matrix consisted of kamacite (6.5% Ni) and taenite (18% Ni)31 containing troilite graphite oval nodules that are commonly rimmed by cohenite and a small quantity of a mixture of schreibersite and djerfisherite. In addition to troilite-graphite, the nodules contained several silicate minerals, including pyroxenes (enstatite, kosmochlor-augite solid solution and kosmochlor), olivines, alkali feldspars and quartz.[32,33]

The Muonionalusta iron meteorite assigned to the IVA group, which was found in 1906 in an iron-rich mass located in a forest 3-km southwest the village of Kitkiöjärvi, Sweden. The analysis of this strongly shock-metamorphosed iron meteorite showed a content of 8.4 %wt Ni and trace amounts of rare elements, including 0.33 μg/g Ga, 0.133 μg/g Ge and 1.6 μg/g Ir. The meteorite contained the minerals chromite, daubréelite, schreibersite, akaganéite and inclusions of troilite.[34]

The North West Africa 11104 meteorite[35], which originated from an unknown place in Morocco, was classified as an iron IAB-anomalous meteorite. Polished surfaces displayed abundant inclusions of silicates and sulfides. Etched surfaces featured few kamacite lamellae ~0.5 mm across; otherwise, kamacite and taenite grains had irregular shapes and were interstitial between the inclusions. The elemental composition measured by ICP-MS[35] was: Ni 10.8 wt%, Co 0.46 wt%, Ir 3.3 μg/g, Ga 73 μg/g, Ge 322 μg/g, W 0.9 μg/g, Re 0.3 μg/g, Os 1.3 μg/g, Cu 471 μg/g, Ru 4.9 μg/g, Pd 4.2 μg/g and Pt 5.4 μg/g. Although the Ni content was intermediate for the IAB complex,[36] the Ga and Ge contents were too high to be assigned to the sLM subgroup (Low Au, Medium Ni).

The North West Africa 4051 meteorite, which was found in 2004 in the Kem Kem Basin, Algeria, was classified as an anachondrite, eucrite.[37] It consisted of large pyroxene and plagioclase grains in a matrix of the same mineral composition but smaller grain size. Minor phases included chromite, ilmenite, silica, iron sulfide and a Ni-rich metal Fe phase. The composition of pyroxenes ranged from pigeonite to subcalcic augite, the plagioclase was essentially anorthite, and other mineral phases consisted of oxides as chromite and ilmenite.

2.2. Unclassified meteorites

The unclassified meteorites investigated in this work consisted of 6 stony samples, 3 iron samples and 3 stony-iron samples, all collected in the field in the Bouarfa and Agoudal area in Morocco which were grouped only on the basis of preliminary visual inspection. Three stony samples were analysed both as whole bulk fragments and petrographic thin sections.

2.3. Pig iron, hematite nodule and meteor-wrong samples

A number of different samples of unusual shape and density higher than that of common rocks, which could be easily mistaken for meteorites but were actually meteor-wrongs, were tested in the model. These included: an iron oxide nodule of hematite formed in sedimentary rocks by oxidation of pyrite or marcasite (iron sulfide) crystals, two pig iron samples generated from industrial and manufacturing plants and two samples visually identified in the field as iron meteorites but suspected "meteor-wrongs".

3 Methods

3.1. Handheld LIBS instrumentation and analytical procedure

A portable, handheld LIBS instrument (B&W Tek, Newark, DE, USA) consisting of a miniature-diode-pumped, solid-state, short-pulsed laser emitting at the wavelength of 1064 nm and operating at a high repetition rate, i.e. between 1 and 5 kHz, was used to perform the measurements. The maximum laser output was 300 mW, the pulse duration 500 ps and the maximum pulse energy 150 μJ. The spectra were recorded by a compact spectrometer in the non-gated mode with an overall resolution of 0.4 nm for the entire wavelength acquisition range from 180 to 800 nm, i.e. from UV to visible. The integration time of the spectrometer was set at 4 ms, which corresponded to twelve laser pulses for a 1-kHz laser. A liquid crystal display (LCD) touch screen panel controlled the spectrometer and the acquisition settings.

As in the handheld LIBS instrument the laser energy was lower than that in conventional LIBS, and produced a smaller-sized plasma with lower continuum radiation, the spectrometer was used in the non-gated mode and twelve ablations occurred during one gate time. Thus, despite the low intensity and short time period of the continuum, the background signal was unavoidable. Further, the matrix effects resulted in a strong background signal in addition to the poorer resolution of the miniaturized spectrometers with respect to standard spectrometers.

To perform the measurements the instrument was placed against the sample surface and then the analysis was started via a trigger. The system was equipped with a rastering beam that covered an area of about 300-μm diameter. As the samples were not homogeneous, 10 different positions were analysed and a total of 150 spectra corresponding to 1800 laser shots were acquired and averaged to obtain a single spectrum for each position. Each measurement lasted about 6 sec. The average LIBS spectra were stored in the analyzer and accessed by a Linux operating system with graphic user interface (GUI) by downloading via Wi-Fi or Ethernet. Then, the spectra were analysed qualitatively to identify the emission lines using the NIST database.

For safety reasons, a sensor installed in the analytical head controlled automatically the laser output, so that the laser could operate only when the analytical head was in contact with the sample. As the NanoLIBS instrument was designed to operate primarily as a handheld device outdoor in the field, it was provided with a rechargeable Li-ion battery allowing up to 8 h of continuous operation. However, the instrument could also operate indoor while connected to AC power. The whole system was enveloped in a lightweight and compact (mass and dimensions of approximately 1.8 kg and 26x10x30 cm, respectively) handheld body.

3.2. Data acquisitions and algorithms

3.2.1. PCA

PCA consists in a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a smaller number of uncorrelated variables called principal components. This transformation is defined so that the first principal component has the largest possible variance, and each successive component accounts for as much of the remaining variability as possible. PCA is mostly used as a tool in exploratory data analysis, because it allows an effective reduction of dimensionality of large datasets (feature selection), preserving the largest part of information of the data derived from variance. In this work we have applied the *pca* function implemented in the mixOmics R package.[38]

3.2.2. PLS - DA

PLS-DA is a mathematical method able to extract the principal components from a dataset, considering both the largest possible variance and a known classification of the input sample. This procedure allows to define the components, so that a maximum separation among classes is achieved, and to understand which variables carry the information separating the classes. Thus, the method yields a model able to classify the samples, and also usable to predict the class of new samples. In this work we have applied the *plsa-da* function implemented in the mixOmics R package.[38] The accuracy of the model was evaluated with the leave-one-out cross validation procedure.

3.2.3. Fuzzy classifier and feature selection with the genetic algorithm

The fuzzy logic computational approach proposed to build a classifier able to distinguish different types of meteorites includes a genetic algorithm for feature selection and a fuzzy-rule-based system for producing the inference and achieving the classification. The method is able to solve non-linear classification problems by selecting automatically the most important variables in the samples. Further, the result is evaluated not only in terms of accuracy - compared with PLS-DA - but also provides a human-friendly set of rules that explains how the final model classifies the samples.

The dataset of this study consisted of 3496 features of 26 samples, thus it was necessary to reduce its dimensionality in order to control the complexity of the classification model. Differently from PCA, which is often used to solve this task by transforming the data using a linear combination of the features so complicating the interpretation of results, in this fuzzy method a genetic algorithm is used to select the features, i.e. the most significant wavelengths able to discriminate the samples. In this algorithm, which is inspired by the biological concepts of generation of individuals, chromosome crossover and natural selection, the individuals consist of a random selection of features. The fitness function for individual selection is the accuracy of the fuzzy inference system computed on a subset of features, and the generation of new individuals is implemented as the recombination of portions from two individuals selected through roulette wheel extraction (individuals with better fitness are selected with higher probability). The fuzzy inference system is included both in the computation of the fitness of the genetic algorithm and in the final step, i.e. the construction of the final classification model.

In a fuzzy-rule based system, each variable of the dataset is associated to some fuzzy sets, which are functions that assign a value to a set by allowing a partial membership.[39,40] This method allows to manage uncertainty and to translate numeric values into interpretable labels such as small, medium, high. The fuzzy logic provides some mathematical operators to apply logical reasoning to the fuzzy sets exploited in the inference system by the definition of some IF-THEN rules.[41] Each rule describes the characteristics of the samples associated to a specific class using some linguistic labels. For example, the rule "IF feature1 IS high AND feature2 IS small and feature3 IS medium THEN class is A" provides both a rigorous system to compute if a sample belongs to a class A and a description of the reasoning that is clearly interpretable by humans. The whole set of rules identified to classify all the samples of a dataset is called the fuzzy-rule base.

A fuzzy-rule base can be either manually defined by a domain expert or automatically derived using an algorithm that extracts the rules (supervised learning) from a set of samples with known class label (training set).[42] The selected rules must be as compact as possible to ensure their readability, but to reach an optimal accuracy the model should involve a large number of rules including many variables so resulting hard to interpret. Thus, the best solution is a compromise between accuracy and readability using a compact set of fuzzy rules defined only on the most important variables. As the selection of variables is a crucial step for the definition of a compact fuzzy-rule base, it is also possible to adjust manually the model with the intervention of a domain expert. Hatch et al.[43] were the only authors to present a hybrid approach for LIBS spectra classification of steel alloys using fuzzy-rule based on the manual selection of the variables of interest and automated definition of the rules with K-means clustering.[43]

The final model is evaluated first using the features selected by the genetic algorithm, then is pruned further in order to select only the essential rules and features. Pruning operates by removing iteratively a feature, i.e. if the accuracy of the model remains the same, the feature is removed, otherwise it is preserved.

In this work a completely automated workflow (Figure 1) has been developed by "the environment for statistical computing R"[44] to define a compact set of fuzzy rules for the classification of meteorites. In particular, the package *frbs*[45] has been used for the training and testing steps of the fuzzy inference system. The selection of the most significant variables was performed using a genetic algorithm[46,47] that executed a random search of the most important variables in the datasets based on random generation, crossover and evaluation of subsets of variables.

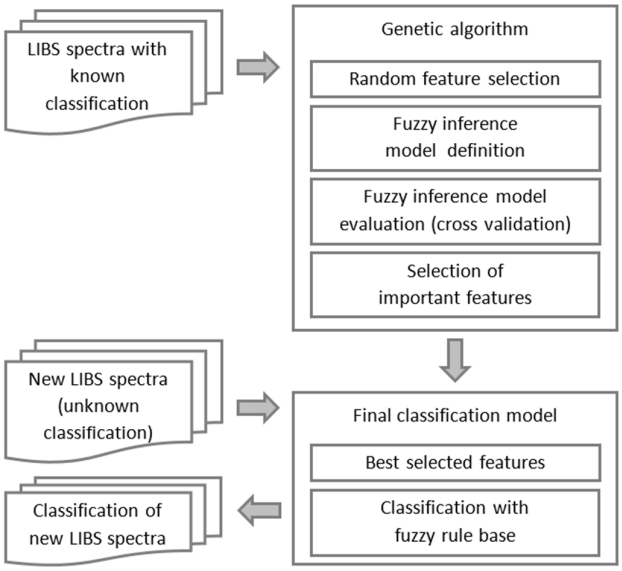

Figure 1 - Workflow of the classification method based on fuzzy sets and genetic algorithms.

3.2.4. Accuracy validation

The accuracy of the model was evaluated by the leave-one-out cross validation procedure, i.e. the model was trained repeatedly on all the dataset samples but one, and was tested on the excluded data. The overall accuracy of the model was then evaluated as the mean of resulting accuracies. The results so obtained were compared with PLS-DA and PCA results.

4 Results and discussion

4.1. Meteorites classification

Figure 2 shows some examples of the broadband emission spectra acquired on six samples. A full broadband average emission spectrum could provide the "spectral fingerprint" including simultaneously all elements in a meteorite so representing its "geochemical fingerprint"[17]. The use of a full broadband LIBS spectrum, which contains all the chemical information about the sample, is fundamental for qualitative applications related to rapid material discrimination and identification in the field.

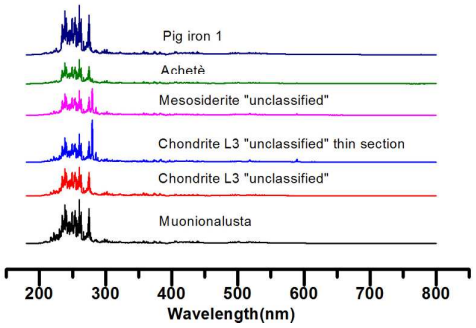

Figure 2 - Typical LIBS spectra of Muonionalusta, "unclassified" Mesosiderite, whole and petrographic thin section of Chondrite L3 "unclassified", suspected meteorite fragment Achetè and pig iron 1 sample.

The LIBS dataset obtained was analysed by PCA and PLS-DA methods and the approach based on fuzzy sets and genetic algorithms. The PCA results showed a very high correlation of features, i.e. more than 90% of variance was explained by the first two principal components. However, the linear combinations computed by PCA were not able to distinguish properly the dataset classes, as shown in Figure 3.

Similar results were obtained by testing the dataset with the PLS-DA model with up to 10 components. The best accuracy (69.23%) evaluated by the leave-one-out cross validation procedure was achieved using 6 components, i.e. only 18 out of 26 samples were properly classified, whereas 4 iron, 1 stony and 3 stony-iron meteorites were misclassified (confusion matrix in Table 1). The accuracy obtained using only the first component was 61.54%. The bad performance of the PLS-DA classifier is illustrated in Figure 4 that shows the clusters obtained by the first two principal components as too much overlapped, which do not allow to distinguish correctly the various classes.

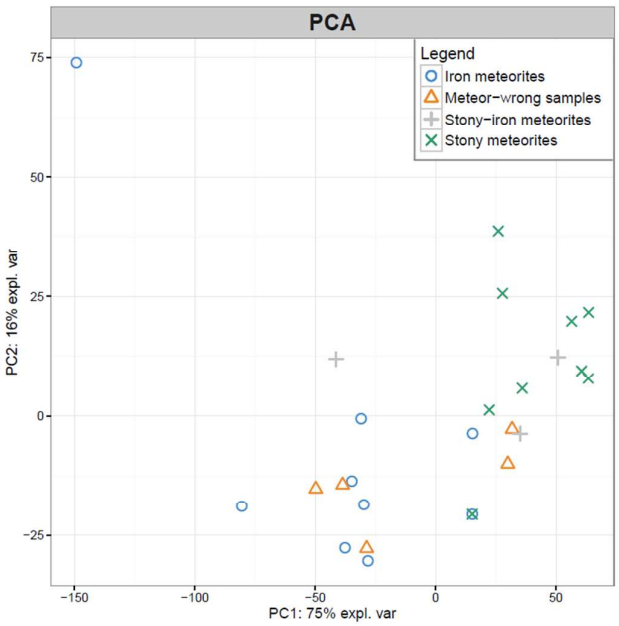

Figure 3 - Representation of the dataset with the first two components computed by PCA.

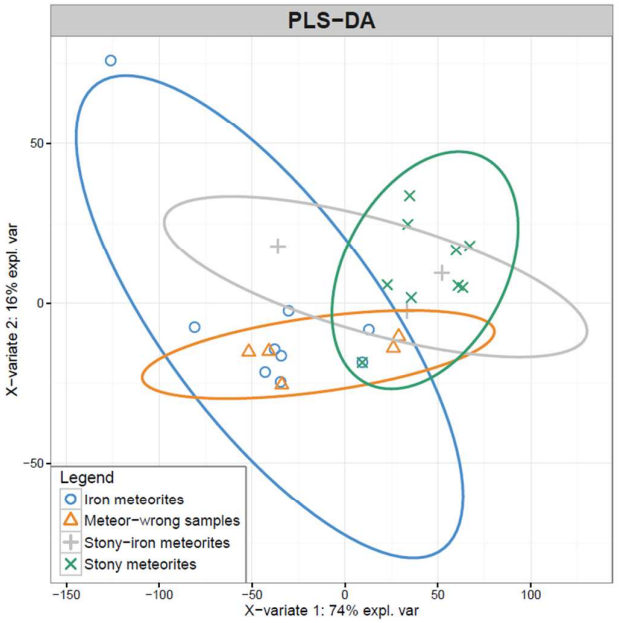

Figure 4 - Representation of the dataset with the first two components computed by PLS-DA. The ellipses represent the clusters for the four classes. One iron sample is so far from the center of the cluster that is represented outside the ellipse.

Confusion matrix for PLS-DA					
True label	Iron	5	0	1	3
	Stony	1	8	0	0
	Stony-Iron	1	0	2	0
	Meteor-Wrong	0	0	0	5
		Iron	Stony	Stony-Iron	Meteor-Wrong
		Predicted label			

Table 1 – Confusion matrix obtained for PLS-DA classification by leave-one-out cross validation procedure (total accuracy: 69.23%).

Confusion matrix for Fuzzy Classifier					
True label	Iron	9	0	0	0
	Stony	1	8	0	0
	Stony-Iron	0	0	3	0
	Meteor-Wrong	0	0	0	5
		Iron	Stony	Stony-Iron	Meteor-Wrong
		Predicted label			

Table 2 – Confusion matrix obtained for Fuzzy model classification by leave-one-out cross validation procedure (total accuracy: 96.15%).

The apparent performance limitations of the linear models described above could be solved by using the method proposed in this work, i.e. the fuzzy-rule based method, which could implement a non-linear classification model that did not need to transform the original features in principal components. A workflow was defined that could automatically perform a feature selection and supervised learning, with the advantage of producing a set of IF-THEN rules for the classification of meteorites based on their broadband LIBS spectra. These rules were of easy interpretation as they associated a list of features, i.e. selected peak wavelengths, to some fuzzy sets that were defined over the signal intensity domain.

The workflow method was tested preliminarily with different parameters in order to tune each step and obtain the best results. The preliminary analysis of LIBS dataset by this method showed that no significant increase of accuracy was achieved after data normalization to the most intense peak and their normalization to the sum of intensities. Thus, the entire raw LIBS dataset was used in the subsequent tests.

To avoid the construction of very complex IF-THEN fuzzy rules, the maximum number of features selected randomly by the genetic algorithm was imposed at a maximum of 20, which cover the whole length of each selected individual. Further, each fuzzy rule set obtained during the tests was pruned further in order to remove those features that were not essential for the computation. The resulting models never requested more than 10 features to reach the best accuracy for sample classification.

By testing the method with 3, 5, 7 and 9 fuzzy sets per feature, the best compromise between accuracy and interpretability was obtained using 5 fuzzy sets associated to the following labels: very small, small, medium, large and very large. The final model implemented with the previously described parameters yielded an accuracy of 96.15%, i.e. 25 out of 26 samples from the training dataset could be classified by the leave-one-out cross validation procedure, with only one sample resulting misclassified, and produced 15 fuzzy rules over 7 features (confusion matrix in Table 2).

The misclassified sample was chondrite L3, which was wrongly classified as an iron instead of a stony meteorite. In particular, the fuzzy rule method assimilated the chondrite L3 sample to the altered fragment of Agoudal and to Dronino meteorite that also showed evidence of weathering. The misclassification could thus be attributed to the laser raster having specifically hit a metal phase of the chondrite during the analysis. Besides visual inspection, which showed clearly that this sample was a stony meteorite, its petrographic thin section confirmed its correct assignment to the stony meteorite group. Further, the similarity of chondrite L3 to the altered fragments of Dronino and Agoudal meteorites could be ascribed to the total/partial substitution of the metallic phase of weathered iron by a variety of secondary minerals such as silicates and various oxides.

Another interesting result of the model was that all the pig iron samples fell in the same rule. Further, one of the two samples visually identified in the field as iron meteorites but suspected “meteor-wrongs”, could be identified by the model as a pig iron.

4.2. Meteorite intragroup variations

Besides being able to assign correctly unclassified meteorites to their macro-group, LIBS analysis coupled with the fuzzy model was tentatively used to evaluate and identify intragroup variations, for example within the iron meteorites class. Samples that obeyed to slightly different but similar (medium-large and small-very small are adjacent fuzzy sets) fuzzy-rules, thus should belong to the same iron-subgroup, were Morasko, Agoudal and North West Africa 11104 meteorites. In particular, Morasko and Agoudal meteorites differed only for the feature at 248.72 nm identified as Co II, while the North West Africa 11104 meteorite differed from Agoudal in the feature at 293.78 nm assigned to Mn II.

Similarly, the model associated two iron meteorites, i.e. the classified Muonionalusta meteorite and one not yet classified, provisionally named Tasseraft, as obeying to the same rules, so they would belong to the same iron-subgroup. However, this hypothesis should be verified by performing further chemical analysis.

For iron meteorites the probability of striking and analysing a representative metallic phase (Fe-Ni alloy) would result higher then

for stony meteorites the probability decreases as a function of decreasing iron content, e.g. it is about the half for stony-iron meteorites. Conversely, for stony meteorites the probability of striking on different silicate minerals/oxides/sulfides increases, thus the typical elements of these minerals, which commonly show a lithophile behaviour, should be considered. In particular, stony-iron meteorites commonly include representative minerals such as olivine and pyroxenes in a Fe-Ni matrix.

As the model could operate by defining 5 fuzzy sets associated to linguistic labels, it was possible also to draw some comments on the features selected by the final model. For example, the wavelengths at 248.44 and 274.09 nm, which corresponded to the element Fe, were present at different intensity in all four groups. The model associated them properly to the iron meteorite and meteor-wrong groups at large intensity, to the iron meteorite, meteor-wrong and stony-iron meteorite groups at medium intensity, and to the stony-iron and stony meteorite groups at small intensity (Figure 5), thus the contents of Fe were congruent with the different groups. In particular, a meteor-wrong sample provisionally named Achetè and the hematite nodule sample showed similar contents of Fe, which suggested a possible advanced state of alteration of the meteor-wrong Achetè.

Further, the wavelengths at 279.55, 280.27 and 285.21 nm, which corresponded to the element Mg, were present mainly in two groups that were correctly associated to the stony meteorite group at very large intensity and to the stony-iron and stony meteorite groups at large and medium intensity, whereas at small intensity the samples were assigned to the iron meteorite and meteor-wrong groups (Figure 6). The inter-class variation of Mg was congruent with the contents of Mg in the different groups, and was ascribed mainly to the presence of olivines and pyroxenes. However, the presence of Mg in the meteor-wrong samples, might be ascribed to terrestrial contamination and cation exchange processes.

Other features significant in providing specific inputs to the model were identified at 216.56 and 217.52 nm for Ni (Figure 7), 220.81 nm for Ir (Figure 8) and 225.50 for Ga I (Figure 9). These features were identified mainly to distinguish the iron meteorite group from the other groups but also added further information about intragroup variability in the iron meteorite group. In particular, the content of Ni was compatible with that expected in the different groups of meteorites, i.e. it was low in stony and stony iron meteorites, and undetectable in meteor-wrong samples. Further, the wavelengths associated to Ga, Ir and Ni showed the same intraclass and intergroup trend. In particular, within the iron meteorite group a positive correlation of these elements was observed in each sample.

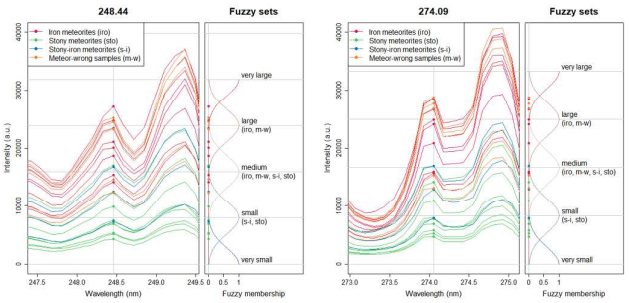

Figure 5 - LIBS wavelengths of Fe at 248.44 and 274.09 nm. Each spectrum plotted on the left represents a sample of the dataset. The intensities of the wavelengths associated to the Gaussian fuzzy sets automatically defined by the model are projected on the right.

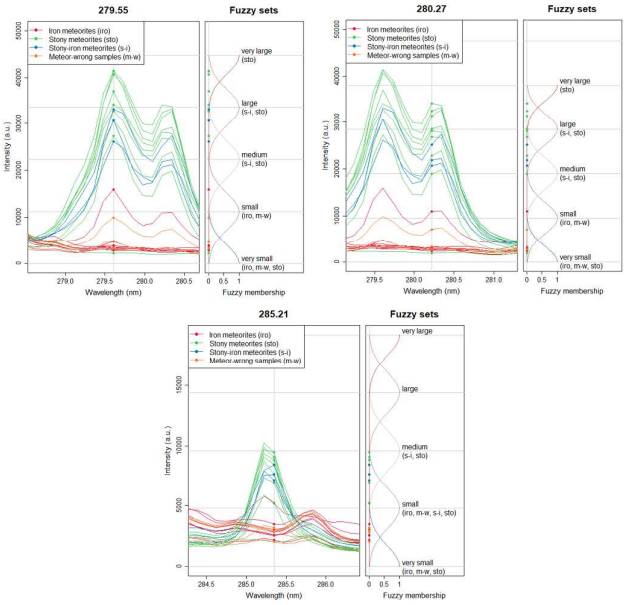

Figure 6 - LIBS wavelengths of Mg at 279.55, 280.27 and 285.21. Each spectrum plotted on the left represents a sample of the dataset. The intensities of the wavelengths associated to the Gaussian fuzzy sets automatically defined by the model are projected on the right.

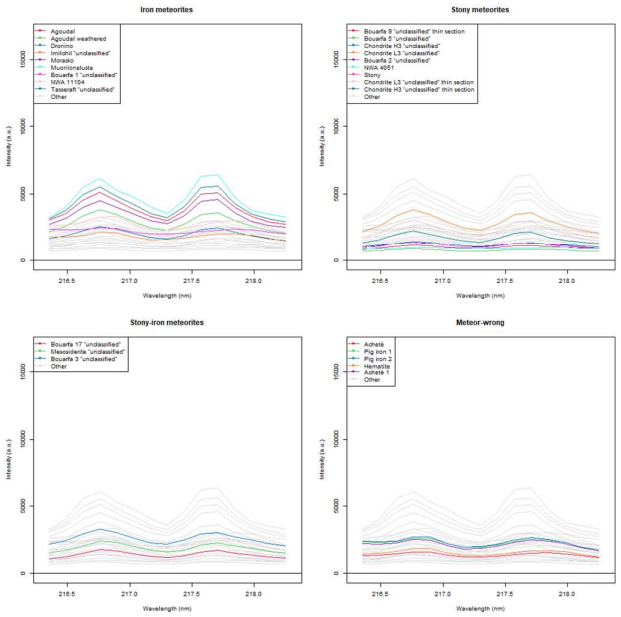

Figure 7 - Intensities of Ni wavelengths at 216.56 and 217.52 nm measured on each sample and grouped according the corresponding class.

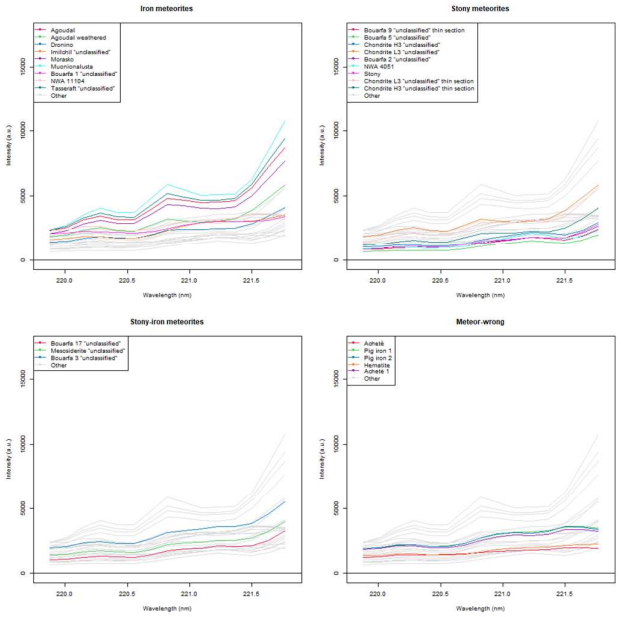

Figure 8 - Intensities of Ir wavelength at 220.81 nm for Ir measured on each sample and grouped according the corresponding class.

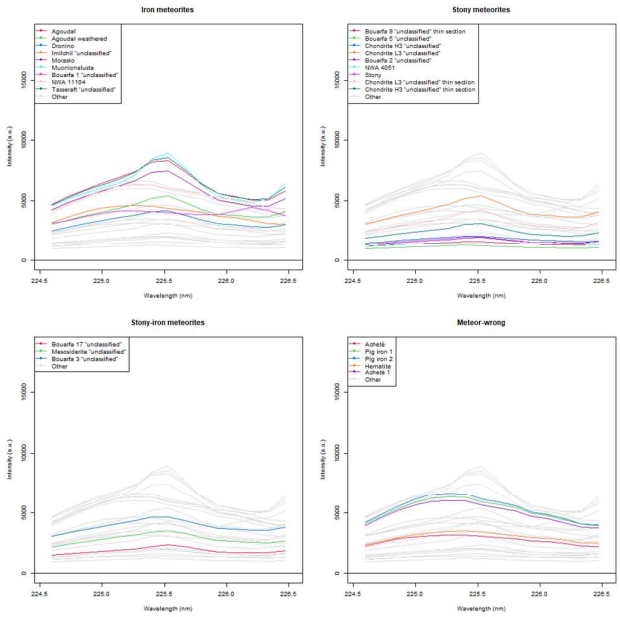

Figure 9 - Intensities of Ga wavelength at 225.50 nm for Ga measured on each sample and grouped according the corresponding class.

Conclusions

In this work an innovative strategy based on broadband LIBS spectra measured by a handheld instrument associated to a fuzzy logic-based computational method has been developed and applied successfully to meteorite classification and discrimination. The approach used an integrated automated feature selection and supervised learning in a fuzzy-rule based classifier, which allowed to achieve a high degree of success in the correct assignment of each unclassified meteorite sample to its macro-group and in identifying the most relevant intragroup variations.

The final model was able to classify correctly 25 out of 26 samples using a set of fuzzy rules that could be easily interpreted by humans, while PLS-DA recognized only 18 samples. The accuracy achieved by the final fuzzy-rule based model applied to the dataset examined, compared to that obtained from the PLS-DA method, confirmed that the classification of LIBS spectra of meteorites is a non-linear task and that a high accuracy cannot be obtained using methods able to evaluate only linear combinations of the input variables, i.e. PCA and PLS-DA. Differently, the fuzzy-rule based approach allows to classify meteorites with very high accuracy, and also provides a model usable for sample discrimination and assignment to the correct group.

This study demonstrates that even chemically similar meteorites could be readily discriminated using broadband LIBS analysis combined with appropriate multivariate chemometric data processing. Further, the identification of the elements yielding the 6 wavelengths selected would allow the fuzzy rules to be exploited for further chemometrics studies. Finally, handheld LIBS instruments appear particularly

promising for performing in-situ measurements to preliminarily identify and classify meteorites, and so to decide about any further laboratory analysis.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors acknowledge Davide Manzini from MADAtec s.r.l. and B&W Tek for providing the NanoLIBS handheld instrument used in this work and are grateful to Dr. A. Bove, head of Raw Materials Quality Control Laboratories at Ilva Taranto Steelworks – ILVA GROUP for providing the pig iron samples. The authors acknowledge Prof. Abderrahmane Ibhi and Dr. Ahmed Ait Touchnt for providing three unclassified meteorite samples from the Agoudal area.

References

1 M. K. Weisberg, T. J. McCoy, A. N. Krot, in *Meteorites and the Early Solar System II*, University of Arizona Press, 2006, pp. 19-52.
2 G. K. Benedix, H. Haack, T. J. McCoy, in *Meteorites and cosmochemical processes*, Elsevier, 2014, pp. 267-285.
3 G. S. Senesi, *Earth Sci. Rev.*, 2014, **139**, 231-267.
4 A. Cousin, V. Sautter, C. Fabre, S. Maurice, R. C. Wiens, *J. Geophys. Res.*, 2012, **117**, E10002.
5 M. Dell'Aglio, A. De Giacomo, R. Gaudiuso, O. De Pascale, G. S. Senesi, S. Longo, *Geochim. et Cosmochim. Acta*, 2010, **74**, 7329–7339.
6 M. Horňáčková, J. Plavčan, J. Rakovsky, V. Porubčany, D. Ozdin, P. Veis, *Eur. Phys. J- Appl. Phys.*, 2014, **66**, 10702.
7 J. R. Thompson, R. C. Wiens, J. E. Barefield, D. T. Vaniman, H. E. Newsom, S. M. Clegg, *J. Geophys. Res.*, 2006, **111**, E05006.
8 G. S. Senesi, G. Tempesta, P. Manzari, G. Agrosì, *Geostand. Geoanal. Res.*, 2016, **40**, 533-541.
9 G. Tempesta, G. S. Senesi, P. Manzari, G. Agrosì, *Spectrochim. Acta, Part B*, 2018, **144**, 75-81.
10 R. C. Wiens, P. -Y. Meslin, D. F. Wellington, J. R. Johnson, A. Fraeman, O. Gasnault, S. Maurice, O. Forni, P. Beck, B. A. Cohen, H. Newsom, J. C. Bridges, V. Sautter, P. Gasda, N. Lanza, A. Ollila, S. E. Johnstone, A. Fairen, 80th Annual Meeting of the Meteoritical Society, 2017, **1987**, 6168.
11 P. -Y. Meslin, J. R. Johnson, O. Forni, P. Beck, A. Cousin, J. Bridges, W. Rapin, B. Cohen, H. E. Newsom, V. Sautter, E. Lewin, M. Nachon, R. C. Wiens, V. Payré, O. Gasnault, S. Maurice, A. G. Fairén, S. Schröder, N. Mangold, N. Thomas, Lunar and Planetary Science XLVIII, 2017, 2258.
12 N. L. Lanza, P. J. Gasda, A. M. Olilla, R. C. Wiens, S. M. Clegg, D. Delapp, M. Bodine, C. Agee, P. -Y. Meslin, P. Beck, H. E. Newsom, S. Maurice, 80th Annual Meeting of the Meteoritical Society 2017, **1987**, 6402.
13 G. S. Senesi, P. Manzari, G. Tempesta, G. Agrosì, A. A. Touchnt, A. Ibhi, O. De Pascale, *Geostand. Geoanal. Res.*, 2018, in press. DOI: 10.1111/ggr.12220.
14 G. Galbács, Anal. Bioanal. Chem., 2015, 407, 7537-7562.
15 S. M. Clegg, E. Sklute, M. D. Dyar, J. E. Barefield, R. C. Wiens, *Spectrochim. Acta, Part B*, 2009, **64**, 79-88.
16 J. L. Gottfried, R. S. Harmon, F. C. De Lucia Jr, A. W. Miziolek, *Spectrochim. Acta, Part B*, 2009, **64**, 1009-1019.
17 R. S. Harmon, R. R. Hark, C. S. Throckmorton, E. C. Rankey, M. A. Wise, A. M. Somers, L. M. Collins, *Geostand. Geoanal. Res.*, 2017, **41**, 563-584.
18 J. Lasue, R. C. Wiens, T. F. Stepinski, O. Forni, S. M. Clegg, S. Maurice, *Anal. Bioanal. Chem.*, 2011, **400**, 3247-3260.
19 N. J. McMillan, C. Montoya, W. H. Chesner, *Appl. Opt.*, 2012, **51**, B213-B222.
20 J. J. Remus, J. L. Gottfried, R. S. Harmon, A. Draucker, D. Baron, R. Yohe, *Appl. Opt.*, 2010, **49**, C120-C131.
21 M. Barker and W. Rayens, *J. Chemom.*, 2003, **17**, 166-173.
22 T. M. Cover, *IEEE Trans. Electron. Comput.*, 1965, **3**, 326-334.
23 O. Linda, M. Manic, T. R. McJunkin, Proc. 4th International Symposium on Resilient Control Systems, 2011, pp. 35-41.
24 R. A. Putnam, Q. I. Mohaidat, A. Daabous, S. J. Rehse, *Spectrochim. Acta, Part B*, 2013, **87**, 161-167.
25 H. Chennaoui Aoudjehane, L. A. J. Garvie, C. D. K. Herd, G. Chen, M. Aboulahris, 76th Meeting of the Meteoritical Society, 2013, 5026.
26 Meteoritical Bulletin Database, entry for Agoudal: http://www.lpi.usra.edu/meteor/metbull.php?code=57354
27 M. Schmieder, H. Chennaoui Aoudjehane, E. Buchner, E. Tohver, *Geol. Mag.*, 2015, **152**, 751–757.
28 V. I. Grokhovsky, V. F. Ustyugov, D. D. Badyukov, M. A. Nazarov, Lunar and Planetary Science XXXVI, 2005, 1692.
29 W. T. J. Stankowski, Seria Geologia, 2008, **19**, 91.
30 A. S. Pilski, J. T. Wasson, A. Muszyński, R. Kryza, Ł. Karwowski, and M. Nowak, *Meteorit. & Planet. Sci.*, 2013, **48**, 2531-2541.
31 W. T. J. Stankowski, *Planet. Space Sci.*, 2001, **49**, 749–753
32 Ł. Karwowski and A. Muszynski, *Mineral. Pol., Special Papers*, 2006, **2**, 140–143.
33 Ł. Karwowski, A. S. Pilski, A. Muszynski, S. Arnold, G. Notkin, and A. Gurdziel, *Meteorites*, 2011, **1**, 21–28.
34 R. Schaudy, J. T. Wasson and V. F. Buchwald, *Icarus*, 1972, **17**, 174-192.
35 Meteoritical Bulletin, no. 106, (2017)
36 J. T. Wasson and G. W. Kallemeyn, *Geochim. Cosmochim. Acta*, 2002, **66**, 2445-2473.
37 Meteoritical Bulletin, no. 91, MAPS 42, 413-466 (2007)
38 F. Rohart, B. Gautier, A. Singh, K. A. Le Cao, *PLoS Comput. Biol.*, 2017, **13**, e1005752.
39 L. A. Zadeh, *Fuzzy Set Syst.*, 1978, **1**, 3-28.
40 A. Bardossy and L. Duckstein, *Fuzzy Rule-Based Modelling with Application to Geophysical, Biological and Engineering Systems*, CRC Press, 1995.
41 J. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice Hall, 2001.
42 Z. Chi, H. Yan, T. Pham, *Fuzzy algorithms: with applications to image processing and pattern recognition (Vol. 10)*, World Scientific Publishing, 1996.
43 J. J. Hatch, T. R. McJunkin, C. Hanson, J. R. Scott, *Appl. Opt.*, 2012, **51**, B155-B164.
44 R Development Core Team (2011), R: A Language and Environment for Statistical Computing. Vienna, Austria: the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/.
45 L. S. Riza, C. Bergmeir, F. Herrera, J. M. Benítez, *J. Stat. Softw.*, 2015, **65**, 1-30.
46 R. A. Fisher, *The genetical theory of natural selection: a complete variorum edition*, Oxford University Press, 1999.
47 M. Mitchell, *Complexity: A Guided Tour*, Oxford University Press, 2009.

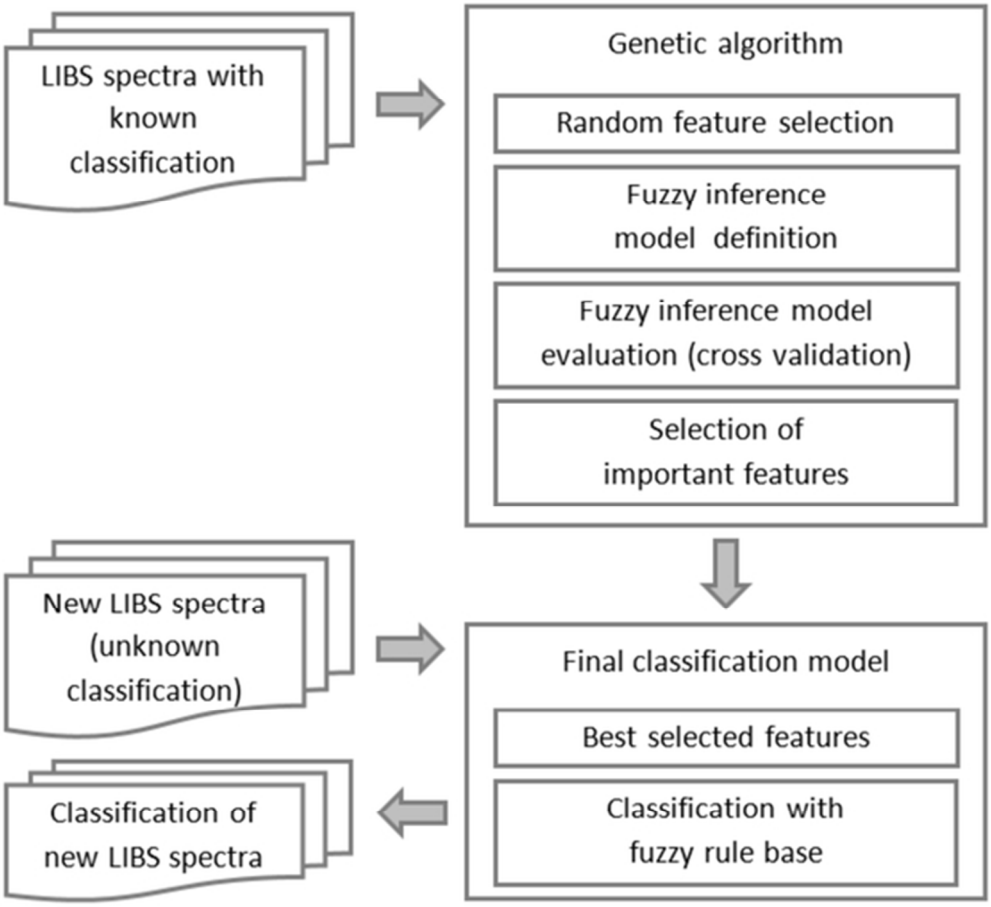

Figure 1 - Workflow of the classification method based on fuzzy sets and genetic algorithms.

52x48mm (300 x 300 DPI)

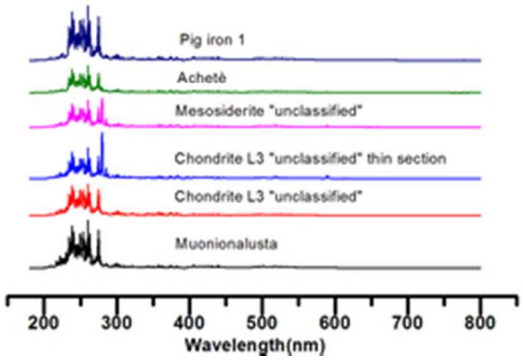

Figure 2 - Typical LIBS spectra of Muonionalusta, "unclassified" Mesosiderite, whole and petrographic thin section of Chondrite L3 "unclassified", suspected meteorite fragment Achetè and pig iron 1 sample.

30x17mm (300 x 300 DPI)

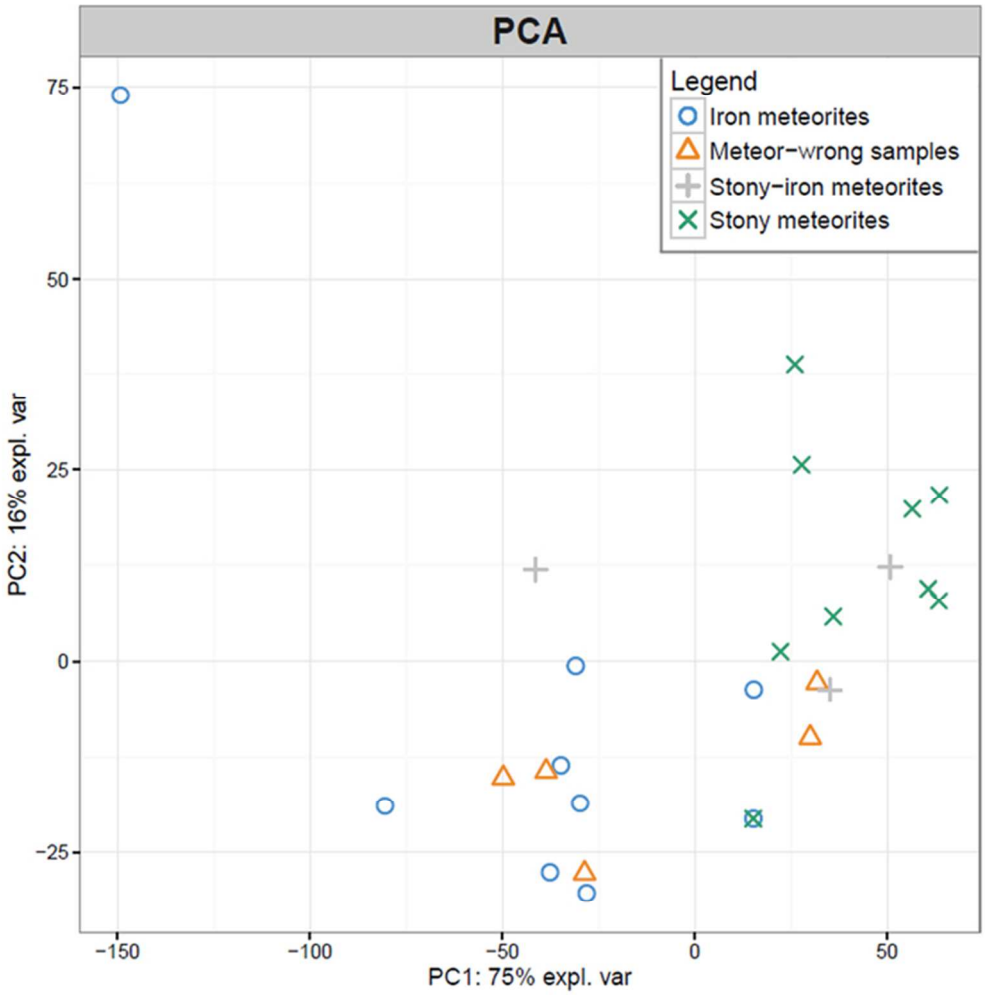

Figure 3 - Representation of the dataset with the first two components computed by PCA.

60x60mm (300 x 300 DPI)

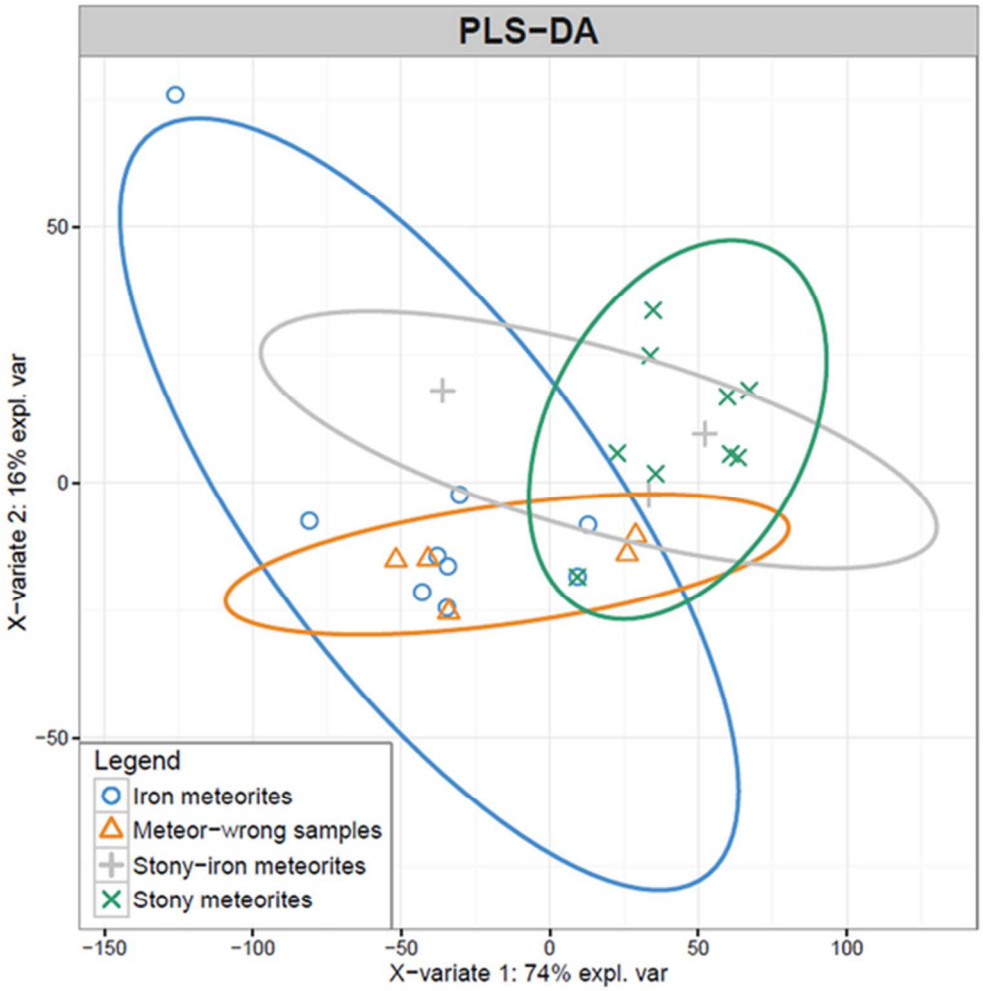

Figure 4 - Representation of the dataset with the first two components computed by PLS-DA. The ellipses represent the clusters for the four classes. One iron sample is so far from the center of the cluster that is represented outside the ellipse.

50x50mm (300 x 300 DPI)

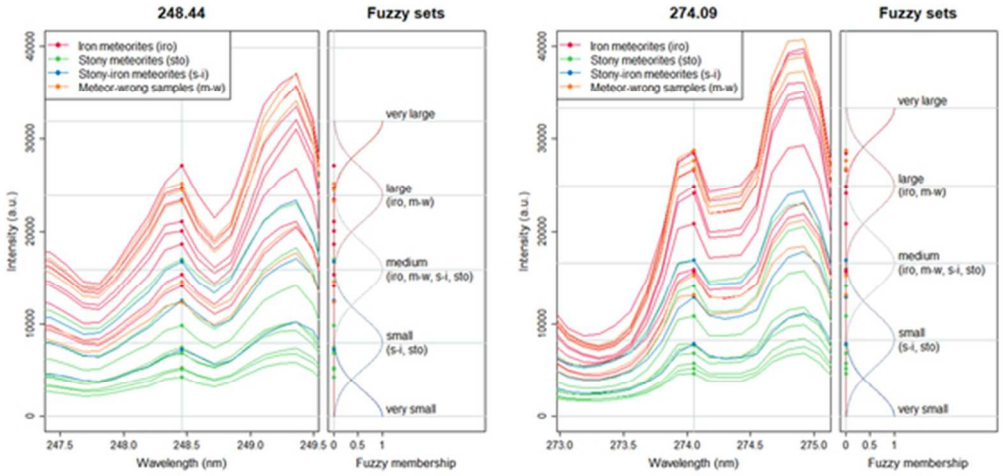

Figure 5 - LIBS wavelengths of Fe at 248.44 and 274.09 nm. Each spectrum plotted on the left represents a sample of the dataset. The intensities of the wavelengths associated to the Gaussian fuzzy sets automatically defined by the model are projected on the right.

51x24mm (300 x 300 DPI)

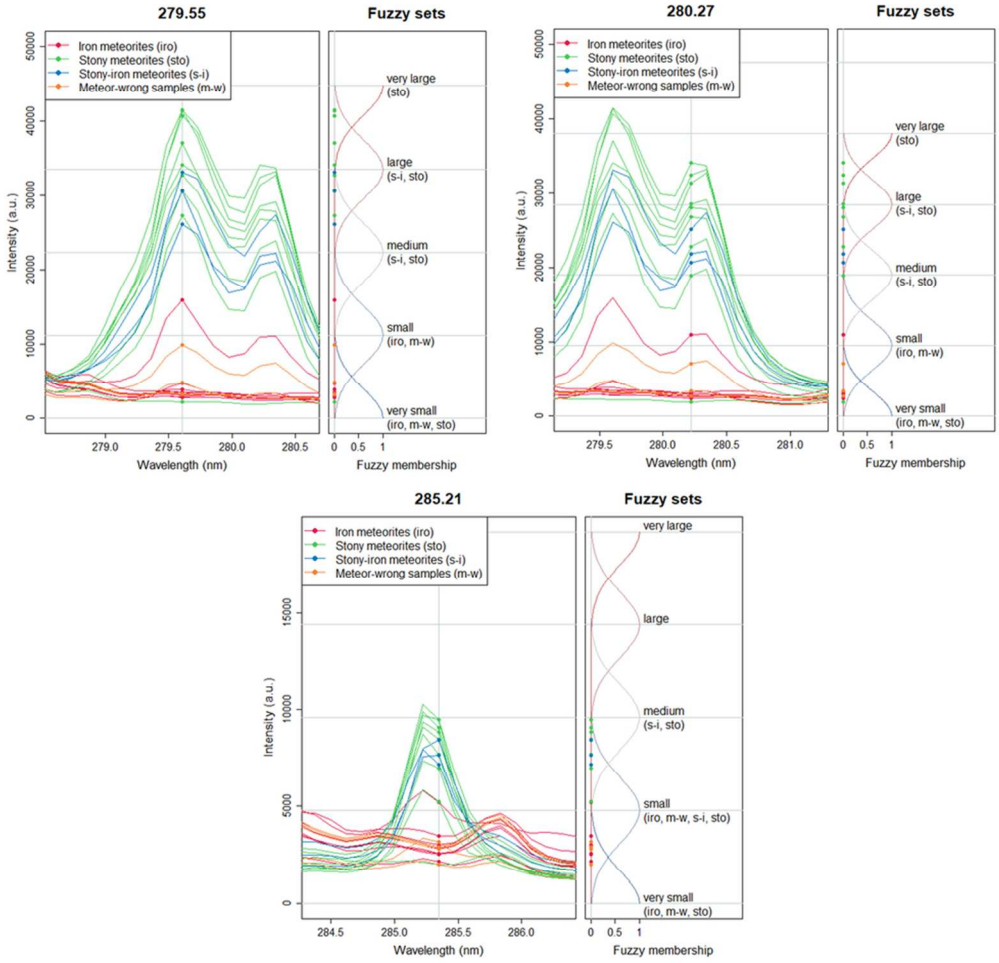

Figure 6 - LIBS wavelengths of Mg at 279.55, 280.27 and 285.21. Each spectrum plotted on the left represents a sample of the dataset. The intensities of the wavelengths associated to the Gaussian fuzzy sets automatically defined by the model are projected on the right.

83x80mm (300 x 300 DPI)

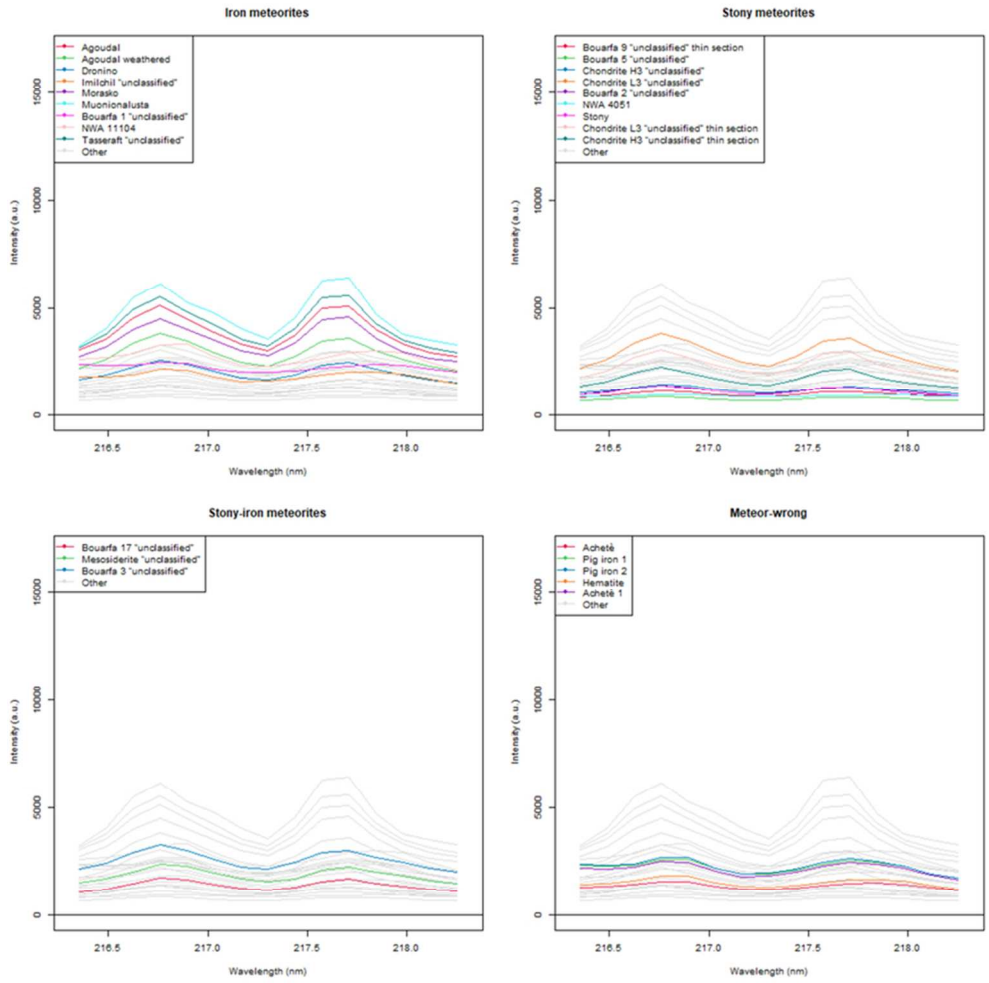

Figure 7 - Intensities of Ni wavelengths at 216.56 and 217.52 nm measured on each sample and grouped according the corresponding class.

74x74mm (300 x 300 DPI)

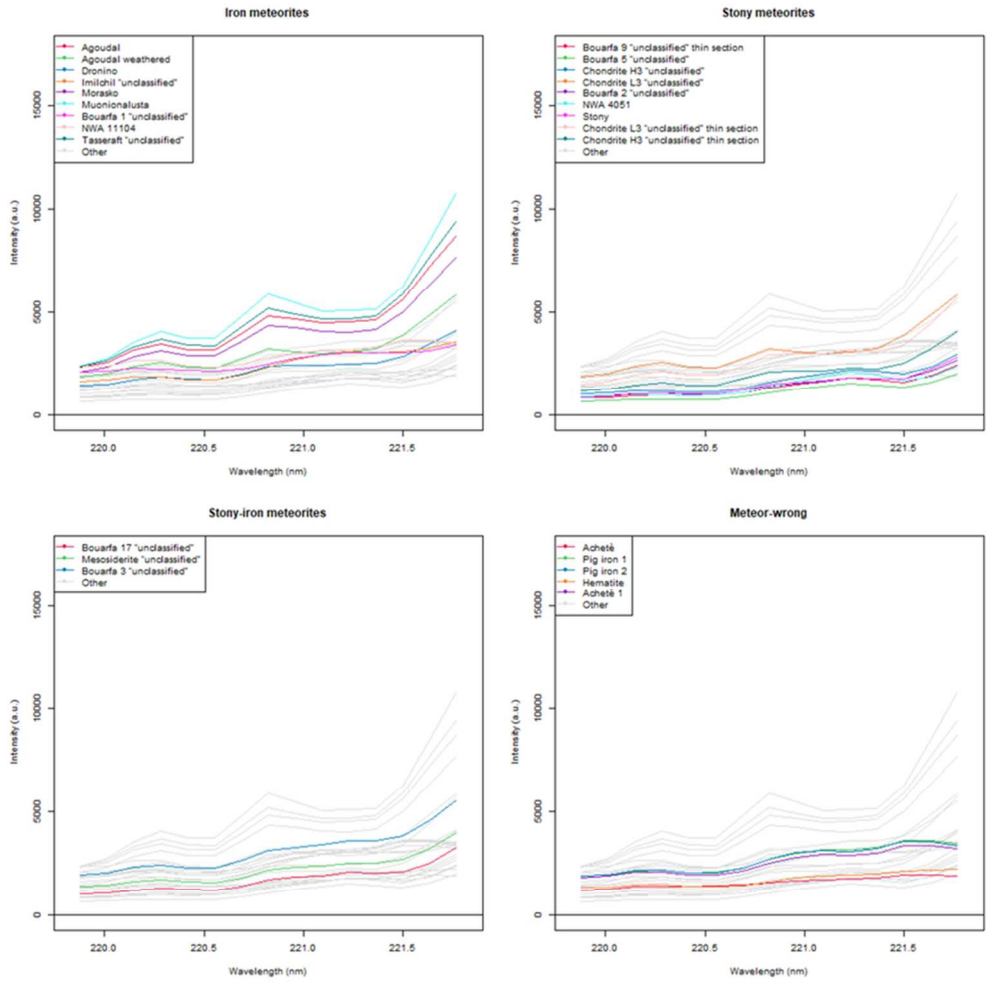

Figure 8 - Intensities of Ir wavelength at 220.81 nm for Ir measured on each sample and grouped according the corresponding class.

74x74mm (300 x 300 DPI)

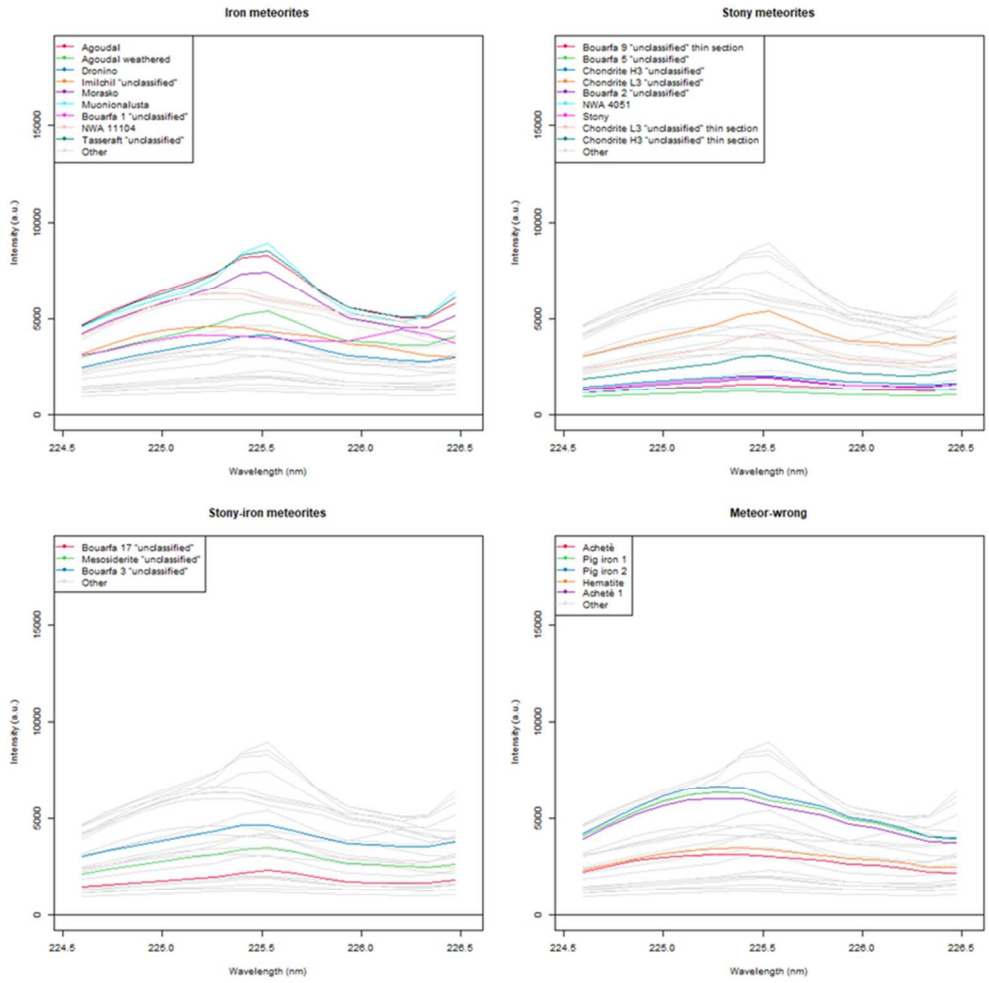

Figure 9 - Intensities of Ga wavelength at 225.50 nm for Ga measured on each sample and grouped according the corresponding class.

74x74mm (300 x 300 DPI)