The Use of Electron Spin Resonance Spectroscopy for Determining the Provenance of Classical Marbles

D. Attanasio

Istituto di Chimica dei Materiali, CNR, Roma, Italy

Abstract. The possibility of identifying the provenance of classical marbles and solving related questions, such as the joining fragments problem, via electron spin resonance spectroscopy has been re-examined. The method is based on characterization of the Mn^{2+} impurity ubiquitously present in marbles. Six different, although correlated, spectroscopic variables, such as intensity, linewidth, metal hyperfine splitting and others have been measured, together with some petrographic properties, for over 500 samples belonging to 14 different Italian, Greek and Turkish quarrying sites. The work, still in progress, is aimed to establish a marble database including all the historically relevant sites within the Mediterranean basin. The experimental data matrix has been analyzed with the aid of multivariate statistical techniques, primarily linear and quadratic discriminant analysis, and the ability of the method to classify correctly unknown samples has been estimated through standard techniques (resubstitution, jackknife), but also employing control "unknown" samples. The essential result is that, although the all-variables approach may describe the data set very accurately, the predictive power is due to four variables only, which are a suitable combination of spectroscopic and petrographic information. In this way 82.4% of the control samples could be assigned correctly, whereas inclusion of additional variables in the classification rule may result in substantially poorer performance. The conclusion is that ESR spectroscopy, although not providing a complete and general solution for the marble provenance problems, is probably, at the moment, the most developed methodology for identifying marbles. Its results, coupled with artistic historical information, deal correctly with a number of relevant archaeometric problems.

1 Introduction

The applications of electron spin resonance spectroscopy in the field of cultural heritage, which includes problems of predominant artistic, historic or archaeological interest, are rather scarce. Among the few contributions reported so far, the most relevant and best known is, certainly, ESR dating. The method, which has given essential support to our present knowledge of the origin and evolution of modern humans [1] measures the concentration of radiation-induced paramagnetic centres and operates in a time range where other dating techniques cannot be applied or are unreliable [2].

However, the capabilities of ESR are potentially much broader and refer to widely different materials and artifacts. Such expectation stems from the obvious observation that most diamagnetic materials, from glasses to wooden objects, to potteries, pigments, tissues, organic remains and others, contain measurable concentrations of paramagnetic impurities and are, therefore, amenable to ESR spectroscopy.

Such impurities and their spectroscopic details are more or less characteristic of the specific sample and may be used as probes for identification, provenance studies, detection of later modifications or forgeries, as well as for monitoring reassembling and restoration of fragments. More importantly the paramagnetic impurities often undergo time evolution or behave as active centres and catalysts for the onset of various chemical processes. This means that they can, in some way, trace the history of the artifact, define its conservation state on a physicochemical basis and, sometimes, even suggest possible procedures for the restoration or, at least, for avoiding or slowing down further degradation.

Unfortunately most of the work shortly mentioned above is reported in the specialized literature and is probably unfamiliar to most ESR spectroscopists. This is a great inconvenience since many more ESR applications in the field of cultural heritage would be possible simply if knowledge of the existing problems were more diffused. For this reason we thought that our work, dealing with the characterization of marbles used in classical architecture and sculpture and the determination of their provenance, although making use of fully conventional ESR techniques and not claiming any spectroscopic novelty, could be of some interest to many researchers in the field.

2 The Problem

White and colored marbles have been used for sculpture and architectural purposes for several thousands years. In some periods, notably during the Roman imperial period, this material acquired enormous importance; it was the visible sign of social and political prestige and its use, with the aim of celebrating the imperial power, had also precise ideological significance. The most important quarries were of imperial property and a complex organization was developed for the exploitation, transportation and commercialization of marble [3–7]. As a consequence, investigation of the material marble is a multifaceted problem concerning not only a deeper knowledge of the artworks we inherited from the past, but also contributing to a detailed picture of ancient trades, commercial routes and economical relationships between different countries.

Within this context, the characterization of marbles and the determination of their provenance from one of the several quarrying sites exploited in the past is a point of primary relevance. However, while the place of origin of colored stones is often relatively simple to be determined on the basis of their morphological properties, the problem of tracing the provenance of white marbles· represents a challenging archaeometrical problem.

To this purpose many analytical techniques have been developed, the most successful and widespread being neutron activation trace element analysis [8, 9] and the determination of the carbon and oxygen isotopic ratios [10, 11]. However, it is generally recognized that effective discrimination requires the combined use of the existing methods, as well as the exploitation of new ones. At present a combination of petrographic, physicochemical and artistic or historic information is often used with the aim of reducing the level of uncertainty and achieve reliable assignments for artifacts of unknown origin [12–14].

In this field use of ESR spectroscopy, based primarily on detection of the ubiquitous Mn^{2+} impurity, was introduced several years ago by Cordischi et al. [15] and showed immediately promising results. Subsequently scattered reports appeared in the literature [16–20], but the method never gained wide popularity.

In our opinion this is due to a number of reasons, which are partly specific of the technique or, at least, of the way it was developed, but also common to many other provenancing techniques. Easy-to-use databases are still to come and the assignment of unknown samples on the basis of literature data often turns out to be a rather naive procedure.

Beside the determination of provenance an additional problem, frequently posed by archaeologists, concerns the possibility of identifying joining fragments. This implies careful examination of the parameter variability within blocks of specified dimensions, so as to be able to discriminate between such variability and that exhibited by the entire quarry.

Being involved in a long-term historical and archaeological program on the provenance of white marbles used in Rome and Southern Italy during the imperial period, we devoted some time to re-examine the real usefulness of ESR spectroscopy as a tool for characterizing and provenancing white marbles [21–23], (Attanasio and Platania, unpubl.), (Attanasio et al., unpubl.), (see http://mlib.cnr.it/marble/main.html).

3 Experimental

3.1 Materials

At present more than 500 marble samples have been measured. Most of them have been collected by us during the last two years, whereas a small "starting set" came from a marble collection existing at the University of Rome. The samples belong to 14 different quarrying sites which include the most important Italian and Greek locations known to have been regularly exploited in the antiquity. Anatolian marbles, which are equally important, are represented only by the quarries of Aphrodysias, Altintaš and Afyon (Docimium marble). Collection of samples from other Turkish sites is still under way, with the final aim of setting up an ESR marble database comprising all the relevant historical quarries existing within the Mediterranean basin. Table 1 summarizes the number and provenance of the samples, whereas the approximate geographical location of the quarries is given in Fig. 1.

Table 1. Number and provenance of valid samples in the electron spin resonance marble database.[*]

No.	Site	Samples
1	Carrara	88
2	Doliana	10
3	Filippi	8
4	Hymetto	43
5	Lesbos	9
6	Mani	23
7	Naxos	47
8	Paros	66
9	Pentelicon	87
10	Thasos	33
11	ThasosD	19
12	Tinos	5
13	Afyon	76
14	Altintas	13
15	Aphrodysias	22
	Total	549

[*] 15 groups are listed because the island of Thasos includes the largely different dolomitic (ThasosD) and nondolomitic (Thasos) marbles.

Fig. 1. Approximate geographic location of some important Mediterranean quarrying sites. Abbreviations are as follows: Afy – Afyon or Docimium marble; Alt – Altintaš; Aphr – Aphrodysias; Ca – Carrara; Dol – Doliana; Fi – Filippi; Hy – Hymettos; Le – Lesbos; Ma – Mani; Pro – Marmara or Proconnesian marble; Nx – Naxos; Pa – Paros; Pe – Pentelicon; Th – Thasos; Us – Ušak.

3.2 Petrography

Four simple petrographic characteristics of the samples have been measured with the aim of increasing the discrimination provided by ESR alone. They are the typical foetid *ODOUR* that some marbles emit upon fracture, the maximum grain size (*MGS*), the *COLOUR* and the colour homogeneity or *STDCOL*. At present the variable *ODOUR* has not been quantified and is given on a simple threefold scale (1-2-3) corresponding to the odour being absent, detectable or strong. The variable *MGS* was obtained by microscopic examination of a polished surface, after etching with diluted hydrochloric acid (0.2 N) for about 30 s, the *COLOUR* and the *STDCOL* were the result of image analysis of the same surface (before etching) carried out after scanning with a conventional HP Scanjet 2cx scanner. *MGS* is given in millimeter, whereas the variable *COLOUR* is measured on a 8-bit gray scale where 0 represents black and 255 white, after selecting a fixed area corresponding to ca. 5000 pixels and standardizing the scanner with a Kodak gray scale. The *COLOUR* of a sample is defined as the mean value of the distribution obtained by analyzing by pixel the scanned image with the aid of standard software (Adobe Photoshop). The corresponding standard deviation is the variable *STDCOL*.

3.3 ESR Measurements

The spectra were recorded at room temperature, at the X-band frequency, with a standard, computer-controlled Varian E-9 spectrometer, equipped with a multi-purpose rectangular cavity. When necessary, the field and frequency were measured with a Bruker BNM-20 gaussmeter and with a Systron-Donner 6246A counter. To obtain a higher filling factor and a slightly better sensitivity all the spectra were recorded with the variable-temperature quartz dewar inserted into the cavity. All the instrumental settings were kept constant and were as follows: frequency ca. 9.10 GHz, microwave power 12 db (12.5 mW), modulation amplitude 1 G, time constant 0.1 or 0.3 s with corresponding recording times of 4 or 8 min. Only the spectrometer gain, which exhibits a quasilinear behavior, was varied (between 100 and 8000) according to the actual sample intensity.

The spectra were recorded on weighted amounts (ca. 30 mg) of finely ground marble powder packed in constant diameter (2.8 mm) quartz tubes. The sample height, lower than 5 mm, allows the powder to be fully contained within the cavity region possessing maximum and constant sensitivity. Proper sample location, corresponding to maximum signal intensity, was found by trial and error and then maintained constant with the aid of a simple teflon spacer.

Usually the marble ESR spectrum is the standard spectrum given by Mn^{2+} substitutionally diluted into the calcite lattice and presented in Fig. 2. Occasionally dolomitic marbles, or marbles containing variable amounts of dolomite are found. In this case the dolomitic spectrum or clearly detectable satellite lines appear because of a rather different zero-field splitting parameter and allow quantitative estimation of the dolomitic content.
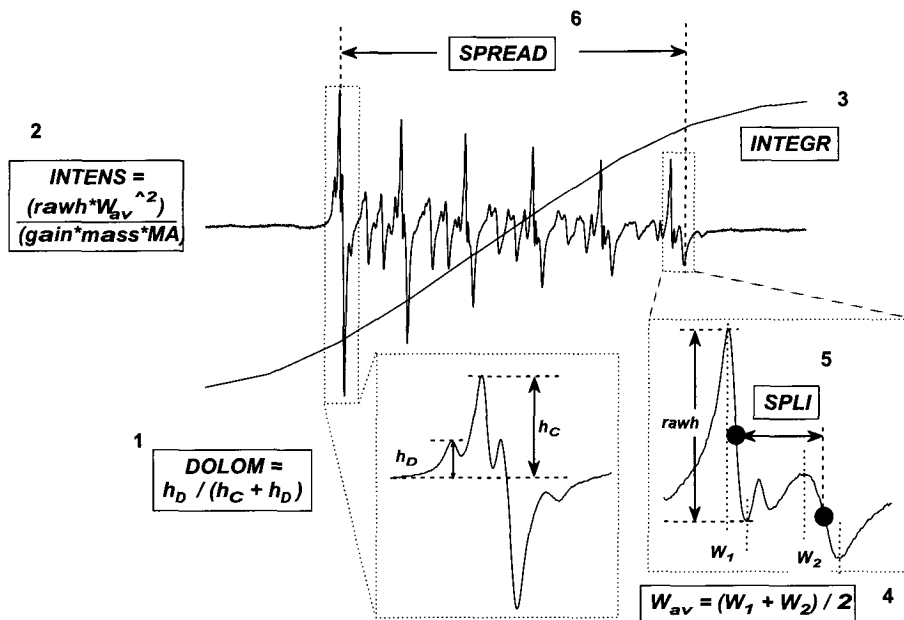
Fig. 2. The room-temperature, X-band Mn^{2+} spectrum of marble. Satellites due to the presence of dolomite are clearly seen. The definition and derivation of the six ESR variables is schematically given.

Six different, although correlated, variables were obtained from the spectra. They are: the dolomitic content (*DOLOM*), the spectral intensity (*INTENS*), the spectral integral (*INTEGR*), the linewidth (W_{av}), the doublet splitting (*SPLI*) and the total extension of the spectrum (*SPREAD*). Some of these variables do not correspond to conventionally defined spectroscopic quantities. However, they were conveniently used in the present context. Their definition and derivation are shortly given in Fig. 2. Additional experimental details can be found in a specifically designed web site (http://mlib.cnr.it/marble/main.html).

However, a few comments are due here. Dolomite is the double carbonate $MgCa(CO_3)_2$ and manganese substitutes not only Mg^{2+}, but Ca^{2+} as well, the substitution ratio being variable. This means that the variable *DOLOM*, although related to the amount of dolomite present in the sample, does not coincide with it. An additional reason for stressing this point is that the variable value is obtained by measuring peak intensities and not integrals, as should be done.

Similar problems arise with the variable *INTENS*. In this case the choice of the low-field line of the high-field doublet has been made simply on the basis of better line resolution and higher and easily measurable linewidth. The aim was to keep the data collecting step as simple as possible. For this reason we chose not to measure absolute concentrations, but only easily measurable, comparable, concentration-dependent parameters.

Finally the measurements were normalized with two different ·standards, one for the signal intensity and integral the second for the field linearity. The former

was a freshly prepared 100 ppm toluene solution of DPPH (Aldrich, 95%). For the field linearity we used a standard dolomitic limestone with a known Mn^{2+} content (N368 BCS, Bureau of Analysed Samples, Newham Hall, Newby, Middlesbrough, Cleveland, UK). Again a detailed description of the standardizing procedures can be found at the web address cited above.

3.4 Statistical Data Analysis

Most of the analyses reported below have been performed with the aid of the commercial package STATISTICA (release 5.1, '97). The jackknifed calculations and the quadratic discriminant analyses were carried out by the SYSTAT 7.0 (SPSS Inc.) package.

4 Results and Discussion

4.1 The Variables and the ESR Database

Traces of Mn^{2+} are the only systematically detectable paramagnetic impurity present in white marbles. Sometimes additional weak lines may be observed. They are sharp signals in the $g = 2$ region, which have been ascribed to radiation-induced defects and to traces of organic radicals [19], or broad, featureless lines probably due to Fe^{3+} impurities. Due to their occasional presence none of these signals may form the basis for systematic marble characterization. As far as the spectrum of Mn^{2+} is concerned, its sharp lines allow rather easy and reliable intensity measurements, whereas the low anisotropy is a disadvantage for the use of spin-Hamiltonian parameters as probes of small, provenance-dependent crystal field variations of intensity and symmetry.

A possible, exhaustive choice for the variables to be measured could simply be to use the intensity, width and position of every spectral line. However, for the reasons already discussed, we preferred to use the fewer variables specified in Section 3. Some of them (*INTENS, INTEGR, W_{av}*) are concentration-dependent, others (*SPLI, SPREAD*) are related to spin-Hamiltonian parameters such as the metal hyperfine constant or the zero-field splitting parameter, one (*DOLOM*) reflects the sample composition in terms of dolomitic content. Obviously extensive correlation exists between these variables, although each of them carries in some additional information. Subsequent data analysis will clarify the effective significance of each variable, as well as the presence of any redundancy in the database.

This same analysis has shown, as expected, that the ESR variables alone do not always provide satisfactory discrimination. In particular it turned out that quarrying sites showing severe ESR superposition may be easily distinguished on the basis of different properties. This is the reason which led us to introduce the petrographic variables.

Table 2. Mean values, given by quarry, for the ten variables included in the marble database. The standard deviations are given in parentheses.[a]

		DOLOM[b]	INTENS	INTEGR	W_{av} (G)	SPLI (G)	SPREAD (G)	MGS (mm)	COLOUR	STDCOL	ODOUR
1	Afyon	0.0 (.0)	2.21(1.2)	13.21 (9.5)	3.44(.29)	15.24 (.23)	489.4(1.6)	0.91 (.30)	195.8(24.6)	8.00(3.8)	1.0 (.0)
2	Altintas	0.088 (.15)	0.37 (.35)	2.70 (2.0)	2.70(.25)	15.28 (.23)	486.8(1.4)	0.98 (.78)	198.4 (9.4)	6.93(2.7)	1.0 (.0)
3	Aphrodysias	0.019 (.04)	0.74 (.66)	3.36 (3.0)	3.64(.31)	15.42 (.23)	488.4(2.2)	1.90 (.50)	198.8(22.9)	10.9 (4.7)	1.18(.6)
4	Carrara	0.015 (.04)	0.84 (.37)	3.14 (1.37)	4.15(.22)	15.04 (.13)	486.5(3.1)	0.74 (.17)	209.9(19.3)	6.19(3.3)	1.06(.2)
5	Doliana	0.125 (.03)	0.128(.07)	0.623 (.67)	3.43(.24)	15.30 (.15)	488.0 (.61)	—	—	—	1.0 (.0)
6	Filippi	0.0 (.0)	0.132(.06)	1.06 (.65)	2.92(.33)	15.40 (.19)	487.0 (.93)	—	—	—	2.25(.9)
7	Hymetto	0.307 (.04)	0.13 (.15)	0.82 (1.1)	2.98(.26)	15.48 (.14)	488.0(2.0)	0.69 (.20)	181.9(24.3)	8.27(3.4)	2.37(.5)
8	Lesbos	0.070 (.07)	0.297(.4)	13.86(19.2)	2.88(.17)	15.47 (.16)	486.5(1.02)	—	—	—	1.0 (.0)
9	Naxos	0.0 (.0)	0.92 (.89)	5.16 (6.2)	3.22(.52)	15.36 (.22)	487.7(1.2)	5.01(1.7)	203.5(13.4)	10.3 (2.5)	1.93(.9)
10	Paros	0.023 (.04)	0.13 (.09)	0.84 (0.48)	3.17(.29)	15.53 (.16)	489.0(2.5)	1.87 (.62)	206.0(16.6)	6.54(1.8)	1.17(.5)
11	Pentelicon	0.0 (.0)	2.03(1.04)	13.72 (8.8)	3.68(.31)	15.29 (.23)	488.2(2.9)	0.99 (.21)	228.0(13.5)	5.50(1.9)	1.0 (.0)
12	Thasos	0.0017(.01)	1.49(1.2)	7.99 (9.26)	3.58(.31)	15.28 (.13)	484.2(3.5)	4.33(1.0)	191.7(13.9)	10.9 (1.4)	2.0 (.4)
13	ThasosD	1.0 (.0)	0.35 (.25)	2.86 (2.16)	3.31(.45)	56.53 (2.6)	523.0(2.4)	2.21 (.72)	240.4 (9.5)	4.84(1.4)	1.71(.5)
14	Tinos	0.334 (.21)	0.346(.25)	3.28 (2.1)	3.26(.21)	15.20 (.10)	485.6 (.72)	—	—	—	1.0 (.0)

[a] *INTENS* and *INTEGR* are given in arbitrary units. The units of the petrographic variables are discussed in Section 3.2.
[b] The variable *DOLOM* does not coincide with the dolomitic content. See Section 3.3.

To summarize, all the above experimental results are conveniently organized in a database which has the form of a 549×10 data matrix containing the values of the ten variables (6 ESR and 4 petrographic) for all the measured marble samples, which, in turn, are grouped into 14 quarrying sites or 15 different classes, remembering that the Thasos marbles are split into two sets, depending if they are dolomitic or not. Table 2 gives the mean values and standard deviations for all the variables by class.

4.2 The Statistical Approach

Multivariate statistical analysis of the information contained in the database defined above may allow one to set up the classification rule to be used for assigning samples of unknown origin.

To this purpose both so-called unsupervised and supervised methods [24] have been used in the literature. However the former, which are essentially clustering techniques [25] and seek for an intrinsic grouping structure of the data, have proven not particularly suitable for dealing with the marble problem. Their use has been constantly decreasing and they will not be mentioned any further. In addition the problem we are dealing with is a typical allocation problem, where the quarries are well prespecified groups which can be sampled in detail. In this sense it is a problem particularly amenable to supervised techniques.

Among these latter, which use the provenance information contained in the database to set up the classification rule, discriminant analysis [26, 27] is certainly the most popular and widespread. However, other techniques, such as classification trees or artificial neural networks have been developed and have attracted considerable or, sometimes, enormous interest [26, 28, 30]. Due to increased flexibility their performance in discriminating marbles is worth to be tested. However, the following discussion is focussed on discriminant analysis alone, whereas detailed analysis of these alternative approaches is left for the future.

There are two main aspects of the discriminant analysis process which need to be stressed and these are the choice of an effective discriminant space capable to maximize group separation and the definition of a proper measure of distance to be used in this space.

The first point is accomplished by linear transformation of the variables coordinates. The purpose is to maximize the so-called between-the-groups variance (i.e., the distance between the group means) with respect to the variance within the groups. In other words, we look for composites of the original variables exhibiting maximum correlation with the grouping variable. The linear discriminant functions are those combinations of the original variables which maximize such correlation or the ratio of the between over within variances.

Geometrically speaking, this corresponds to transforming the original coordinate system so as to direct the new axes along the principal axes of the between the groups ellipsoids. Algebraically the problem reduces to diagonaliza-

tion of the **B** (between) covariance matrix standardized against the **W** (within) matrix, i.e., to finding the eigenvalues and eigenvectors of the $\mathbf{W}^{-1}\mathbf{B}$ matrix. The eigenvectors, then, are precisely the required linear combinations of the original variables, also called linear discriminant functions.

The problem of using a suitable distance measure is also a crucial one in that it can be shown that if the choice is made properly the problem of finding the highest probable group for a sample reduces to a problem of closeness [27].

4.3 Discrimination and Assessment

4.3.1 The Quarries Discrimination

Although the general problem of being able to discriminate a large number of quarries is certainly a relevant question to be posed, it is also true that many of the source sites are only of local interest or can be excluded a priori on the basis of historical information, knowledge of the periods the quarries were in operation, investigation of the types of workmanship and so on. Therefore, focussing on the problem of the marbles used in Rome we may reasonably select a subset of the entire database comprising only 6 sites (7 groups) which are: Afyon, Aphrodysias, Carrara, Paros, Pentelicon, Thasos and ThasosD. This set, although incomplete for the absence of some important anatolian marbles, includes all the relevant sites actually present in the database; the purpose of the following discussion will be to verify if satisfactory discrimination may be obtained within this reduced database.

4.3.2 Evaluating the Classification Rule

The crucial point to be discussed is the criterion we use to estimate the level of success of the method and to validate it. After performing a discriminant analysis on the above database (374 valid samples, 7 groups) the hit rate or rate of success, based on the so-called resubstitution method, is 90.4%. This means that the data points are first used for setting up the discriminant rule and then treated as "unknowns" to see if they are reassigned correctly, the result being the percentage of correct classification. Obviously if the system is sufficiently flexible (i.e., if it uses a high number of variables) almost perfect "fitting" of the database may be obtained, but this does not give any certainty about the capability of the method to generalize correctly, i.e., to perform equally well with real unknowns. On the contrary, we expect the resubstitution result to be more or less positively biased simply because the data set is only an approximate representation of the real marble population distribution.

To overcome this problem the database may be split into a design set, used to set up the classification rule, and a test set used to verify its validity. However, in order to avoid the disadvantage of a design set of reduced size, various

cross validation methods have been developed. In the following the so-called jackknife technique will be used. The essential idea of these cross validation procedures [26] is to extract subsets of the data to test the performance of the classification built by the remainder, to repeat this for different subsets and to average the results. In this way all the experimental data points contribute to the definition of the classification rule.

The results obtained following the above lines are summarized in Table 3. Figures 3 and 4 give a graphical representation of the same results by plotting the data points (and the corresponding 90% probability group ellipses) as a function of the two first (most important) discriminant coordinates. As already noted, these latter are suitable linear combinations of the original variables obtained by diagonalization of the covariance matrixes. The discriminant plots obtained in this way represent the best bidimensional representation of a multidimensional reality. However, they have only illustrative purposes and are not used for assigning samples. This piece of information is obtained by inserting the variables values of the unknown samples into the database and calculating the distances from the various group centroids in the 10-dimensional variable space. This allows one to obtain not only the most probable groups, but also the following choices, each with its corresponding probability.

The results obtained so far may be improved in, at least, two ways. The classification procedure we have used, termed linear discriminant analysis, is based on the assumption that all the groups have similar dispersion characteristics and makes use of a single (pooled) within-the-groups covariance matrix. It is obvious that the equal-variances assumption is rarely true and using for each group its actual dispersion characteristic, or covariance matrix, means to exploit this additional information to obtain improved classification performance. To this. purpose so-called quadratic discriminant techniques are employed.

Table 3. Percent success rate, given by quarry, for the marble database by linear or quadratic discriminant function analysis with 10 raw or logarithmically transformed variables.[a]

| | LDA | | QDA[b] | |
	Raw var.	Log transf.	Raw var.	Log transf.
Afyon	86.4 (86)	87.9 (86)	91 (85)	94 (89)
Aphrodysias	63.6 (54)	63.6 (50)	73 (73)	77 (69)
Carrara	96.6 (94)	96.6 (96)	98 (96)	98 (98)
Paros	88.3 (81)	90.0 (84)	97 (92)	87 (86)
Pentelicon	89.7 (90)	93.6 (93)	96 (95)	93 (88)
Thasos	100 (100)	97.0 (97)	100 (97)	100 (97)
ThasosD	100 (100)	100 (100)	100 (100)	100 (100)
Total	90.4 (88)	91.5 (89)	95 (92)	93 (90)

[a] Resubstitution and jackknifed (bracketed values) results are given.
[b] The *DOLOM* and *ODOUR* variables give rise to group matrix singularities and are not used in quadratic analysis. Therefore this latter is based on eight variables only.
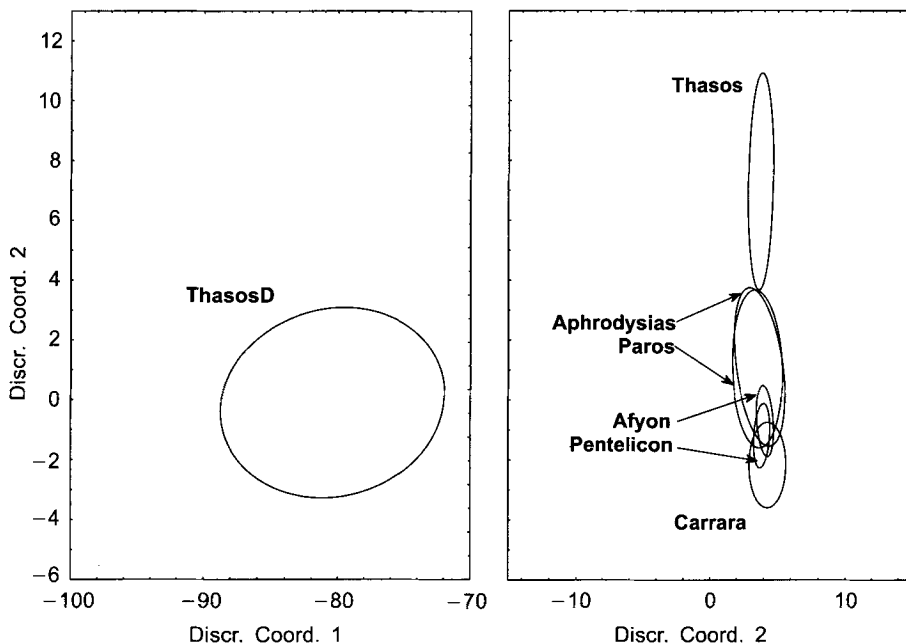
Fig. 3. Discriminant plot of the ESR and petrographic marble database. The 90% probability ellipses are given for the seven most important quarrying sites in the space defined by the first two discriminant variables. These are suitable linear combinations of the ten variables defined in the text. ThasosD includes dolomitic samples. This explains the very large separations with respect to all the other, calcitic, marbles.

In addition both linear and quadratic discriminant analyses are based on the hypothesis that the data points, within each group, are normally distributed. Again the data seldom comply with this assumption and any transformation which improves the normality of their distribution may also improve the final performance. A simple and well known way to accomplish this result is to use a logarithmic transformation of the data. The results obtained adopting both quadratic analysis and logarithmic transformation have been calculated via the resubstitution and jackknife methods and are again summarized in Table 3. The conclusion is that, in this case, quadratic analysis carried out using the raw variables gives the best possible results.

4.3.3 Validation and Variables Selection

The classification rule developed so far still cannot be considered operational. In fact, the validation procedures have been carried out always using samples belonging to the data set. In general, and even more specifically in the case of marbles, we expect the database samples to exhibit some sort of "added similar-
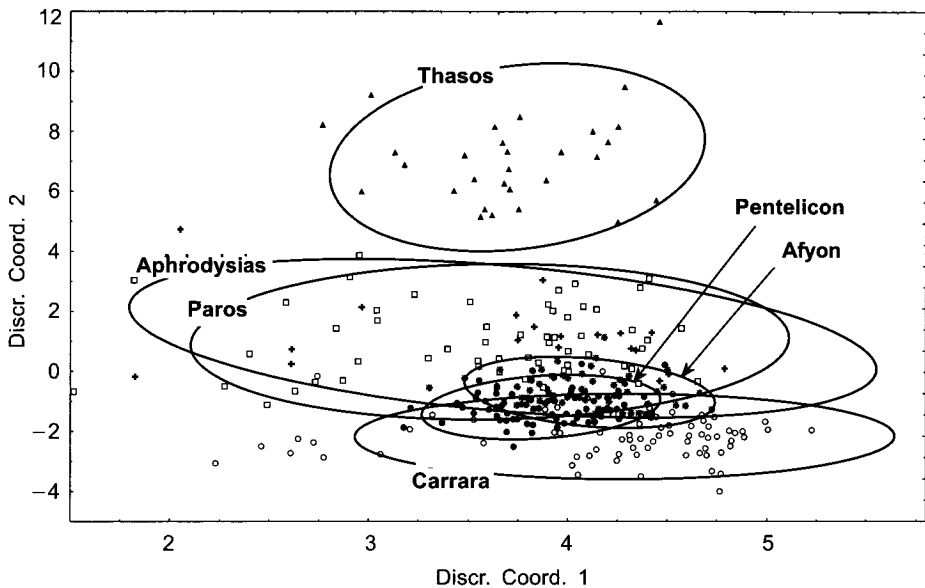
Fig. 4. Detail of the region of calcitic marble in Fig. 3, showing individual location of the experimental data points.

ity" which does not extend to real unknowns. There are several reasons for this, the most important being:

– quite commonly modern marbles are measured and used to assign antique ones,
– sampling is usually carried out following criteria which often are not as general as we may think,
– despite of the standardization procedures, the database measurements, performed by a single operator at a certain time, are certainly more comparable than measurements carried out separately, at a later stage.

For these reasons, controls carried out using samples of known provenance, not belonging to the data set but collected under different circumstances or originating from artifacts, seem indispensable to estimate the effective capability of the method to generalize correctly.

As we shall see, this kind of validation may also help in discovering the redundancies which are certainly present in our classification rule. It is well known that including in the method variables which do not contribute sizably to the discrimination is not only useless, but may also lead to substantially poorer performance.

In the following the assignment of a control set containing 17 samples of the type defined above is discussed, also presenting the results of stepwise discriminant analysis which allows us to introduce the variables one by one (starting with the most important) and to verify the results obtained at each stage.

In Fig. 5 the performance of the analysis is presented giving the value of the Wilks λ as a function of the added variables. The Wilks λ is defined as the ratio
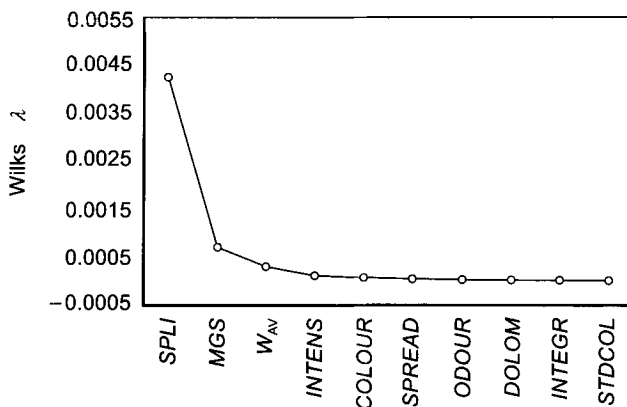
Fig. 5. The Wilks λ as a function of the number and type of variables included in the analysis. The plot shows that the first four or five variables provide most of the discrimination.

between the determinants of the within-the-groups and between-the-groups covariance matrixes, it ranges from 1 to 0, a low value indicating better discrimination. On this basis Fig. 5 clearly indicates that only the first four, or perhaps five, variables give a sizeable contribution to discrimination, the presence of the others being almost irrelevant.

However, two points must be noted. The first is that stepwise analysis (at least as implemented in commercial software) does not perform an exhaustive choice for the best discriminant set. In other words, taking the first n variables selected stepwise does not mean to select the best possible subset of size n. The second point is that the Wilks λ is a criterion related to group separation, whereas we are interested in classification accuracy and the two concepts are by no means equivalent.

For these reasons we carried out a series of calculations, testing different sets of reduced size and checking the classification performance in various ways. In agreement with the conclusion reached above, the resubstitution and jackknife results indicate that sizeable contributions to discrimination are provided only by four out of ten variables, see Table 4. These are the splitting, the intensity, the grain size and the colour, while the other six have no apparent influence on the results. However the results obtained with the control set are quite different and demonstrate that the effect of the less discriminating variables may severely worsen the performance (see Table 5 and Fig. 6).

Referring to the jackknife results the classification accuracy may reach 92% for the data set (jackknifed quadratic analysis with 8 raw variables) but drops to only 64.7% if the same method is applied to the control samples. In this case the best result is 82.4% and is obtained with a rather different, and simpler, type of analysis (linear analysis with 4 raw variables). It turns out that in "real" applications some variables (*SPLI, INTENS, MGS, COLOUR*) exhibit high discriminating ability, others are irrelevant (*SPREAD, DOLOM, ODOUR*), some may have a strong negative impact on classification accuracy (*INTEGR*, W_{av}, *STDCOL*).

Table 4. Comparison of the total percent success rate obtained for the marble database by the all-variables or selected-variables approaches.

Validation method	LDA, 10 var	LDA, 4 var[a]	QDA, 8 var[b]	QDA, 4 var[a]
Resubstitution	90.4	77.5	95	87
Jackknife	88	76	92	85
Control samples	58.8	82.4	64.7	70.6

[a] The 4 variables selected are: *SPLI, INTENS, MGS* and *COLOUR.*
[b] The *DOLOM* and *ODOUR* variables give rise to group matrix singularities and are not used in quadratic analysis. This latter is based on eight variables only.

This kind of behavior has been determined in a purely empirical way and its general validity may be questioned. However, it has been tested in a variety of applications, with databases including different quarrying sites, and found to retain its validity to a remarkable extent (Attanasio D., Platania R., unpubl.).

Intuitively we may note that the most important variables are those which describe basic marble characteristic (intensity, colour, grain size), whereas others (splitting) are important simply to discriminate between dolomitic and non-

Table 5. Assignment of 17 control samples by different types of linear and quadratic discriminant analysis, with the raw variables.[a]

Sample	Provenance	Assignment LDA, 10 var	Assignment LDA, 4 var	Assignment QDA, 8 var[b]	Assignment QDA, 4 var
X1	Carrara	Penteli*	Afyon*	Afyon*	Aphrody*
X2	Aphrody	Afyon*	Aphrody	Aphrody	Aphrody
X4	Penteli	Penteli	Penteli	Penteli	Penteli
X5	ThasosD	ThasosD	ThasosD	ThasosD	ThasosD
X7	Carrara	Penteli*	Carrara	Carrara	Penteli*
X9	Aphrody	Paros*	Paros*	Paros*	Paros*
X10	Carrara	Carrara	Carrara	Carrara	Carrara
X12	Penteli	Penteli	Penteli	Penteli	Penteli
X13	Aphrody	Penteli*	Aphrody	Penteli*	Aphrody
X15	ThasosD	ThasosD	ThasosD	ThasosD	ThasosD
X16	Aphrody	Paros*	Aphrody	Aphrody	Aphrody
X19	Carrara	Penteli*	Carrara	Carrara	Penteli*
X20	Paros	Paros	Paros	Aphrody*	Paros
X21	Paros	Paros	Paros	Paros	Paros
X22	Paros	Paros	Aphrody*	Paros	Paros
X23	ThasosD	ThasosD	ThasosD	ThasosD	ThasosD
X24	Carrara	Carrara	Carrara	Aphrody*	Afyon*

[a] Misclassifications are marked by asterisks.
[b] The *DOLOM* and *ODOUR* variables give rise to group matrix singularities and cannot be used in quadratic analysis. This latter is based on eight variables only.
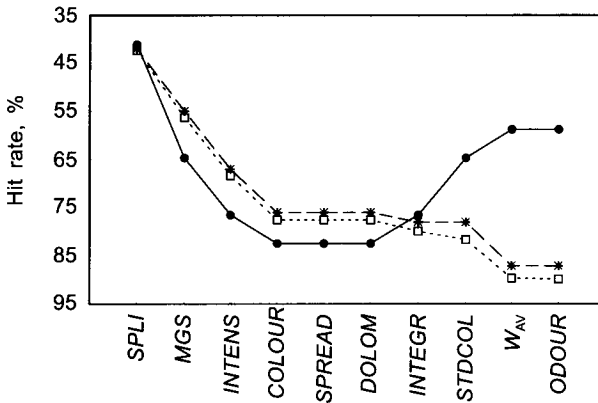
Fig. 6. Three different estimates of the percentage of correct classification given as a function of the number and type of variables included in the analysis. As discussed in the text, the control samples set shows that a simple four-variables analysis provides the best performance. Inclusion of additional variables is irrelevant or leads to worse results. □ resubstitution, * jackknife, ● control samples.

dolomitic marble. Additional variables describe marginal properties or are affected by rather large errors or by the conservation state of the stone. They prove to be useful in describing details and oddities of the data set, but turn out to be misleading when applied to the identification of unknown samples.

4.4 The Variability of the ESR Parameters with Linear Dimensions

As already mentioned, this problem is related to the possibility of identifying joining fragments or recognizing restorations and forgeries and will be shortly summarized here. Intuitively we expect the variability of the physicochemical parameters to be smaller when the samples belong to a single block of marble of specified dimensions than when they come from the entire quarry or the whole quarrying site. In this way, i.e., comparing variabilities, it should be possible to distinguish joining fragments from samples which, although belonging to the same quarry, are not taken from the same single block.

To test this hypothesis we have compared six different groups of samples, in order of descending dimensions. These are one of the three districts of the Carrara site, namely the Torano valley, one of its quarries, i.e., Polvaccio, and three approximately cubic, single blocks from this quarry with edges of ca. 150, 50 and 25 cm. Finally the reproducibility has been tested with the so-called single point, i.e., a repeatedly measured single sample. The petrographic variables and the dolomitic content were found to be unsuitable for this variability study and were discarded, so that we are left with the remaining five ESR variables.

Referring to the single-variable results of Figs. 7 and 8 we note that the situation is not as simple as anticipated. Usually a block, even a small one, exhibits a single-variable variability which is a substantial fraction of the entire
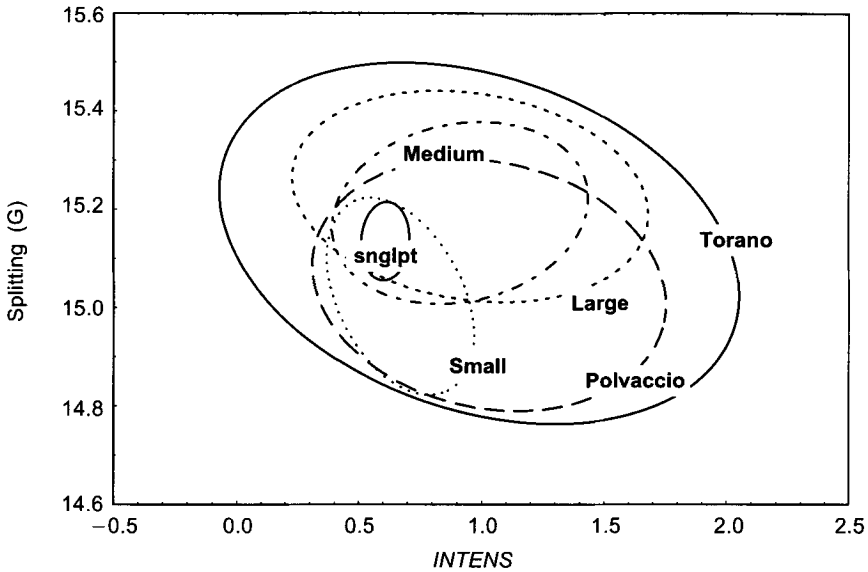
Fig. 7. One of the possible two-dimensional plots showing the variability of marble parameters at various scales with the aid of 90% probability ellipses. Six sets of marble samples are considered. From large to small they are: Torano (one of the three districts of the Carrara site); Polvaccio (a quarry in Torano); the large, medium and small blocks from Torano with edges of ca. 150, 50 and 25 cm; single point (*snglpt*) was obtained by remeasuring several times the same sample.

quarry. In addition, Fig. 8 shows the presence of many inconsistencies, possibly due to experimental errors, inadequate sampling and similar reasons. The standard deviations measured for small blocks are often roughly equal, or even higher, than those found for larger ones or for the whole quarry. On this basis identification of samples belonging to the same block appears to be, at least, problem-
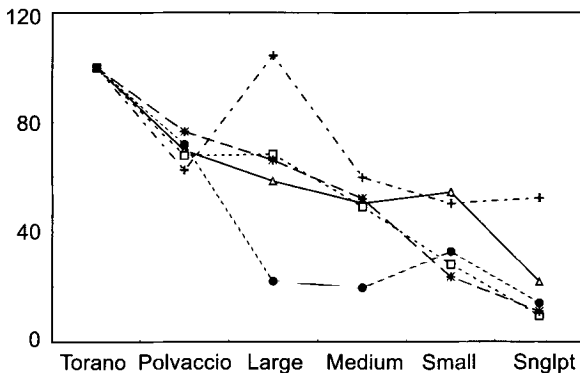

Fig. 8. The single-variable variability of the district, quarry and blocks expressed by the standard deviation of five ESR variables. △ *SPLI*, □ *INTENS*, * *INTEGR*, + W_{av}, ● *SPREAD*.
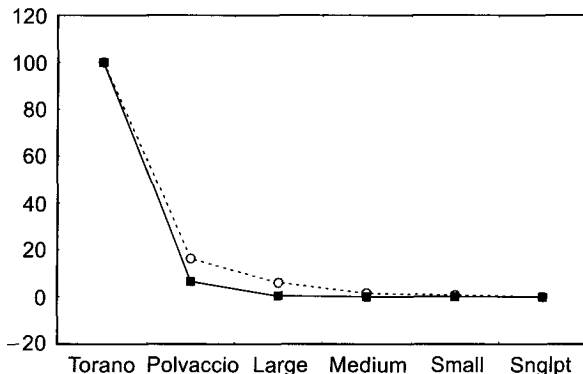
Fig. 9. The total variability of the district, quarry and blocks expressed by the product and general standard deviations (see text for the definitions).

atic. As a matter of fact, similar results are not new and have been already reported from measuring the shifts of the carbon and oxygen isotopic ratios [30, 31].

In order to overcome this difficulty we must take into account all the experimental variables simultaneously, taking advantage, again, of the multivariate nature of the problem. The global, all-variable variability of the samples is expressed by their covariance matrix, a so-called general standard deviation being defined as the square root of its determinant [26]. In the case of uncorrelated variables such parameter simply reduces to the product of the single-variable standard deviations, the relationship between the two quantities being:

$$\text{general } \sigma = \text{product } \sigma \pm \text{off-diagonal terms .}$$

Therefore, for weakly correlated variables the simple product σ may be used, at least as first approximation. The results of Fig. 9 are encouraging, in the sense that the two new parameters exhibit the expected dependence from the block dimensions and approach a simple exponential behavior. On going from the quarry to the blocks the general σ reduces by about two orders of magnitudes and provides the basis for confident identification of joining fragments.

Additional work has shown that such behavior is, qualitatively, quite general. However, more experimental data are necessary to verify whether a general mean variability model may be adopted or variability data must be collected for each single quarry.

5 Conclusions

The work discussed above indicates that electron spin resonance spectroscopy, although not providing a complete and general answer for the marble problem, may yield a remarkable amount of information concerning its characterization and

identification, and appears a valid alternative to other, more widespread physicochemical methods. The technique, especially when coupled with petrographic measurements, seems capable of answering a number of important archaeometrical problems, which refer not only to the provenance of samples but also to their homogeneity and identification as joining fragments. Quite relevant is also the fact that ESR measurements are fast, require limited sample amounts and almost no preliminary manipulation.

An essential, but still overlooked, point is that full exploitation of the available experimental information requires extensive application of the several multivariate statistical techniques available. From this point of view much work remains to be done.

Finally, it must be kept in mind that fruitful application of the techniques illustrated in this paper, as well as of the many others developed in the literature, requires preselection of the probable sources based on historical information, knowledge of the quarries operating times, investigation of the types of workmanship and of the known trading routes.

This approach appears indispensable for reducing uncertainties and is the only way for obtaining correct assignments with high chances of success.

Acknowledgements

References

1. Grün R., Stringer C.B.: Archaeometry **33**, 153–199 (1991)
2. Grün R.: Quaternary International **1**, 65–109 (1989)
3. Gnoli R.: Marmora Romana. Roma: Edizioni dell'Elefante 1988.
4. Borghini G. (ed.): Marmi Antichi. Roma: Edizioni De Luca 1997.
5. Herz N., Waelkens M. (eds.): Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E **153** Dordrecht: Kluwer 1988.
6. Waelkens M., Herz N., Moens L. (eds.): Ancient Stones: Quarrying, Trade and Provenance. Leuven: Acta Archaeologica Lovaniensia, Leuven University Press 1992.
7. Maniatis Y., Herz N., Basiakos Y. (eds.): The Study of Marble and Other Stones Used in Antiquity. London: Archetype Publications 1995.
8. Matthews K.J.: Archaeometry **39**, 321–332 (1997)
9. Meloni S., Oddone M., Mello E., Monna D. in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E **153** (Herz N., Waelkens M., eds.), pp. 293–301. Dordrecht: Kluwer 1988.
10. Herz N.: Archaeometry **29**, 35–43 (1987)
11. Herz N. in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E **153** (Herz N., Waelkens M., eds.), pp. 305–314. Dordrecht: Kluwer 1988.
12. Jongste P.F.B., Jansen J.B., Moens L., De Paepe P. in: The Study of Marble and Other Stones Used in Antiquity (Maniatis Y., Herz N., Basiakos Y., eds.), pp. 143–150. London: Archetype Publications 1995.

13. Matthews K.J., Leese M.N., Hughes M.J., Herz N., Bowman S.G.E. in: The Study of Marble and Other Stones Used in Antiquity (Maniatis Y., Herz N., Basiakos Y., eds.), pp. 171–180. London: Archetype Publications 1995.
14. Moens L., de Paepe P., Waelkens M. in: Ancient Stones: Quarrying, Trade and Provenance (Waelkens M., Herz N., Moens L., eds.), pp. 247–252. Leuven: Acta Archaeologica Lovaniensia, Leuven University Press 1992.
15. Cordischi D., Monna D. Segre A.L.: Archaeometry 25, 68–76 (1983)
16. Lloyd R.V., Smith P.W., Haskell H.W.: Archaeometry 27, 108–116 (1985)
17. Lloyd R.V., Tranh A., Pearce S., Cheeseman M., Lumsden D.N. in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E 153 (Herz N., Waelkens M., eds.), pp. 369–377. Dordrecht: Kluwer 1988.
18. Cordischi D., Monna D., Passariello B., Pensabene P. in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E 153 (Herz N., Waelkens M., eds.), pp. 453–462. Dordrecht: Kluwer 1988.
19. Maniatis Y., Mandi V., Nikolau A. in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E 153 (Herz N., Waelkens M., eds.), pp. 443–452. Dordrecht: Kluwer 1988.
20. Mandi V., Maniatis Y., Bassiakos Y., Kilikoglou V. in: Ancient Stones: Quarrying, Trade and Provenance (Waelkens M., Herz N., Moens L., eds.), pp. 213–222. Leuven: Acta Archaeologica Lovaniensia, Leuven University Press 1992.
21. Armiento G., Attanasio D., Platania R.: Archaeometry 39, 309–319 (1997)
22. Armiento G., Attanasio D., Platania R. in: Materials Issues in Art and Archaeology V, vol. 462, pp. 331–336. Boston: Materials Research Society 1997.
23. Attanasio D., Armiento G., Platania R.: J. Archaeol. Sci. (in press)
24. Hand D.J.: Discrimination and Classification. New York: Wiley-Interscience 1981.
25. Romesburg H.C.: Cluster Analysis for Researchers. Malabar: Krieger Publ. Co. 1990.
26. Hand D.J.: Construction and Assessment of Classification Rules. New York: John Wiley & Sons 1995.
27. Huberty C.J.: Applied Discriminant Analysis. New York: John Wiley & Sons 1994.
28. Ripley B.D.: Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press 1996.
29. Breiman L., Friedman J.H., Olshen R.A., Stone C.J.: Classification and Regression Trees. Monterey: Wadsworth & Brooks/Cole 1984.
30. Matthews K.J.: in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E 153 (Herz N., Waelkens M., eds.), pp. 339–346. Dordrecht: Kluwer 1988.
31. Wenner D.B., Havert S., Clark A.: in: Classical Marble: Geochemistry, Technology, Trade. NATO ASI. E 153 (Herz N., Waelkens M., eds.), pp. 325–338. Dordrecht: Kluwer 1988.

Author's address: Dr. Donato Attanasio, Istituto di Chimica dei Materiali, CNR, P.O. Box 10, 00016 Monterotondo Staz., Roma, Italy