Available online at ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Journal of Cultural Heritage

Journal of Cultural Heritage 10 (2009) 73–81

Original article

# Principal Component Analysis in monument conservation: Three application examples

Antonia Moropoulou [a,1], Kyriaki Polikreti [b,*]

[a] National Technical University of Athens, Materials Science and Engineering Department, Zografou Campus, Athens 15780, Greece
[b] Hellenic Ministry of Culture, Directorate of Conservation of Ancient and Modern Monuments, Department of Applied Research, 81 Peiraios Street, Athens 10553, Greece

Received 3 July 2007; accepted 26 March 2008

**Abstract**

Multivariate statistics is a well-known and invaluable tool in archaeological science but its use is limited in monument restoration. The aim of this work is to demonstrate the effectiveness of Principal Component Analysis (PCA) on the characterization, technology and weathering condition investigation of building materials from historical monuments.

Towards this aim, three examples are given:

The first one is a provenance and technology investigation of the Aghia Sophia (Istanbul, Turkey) bricks, some of which had to be replaced due to weathering, during recent restoration works. It was proved by PCA that the original clay, used for the construction of the bricks, is not similar to the clay of other contemporary constructions in Istanbul but presents high similarity to the raw material of the bricks from a contemporary church in the island of Rhodes (Dodecanese, Greece). Additionally, the technology of the bricks was studied by mercury intrusion porosimetry, strength tests and Scanning Electron Microscopy. The use of PCA gives a very comprehensive way to present the difference in the technology of the dome bricks.

The second presents a classification of mortars from medieval (Byzantine) monasteries, based on their microstructural characteristics (porosity, reverse hydraulicity ratio) and strength measurements. The PCA grouping gives an illustrative diagram depicting the correlation between mortar syntheses and resulting characteristics.

The third case shows an example of the correlation between environmental pollution data and data from the weathering layers of marble surfaces (patina composition, orientation of the monument surface, etc.).


*Keywords:* Multivariate statistics; Principal Component Analysis; Monument restoration; Brick; Historical mortar; Marble patina

## 1. Research aims

The conservation of monuments has been developed into a large interdisciplinary scientific field where the demand for quantified information and values of tested precision and reliability rise more and more in recent years. Monument conservation and restoration includes the study of the original materials, the historical information on their construction as well as the know-how of the architects and mechanics of the specific time period at which the monument was built. Besides, the more information we get on the original material the more easily we can restore the damaged parts by others made on compatible materials.

Statistics, with its unique flexibility to apply its mathematical theory to various and totally different problems, has already been proved invaluable in materials science. Furthermore, the use of statistics in archaeological science, that is as a tool for characterizing, classifying and provenancing archaeological objects, mainly pottery but also stone, pigments

* Corresponding author. Tel.: +30 210 3218475.
*E-mail addresses:* amoropul@central.ntua.gr (A. Moropoulou), kpolikre@ucy.ac.cy (K. Polikreti).
[1] Tel.: +30 210 7723276.

1296-2074/$ - see front matter 
doi:10.1016/j.culher.2008.03.007

and organic material, is well established [1–3]. However, the use of statistics in selecting building materials compatible to the original ones for monument restoration purposes is scarcely been demonstrated.

This paper aims to give some examples of how statistics and especially Principal Component Analysis can be a powerful tool in conservation and restoration problems. The interdisciplinary nature of conservation results in a large number of measured variables, where simple statistical treatments are very poor to reveal hidden correlations of the material properties or give meaningful, illustrative graphs depicting the classification results for large number of objects or samples.

Three different but complementary examples on three different building materials, bricks [4], mortars [5] and stones [6], will be presented here.

## 2. The Aghia Sophia bricks

The first example concerns a detailed study on the bricks of the basilica of Aghia Sophia (Istanbul, Turkey). *The Great Basilica of Aghia Sophia* (532–537 A.D.) is considered as an architectural and artistic prototype of the Eastern Roman Empire [7]. In 2000, Princeton University, Bogazici University and National Technical University of Athens [8] collaborated for providing restoration materials for Aghia Sophia, which on one hand would be compatible with the original ones and on the other hand would show good durability to physical weathering and earthquakes.

We have to keep in mind that, in the case of Aghia Sophia the risk of an earthquake was a critical factor in making decisions on the building materials from the date of its first construction. Istanbul is located at a high seismic risk area and the basilica suffered several damages and collapses of different construction parts during its life. According to old Byzantine myths, the masonry and especially the 6th century dome were constructed with special bricks, originally designed to survive earthquakes [7]. They were imported to Istanbul from the island of Rhodes (Dodecanese, Greece) and weighed one-twelfth the weight of the common ones. Therefore, during the dome restoration works, several questions had to be answered. Is it true that the raw material of the original bricks was not local? Is it true that the technology of the original dome bricks was special? And finally, what should be the best type of brick for replacing the weathered original ones? In an attempt to investigate the provenance of the raw material and the production technology of the bricks, five bricks from the Great Dome, at locations corresponding to different historic periods (6th century A.D. and 10th century A.D.), as well as four bricks from the Basement and the main Entrance were sampled. We also sampled the brickwork at the early Christian Great Basilica of Rhodes (6th century) for comparison purposes. Eight samples (from masonry bricks and roof tiles) were collected. The two churches were studied in parallel, since the historical documentation refers directly to Rhodes as the place of origin of the Aghia Sophia bricks.

### 2.1. PCA and Hotelling's $T^2$ statistics used in clay provenance investigation

Geochemical investigation of ceramic samples may permit to locate the source of the raw material used, by comparing chemical data of the examined specimens with that of materials of more precisely known provenance or workshop. The chemical data of these "reference groups" are next subjected to a variety of multivariate statistical transformations such as Principal Component Analysis. The most widely used tool for obtaining the necessary series of chemical data is Neutron Activation Analysis (NAA), a technique that determines major and trace elements in ancient ceramics [9].

Although a dozen of sites are known to be pottery production workshops for Late Byzantine period (10th–14th century A.D.) [10], for Early Byzantine (4th–9th century A.D.) workshops, almost no information can be found in literature. Hence, in our case, no 'reference groups' are available. The only solution to this problem was to use brick samples from other contemporary to Aghia Sophia structures in Istanbul and use them as 'reference groups'. The role of "reference group samples" thus played ten samples from three 5th and 6th century constructions in Istanbul, the Theodosian City Walls (early 5th century), the church of Aghia Irene (532 A.D.) and the church of Sts. Sergius and Bacchus (527–536 A.D.).

The mean values of the element concentrations calculated by NAA for all the analysed ceramic samples are presented in Fig. 1 (full data in [4]). Principal Component Analysis (PCA) was then performed on the data.

PCA is a technique for simplifying a data set, by reducing the number of variables [11]. When the number of the examined variables is large it is very likely that subsets of variables are highly correlated with each other, maybe because they actually measure the same physical parameter. In that case, a number of variables are redundant, patterns in data are hard to find and the usually very enlightening graphical representation is not available. PCA may then be employed as a way of reducing the initial number of variables, identifying patterns within the data and expressing the data in such a way

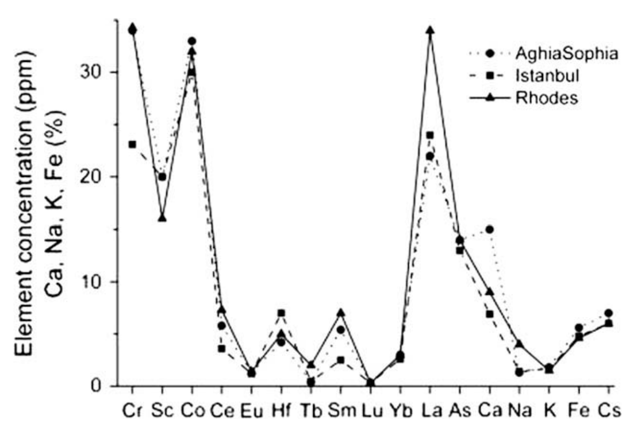


Fig. 1. Mean values of trace element concentrations for the examined bricks (measured by NAA).

as to highlight their similarities and differences. We have to note that because the goal of PCA is to ‘summarize’ the data, it is not considered a clustering tool. However, it reveals similarities and differences and is usually the first step to search for groups in large data sets.

In terms of mathematics, PCA is an orthogonal linear transformation that transforms the data to a new coordinate system [11]. Each coordinate, the principal component, is a linear combination of the observed variables in such a way that it accounts for a maximal amount of variance in the data set. The first PC is the direction through the data that explains the most variability in the data. The second PC is orthogonal to the previous one and describes the maximum amount of the remaining variability. The number of components extracted in a PCA is equal to the number of observed variables but usually only the first few components account for meaningful amounts of variance, so only these are retained, interpreted and used in subsequent analyses. Finally, the results are given in the form of a scatter plot, where the data are plotted in two axes representing two of the principal components. Along these axes, where the variance is maximum we expect to have the maximum differentiation (if there is any) between different observations and consequently detect clusters.

Principal Component Analysis of course works best when there is substantial correlation between the variables. When variables are uncorrelated, no reduction in dimensionality can occur whatever transformation is performed. Since the variables are uncorrelated, all of them have to be retained.

One of the most important points in PCA is the selection of parameters. In the case of the NAA results for the Aghia Sophia bricks for example, Cs, Na and K concentrations were excluded from the statistical treatment because they are well known to be unreliable due to contamination and As and Tb due to either pure counting statistics or missing values. Ca may depend on technological characteristics and may have been intentionally added to the clay mix. Finally, Cr, Sc, Ce, Sm, La, and Fe were selected for statistical analysis.

Fig. 2 is a bivariate plot, where the two axes represent two of the PCs derived from the data from Rhodes and Istanbul reference group. PC1 is a linear combination of mainly Sm (45%), Ce (39%) and La (19%), while PC2 is mainly loaded with Cr (92%). These two principal components together account for 88.6% of the total variation. The samples from Istanbul reference group are clearly discriminated from those from Rhodes. They form a well-defined, tight group. This fact implies that the samples from Sts. Sergius and Bacchus, Aghia Irene and the Theodosian walls come from a common local workshop or at least are made by the same clay source, which is most probably local. The Rhodian samples though do not form a very tight group, mainly due to R3 and R4.

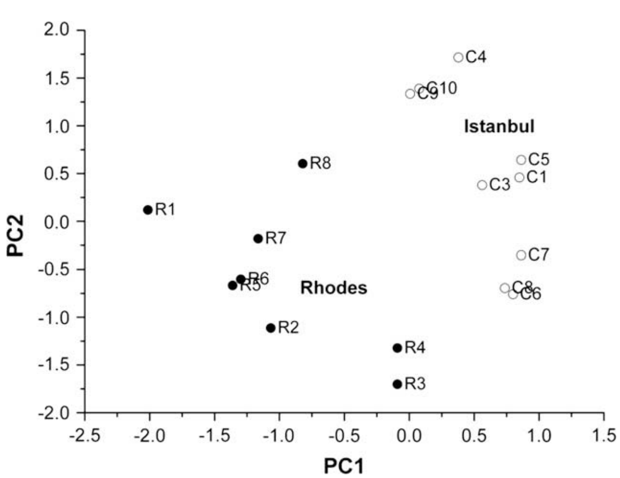


Fig. 2. Principal Component Analysis discriminating “Rhodes” from “Istanbul Reference” group.

In order to quantify the degree of similarity of the Aghia Sophia samples with the two groups, Hotelling’s $T^2$ statistics [12] was used to estimate the probability of each sample from Aghia Sophia to belong to “Rhodes” group or “Istanbul reference” group.

Hotelling’s $T^2$ statistic is a generalization of the popular Student’s $t$ statistic that is used in multivariate hypothesis testing [11]. The Student’s $t$-test is used for the statistical significance of the difference between two sample means and for confidence intervals for the difference between two population means. The results are given in Table 1. The probabilities of the Aghia Sophia samples to belong to “Istanbul reference” group are negligible (∼1%). The 6th century Dome samples (AS9 and AS10) show the greatest similarity with those from Rhodes (probabilities 91.3 and 97.2%). The 10th century Dome samples (AS11, AS12 and AS13) show a great probability to belong to “Rhodes” group too (from 50 to 85%) as well as the samples from the Basement (AS3 and AS4, probabilities 63 and 53% correspondingly). AS2 is a sample taken from the main entrance of the basilica and its probability to belong to the “Rhodes” group is rather small (18%). This can be explained by the fact that the main entrance is a part of the monument that has been restored many times and the sample is most probably modern.

According to these observations, the bricks from the dome (and most probably the masonry) of Aghia Sophia were not made from the same clay mix with the bricks of other contemporaneous churches and buildings in Istanbul. Furthermore, the raw materials used show a great similarity with those used for the bricks of the Great Basilica of Rhodes. The validity of deducing common origin will be enlightened by collecting more NAA data from other Byzantine pottery manufacturing centers in the future.

Table 1
Probabilities, calculated by Hotelling’s $T^2$ statistic, of the Aghia Sophia samples belonging to the Istanbul or Rhodes groups.

| Samples | Probability (%) | |
|---|---|---|
| | Istanbul | Rhodes |
| AS2 | 0.5 | 17.7 |
| AS3 | 1.1 | 63.1 |
| AS4 | 1.1 | 53.7 |
| AS9 | 1.2 | 91.3 |
| AS10 | 1.2 | 97.2 |
| AS11 | 0.8 | 85.2 |
| AS12 | 1.0 | 50.3 |
| AS13 | 1.1 | 75.5 |

### 2.2. PCA used in technology investigation

Mercury intrusion porosimetry, Thermogravimetric Analysis, Differential Thermal Analysis, Optical and Scanning Electron Microscopy, Tensile Strength tests were used for investigating the technology characteristics of the bricks. In Table 2, seven parameters are given, containing the microstructural and morphological characteristics of the bricks. In [4] an extended discussion explains the differences between the three brick groups. However, PCA can be used in order to give a simple but very illustrative graph (Fig. 3), which reviews the data consideration and comparisons between group properties.

The PCA results are given in Fig. 3a. The vectors of the components of all the samples along PC1 and PC2 are commonly known in statistics as the ‘‘scores’’ of PC1 and PC2. The PCs of Fig. 3a were extracted by three parameters: total porosity, tensile strength and pore volume for pore-sizes between 0.3 and 0.8 μm. These three were found to be the most useful ones to group the samples on one hand and reveal the differences between groups on the other hand. If we also plot the projection of the corresponding PCs of each sample along the vectors in Fig. 3b, representing the original variables in the page plane, we can have an estimation of which is the most critical parameter in group differentiation. The pore volume for pore-sizes between 0.3 and 0.8 μm seems to be the most important difference of the Aghia Sophia group from the other two. If the original clay is the same, as we accepted at the previous section, the very narrow pores, of an almost standard diameter, are probably the result of a fine sieving of the clay mix or pressing of the brick moulds, or of levigation or grinding. Tensile strength is what differentiates the two bricks of the 6th century dome from all the other bricks and makes them fall far away from all samples in Fig. 3a. The only differentiation of the dome bricks from the others is the size of the quartz temper inclusions (<100 μm in the dome bricks while <250 μm in the masonry bricks). This difference could explain the unexpected differentiation observed in strength. Tempering is known to increase toughness, but, above a certain volume fraction, it decreases strength to dangerously low levels. The presence of larger inclusions, in a high volume fraction, in the masonry samples, may decrease transverse rapture strength (TRS), while porosity remains at the same level.

## 3. Historical mortars from Byzantine monasteries

Historical mortars have been extensively studied during the last twenty years, not only due to their value as remnants of past production technologies, but also as unique examples of high durability building materials, the synthesis of which may serve as the best option for restoring the original ones. All

Table 2
Microstructural characteristics, tensile strength, inverse hydraulicity ratio and firing temperatures of the examined brick samples.

| Sample code | Total porosity (%) | Pore radius average (μm) | Pore volume for pore-size 0.3–0.8 μm | Density (g/cm$^3$) | Tensile strength (MPa) | $CO_2$/SBW | Firing temperature (DTA) |
|---|---|---|---|---|---|---|---|
| AS1 | 41.460 | 0.528 | 80.0 | 1.550 | 0.403 | 0.38 | 740 |
| AS2 | 47.210 | 0.467 | 81.2 | 1.650 | 0.645 | 0.70 | 740 |
| AS3 | 43.670 | 0.697 | 78.5 | 1.710 | 0.527 | 0.46 | 740 |
| AS4 | 52.390 | 0.422 | 77.3 | 1.520 | 0.143 | 0.48 | 740 |
| AS5 | 44.700 | 0.411 | 87.4 | 1.500 | 0.593 | 0.29 | 740 |
| AS6 | 51.330 | 0.422 | 88.6 | 1.480 | 0.463 | 0.33 | 740 |
| AS7 | 31.460 | 0.718 | 80.6 | 1.900 | 0.955 | 0.23 | 740 |
| AS8 | 40.900 | 0.458 | 80.4 | 1.680 | 0.195 | 0.41 | 740 |
| AS9 | 45.540 | 0.492 | 80.8 | 1.620 | 1.328 | 0.50 | 750 |
| AS10 | 45.620 | 0.734 | 86.2 | 1.620 | 1.405 | 0.34 | 750 |
| AS11 | 44.140 | 0.730 | 85.7 | 1.590 | 0.256 | 0.42 | 750 |
| AS12 | 40.710 | 0.543 | 87.8 | 1.750 | 0.309 | 0.20 | 750 |
| AS13 | 35.700 | 0.686 | 84.3 | 1.520 | 0.472 | 0.05 | 740 |
| C1 | 40.290 | 0.306 | 43.5 | 1.760 | 0.520 | 0.43 | 740 |
| C2 | 36.570 | 0.625 | 42.3 | 1.750 | 0.738 | 0.36 | 740 |
| C3 | 42.130 | 0.249 | 63.2 | 1.630 | 0.410 | 0.25 | 740 |
| C4 | 37.830 | 0.731 | 47.9 | 2.020 | 0.601 | 0.28 | 740 |
| C5 | 42.180 | 0.407 | 59.4 | 1.580 | 0.376 | 0.34 | 740 |
| C6 | 41.600 | 0.446 | 42.8 | 1.850 | 0.473 | 0.26 | 740 |
| C7 | 32.660 | 0.664 | 64.3 | 1.850 | 0.695 | 0.25 | 740 |
| C8 | 36.070 | 0.673 | 58.2 | 1.780 | 0.624 | 0.29 | 740 |
| C9 | 36.040 | 1.397 | 55.6 | 1.730 | 0.582 | 0.38 | 740 |
| C10 | 36.640 | 0.861 | 45.2 | 1.750 | 0.650 | 0.47 | 740 |
| R1 | 42.890 | 0.785 | 10.2 | 1.540 | 0.453 | 1.04 | 850 |
| R2 | 26.850 | 0.315 | 14.7 | 2.010 | 1.124 | 1.86 | 960 |
| R3 | 28.550 | 0.158 | 18.6 | 1.920 | 0.937 | 1.96 | 850 |
| R4 | 29.860 | 0.158 | 15.3 | 1.890 | 1.020 | 1.48 | 850 |
| R5 | 45.700 | 0.984 | 12.8 | 1.500 | 0.328 | – | 800 |
| R6 | 54.640 | 1.525 | 12.5 | 1.340 | 0.267 | 0.67 | 750 |
| R7 | 27.550 | 2.657 | 14.6 | 1.920 | 0.892 | 0.40 | 730 |
| R8 | 40.820 | 0.622 | 15.3 | 1.570 | 0.502 | 1.94 | 860 |

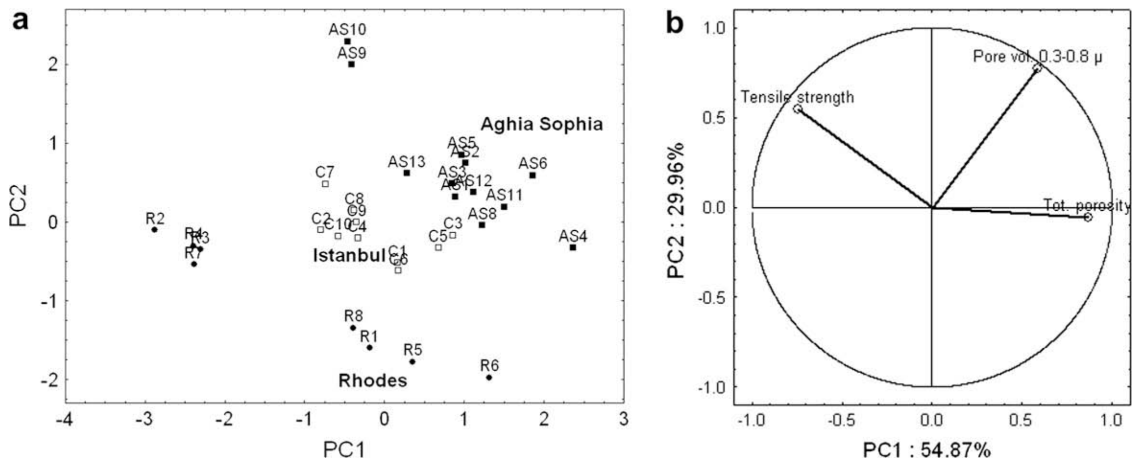


Fig. 3. a) PC scores for the Aghia Sophia, Istanbul and Rhodes samples. The statistical treatment is based on the technological characteristics of the sample sets, b) projection of the initial variables on the PC plane.

types of historical mortars, from the ancient to post-medieval times, consist of hydraulic or aerial lime (binder), a variety of inert materials (silicate, carbonate, or dolomite sand) and some additives for improving adhesion, workability, strength or durability (finely ground bricks, volcanic pozzolans, etc.). After more than 10 years of collecting and studying mortars from Byzantine and post-Byzantine monasteries and churches in Greece [13], the Materials Science and Technology Laboratory of the NTUA keeps a very large sample and databank on mortars of the specific period. Four main types of mortars were found among these samples: ‘‘typical lime’’ mortars (mixtures of calcite (∼80%) and quartz), ‘‘crushed brick-lime’’ (mixtures of a calcitic, binding material with finely ground bricks), ‘‘cementitious’’ (a conglomerate of gravel with sand, lime and pozzolana) and ‘‘portlandite’’ (with the characteristic presence of portlandite) mortars.

The ultimate goal of the majority of mortar studies is to prepare appropriate syntheses for compatible restoration mortars. However, the production of restoration mortars is very difficult to be standardised due to the large number of involved factors: the raw material mixture, the type of the building material, the environmental weathering factors and the aesthetic criteria of the building to be restored. The difficulty in defining guidelines for compatible restoration mortars together with the high durability of historical mortars results in the following conclusion: the more we know about the original mortars, the more easily we can reach a decision on the best restoration mortar syntheses. Towards this direction, multivariate statistics can serve a very useful role. PCA for example can help us to answer a number of critical questions: Can we ascribe certain properties to each one of the four previously mentioned mortar types? How strict and reliable is this correlation between ‘‘type’’ and ‘‘properties’’?

We have to note here, that sampling is very critical when statistical analysis is involved. In this particular case, we have to ensure that all the collected samples are mortars for the same technical application (mortars for plasters, masonry mortars, mortars for facings, etc.). Here, fifty-five (55) samples were selected for the analysis and they have all been taken from building mortars for brickwork (from material found between bricks). We also note that even the slightest weathering indication should lead to discarding of the suspect sample because weathered samples could alter the final result or diminish all possibly existing group differences. Another important factor in order for the correspondence between mortar type and physicochemical properties to be meaningful and reliable is the grain size distribution of the examined samples. An attempt to classify mortars of random grain size distributions would be meaningless and doomed to fail. Therefore, all the examined samples show a more or less symmetric distribution around the range of 0.25 mm, a fact that, in terms of weight percentage, gives about 25% of the total weight to the fine fraction (<0.038 mm). The thermal results and the porosity of the 55 samples together with their tensile strength values are given in Table 3.

The variables given in Table 3 were used for PCA. Various attempts with different sets of parameters were done but the best results were given by using only three of them: inverse hydraulicity ratio, porosity and strength. Three components were extracted by PCA (Table 4, Fig. 4b), from which only the first one, with eigenvalue higher than 1, was kept, which carries 68.8% of the total variance. The best discrimination between different groups was obtained with PC1 and tensile strength, as given in Fig. 4.

The standard probability ellipses of Fig. 4a were drawn for each group with 90% confidence limits. If the data for a given variable can be assumed to represent a random sample from a bivariate normal distribution, then probability ellipses can be used to show areas within which given proportions of the samples are expected to lie. Here, the ellipses are expected to include 90% of the samples.

As theoretically expected, lime mortars form a well-defined and compact group, totally different from the other two [14]. The observed differentiation of the lime mortars group can obviously be ascribed to differences in all parameters. Hydraulicity, porosity and tensile strength, as theoretically expected, are the critical properties for discriminating between lime mortars and crushed brick or cementitious mortars. As for

Table 3
Thermal analysis, porosimetry and strength test results for the examined mortar samples.

| Sample code | SBW (%) | $CO_2$/SBW | $CO_2$ (%) | Porosity (%) | Tensile strength (MPa) |
|---|---|---|---|---|---|
| *Lime mortars* | | | | | |
| Agarathos1 | 3.02 | 11.1 | 33.46 | 44.46 | 0.32 |
| Arkadi1 | 3.13 | 11.6 | 36.31 | 32.70 | 0.32 |
| Toplu5 | 2.86 | 13.3 | 38.01 | 46.20 | 0.14 |
| Metochi4 | 2.90 | 11.1 | 32.16 | 38.55 | 0.30 |
| Rhodes1 | 2.82 | 10.8 | 30.5 | 43.90 | 0.27 |
| Rhodes2 | 2.99 | 10.4 | 31.21 | 33.40 | 0.33 |
| Rhodes3 | 2.93 | 11.2 | 32.82 | 45.75 | 0.37 |
| Rhodes4 | 2.02 | 16.1 | 32.45 | 40.60 | 0.25 |
| Agarathos3 | 3.40 | 11.9 | 40.34 | 34.69 | 0.20 |
| Agarathos4 | 2.76 | 15.9 | 43.78 | 42.10 | 0.27 |
| Agarathos7 | 3.21 | 11.6 | 37.11 | 39.30 | 0.31 |
| *Portlandite mortars* | | | | | |
| Metochi6 | 14.05 | 2.72 | 17.14 | 27.53 | 0.67 |
| Rhodes34 | 12.06 | 2.46 | 17.97 | 30.62 | 0.61 |
| *Cementitious mortars* | | | | | |
| Kerkyra1 | 6.51 | 3.1 | 20.21 | 42.84 | 0.45 |
| Kerkyra2 | 12.67 | 1.0 | 12.08 | 29.80 | 0.52 |
| Kerkyra3 | 6.61 | 2.0 | 13.10 | 28.50 | 0.50 |
| Arkadi2 | 8.83 | 1.7 | 14.92 | 41.56 | 0.45 |
| Preveli1 | 5.40 | 2.3 | 12.67 | 33.16 | 0.55 |
| Preveli2 | 5.87 | 5.2 | 30.40 | 33.27 | 0.44 |
| Preveli3 | 8.04 | 2.1 | 17.23 | 41.54 | 0.37 |
| Chrysopighi1 | 11.38 | 1.3 | 14.79 | 32.48 | 0.54 |
| Agarathos2 | 3.77 | 6.1 | 22.88 | 40.04 | 0.47 |
| Toplu6 | 5.38 | 3.3 | 17.65 | 32.84 | 0.51 |
| Rethymno1 | 5.56 | 2.7 | 15.24 | 23.79 | 0.52 |
| Rethymno2 | 4.13 | 2.7 | 11.23 | 34.16 | 0.53 |
| Rethymno3 | 6.37 | 4.0 | 25.66 | 25.46 | 0.54 |
| A. Sophia1 | 7.56 | 1.5 | 11.67 | 35.49 | 0.53 |
| Rhodes5 | 10.51 | 1.9 | 19.65 | 42.76 | 0.42 |
| Rhodes6 | 15.59 | 0.8 | 12.32 | 38.18 | 0.51 |
| Rhodes7 | 15.04 | 0.9 | 12.93 | 21.52 | 0.51 |
| Aghio Oros6 | 12.93 | 1.1 | 14.32 | 36.46 | 0.45 |
| Aghio Oros1 | 10.12 | 1.2 | 11.97 | 27.76 | 0.53 |
| Aghio Oros2 | 4.34 | 3.9 | 17.10 | 31.12 | 0.50 |
| Aghio Oros3 | 6.32 | 1.8 | 11.52 | 34.71 | 0.46 |
| *Crushed brick* | | | | | |
| Toplu7 | 6.47 | 1.6 | 10.25 | 25.13 | 0.65 |
| Metochi5 | 6.31 | 3.6 | 22.65 | 45.19 | 0.57 |
| Rhodes10 | 3.74 | 9.1 | 33.85 | 38.79 | 0.57 |
| Rhodes11 | 4.85 | 6.2 | 30.14 | 40.20 | 0.62 |
| Rhodes12 | 4.40 | 6.2 | 27.09 | 42.80 | 0.55 |
| Rhodes14 | 5.20 | 6.0 | 31.12 | 43.52 | 0.57 |
| Rhodes15 | 4.91 | 4.6 | 22.73 | 46.10 | 0.64 |
| A. Sophia1.2 | 4.78 | 4.1 | 19.45 | 35.62 | 0.59 |
| A. Sophia2.2 | 4.75 | 2.3 | 11.01 | 41.67 | 0.52 |
| A. SophiaW2 | 3.52 | 4.0 | 13.99 | 37.90 | 0.59 |
| A. Sophia3.1 | 5.18 | 2.4 | 12.67 | 35.39 | 0.63 |
| A. Sophia3.2 | 4.31 | 2.6 | 11.18 | 32.71 | 0.57 |
| A. Sophia4 | 3.79 | 3.3 | 12.35 | 42.12 | 0.57 |
| Toplu8 | 5.21 | 3.3 | 17.00 | 26.94 | 0.58 |
| Rhodes16 | 6.12 | 2.6 | 15.98 | 29.67 | 0.52 |
| Rhodes17 | 5.10 | 3.0 | 15.20 | 32.16 | 0.64 |
| Rhodes20 | 4.76 | 5.5 | 26.23 | 38.60 | 0.52 |
| Rhodes21 | 5.67 | 4.3 | 24.48 | 42.37 | 0.53 |

Table 4
Eigenvalues and total variances of the extracted principal components.

| Eigenvalue | % Total variance | Cumulative eigenvalue | Cumulative total variance % |
|---|---|---|---|
| 2.048860 | 68.29533 | 2.048860 | 68.2953 |
| 0.706127 | 23.53756 | 2.754987 | 91.8329 |
| 0.245013 | 8.16711 | 3.000000 | 100.0000 |

the crushed brick and cementitious mortar samples, they lay in overlapping probability ellipses. It is obvious that the only parameter which partly differentiates these ellipses is tensile strength. The crushed brick mortar samples show higher values of tensile strength than the cementitious ones. Mortar is a two-phase material and in a first approximation, its strength is proportional to the strength of the weaker component, i.e., the binder matrix. The strength of the binder matrix depends on the type of the binder, its theoretical strength and the porosity of the matrix. The distinct increase in strength observed along the tensile strength axis of Fig. 4a mainly reflects the increase in the amount of hydrated phases in the binder. The variations within each group are expected to result from the differences in the binder/inert ratios between different samples.

Fig. 4a demonstrates a very illustrative way to show that we can ascribe a specific range of permissible property values to each one of the four mortar types. Adding more data to the plot will stabilise the ellipses and make the discrimination more strict and reliable. This correlation between "type" and "properties" on one hand reveals the specifications demanded by the mortar producers of the Byzantine and post-Byzantine epoch and on the other hand offers an invaluable tool for selecting proper syntheses for restoration mortars.

## 4. The marble patinas of the archaeological site of Eleusis, Greece

The archaeological site of Eleusis hosts structures that were founded from 1500 B.C. to 400 B.C. The surrounding area was rural from antiquity till the post-war period, when it became heavily industrial. The large metallurgical and chemical manufacturing sectors located in the vicinity of the site together with the proximity of the sea are considered to be the main causes of marble deterioration. The main structural material of the site is white, calcitic, well-crystallised marble, with grains of 1–2 mm. Four patina types have been recognized at the site [15]: a) yellowish patinas (Y), which typically occur on washed-out surfaces. Recrystallised calcite is the main constituent of these patinas and iron particles are responsible for their overall yellowish color. b) Black-gray patinas (B), on areas sheltered from direct rainwater contact. These patinas consist of layers, an amorphus one with deposits rich in S, Si and Fe and a crystalline one with large amounts of recrystallised calcite. c) Loose, black depositions (L) mainly consisting of gypsum and fly ash particles. d) Cementitious patinas (C) which are "crust type" patinas with intense pitting. During the conservation and restoration works at the site, an attempt was made to understand the specific

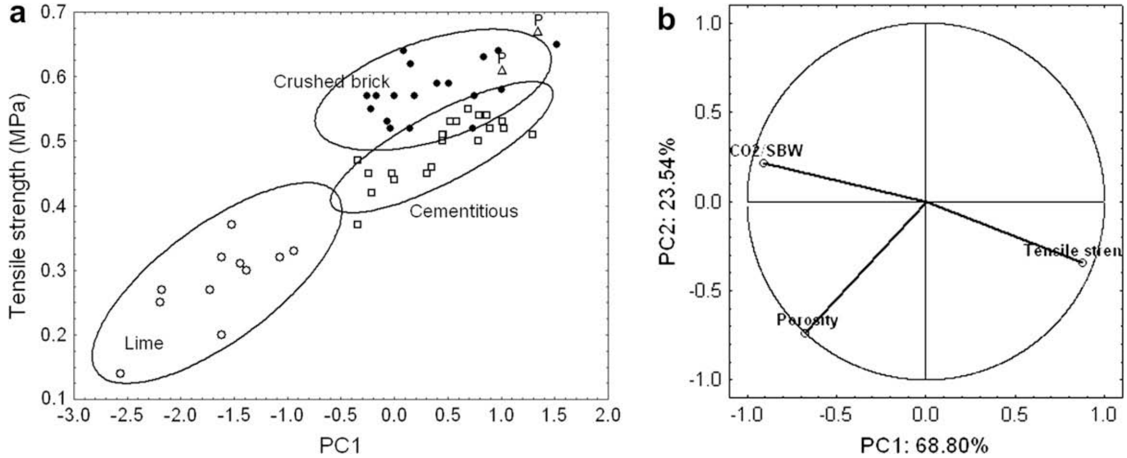


Fig. 4. a) Biplot of PC1 scores versus tensile strength values for the studied mortars, b) projection of the initial variables on the PC plane.

weathering mechanisms of the ancient marbles and relate the composition of the patinas to the environmental conditions that prevail at the exact point where different patina types develop [16].

Sixteen patina samples were analysed under the Scanning Electron Microscope (SEM) by Energy Dispersive X-Ray (EDAX) Fluorescence spectroscopy and porosity was measured by digital image processing of the SEM photographs ([6], Table 5).

In this case, we introduce in the variable set, not only numeric but also alphanumeric variables which have been transformed to numeric. This flexibility of PCA makes it an even more attractive tool, which in this particular case may be used to combine analytical results with micro-climate data. The choice of the particular numerical values used to express qualitative properties of the examined surfaces is a rather difficult task. It is quite reasonable of course to assign the values 0 and 1 to extreme conditions (for example, no or full exposure to rain). However, for intermediate conditions the choice of the numeric value has to be justified by the nature of the data. We have to note here, that the assignment of numerical values to the alphanumeric data was done during the sampling procedure or the examination of the samples under the SEM and not when the data collection was completed. The transformation, in our case, was done according to the following:

a) Exposure to rain = REXP = Value “1” for washed-out areas, value “0” for sheltered areas and “0.5” for areas where water percolates.
b) Surface orientation = SO = Value “1” for horizontal areas, value “0” for perpendicular areas and “0.25” for inclined areas. This intermediate value was chosen because the inclination of all surfaces was less than 45°. If it was higher than 45°, “0.75” would be a better choice.
c) Presence of recrystallised calcite = RCa = Value “1” for samples with almost intact films of secondary calcite, value “0” for samples with no trace of recrystallised calcite and “0.5” for samples partially covered by recrystallised calcite.
d) Degree of gypsum formation = G = Value “1” for extended gypsum formations, value “0” for absence of gypsum and “0.25” for samples with limited gypsum formation. The intermediate value of “0.25” was chosen

Table 5
Compositional data, porosity and micro-climate characteristics of the marble patinas of Eleusis archaeological site.

| Patina code | Si | Mn | Cl | Fe | Sr | Ca | REXP | SO | RCa | G | Porosity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Y | 9300 | 195 | 3300 | 65,525 | 577 | 387,000 | 1 | 0 | 0.5 | 0.25 | 35.77 |
| Y | 19,000 | 360 | 1600 | 237,057 | 194 | 363,000 | 1 | 0 | 0.5 | 0.25 | 35.77 |
| Y | 29,600 | 186 | 2700 | 93,459 | 220 | 373,000 | 1 | 0 | 0.5 | 0 | 33.5 |
| Y | 14,000 | 380 | 2200 | 1800 | 880 | 280,000 | 1 | 0 | 0.5 | 0.25 | 35.77 |
| B | 31,000 | 220 | 1800 | 2002 | 242 | 263,000 | 0.5 | 0.25 | 1 | 0.25 | 28.62 |
| B | 35,000 | 323 | 3300 | 4000 | 216 | 292,000 | 0.5 | 0.25 | 1 | 0.25 | 28.59 |
| B | 52,000 | 197 | 4300 | 1170 | 230 | 337,000 | 0.5 | 1 | 0 | 0.25 | 30.1 |
| B | 46,000 | 380 | 2800 | 3300 | 202 | 410,000 | 0.5 | 0.25 | 1 | 0.25 | 28.62 |
| L | 31,000 | 167 | 5230 | 2239 | 242 | 280,000 | 0 | 0 | 0 | 1 | 31.26 |
| L | 38,000 | 159 | 5480 | 2880 | 230 | 380,000 | 0 | 0 | 0 | 1 | 31.28 |
| L | 26,000 | 148 | 5100 | 2540 | 203 | 350,000 | 0 | 0 | 0 | 1 | 33.28 |
| L | 18,000 | 131 | 5600 | 3680 | 182 | 310,000 | 0 | 0 | 0 | 1 | 31.28 |
| C | 38,000 | 385 | 2500 | 4070 | 204 | 328,000 | 1 | 1 | 0 | 0 | 25.51 |
| C | 42,000 | 397 | 3800 | 18,700 | 230 | 238,000 | 1 | 1 | 0 | 0 | 25.51 |
| C | 48,000 | 431 | 4050 | 11,000 | 330 | 380,000 | 1 | 1 | 0 | 0 | 25.51 |
| C | 31,000 | 330 | 5000 | 8000 | 160 | 140,000 | 1 | 1 | 0 | 0 | 25.51 |

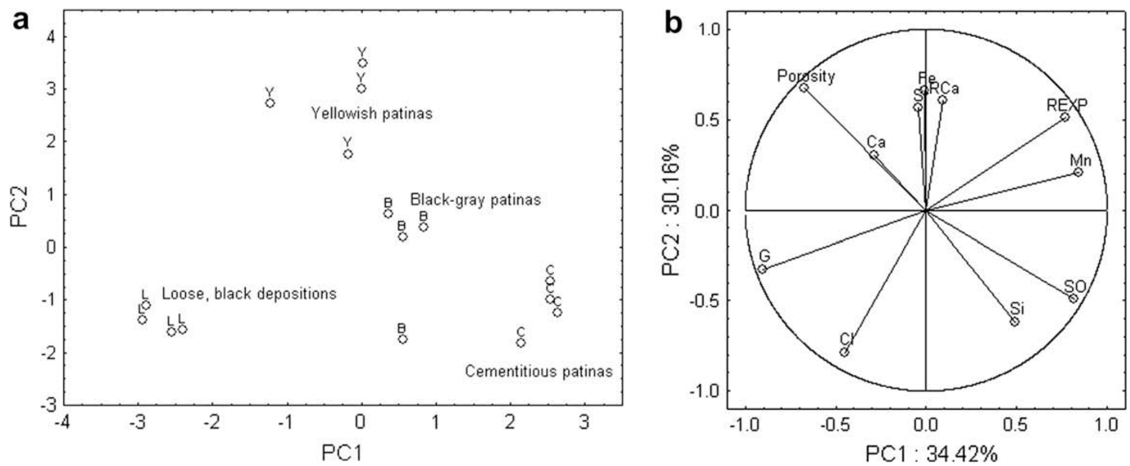


Fig. 5. a) Principal Component Analysis discriminating between the four patina types: yellowish, black-gray, loose depositions and cementitious. b) Projection of the initial variables on the PC plane.

according to the distribution of gypsum crystals in the patina layers (identified under the Scanning Electron Microscope). At the beginning, the thicker and denser formation was assigned the value "1" while the thinner formation was assigned the value "0". The value "0.25" expresses the fact that the extent of the gypsum crystal growth on the patinas of the intermediate samples is closer to that of the samples characterized by "0".

The PCA results are given in Fig. 5. Different types of patinas are obviously separated and form compact groups. Gypsum and chlorine are mainly related to loose depositions while exposure to rain is not related at all with this patina type. Yellowish patinas are related as expected, to high porosities due to the intense presence of recrystallised calcite, iron which is responsible for the yellow color and the presence of calcium and strontium, the latter being typical substitutional element in natural calcite crystals. Cementitious patinas are characterized by high concentration of Si and horizontal surfaces, which favor the deposition of large suspended particle agglomerations. Finally, black-gray patinas are characterized by percolating water as theoretically expected and the presence of Mn, which is not a very straightforward result.

A similar methodology, like the one in the third example, has been applied on the building material of the medieval city of Rhodes [17]. In that case, in order to reveal the correlation between environmental conditions and decay patterns, the results of ion chemical analysis of soluble salts, along with data on the sea and sun exposure and air flow, have been included in the statistical analysis.

## 5. Conclusions

The use of Principal Component Analysis proves to be a very powerful tool in the study of building materials properties and weathering condition. PCA can be used in the provenance investigation of raw materials, in order to replace the original building elements with other aesthetically and physicochemically similar ones. It can also be used to reveal different groups existing in sample sets, based on the technological properties of the samples. The capacity of PCA to provide simple, illustrative and comprehensive diagrams is unique. Unique is also the capacity of the methodology to limit the number of variables and reveal the hidden correlation between them. As such, its use is recommended not only to cultural heritage monuments but also to buildings or structures of any type. Furthermore, PCA could provide a technique for predicting what type of weathering would be expected according to the micro-environmental conditions at various locations. Hence, the directions of a conservation plan could be ascribed and evaluated accordingly.

## Acknowledgements

We are indebted to E. Kollias, M. Michailidou, Prof. M. Erdik, A.M. Isikara, Prof. Th. Tassios, Prof. R. Van Grieken, Dr K. Torfs, K. Bisbikou, A. Bakolas, P. Michailidis, Th. Tsiourva, D. Symwonis, N. Avgerinos and N. Tsantanis.

## References

[1] M.J. Baxter, Exploratory Multivariate Analysis in Archaeology, Edinburgh University Press, Edinburgh, 1994.
[2] R.D. Drennan, Statistics for Archaeologists: a Commonsense Approach. Interdisciplinary Contributions to Archaeology, Plenum Press, New York, 1996.
[3] S. Shennan, Quantifying Archaeology, Edinburgh University Press, Edinburgh, 1997.
[4] A. Moropoulou, A. Cakmak, K. Polikreti, Provenance and technology investigation of the Agia Sophia bricks, Istanbul, Turkey, J. Am. Ceram. Soc. 85 (2002) 366−372.
[5] A. Moropoulou, A. Bakolas, C. Michaelides, K. Polikreti, Correlation of physicochemical and mechanical properties of historical mortars and classification by multivariate statistics, Cement Concrete Res. 33 (2003) 891−898.
[6] A. Moropoulou, K. Bisbikou, R. Van Grieken, K. Torfs, K. Polikreti, Correlation between aerosols, deposits and weathering crusts on ancient marbles, Environ. Technol. 22 (2002) 607−618.

[7] C. Mango, in: R. Mark, A.S. Cakmak (Eds.), Byzantine Writers on the Fabric of Agia Sophia: Agia Sophia from the Age of Justinian to the Present, Cambridge University Press, Cambridge, 1992.
[8] Trilateral Protocol of Princeton University (PU), Bogazici University (BU) and National Technical University of Athens (NTU-A), Earthquake Protection of Historical Monuments, undersigned by Prof. A.S. Cakmak, Prof. M. Erdik, Prof. A. Isikara and Prof. A. Moropoulou, Istanbul, March 19, 1994.
[9] P.M. Day, E. Kyriatzi, A. Tsolakidou, V. Kilikoglou, Group therapy in Crete: a comparison between analyses by NAA and thin section petrography of Early Minoan pottery, J. Archaeol. Sci. 26 (1999) 1025−1036.
[10] D. Papanikola-Bakirtzis, E. Maguire, H. Maguire, Ceramic Art from Byzantine Serres, University of Illinois Press, Urbana and Chicago, 1992, p. 12.
[11] J.C. Davies, Statistics and Data Analysis in Geology, John Wiley and Sons, New York, 1986.
[12] V. Kilikoglou, M. Vassilaki-Grimani, Y. Maniatis, A.P. Grimanis, A study of ancient roof tiles found in Pella, Greece, Material Res. Soc. Symp. Proc. 123 (1988) 117−122.
[13] A. Moropoulou, A. Bakolas, M. Michailidis, M. Chronopoulos, Ch. Spanos, Traditional technologies in Crete providing mortars with effective mechanical properties, in: C.A. Brebbia, B. Leftheris (Eds.), Structural Studies of Historical Buildings IV, vol. 1, Computational Mechanics Publications, Southampton, 1995, pp. 151−161.
[14] J. Shafer, H.K. Hilsdorf, Ancient and new lime mortars—the correlation between their composition, structure and mechanical properties, in: M.-J. Thiel (Ed.), Proceedings of the International Congress on Conservation of Stone and Other Materials, E & FN Spon, Paris, 1993, pp. 605−612.
[15] A. Moropoulou, K. Bisbikou, K. Torfs, R. Van Grieken, F. Zezza, F. Macri, Origin and growth of weathering crusts on ancient marbles in industrial atmosphere, Atmos. Environ. 32 (1998) 967−982.
[16] A. Moropoulou, K. Bisbikou, K. Torfs, R. Van Grieken, Atmospheric deposition as decay factor for the Demeter Sanctuary ruins in the industrial atmosphere of Eleusis, Greece, Environ. Technol. 19 (1998) 1039−1047.
[17] A. Moropoulou, P. Theoulakis, T. Chrysofakis, Correlation between stone weathering and environmental factors in marine environment, Atmos. Environ. 29 (2005) 895−903.