



INSTITUTE OF ARCHEOLOGY  
AND ART HISTORY OF ROMANIAN  
ACADEMY CLUJ-NAPOCA



UNIVERSITATEA TEHNICĂ  
DIN CLUJ-NAPOCA

JAHA  
JOURNAL OF ANCIENT HISTORY  
AND ARCHAEOLOGY

editura  
MEGA

---

# Journal of Ancient History and Archaeology

DOI: <http://dx.doi.org/10.14795/j.v13i1>

ISSN 2360 266x

ISSN-L 2360 266x



Scopus®



Clarivate  
Analytics



Central and Eastern European Online Library

EBSCO



DOAJ  
DIRECTORY OF  
OPEN ACCESS  
JOURNALS

No. 13.1/2026



# CONTENTS

## STUDIES

### ANCIENT HISTORY

#### Helena LÓPEZ GÓMEZ

FROM THE MOTHER OF THE GRACCHI TO THE SISTER OF AUGUSTUS: REIMAGINING CORNELIA'S LEGACY IN THE LITERARY AND MATERIAL DEPICTIONS OF OCTAVIA ..... 5

### ARCHAEOLOGY

#### Sergey YARYGIN, Sergazy SAKENOV, Ainagul GANIYEVA

NOMADS OF ZERENDY AND SANDYKTAU AT THE TURN OF ERAS (NORTHERN KAZAKHSTAN) ..... 14

#### Hakan GÖNCÜ

A MONOLITHIC DORIC ENTABLATURE FROM IASOS ..... 32

#### Vitalie BÂRCĂ, Alpár DOBOS, Cristinel PLANTOS, Szilárd Sándor GÁL

TWO GEPIDIC AGE BURIALS RECENTLY DISCOVERED AT ALBA IULIA. ARCHAEOLOGICAL AND ANTHROPOLOGICAL EVALUATION ..... 48

#### Boaz ZISSU, Shemesh YA'ARAN

KHIRBET BEIT 'ATÁB AND THE MÜGHÂRET BIR EL HASÛTAH CAVE COMPLEX: A MULTI-PERIOD ARCHAEOLOGICAL SITE IN THE JERUSALEM HILLS ..... 64

### ARCHAEOLOGICAL MATERIAL

#### Dan George ANGHEL, Sergiu-Traian SOCACIU, Sorin Ilie COCIŞ

THE GLAZED EARTHENWARE OF SUTOR/OPTATIANA ..... 95

#### Kübra YAZAR, Berna KAVAZ KINDİĞİLİ

ROMAN AND LATE ANTIQUE GLASS BOTTLES IN THE ADANA MUSEUM COLLECTION: AN EVALUATION OF TYPOLOGY AND PRODUCTION TECHNIQUES ..... 113

#### Burak ARSLAN

STUCCOWORK FROM METROPOLIS: A COMPREHENSIVE RESEARCH ON STYLISTIC FEATURES OF THE STUCCOES IN THE ROMAN BATH-ALAEŞTRA COMPLEX ..... 132

#### Vlad VORNIC, Sergiu MATVEEV

ISOLATED MIGRATION-PERIOD FINDS: THE BRONZE FIBULA FROM RADOAIA (SANGEREI DISTRICT, REPUBLIC OF MOLDOVA) ..... 143

#### Alexander HARIZANOV

FOLLOWING THE LEAD: *PLUMBATAE* IN *SCYTHIA* AND *MOESIA SECUNDA* ..... 153

### ARCHAEOOMETRY

#### Hakan AYCAN, Barış SEMİZ, Onur ZUNAL

CHARACTERIZATION OF EASTERN SIGILLATA C (ESC) FROM THE YK SECTOR AT KLAROS: AN ARCHAEOLOGICAL AND ARCHAEOMETRIC APPROACH ..... 168

#### Wahidullah ENAYAT, Ahmet KÖROĞLU, Davut YİĞİTPAŞA

APPLICATION OF ANALYTICAL TECHNIQUES AND MACHINE LEARNING MODELS FOR THE IDENTIFICATION AND CULTURAL ATTRIBUTION OF ANCIENT POTTERY: A CASE STUDY FROM BOLU CLAUDIOPOLIS, TURKEY ..... 179

### ARCHAEOLOGICAL TOPOGRAPHY

#### Cristian FLOCA, György SIPOS, Florin GOGÂLTAN, Alexandru HEGYI, Tamás BARTYIK, Dávid FILYÓ

SETTLING STRATEGIES OF HUMAN COMMUNITIES IN PANNONIAN HOLOCENE RIVERSCAPE. A MICRO ZONE CASE STUDY FROM BANAT (ROMANIA) .... 193

### EPIGRAPHY AND PAPHYROLOGY

#### Héctor ARROYO-QUIRCE

REVISITING THE CULT OF ARES IN GRECO-ROMAN ASIA MINOR ..... 214

#### Péter KOVÁCS

SOME CONSIDERATIONS ON THE PROVINCIAL POPULATION IN NORTHERN PANNONIA IN THE 6TH CENTURY IN THE LIGHT OF WRITTEN AND EPIGRAPHIC SOURCES ..... 224

### NUMISMATICS

#### Metodi MANOV

AGAIN ABOUT THE COINS OF MOSKON – AN ENIGMATIC RULER FROM DOBRUDJA IN THE HELLENISTIC PERIOD ..... 240

#### Cristian GĂZDAC, Adrian-Daniel STAN

WHY DID THEY NOT RECOVER THE HOARDS? INSECURITY IN THE ROMAN EMPIRE ..... 245

#### Hacer SANCAKTAR, Mehmet ÖZHANLI

COINS AND THE TRANSFORMATION OF A SANCTUARY: NUMISMATIC EVIDENCE FROM THE SANCTUARY OF MEN ASKAENOS IN PISIDIAN ANTIOCH..... 258

#### Stefan KRMNICEK

COINS IN THE WRONG PLACE? CURIOUS DEPOSITIONS, HIDDEN HOARDS, AND ACCIDENTAL AFTERLIVES..... 291

## HISTORIOGRAPHY

### Alin HENȚ

THE PERSISTENCE OF THE PAST: A CRITICAL PERSPECTIVE  
OF ROMANIAN LATE IRON AGE ARCHAEOLOGY  
(AND ITS RELATED FEATURES) IN POST-COMMUNISM  
AND NEOLIBERALISM ..... 297

## CULTURAL HERITAGE PROTECTION

### Silviu MILOIU, Célio Gonçalo MARQUES, Sergiu MUSTEAȚĂ, Marusya SMOKOVA Lucia NOVÁKOVÁ, Vojtěch BLAŽEK, Renato DE LEONE

GAMIFICATION AS A CATALYST FOR COMMUNITY-BASED  
HERITAGE WORK IN LOW-DENSITY TERRITORIES:  
SELECTED BEST PRACTICES FROM EUROPE ..... 317

## REVIEWS

### Victor SAVA

Stanislav Grigoriev, Indo-Aryans in the Bronze Age, Archaeopress  
Archaeology, Bicester 2025, 298 pages, 92 figures, 2 tables. ISBN  
Hardback: 9781805830665, Digital: 9781805830672, DOI  
10.32028/9781805830665 ..... 327

### Ruslan TSAKANYAN

Yervand H. Grekyan. The Urartian Onomasticon: A Prosopographic  
Study, ARAMAZD: Armenian Journal of Near Eastern Studies.  
Editor-in-Chief Aram Kosyan, Vol. XVII, Issue 2. Association for  
Near Eastern and Caucasian Studies. Archaeopress Publishing Ltd:  
Oxford, UK. Yerevan-Oxford 2023. ISSN 1829-1376, ISBN 978-1-  
80327-706-6. 128 pp. <https://doi.org/10.32028/ajnes.v17i2> ... 330

### Ruslan TSAKANYAN

Yervand H. Grekyan (ed.), By God's Grace: Ancient Anatolian  
Studies Presented to Aram Kosyan on the Occasion of his 65th  
Birthday. Ancient Near Eastern Studies Supplement Series 61.  
Leuven – Paris – Bristol, CT.: Peeters Pub & Booksellers, 2023. Pp.  
XX+366. ISBN: 9789042948686; E-ISBN: 9789042948693 ..... 332

### Victor COJOCARU

S. Yu. Monakhov, E. V. Kuznetsova, A. N. Kovalenko, N. B.  
Churekova, A. G. Yazovskikh, Amphorae of the 6th–3rd Centuries  
BC from Museum Collections of Rostov-on-Don [in Russian],  
Saratov: Amirit, 2025, 284 pages, 123 figs, 5 tables, and 199  
catalogue illustrations integrated into the text. ISBN 978-5-00207-  
976-6 ..... 334

### Vitalie BÂRCĂ

O. V. Symonenko, O. S. Dzneladze, D. M. Sikoza, Atlas of Late  
Scythian Sites of the Lower Dnieper [In Ukrainian], Kyiv: Institute  
of Archaeology of the National Academy of Sciences of Ukraine,  
2024. 214 p. ISBN 978-617-7810-43-7 ..... 340

Design & layout: Petru Ureche



EDITURA MEGA | [www.edituramega.ro](http://www.edituramega.ro)  
e-mail: [mega@edituramega.ro](mailto:mega@edituramega.ro)

## ARCHAEOOMETRY

### APPLICATION OF ANALYTICAL TECHNIQUES AND MACHINE LEARNING MODELS FOR THE IDENTIFICATION AND CULTURAL ATTRIBUTION OF ANCIENT POTTERY: A CASE STUDY FROM BOLU CLAUDIOPOLIS, TURKEY

**Abstract:** Understanding the origin, manufacturing technology, and cultural affiliation of ancient pottery is a major objective in interdisciplinary archaeological research. In this study, 20 pottery samples from the ancient site of Bolu Claudiopolis (modern-day Bolu, Turkey) were investigated using a multi-analytical approach. X-ray Fluorescence (XRF), X-ray Diffraction (XRD), and Fourier Transform Infrared Spectroscopy (FTIR) were employed to characterize the chemical, mineralogical, and structural properties of the samples. The resulting datasets were analyzed using several machine learning algorithms, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Random Forest (RF), implemented in Weka software. Principal Component Analysis (PCA) was applied for dimensionality reduction and pattern recognition, while fuzzy logic was used as a complementary framework for approximate cultural attribution. In addition, data fusion strategies were applied to combine information obtained from the different analytical techniques. The results revealed compositional variability among the pottery samples and indicated that most of them were more closely associated with Roman compositional patterns, whereas a limited number showed affinity with Greek-related characteristics. Among the tested models, Random Forest provided the most consistent classification performance. XRD and FTIR results further suggested medium firing temperatures and the presence of silicate-based raw materials with specific mineral phases. Overall, this study demonstrates the value of integrating analytical techniques with machine learning and fuzzy logic for the classification and comparative assessment of ancient pottery.

**Keywords:** Ancient Pottery, XRF Analysis, Machine Learning, Bolu Claudiopolis, Fuzzy Logic.

#### Wahidullah ENAYAT

Ondokuz Mayıs University, Turkey  
enayatwahidullah@gmail.com  
<https://orcid.org/0000-0002-2819-6522>

#### Ahmet KÖROĞLU

Ondokuz Mayıs University, Turkey  
kahmet@omu.edu.tr  
<https://orcid.org/0000-0002-8627-8163>

#### Davut YİĞİTPAŞA

Ondokuz Mayıs University, Turkey  
davut.yigitpasa@omu.edu.tr  
<https://orcid.org/0000-0001-8821-5628>

## INTRODUCTION

The study of ancient pottery, as one of the most widely used sources in scientific archaeology, provides valuable information about the history,

technology, and cultural interactions of past civilizations<sup>1</sup>. The chemical composition, mineralogical properties, and microscopic structures of pottery not only shed light on their geographical origin, but also play a fundamental role in determining the origin of civilization by analyzing manufacturing techniques, firing temperatures, and raw materials<sup>2</sup>. Non-destructive analysis methods such as X-ray Fluorescence (XRF)<sup>3</sup>, X-ray Diffraction (XRD) and Fourier Transform Infrared Spectroscopy (FTIR) are the most important tools in archaeometry<sup>4</sup>, which are used to study the composition of pottery<sup>5</sup>, especially in comparisons between civilizations (Roman, Greek, Egyptian, Persian, etc.)<sup>6</sup>. Also, studies such as the Sah Valley study in Turkey using XRF and XRD have shown that chemical composition in conjunction with mineralogical data can also reveal the technological pattern of manufacture<sup>7</sup>. With the growth of analytical data, the use of machine learning algorithms and multivariate analyses such as PCA, K-means, HCA, and classification algorithms (KNN, SVM, Random Forest) in pottery analysis has rapidly expanded<sup>8</sup>. Several studies in this field have used ML to classify Iberian, Chinese, Greek, and even pre-Columbian pottery samples<sup>9</sup>, showing accuracies between 85 and 96%<sup>10</sup>. In addition, studies such as research on Apulian Red Figure pottery or Roman Tavoliere pottery have also shown that even minor differences in chemical composition can determine the origin of production from a civilizational or workshop perspective<sup>11</sup>. On the other hand, comparative analysis of pottery raw materials in civilizations such as Egypt, Mesopotamia and Anatolia have also shown that production technology plays an important role in understanding cultural exchange and civilizational origin<sup>12</sup>. In the last few decades, studies that have systematically examined the use of machine learning in archaeology suggest that the combination of chemical, mineralogical and image data, along with ML algorithms and fuzzy logic<sup>13</sup>, has provided a new perspective for the classification and identification of the origin of artifacts<sup>14</sup>. Even some researches, such as the project on Jian pottery or Yuan Dynasty studies in China<sup>15</sup>, have achieved very successful classification results using neural networks and end-to-end models<sup>16</sup>. The simultaneous use of scientific techniques, data mining and artificial intelligence makes the present research an example of a new interdisciplinary approach in archaeology<sup>17</sup>.

The present research, aiming to examine 20 pottery samples obtained from the ancient site of Claudiopolis in

the Bolu region<sup>18</sup>, used a combination of XRF, XRD, FTIR analytical methods along with machine learning algorithms (KNN, SVM, RF), PCA analysis and fuzzy logic. This research attempts to use the resulting data to determine the civilizational origin (Roman or Greek) of these samples and, through the analysis of compositional characteristics, attribute them to specific civilizations or cultural periods.

## MATERIALS AND METHODS

### Experimental Samples and Data:

For this study, 20 pottery samples were collected from the Bolu Claudiopolis Stadium archaeological site in Turkey<sup>19</sup>. These samples were subjected to three types of scientific analysis to analyze their physical, chemical, and mineral compositions:

- XRF (X-Ray Fluorescence): To identify the chemical elements present in the samples.
- XRD (X-Ray Diffraction): To identify mineral phases.
- FTIR (Fourier-Transform Infrared Spectroscopy): To analyze functional groups and molecular bonds.

### XRD Analysis

The XRD analysis of 20 ceramic samples fired at temperatures between 800°C and 900°C revealed the presence of two major crystalline phases: Quartz and Anorthite. No significant peaks for Mullite or Hematite were detected in the major diffraction peaks. The strongest diffraction peaks were identified at  $2\theta \approx 26.6^\circ$  corresponding to Quartz<sup>20</sup> and  $2\theta \approx 27.9^\circ$  and  $29.3^\circ$  corresponding to Anorthite<sup>21</sup>.

### FTIR Analysis

The FTIR spectra of all 20 samples revealed the presence of the following functional groups. OH- Stretching ( $3600-3700\text{ cm}^{-1}$ )<sup>22</sup>, indicating residual hydroxyl groups or adsorbed water. H<sub>2</sub>O Bending ( $1600-1650\text{ cm}^{-1}$ )<sup>23</sup>, typical of bound water in the clay structure<sup>24</sup>. Si-O Stretching ( $1000-1100\text{ cm}^{-1}$ ), characteristic of silicate networks in clay minerals<sup>25</sup>. Carbonate (C-O) stretching ( $1400-1450\text{ cm}^{-1}$ ) observed in some samples, possibly due to carbonate impurities<sup>26</sup>. These results suggest that the clay ceramic samples maintained silicate frameworks and some residual hydroxyl and carbonate phases post-firing, consistent with clay mineral transformations at medium firing temperatures.

### Data Preparation and Preprocessing:

Data related to each analysis were entered separately into Python software and merged based on the sample ID. Missing values were removed or replaced. Numerical features

<sup>1</sup> PEACOCK 1970; CECCARELLI *et alii* 2018.

<sup>2</sup> GIANNOSSA/FORLEO / MANGONE 2021; ZAMPIERIN *et alii* 2024

<sup>3</sup> PAPACHRISTODOULOU *et alii* 2006

<sup>4</sup> IKEOKA *et alii* 2018

<sup>5</sup> ERAMO *et alii* 2004

<sup>6</sup> PEACOCK 1970; KAŁUŻNA-CZAPLIŃSKA *et alii* 2017; CHEN/CHEN 2024

<sup>7</sup> BAYAZIT/ KAYNAK/COŞKUN 2023; KAHVECİ/ PEKŞEN 2023

<sup>8</sup> Qi *et alii* 2022

<sup>9</sup> DI ANGELO/DI STEFANO/PANE 2017

<sup>10</sup> LING/DELNEVO/MIRRI 2023; CHEN/CHEN 2024

<sup>11</sup> LONČARIĆ/COSTA 2023; LIU/TIAN/CHEN 2024

<sup>12</sup> SPARAVIGNA 2014

<sup>13</sup> BICKLER 2018

<sup>14</sup> LING *et alii* 2024; JIA/CHEN 2025

<sup>15</sup> SUN *et alii* 2020

<sup>16</sup> MU *et alii* 2019; QI *et alii* 2022

<sup>17</sup> NAKAMURA *et alii* 2016; A/QASIM/BARLA 2023; ZANIER 2024

<sup>18</sup> YİĞİTPAŞA 2023

<sup>19</sup> YİĞİTPAŞA 2024

<sup>20</sup> MOHAMMED/ALMASHHADANI 2023

<sup>21</sup> LIU *et alii* 2023

<sup>22</sup> PAN *et alii* 2011; NAKAMURA *et alii* 2016.

<sup>23</sup> SALISU/ ALMAJIR 2020

<sup>24</sup> WU *et alii* 2020; KEAWSAWASVONG/ LAI 2021

<sup>25</sup> VOLL *et alii* 2001; RASDI *et alii* 2017

<sup>26</sup> MOHAMMED/ALMASHHADANI 2023

were normalized by StandardScaler. Initial analysis included statistical summary, distribution diagram, and dominant feature analysis.

**Machine learning algorithms used:**

**Unsupervised Learning:**

- PCA (Principal Component Analysis): Dimension reduction and graphical display of clusters.
- K-Means Clustering: Grouping samples based on statistical similarity.
- Hierarchical Clustering: Analyzing the structure of clusters in a hierarchical manner.

**Supervised Learning:**

- K-Nearest Neighbors (KNN)
- Support Vector Machine (SVM)
- Random Forest (RF)

All models were evaluated using a random split method (Train/Test Split: 70/30) and Cross-validation.

**Data Fusion Techniques:**

To effectively combine the information obtained from the three analytical techniques, two fusion methods were used in machine learning. The descriptions of these methods are as follows:

**Table 1.** Fusion method description.

| FUSION METHOD       | DESCRIPTION  |
|---------------------|--|
| <b>EARLY FUSION</b> | All features from XRF, XRD, and FTIR are combined into a single dataset before model training.                     |
| <b>LATE FUSION</b>  | Separate models are trained for each analysis (XRF, XRD, FTIR), and their outputs are fused using majority voting. |

**Civilization Labeling:**

After examining the chemical properties and comparing them with the reference values of civilization elements, the samples were roughly labeled to the following civilizations:

- Roman
- Greek

The labeling algorithm was based on the proximity of the percentages of key elements such as SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, CaO, and K<sub>2</sub>O to the defined intervals of the civilizations.

**Using Fuzzy Logic to Predict Civilizations:**

A Fuzzy Inference System was used to estimate the probable civilization of each ceramic sample. This system is designed based on comparing the percentage of chemical compositions (major elements) of each sample with the reference ranges of elements of ancient civilizations. Instead

of definitively assigning a class, fuzzy logic allows for an approximate and weighted assessment of the sample's belonging to different civilizations. For example: a sample can be attributed to the Roman civilization with a probability of 70% and to the Greek civilization with a probability of 30%.

**Steps to perform fuzzy analysis:**

1. Define the range of elements (Input variables) for different civilizations based on the reference data.
2. Determine the fuzzy membership functions for major elements such as SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, CaO, K<sub>2</sub>O.
3. Create fuzzy rules to match the amount of elements in each sample with civilizations.
4. Calculating the final civilization score using Mamdani fuzzy inference and assigning the dominant civilization.

**Features used in fuzzy analysis**

The application of fuzzy logic is that the fuzzy system acted as a complementary tool to machine learning. Matching the fuzzy results with the results of ML models increased the validity of the classifications (Table 1). In particular, this method provided useful insight in cases where the ML models were doubtful or the data was incomplete.

**Table 2.** Features used in fuzzy analysis.

| Element                                     | Role in Fuzzy Logic                                      |
|---|--|
| <b>SiO<sub>2</sub></b>                      | Indicator of quartz-based structure and firing temp      |
| <b>Al<sub>2</sub>O<sub>3</sub></b>          | Reflects clay mineral content (kaolinite, illite)        |
| <b>Fe<sub>2</sub>O<sub>3</sub></b>          | Color agent, indicates firing and iron-rich environments |
| <b>CaO</b>                                  | High levels linked with lime-based wares (e.g. Roman)    |
| <b>K<sub>2</sub>O, MgO, Na<sub>2</sub>O</b> | Trace indicators, refinement or impurities               |

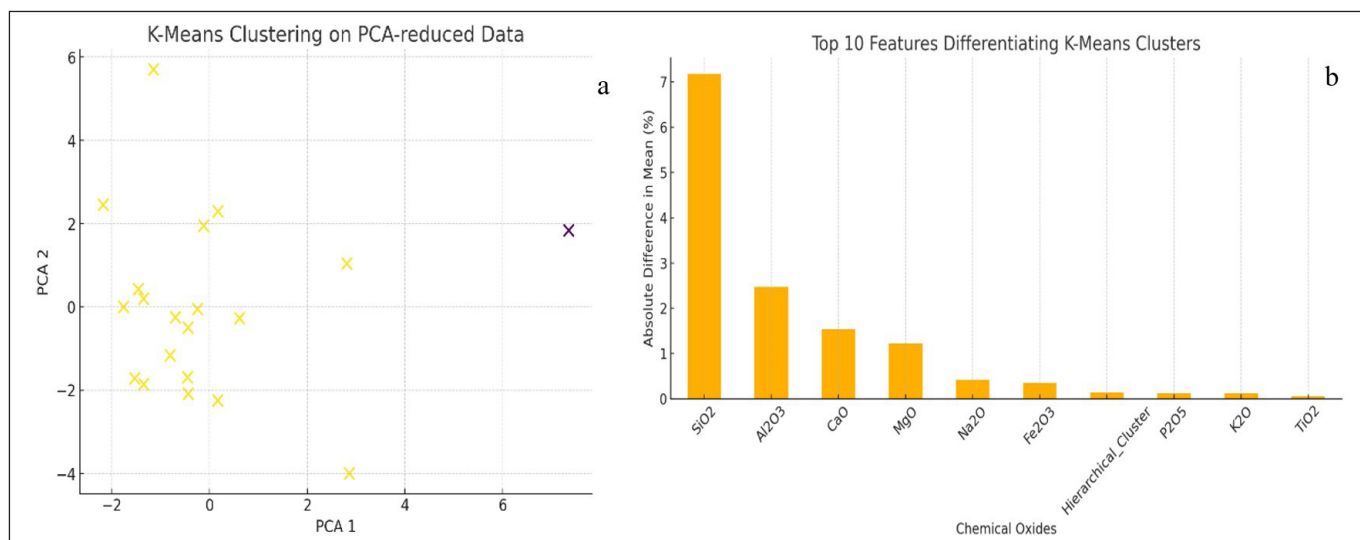
**Weka Analysis**

Weka software was also used to validate and compare the performance of different machine learning models. In this analysis, three models including K-Nearest Neighbors (IBk), Support Vector Machine (SMO) and Random Forest were applied to ceramic data. The aim was to examine the quality of civilization class prediction using chemical and physical properties of the samples. For each model, validation was performed using the Cross-validation method with 10 iterations.

**RESULTS AND DISCUSSIONS**

Unsupervised learning and clustering of ceramic samples based on XRF, XRD and FTIR data.

Unsupervised machine learning methods were used to investigate natural patterns and identify cultural affiliations of ceramic samples without specific civilization labels. In this regard, chemical (XRF), mineralogical (XRD) and



**Fig. 1.** a) PCA plot after clustering with the K-Means method, b) Chemical features distinguishing the clusters of the K-Means model. In this graph, the 10 features with the largest mean difference between the two clusters created by the K-Means model are shown. These differences are calculated as a percentage of the difference in the mean values of each feature between the clusters.

infrared spectroscopy (FTIR) data were combined and, after preparation, subjected to clustering analysis. XRF: Chemical data included the percentage of major oxides and trace elements such as SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, MgO, etc. These data were directly entered into the model as numerical features. XRD: Several peaks were extracted from the XRD powder data for each sample. Then, three important statistical features were calculated from them (average peak intensity, highest intensity and number of peaks with an intensity greater than 5000). FTIR: The spectral data were in text form including various functional groups such as “Carbonate (C-O)”, “OH Stretching”, “Si-O Stretching”. These key terms were extracted and converted into binary features (1: presence, 0: absence). After combining these three types of data, a rich set of combined features was provided for the final clustering.

**Analysis of clustering results of samples.**

In this study, to investigate the natural patterns among ceramic samples, unsupervised clustering was performed using two methods, K-Means and Hierarchical Clustering. The following table shows the clustering results for samples B1 to B5 (Table 3).

**Table 3.** Assignment of ceramic samples B1 to B5 to clusters in two different models.

| Sample | K-Means Clustering | Hierarchical Clustering |
|--------|--------------------|-------------------------|
| B1     | 1                  | 0                       |
| B2     | 1                  | 0                       |
| B3     | 1                  | 0                       |
| B4     | 1                  | 1                       |
| B5     | 1                  | 1                       |

The K-Means model placed all five samples (B1 to B5) in a single cluster (cluster 1). This result shows that the model did not observe any significant differences between the chemical, mineral and spectral properties of these samples

and considered them to be statistically very similar. At this level of data, K-Means had a homogeneous and integrated view of this set of samples. Unlike K-Means, the hierarchical model placed the first three samples (B1, B2, B3) in cluster 0 and the other two samples (B4 and B5) in cluster 1. This separation shows that the hierarchical model was able to identify the relative differences between the samples. B4 and B5 seem to have characteristics that distinguish them from the first three samples; this could be due to differences in raw material composition, manufacturing method, or soil source. The K-Means model in this set considered all samples as a general group. While the hierarchical model was able to separate two distinct samples (B4 and B5) and provide a more accurate classification. Therefore, it can be said that the hierarchical model would be a more appropriate choice when the difference between samples is minor and subtle.

**Dimensionality reduction with PCA.**

To simplify and enable graphical display of the data, the principal component analysis (PCA) method was used. The multidimensional data was reduced to two principal components that retain the most variance possible. These two components were used as horizontal and vertical axes in the clustering analysis.

**PCA Graph Analysis with K-Means Clustering**

In this study, in order to investigate hidden patterns in ceramic composite data, after merging XRF, XRD and FTIR data, Principal Component Analysis (PCA) was used for dimensionality reduction (Fig. 1). The aim of this analysis was to compress the data into two principal components to allow for graphical display and clustering analysis while preserving as much variance as possible.

Each point represents a ceramic sample. The color of the points is determined by the K-Means model and indicates the cluster assigned to that sample. The horizontal and vertical axes represent the first component (PC1) and the

second component (PC2), respectively, which retained the most information (variance) of the data. The PCA plot showed that most of the samples (B1, B2, B3) were clearly concentrated in one area. Two samples (B4 and B5) were slightly away from this concentration, indicating a slight difference in their feature composition compared to the others. This relative separation was also statistically confirmed by the hierarchical model, which placed B4 and B5 in a separate cluster. Although the K-Means model placed all samples in one cluster, the visual analysis of the PCA was able to show the possible differentiation of B4 and B5. PCA analysis has been a very effective tool for understanding the distribution and internal structure of multivariate data. Despite the limited sample size, this diagram was able to highlight the differences between groups and provide a basis for more precise civilization classification in the future. It is suggested that the same method be used to more accurately identify production or cultural groups in a larger sample.

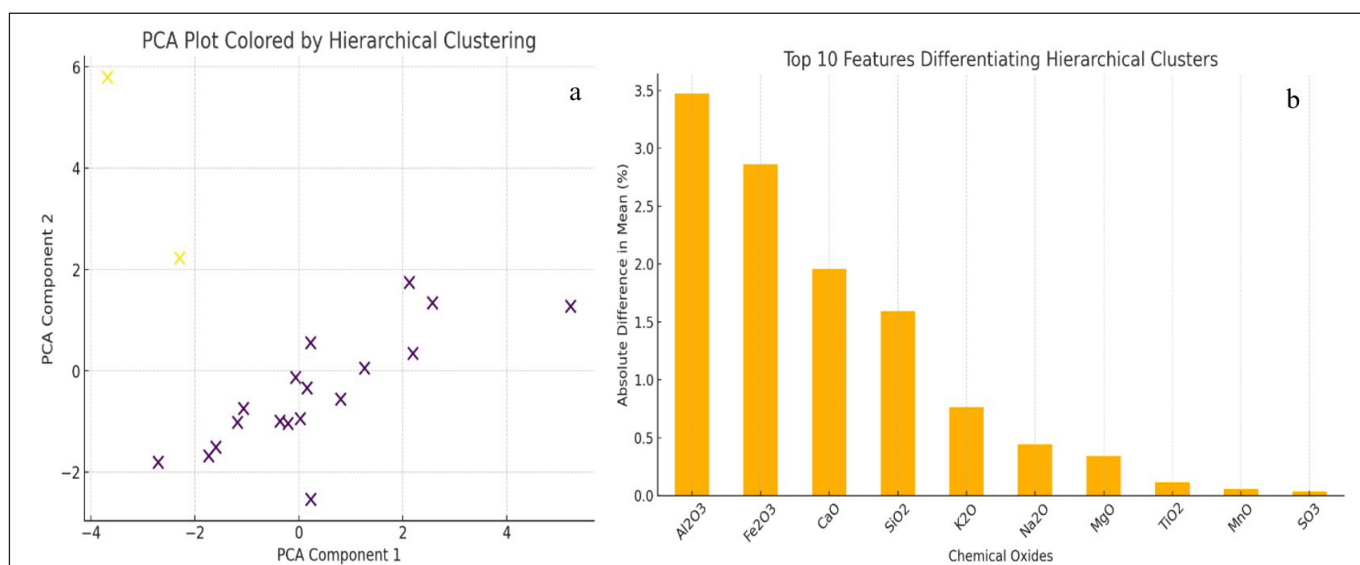
Analysis of the mean features in the two K-Means clusters showed that there were significant differences between them. The following table lists the 10 features with the greatest differences between the clusters (Fig. 2 b). The largest difference between the two K-Means clusters was related to SiO<sub>2</sub> (silica), which, with a difference of 7.18% between the clusters, shows that it played a major role in separating the groups. This may be related to the use of soils richer or poorer in silica in the initial composition of the pottery. Al<sub>2</sub>O<sub>3</sub> and CaO also had high differences and could be an indication of differences in the type of clay or mineral additives in the samples. Other oxides such as MgO, Na<sub>2</sub>O, Fe<sub>2</sub>O<sub>3</sub> and K<sub>2</sub>O also played a minor but effective role in separating the clusters. The K-Means model was able to place the samples into two distinct clusters based on chemical differences, especially in silica and aluminum. Combining this analysis with the PCA plot allows for a better interpretation of the data structure and their relationship to possible civilizations. SiO<sub>2</sub> (silica) played the largest role in separating the clusters, with a difference of more than 7% between the two groups. This

amount of difference usually indicates a difference in the soil source or the initial composition of the clay used in the production of ceramics. Al<sub>2</sub>O<sub>3</sub> (aluminum) and CaO (calcium) also had significant differences between the clusters. These oxides can indicate the use of specific raw materials or different firing temperatures. Differences in the values of MgO, Na<sub>2</sub>O, Fe<sub>2</sub>O<sub>3</sub> and K<sub>2</sub>O also indicate changes in the soil composition or additives in the production process of the samples. Fig. 4 shows that the K-Means model successfully classified the ceramic samples into two distinct clusters using significant differences in chemical composition. This separation can be a basis for civilizational or manufacturing technology analyses, especially when direct historical data are not available.

**PCA Plot Analysis with Hierarchical Clustering Colorization**

To examine the internal structure of the combined data (XRF, XRD, FTIR) and identify natural patterns among the ceramic samples, after applying dimensionality reduction with PCA, clustering (Hierarchical Clustering) was applied to the original data and the results were displayed on the PCA plot (Fig. 2 a).

In this plot, each point represents a ceramic sample. The color of each point corresponds to the cluster that the hierarchical model has identified for it. The PCA components on the two principal axes (PC1 and PC2) have preserved most of the variance of the information. The plot shows that some samples are located in different spaces than others, which the hierarchical model has also placed in separate clusters. This separation shows that hierarchical clustering was able to respond to more precise internal differences in the data, even if these differences are seemingly insignificant. Although the K-Means model in the previous examples placed all the samples in one cluster, the hierarchical model has identified two separate groups in the PCA space, which



**Fig. 2.** a) 2D PCA plot colored based on hierarchical clustering output, b) Chemical Features Distinguishing Hierarchical Clusters. The above graph shows the absolute difference in the mean percentage of ten chemical elements/oxides between the two clusters resulting from hierarchical clustering.

can be very valuable. This difference in clustering can be due to differences in chemical composition, intensity of mineral peaks, or the presence of specific spectral groups in the FTIR that the hierarchical model has well recognized. PCA plot combined with hierarchical clustering is a very suitable tool for multivariate analyses and identifying subtle differences in archaeological data. This method can help to separate samples belonging to different civilizations or production centers, especially when direct historical information is not available. Combining this analysis with other methods such as K-Means and dendrogram analysis can provide more accurate and multifaceted results. Based on the mean difference between two clusters, the following 10 features had the most impact on separating clusters (Fig. 2 b). The three main oxides that caused the greatest difference between the clusters were: Al<sub>2</sub>O<sub>3</sub> (aluminum) the largest difference (3.47%), which is usually associated with refractory materials or feldspar-rich soils. Fe<sub>2</sub>O<sub>3</sub> (iron): 2.86% difference, which may indicate soil type and oxidation during the firing process. CaO (calcium): 1.96% difference, which could indicate the use of lime or calcite in the raw materials. Other elements such as SiO<sub>2</sub>, K<sub>2</sub>O, Na<sub>2</sub>O and MgO also showed significant differences, which are likely related to soil source or pottery additives. These

features, especially Al<sub>2</sub>O<sub>3</sub> and Fe<sub>2</sub>O<sub>3</sub>, play a key role in separating civilization or production groups and can be the basis for more precise classification of civilizations in future studies.

**Clustering analysis of chemical composition of ceramic samples.**

In order to investigate the internal differences of ceramic samples based on chemical composition, a/the clustering technique was applied to the combined XRF data. The results showed that the samples were grouped into four clusters (Cluster 0 to Cluster 3). For each cluster, the average of major elements was calculated to analyze the chemical differentiation between the groups.

Cluster analysis of chemical elements showed that the data were well divided into groups with significant chemical differences. These differences could be due to the selection of different soil sources, additives, or differences in ceramic manufacturing technology. In the next step, to visually examine the separation of clusters, principal component analysis (PCA) will be performed and the clusters will be displayed in two-dimensional spaces with distinct colors.

**Table 4.** Clustering analysis of chemical composition of ceramic samples

| ELEMENTS                       | KEY ELEMENT DIFFERENCES BETWEEN CLUSTERS   |
|--------------------------------|--|
| Al <sub>2</sub> O <sub>3</sub> | Cluster 3 has the highest mean (14.67%), which may indicate the use of richer clay.  |
| SiO <sub>2</sub>               | is highest in Cluster 0 (58.93%), indicating a ceramic body with a high silica content   |
| Fe <sub>2</sub> O <sub>3</sub> | is significantly higher in Cluster 3 (9.38%), which may be related to oxidative conditions or pottery color.   |
| K <sub>2</sub> O               | Cluster 2 has the highest potassium content at 2.29%, which could be an indication of the use of feldspar or potassic clay.                          |
| CaO                            | Cluster 1 has the highest lime content (9.63%), which reinforces the possibility of lime or calcareous soil being used in the manufacturing process. |
| MnO                            | MnO is higher in Clusters 2 and 3 than in the other clusters, which could contribute to the final color.   |

**Table 5.** Interpretation ranges

| ORIGIN LIKELIHOOD (%) | SCIENTIFIC INTERPRETATION                                    |
|-----------------------|--|
| >70%                  | High probability for specific origin (completely consistent) |
| 50-70%                | Moderate probability, general and acceptable composition     |
| <50%                  | Low probability, different origin or unusual composition     |

**Table 6.** Samples with the highest probability

| SAMPLE | ORIGIN LIKELIHOOD (%) | INTERPRETATION   |
|--------|-----------------------|--|
| B2     | ~83                   | Very specific chemical origin, very high probability of belonging to a specific civilization |
| B16    | ~75                   | Composition close to known patterns, identifiable  |
| B12    | ~66                   | Close to a specific group, may require compositional investigation                           |

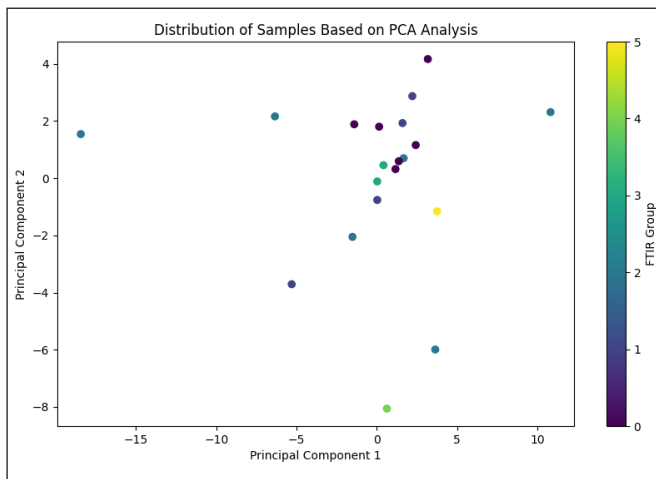
**Scientific and statistical analysis of the output.**

Our model produces an output called Origin Likelihood (%), which ranges from 0 to 100. This number indicates how close the sample is to a particular civilization/origin in terms of chemical composition.

The XRF dataset, based on seven principal oxides, enabled effective compositional differentiation among the ceramic samples. The fuzzy logic model, applied at this stage, functioned as a preliminary classification tool by estimating

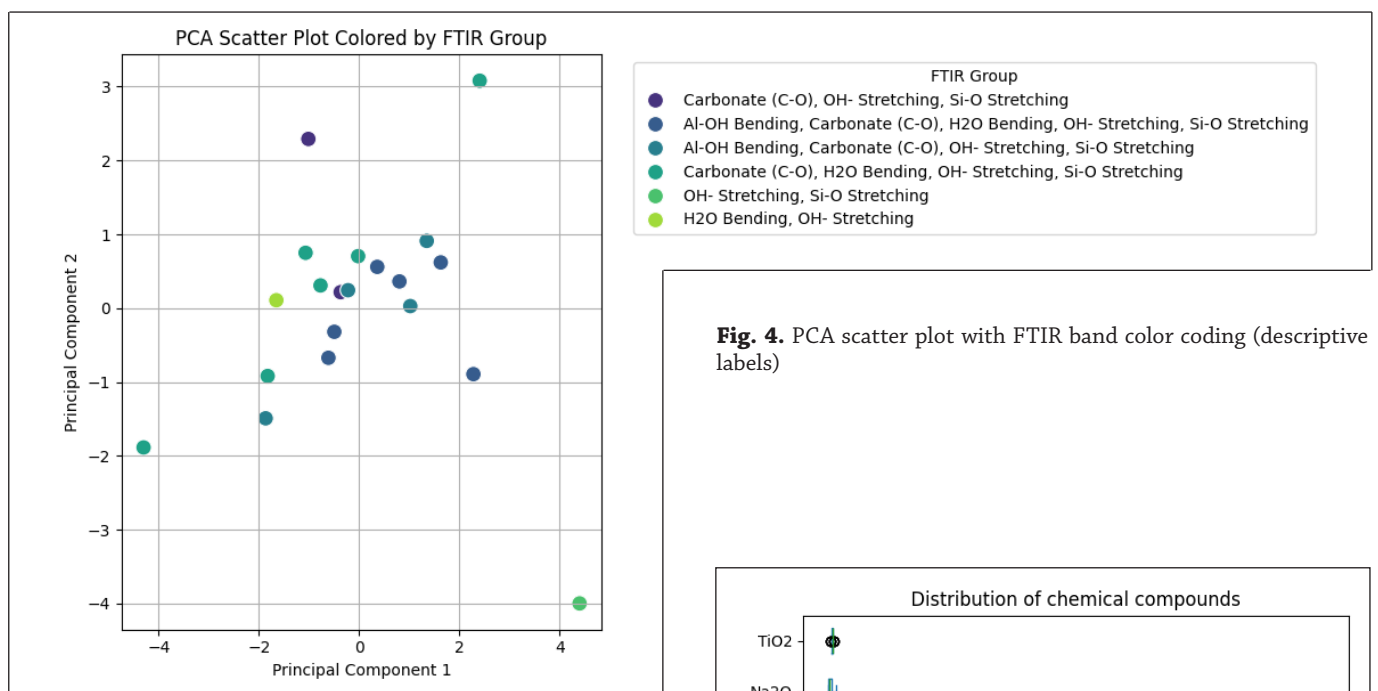
**Table 7.** Samples with low probability:

| SAMPLE | ORIGIN LIKELIHOOD (%) | INTERPRETATION   |
|--------|-----------------------|--|
| B1     | ~43                   | Different composition, may be local or specially crafted   |
| B15    | ~42                   | Probably does not fit the classical origin   |
| B5     | ~45                   | On the verge, requires further investigation into the compositional model or microscopic structure |



**Fig. 3.** Distribution of ceramic samples based on principal component analysis (PCA) of XRF data.

combination of molecular vibrational modes (e.g., OH group stretching, C–O carbonate bonds, Si–O bond stretching, and H<sub>2</sub>O bending). Using the color guide (Legend) allows for a more accurate and understandable interpretation of the spectral bands. This representation enhances the understanding of the relationship between the spectral features and the chemical composition of the samples. Both PCA plots represent chemical similarities or differences between ceramic samples in a reduced two-dimensional space. In both plots, there is a tendency for clusters to form between samples with similar FTIR characteristics. The use of descriptive labels (in Fig. 4) allows the researcher to make logical and scientific connections between spectral features and chemical composition. In contrast, the coded representation (Fig. 3) is more suitable for numerical analysis and computational clustering. In contrast, TiO<sub>2</sub> and Na<sub>2</sub>O have low abundances and low volatility (Fig. 5). The

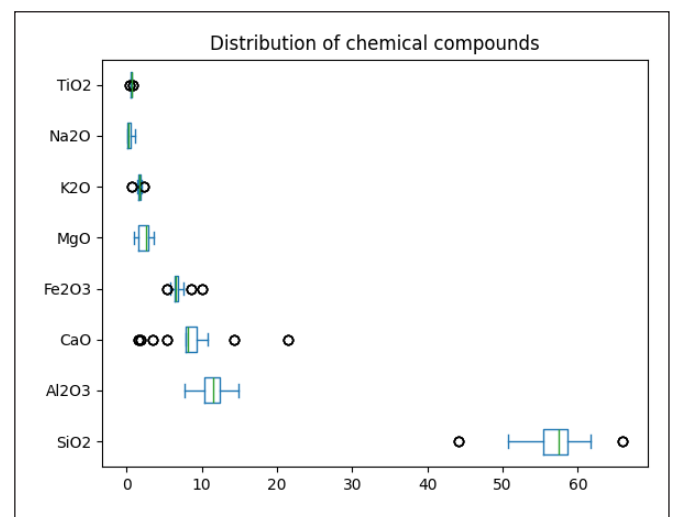


**Fig. 4.** PCA scatter plot with FTIR band color coding (descriptive labels)

the likelihood of each sample’s affiliation with a specific civilization. Samples exhibiting high probability values can be interpreted as strong candidates for cultural attribution, whereas those with lower probability scores may reflect local production, trade interactions, or atypical compositional features. This probabilistic approach complements the clustering and machine learning results, providing an additional layer of interpretative validation.

In Fig. 3, the color of each point represents the FTIR grouping of that sample. It can be seen that samples belonging to specific FTIR groups are mainly located in distinct regions, indicating a correlation between spectral features (FTIR) and chemical composition (XRF). This correlation can be a basis for more precise clustering of samples and investigation of their technological origin (Fig. 4).

The plot in Fig. 4 shows the same PCA results, but the samples are colored based on the full names of the FTIR spectral bands. Each color in the plot represents a specific



**Fig. 5.** Distribution of the main chemical elements in the form of a boxplot. SiO<sub>2</sub> is the most abundant element in the composition of the samples and shows a greater dispersion than the other elements.

presence of outliers in the elements CaO and Fe<sub>2</sub>O<sub>3</sub> could indicate differences in soil source or manufacturing process between civilizations or specific samples.

## Cluster Analysis of Ceramic Samples (XRF + XRD + FTIR)

### 1. K-Means Cluster Analysis

To identify hidden patterns in the integrated XRF, XRD, and FTIR dataset, a K-Means clustering algorithm was applied with  $k=3$  clusters. PCA (Principal Component Analysis) was used to reduce the dimensionality to two components for better visualization (Fig. 6 a). The scatter plot below illustrates the distribution of each ceramic sample based on their clustering. Each point represents a sample categorized by its chemical composition (XRF), mineral phases (XRD), and functional groups (FTIR) in a two-dimensional PCA space (Fig. 6 b).

Cluster 0, samples with diverse FTIR functional groups and higher values of  $\text{SiO}_2$ , and certain oxides in XRF. Cluster 1, simpler samples with fewer FTIR groups and notable variations in oxides like  $\text{CaO}$  and  $\text{MgO}$ . Cluster 2, Similar to Cluster 0 in FTIR diversity but showing distinct XRF elemental patterns. This cluster analysis indicates that the ceramic samples can be broadly divided into three distinct groups based on their physicochemical characteristics. The samples were grouped using KMeans clustering into 4 clusters based on their chemical composition. The dataset was also used to train a Random Forest classifier to predict the 'Group' label (Table 8). The classifier achieved 100% accuracy on the test set.

This document presents the result of predicting the group label for a new ceramic sample based on its chemical composition. The trained Random Forest model was used for this prediction. The predicted group for the sample is OH-Stretching and Si-O Stretching

Supervised learning and clustering of ceramic samples based on XRF, XRD and FTIR data. XRF analysis of ceramic samples

In this study, 20 ceramic fragments were chemically analyzed by X-ray fluorescence (XRF). XRF analysis provided percentage values of major oxides and ppm values of minor elements. The main objective of this study was to investigate the civilizational origin of these samples based on their chemical matching with reference data of ancient civilizations. A range matching method was used to assess the origin of each sample. In this way, a specific range of major element values ( $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{CaO}$ ,  $\text{MgO}$ ,  $\text{K}_2\text{O}$ ,  $\text{Na}_2\text{O}$ ) was considered for each civilization. Then, each sample received a score for each civilization based on its compliance with these ranges. For example, if the  $\text{SiO}_2$  value of a sample falls within the range of Roman and Egyptian civilizations, both civilizations receive one point. Finally, the civilization that scores the highest is considered the probable origin of that sample (Table 9).

Analysis of the results shows that out of the 20 samples examined, 19 samples have the highest match with the ceramics belonging to the Roman civilization, and only

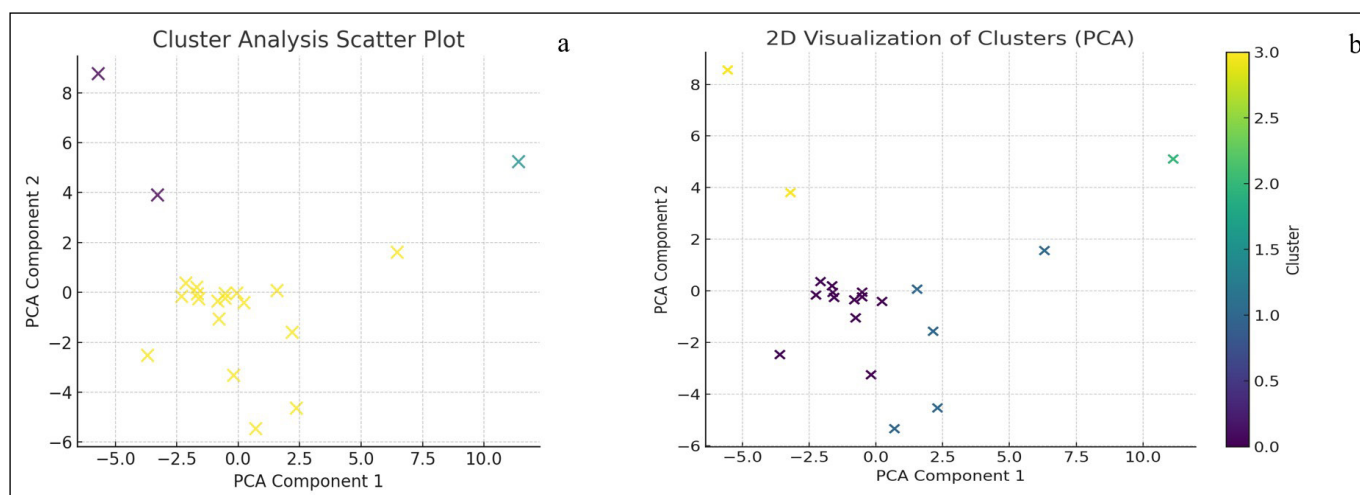


Fig. 6. a) Cluster analysis scatter Plot b) 2D Visualization of Clusters (PCA)

Table 8. Classification Report

| CLASS   | PRECISION | RECALL | F1-SCORE | SUPPORT |
|---|-----------|--------|----------|---------|
| Al-OH Bending, Carbonate (C-O), H <sub>2</sub> O Bending, OH- Stretching, Si-O Stretching | 1.00      | 1.00   | 1.00     | 7.00    |
| Al-OH Bending, Carbonate (C-O), OH- Stretching, Si-O Stretching                           | 1.00      | 1.00   | 1.00     | 3.00    |
| Carbonate (C-O), H <sub>2</sub> O Bending, OH- Stretching, Si-O Stretching                | 1.00      | 1.00   | 1.00     | 5.00    |
| Carbonate (C-O), OH- Stretching, Si-O Stretching  | 1.00      | 1.00   | 1.00     | 1.00    |
| Accuracy  | 1.00      | 1.00   | 1.00     | 1.00    |
| Macro Avg   | 1.00      | 1.00   | 1.00     | 16.00   |
| Weighted Avg  | 1.00      | 1.00   | 1.00     | 16.00   |

**Table 9.** Civilization Scoring for Each Sample (Major Oxides)

| SAMPLE | BEST MATCH | ROMAN | EGYPTIAN | GREEK | ACHAEMENID | ELAMITE | MESOPOTAMIAN | CHINESE | INDIAN |
|--------|------------|-------|----------|-------|------------|---------|--------------|---------|--------|
| B1     | Roman      | 6     | 5        | 5     | 5          | 5       | 6            | 1       | 5      |
| B2     | Roman      | 7     | 7        | 7     | 7          | 7       | 6            | 3       | 6      |
| B3     | Roman      | 7     | 7        | 7     | 7          | 7       | 7            | 2       | 6      |
| B4     | Roman      | 3     | 3        | 3     | 3          | 3       | 3            | 1       | 3      |
| B5     | Roman      | 6     | 4        | 4     | 4          | 5       | 6            | 1       | 4      |
| B6     | Roman      | 6     | 3        | 4     | 3          | 5       | 6            | 1       | 3      |
| B7     | Roman      | 5     | 3        | 3     | 3          | 4       | 5            | 2       | 3      |
| B8     | Roman      | 6     | 6        | 6     | 6          | 6       | 6            | 2       | 6      |
| B9     | Roman      | 5     | 4        | 4     | 4          | 5       | 5            | 1       | 4      |
| B10    | Roman      | 7     | 7        | 7     | 7          | 7       | 7            | 2       | 6      |
| B11    | Roman      | 6     | 6        | 6     | 6          | 6       | 6            | 1       | 5      |
| B12    | Roman      | 7     | 7        | 7     | 7          | 7       | 7            | 3       | 6      |
| B13    | Roman      | 6     | 4        | 5     | 5          | 5       | 6            | 1       | 5      |
| B14    | Roman      | 7     | 7        | 7     | 7          | 7       | 7            | 2       | 6      |
| B15    | Roman      | 6     | 5        | 5     | 5          | 5       | 6            | 1       | 5      |
| B16    | Roman      | 5     | 4        | 4     | 4          | 4       | 5            | 2       | 4      |
| B17    | Roman      | 6     | 5        | 5     | 5          | 5       | 6            | 1       | 5      |
| B18    | Roman      | 6     | 5        | 5     | 5          | 5       | 6            | 1       | 5      |
| B19    | Roman      | 6     | 5        | 5     | 5          | 5       | 6            | 1       | 5      |
| B20    | Greek      | 5     | 5        | 6     | 5          | 6       | 6            | 1       | 5      |

**Table 10.** For samples such as B1 to B3, the key elements are fully aligned with the Roman civilization element range:

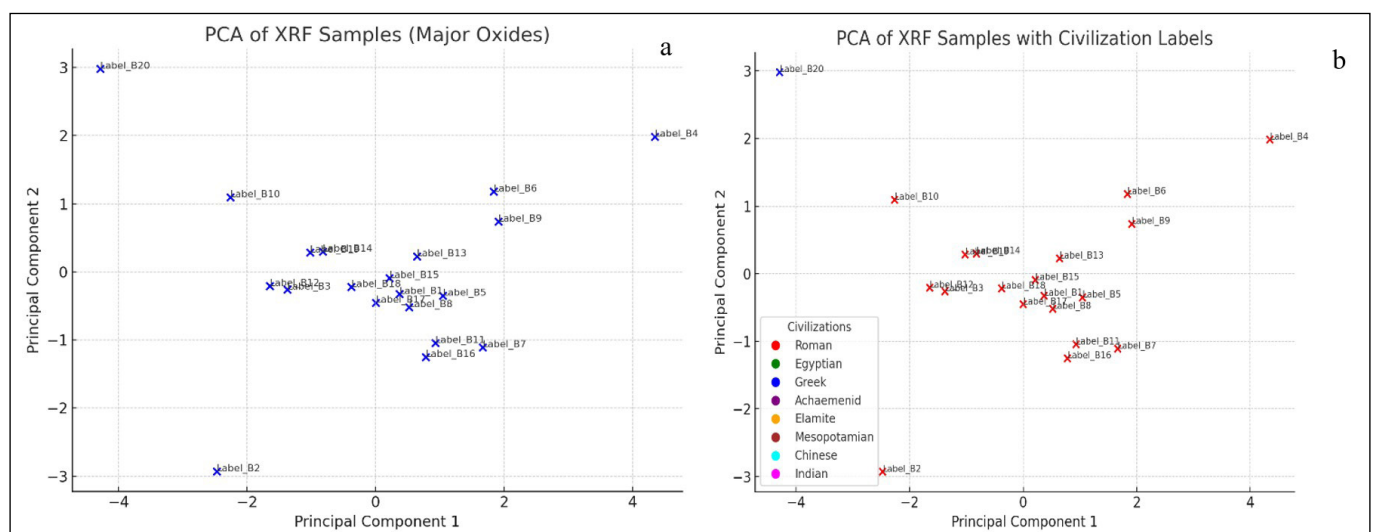
| Sample B1   | Sample B2  | Sample B4  |
|---|--|--|
| SiO <sub>2</sub> = 56.6%, Al <sub>2</sub> O <sub>3</sub> = 11.5%, Fe <sub>2</sub> O <sub>3</sub> = 6.0%, CaO = 8.3%<br>Perfect match with Roman civilization in four major elements | SiO <sub>2</sub> = 66%, Fe <sub>2</sub> O <sub>3</sub> = 5.4%, CaO = 3.4%<br>Perfect match with Roman, Greek, Achaemenid civilizations | SiO <sub>2</sub> is slightly lower, but Fe <sub>2</sub> O <sub>3</sub> and CaO values are still within the Roman civilization range. |

with the Roman civilization is due to the high agreement in the values of the elements SiO<sub>2</sub>, Fe<sub>2</sub>O<sub>3</sub> and CaO.

**Principal Component Analysis (PCA)**

To identify hidden patterns and relationships between samples, a principal component analysis (PCA) was performed based on seven principal oxides. The results of this analysis are as follows: First component (PC1) explaining 48.7% of the variance in the data. Second component (PC2) covering 21.05% of the variation (Fig. 7). Together, the first two components cover more than 70% of the total variance in the data.

sample B20 is closer to the Greek civilization. The high match



**Fig. 7.** a) PCA plot without civilization labels (Major Oxides), b) PCA of XRF Samples with Civilization Labels

In this plot, each sample is colored according to the civilization closest to it. Red is for Roman, blue for Greek, green for Egyptian, and other civilizations are marked with other colors (Fig. 7 b). As can be seen, most of the samples are clustered in the Roman civilization area, which confirms their high correspondence with the chemical composition of Roman ceramics. The findings of this study show that most of the ceramic samples examined have a very high correspondence with the ceramics of the Roman civilization in terms of chemical composition. This correspondence is clearly visible in both numerical examination (score table) and visual analysis (PCA plots). Only one sample was closer to the Greek civilization, indicating limited diversity in the origin of the samples.

**Comparison of Early Fusion and Late Fusion Methods for Civilization Identification**

In this study, two common approaches in machine learning for analyzing combined XRF, XRD, and FTIR data were examined:

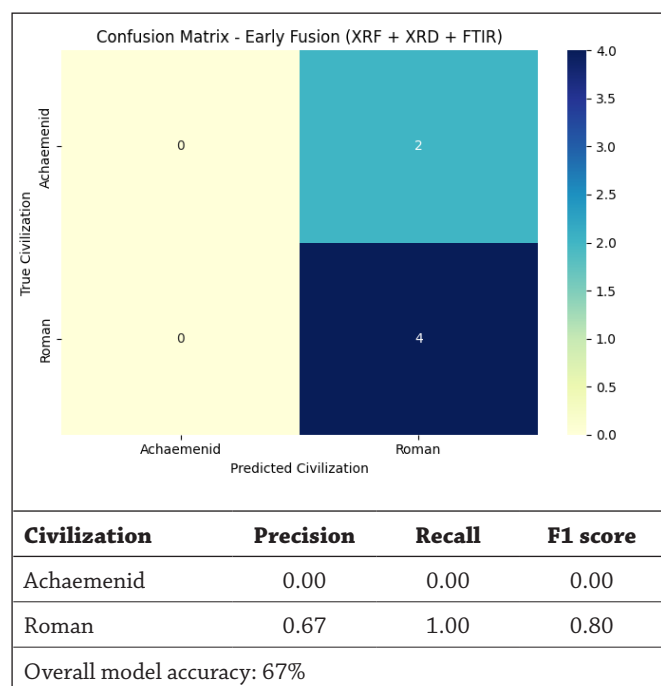
**Early Fusion Method:**

In this method, data related to XRF (elemental), XRD (mineral), and FTIR (spectral) analyses were combined into a table (Fig. 8).

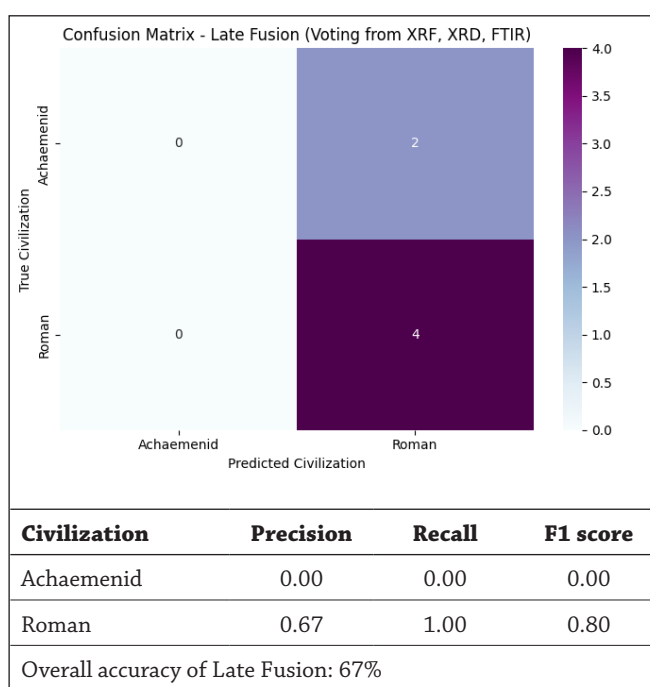
The Roman civilization was identified with very high accuracy by the model. In contrast, the Achaemenid civilization was not predicted by the model due to the small number of samples, which reduced the overall accuracy.

**Late Fusion Method:**

In this method, a separate model was first trained for each type of data (XRF, XRD, FTIR). Then, using the Majority Voting method, the final output of the model was combined.

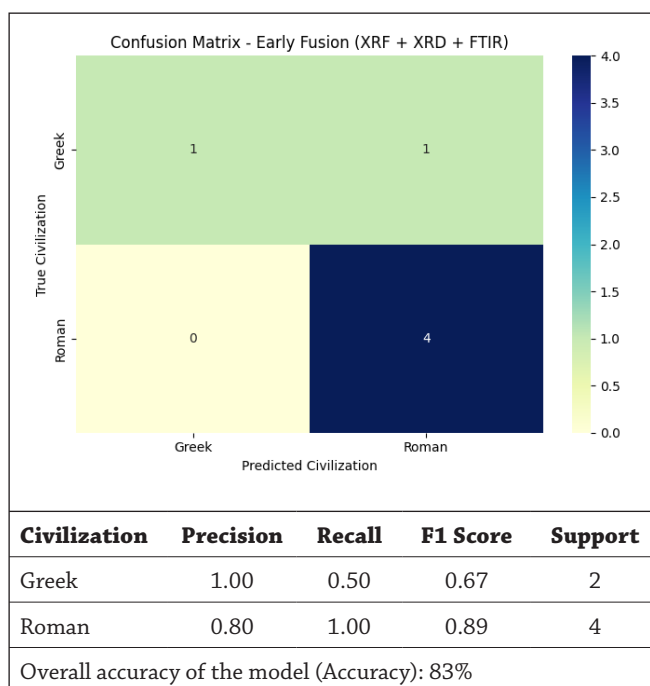


**Fig. 8.** Using the Random Forest model, ancient civilizations were predicted.



**Fig. 9.** Civilization classification results

Late Fusion results were similar to Early Fusion, but its advantage is in analyzing the accuracy of each analysis separately. Also, this method allows to examine the impact of each analysis (XRF, XRD, FTIR) on the final output. Both methods performed similarly in terms of overall accuracy (%67). If the samples were increased and the data were balanced, the accuracy of less represented civilizations such as the Achaemenid would also increase. Late Fusion is more suitable in cases where separate analysis of the data is important. Given the historical and geographical evidence of the area under study, Greek civilization was used as the second class instead of Achaemenid civilization. This



**Fig. 10.** Confusion Matrix, Model evaluation indicators (Classification Report)

is because the Claudiopolis region was influenced by the culture and architecture of Roman and Greek civilizations, and the direct presence of the Achaemenids in this area has not been proven.

Roman civilization was predicted with very good accuracy by the model and all 4 of its samples were classified correctly. Greek civilization had one sample correctly identified but one sample was mistaken for Roman (Fig. 10). Given the small sample size, the 83% performance in Early Fusion is acceptable and promising. After modifying the civilization labels and replacing the Achaemenid civilization with Greek, the machine learning model performed better with the Early Fusion method (Fig. 11). The overall accuracy of the model reached 83% and the prediction accuracy of the Roman civilization was 100%. These results show that correct and historically accurate labeling has a significant impact on improving the model's performance.

Late Fusion performed exactly the same as Early Fusion, demonstrating the stability of the model in multiple data fusion. Using majority voting between three separate models (XRF, XRD, FTIR) was able to identify the Roman class with 100% accuracy. In the case of the Greek civilization, the model made a mistake in identifying one of the two samples, which could be due to the similarity of the characteristics of that sample with the Roman civilization. In the Late Fusion method, after training the models separately for each analysis (XRF, XRD, FTIR) and combining the outputs by voting, the model was able to achieve an accuracy of 83%, which is the same as Early Fusion (Table 11). This result shows that combining the outputs of the models can be as effective as the initial data fusion.

Using combined data from XRF, XRD, and FTIR analyses on 20 ancient ceramic samples, two machine learning

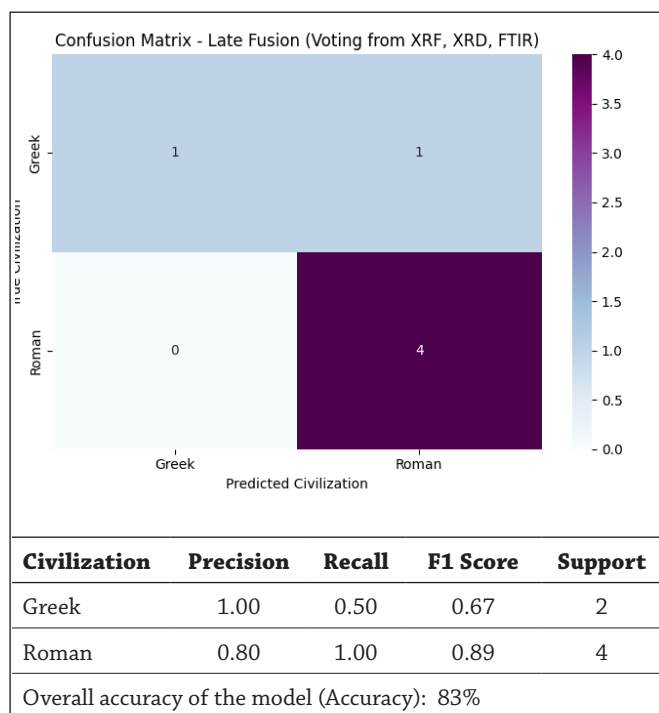


Fig. 11. Confusion Matrix , Late Fusion Evaluation Indicators

Table 11. Final Comparative Table

| CRITERIA                             | EARLY FUSION                          | LATE FUSION   |
|--------------------------------------|---------------------------------------|---|
| <b>Fusion Technique</b>              | Combined features from XRF, XRD, FTIR | Majority voting from separate models (XRF, XRD, FTIR) |
| <b>Overall Accuracy</b>              | 83%                                   | 83%   |
| <b>Precision – Greek</b>             | 1.00                                  | 1.00  |
| <b>Recall – Greek</b>                | 0.50                                  | 0.50  |
| <b>Precision – Roman</b>             | 0.80                                  | 0.80  |
| <b>Recall – Roman</b>                | 1.00                                  | 1.00  |
| <b>Macro F1-Score</b>                | 0.78                                  | 0.78  |
| <b>Model Stability</b>               | Yes                                   | Yes   |
| <b>Easier Feature Interpretation</b> | Yes                                   | No  |
| <b>Advantage</b>                     | Simpler and faster                    | Allows analysis of individual techniques              |

methods, Early Fusion and Late Fusion, were evaluated and compared. Also, to better align with the historical context of the Claudiopolis region, the civilization labels were divided into two groups: Roman and Greek. Both methods provided an overall accuracy of 83% in predicting civilizations, indicating acceptable performance. The models were successful in identifying the Roman civilization with high accuracy (100% Recall). Only one sample of the Greek civilization was incorrectly identified as Roman in both methods. The Early Fusion method is simpler to implement by combining data in a single model, and the interpretation of features is clearer. In contrast, the Late Fusion method allows for the effect of each analysis to be examined separately, which is very useful in multi-source analyses. Overall, both approaches have been able to provide meaningful information about the compositional properties of ceramics, confirming the applicability of machine learning in analytical archaeology. The use of these methods can be developed as an accurate tool for recognizing historical civilizations through the composition of elements, minerals, and spectral structures of ceramics.

RESULT OF WEKA

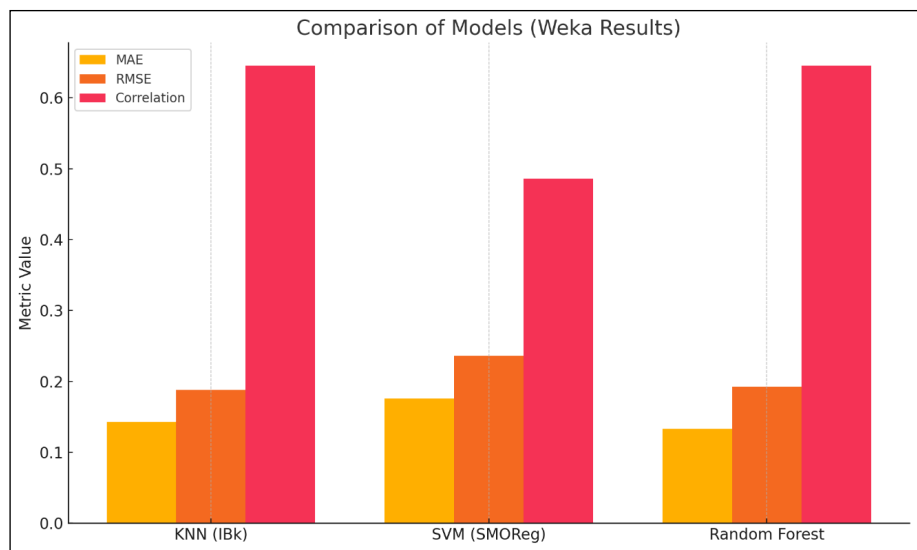
The results of Weka showed that the KNN model with a correlation coefficient of 0.646 provided a good performance compared to other models (Table 12), while the Random Forest model also recorded a similar coefficient. The SVM model had a lower accuracy, which could be due to its sensitivity to small data or kernel settings.

In order to increase the validity and reliability of the applied models, the results of the algorithms were also extracted in Weka software, which showed a relative agreement with the results obtained in the Python environment, and the KNN model showed relatively stable performance in both environments. This is a comparison chart of machine learning models in Weka, which includes,

**Table 12.** Final table of Weka model evaluation

| MODEL         | CORRELATION COEFFICIENT | MAE    | RMSE   | RELATIVE ABS ERROR | RELATIVE SQUARED ERROR |
|---------------|-------------------------|--------|--------|--------------------|------------------------|
| KNN (IBk)     | 0.646                   | 0.1429 | 0.1883 | 32.45%             | 70.43%                 |
| SVM (SMOReg)  | 0.4862                  | 0.1761 | 0.2361 | 66.80%             | 82.32%                 |
| Random Forest | 0.646                   | 0.1331 | 0.1925 | 30.27%             | 70.82%                 |

MAE: Mean Absolute Error – RMSE: Root Mean Squared Error



**Fig. 12.** Comparison of Models, MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) and Correlation Coefficient

MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) and Correlation Coefficient (Fig. 12).

As is clear from the comparison chart, KNN and Random Forest models with high correlation values and lower errors performed better in predicting civilization classes than SVM model. Therefore, these two algorithms are recommended as suitable options for analyzing ceramic composite data in this study.

**CONCLUSION**

This study demonstrated the value of integrating analytical techniques (XRF, FTIR, and XRD) with fuzzy logic and machine learning models for the provenance assessment of ancient ceramic samples from Bolu Claudiopolis. The compositional data, primarily based on major oxides, showed a strong predominance of Roman affiliation, with 19 of the 20 samples displaying the highest membership values within Roman profiles and only one sample (B20) corresponding more closely to Greek characteristics. The extension of the fuzzy model to seven major oxides enhanced the flexibility and reliability of the classification process. Although minor discrepancies were observed between XRF, FTIR, and XRD-based predictions—reflecting differences between bulk, surface, and crystalline sensitivities—the integrated evaluation provided consistent support for a predominantly Roman origin. The combination of majority voting and probability aggregation further strengthened the interpretative framework. Multivariate analysis (PCA

and clustering) confirmed the compositional grouping of samples, while machine learning models implemented in Weka—particularly SVM and KNN—achieved high classification accuracy, supporting the robustness of the proposed methodology. Despite moderate agreement between fuzzy and machine learning outputs, lower concordance between FTIR/XRD and XRF-based predictions highlights the importance of multi-technique integration. Overall, the results indicate that the combined use of chemical, mineralogical, and computational approaches offers a reliable and transparent framework for cultural attribution of archaeological ceramics. This integrative methodology may contribute to more refined provenance studies in future archaeometry research.

**ACKNOWLEDGEMENTS**

This project was supported by the Scientific Research Projects Coordination Unit of Ondokuz Mayıs University with the number PYO.FEN.1904.23.007. We would like to express our special gratitude to the Scientific Research Projects Coordination Unit of Ondokuz Mayıs University for their financial support of this research. We would also like to express our gratitude to the Bolu Archaeological Museum, which played a valuable role in the implementation of this research by providing the samples obtained from the excavation and providing the necessary information. We are indebted to the technical efforts of the research team members and the sincere efforts of all the colleagues who collaborated in different stages of this study.

## REFERENCES

- ARASHEED/QASIM/BARLA 2023  
Rasheed, N. A./Qasim, O. M./Barla, A. K., Selecting the Important Features to Classify the Archaeological Fragments by Using Statistical Tools, *Indian Journal of Computer Science and Engineering* 14/3, 472–9. DOI: [10.21817/indjcs/2023/v14i231403090/3](https://doi.org/10.21817/indjcs/2023/v14i231403090/3).
- BAYAZIT/KAYNAK/COŞKUN 2023  
Bayazit, M./Kaynak, E./Coşkun, N., Chemical And Mineralogical Analyses Of The Late Neolithic Ceramics From Şah Valley (Singuber), Turkey, *Journal Of Scientific Reports-A* 052, 327–351. DOI: [10.59313/jsr-a.1206576](https://doi.org/10.59313/jsr-a.1206576).
- BICKLER 2018  
Bickler, S. H., Machine Learning Identification And Classification Of Historic Ceramics. *Archaeology* 20, 20-32.
- CECCARELLI *et alii* 2028  
Ceccarelli, L./Bellotto, M./Pietro, Caruso, M./ Cristiani, C./Dotelli, G./Gallo Stampino, P./Gasti, G./Primavesi, L., Characterization of clays and the technology of Roman ceramics production, *Clay Minerals* 53/3, 413–429.
- CHEN/CHEN 2024  
Chen, W./Chen, D., Research on the classification of ancient silicate glass artifacts based on machine learning, *Archaeometry*.
- DI ANGELO/DI STEFANO/PANE 2017  
Di Angelo, L./Di Stefano, P./Pane, C., Automatic dimensional characterisation of pottery, *Journal of Cultural Heritage* 26, 118–28.
- ERAMO *et alii* 2004  
Eramo, G./Laviano, R./Muntoni, I. M./Volpe, G., Late Roman cooking pottery from the Tavoliere area (Southern Italy): Raw materials and technological aspects, *Journal of Cultural Heritage* 5/2, 157–65. DOI: [10.1016/j.culher.2003.05.002](https://doi.org/10.1016/j.culher.2003.05.002).
- GIANNOSSA/FORLEO/MANGONE 2021  
Giannossa, L. C./Forleo, T./Mangone, A., The distinctive role of chemical composition in archaeometry. the case of apulian red figure pottery, *Applied Sciences (Switzerland)*, MDPI AG. DOI: [10.3390/app11073073](https://doi.org/10.3390/app11073073).
- IKEOKA *et alii* 2018  
Ikeoka, R. A./Appoloni, C. R./Rizzutto, M. A./Bandeira, A. M., Computed Radiography, PIXE and XRF analysis of pre-colonial pottery from Maranhão, Brazil, *Microchemical Journal* 138, 384–9. DOI: [10.1016/j.microc.2017.12.020](https://doi.org/10.1016/j.microc.2017.12.020).
- JIA/CHEN 2025  
Jia, X./Chen, Y., Machine Learning-Based Comprehensive Analysis of Phase, Composition, and Microstructure of Yuan Dynasty Oil-Glazed Porcelain, *Iranian Journal of Chemistry and Chemical Engineering* 44/2, 583–99.
- KAHVECİ/PEKŞEN 2023  
Kahveci, G./Pekşen, O. Glass Production and Use in Ancient Mesopotamia, Egypt and Anatolia, *Erzurum Teknik Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 17, 91-108. DOI: [10.29157/etusbed.1291241](https://doi.org/10.29157/etusbed.1291241).
- KAŁUŻNA-CZAPLIŃSKA *et alii* 2017  
Kałużna-Czaplińska, J./Rosiak, A./Grams, J./Chałupka, K./Makarowicz, P./Maniukiewicz, W./Szubiakiewicz, E., The studies of archaeological pottery with the use of selected analytical techniques, *Critical Reviews in Analytical Chemistry* 47/6, 490-498. DOI: [10.1080/10408347.2017.1334534](https://doi.org/10.1080/10408347.2017.1334534)
- KEAWSAWASVONG/LAI 2021  
Keawsawavong, S./Lai, V. Q., End bearing capacity factor for annular foundations embedded in clay considering the effect of the adhesion factor, *International Journal of Geosynthetics and Ground Engineering* 7/1, 15. DOI: [10.1007/s40891-021-00261-2](https://doi.org/10.1007/s40891-021-00261-2).
- LING/DELNEVO/MIRRI 2023  
Ling, Z., Delnevo, G./Mirri, S., Finding on Machine Learning Approaches for Classification of Ancient Ceramics-A Systematic Literature Review. DOI: [10.21203/rs.3.rs-3461328/v1](https://doi.org/10.21203/rs.3.rs-3461328/v1).
- LING *et alii* 2024  
Ling, Z./Delnevo, G./Salomoni, P./Mirri, S., Findings on machine learning for identification of archaeological ceramics: A systematic literature review, *IEEE Access* 12, 100167-100185.
- LIU *et alii* 2023  
Liu, C./Wang, W./Wang, H./Zhu, C./Ren, B., A Review on Removal of Iron Impurities from Quartz Mineral, *Minerals* 13/9, 1128. DOI: [10.3390/min13091128](https://doi.org/10.3390/min13091128).
- LIU/TIAN/CHEN 2024  
Liu, Y./Tian, Y./Chen, K., Archaeometric study of the iron objects from the Xuechi sacrificial site and its implication for bloomery iron smelting during early Western Han period in China, *Archaeometry* 66/5, 1050-1062. DOI: [10.1111/arcm.12952](https://doi.org/10.1111/arcm.12952).
- LONČARIĆ/COSTA 2023  
Lončarić, V./Costa, M., Known glass compositions in Iron Age Europe—Current synthesis and emerging questions, *Heritage* 6/5, 3835-3863. DOI: [10.3390/heritage6050204](https://doi.org/10.3390/heritage6050204).
- MOHAMMED/ALMASHHADANI 2023  
Mohammed, R./Almashhadani, H., Synthesis, characterization and thermodynamic study of a polymer nanocomposite from methyl acrylate and zirconium chloride as an anti-corrosion coating for carbon steel. *International Journal of Corrosion and Scale Inhibition* 12, 1180-1191. DOI: [10.17675/2305-6894-2023-12-3-21](https://doi.org/10.17675/2305-6894-2023-12-3-21).
- MU *et alii* 2019  
Mu, T./Wang, F./Wang, X./Luo, H., Research on ancient ceramic identification by artificial intelligence, *Ceramics International* 45/14, 18140-18146. DOI: [10.1016/j.ceramint.2019.06.003](https://doi.org/10.1016/j.ceramint.2019.06.003).
- NAKAMURA *et alii* 2016  
Nakamura, S./Ota, K./Shibuya, Y./Noguchi, T., Role of a water network around the Mn<sub>4</sub>CaO<sub>5</sub> cluster in photosynthetic water oxidation: a Fourier transform infrared spectroscopy and quantum mechanics/molecular mechanics calculation study, *Biochemistry* 55/3, 597-607.
- NAVARRO *et alii* 2021  
Navarro, P./Cintas, C./Lucena, M./Fuertes, J. M./Delrieux, C./Molinos, M., Learning feature representation of Iberian ceramics with automatic classification models, *Journal of Cultural Heritage* 48, 65-73. DOI: [10.1016/j.culher.2021.01.003](https://doi.org/10.1016/j.culher.2021.01.003).
- PAN *et alii* 2011  
Pan, Y. X./Mei, D./Liu, C. J./Ge, Q., Hydrogen adsorption on Ga<sub>2</sub>O<sub>3</sub> surface: a combined experimental and computational study, *The Journal of Physical Chemistry C* 115/20, 10140-10146. DOI: [10.1021/jp2014226](https://doi.org/10.1021/jp2014226).
- PAPACHRISTODOULOU *et alii* 2006  
Papachristodoulou, C./Ioannidou, A./Ioannides, K./Gravani, K., A study of ancient pottery by means of X-ray fluorescence spectroscopy, multivariate statistics and

- mineralogical analysis, *Analytica Chimica Acta* 573, 347-353. DOI: [10.1016/j.aca.2006.02.012](https://doi.org/10.1016/j.aca.2006.02.012).
- PEACOCK 1970  
Peacock, D. P., The scientific analysis of ancient ceramics: a review, *World Archaeology* 1/3, 375-389. DOI: [10.1080/00438243.1970.9979454](https://doi.org/10.1080/00438243.1970.9979454).
- QI *et alii* 2022  
Qi, Y./Qiu, M. Z./Jing, H. Z./Wang, Z. Q./Yu, C. L./Zhu, J. F./Wang, T., End-to-end ancient ceramic classification toolkit based on deep learning: A case study of black glazed wares of Jian kilns (Song Dynasty, Fujian province), *Ceramics International* 48/23, 34516-34532. DOI: [10.1016/j.ceramint.2022.08.033](https://doi.org/10.1016/j.ceramint.2022.08.033).
- RASDI *et alii* 2017  
Rasdi, N. M./Fen, Y. W./Omar, N. A. S./Azis, R. A. S./Zaid, M. H. M., Effects of cobalt doping on structural, morphological, and optical properties of Zn<sub>2</sub>SiO<sub>4</sub> nanophosphors prepared by sol-gel method, *Results in physics* 7, 3820-3825. DOI: [10.1016/j.rinp.2017.09.057](https://doi.org/10.1016/j.rinp.2017.09.057)
- RUSCHIONI *et alii* 2023  
Ruschioni, G./Malchiodi, D./Zanaboni, A. M./Bonizzoni, L., Supervised learning algorithms as a tool for archaeology: Classification of ceramic samples described by chemical element concentrations, *Journal of Archaeological Science: Reports* 49, 103995. DOI: [10.1016/j.jasrep.2023.103995](https://doi.org/10.1016/j.jasrep.2023.103995)
- SALISU/ALMAJIR 2020  
Salisu, B. D./Almajir, I. R., Aflatoxins and aflatoxigenic fungal contamination of common poultry feed products in Katsina State, Nigeria, *Novel Research in Microbiology Journal* 4/1, 653-665.
- SPARAVIGNA 2014  
Sparavigna, A. C., Ancient Technologies: The Egyptian Sintered Quartz Ceramics, *PHILICA.COM* Article number 426.
- SUN *et alii* 2020  
Sun, H./Liu, M./Li, L./Yan, L./Zhou, Y./Feng, X., A new classification method of ancient Chinese ceramics based on machine learning and component analysis, *Ceramics International* 46/6, 8104-8110. DOI: [10.1016/j.ceramint.2019.12.037](https://doi.org/10.1016/j.ceramint.2019.12.037).
- VOLL *et alii* 2001  
Voll, D./Lengauer, C./Beran, A./Schneider, H., Infrared band assignment and structural refinement of Al-Si, Al-Ge, and Ga-Ge mullites, *European Journal of Mineralogy* 13/3, 591-604. DOI: [10.1127/0935-1221/2001/0013-0591](https://doi.org/10.1127/0935-1221/2001/0013-0591).
- WU *et alii* 2020  
Wu, Z./Liu, W./Zheng, J./Li, Y., Effect of methane hydrate dissociation and reformation on the permeability of clayey sediments, *Applied energy* 261, 114479. DOI: [10.1016/j.apenergy.2019.114479](https://doi.org/10.1016/j.apenergy.2019.114479).
- YİĞİTPAŞA 2023  
Yiğitpaşa, D., Re-Assessment of the Claudiopolis stadion rescue excavation in 2008, *Journal of Ancient History and Archaeology* 10/4, 64-85. DOI: [10.14795/j.v10i4.969](https://doi.org/10.14795/j.v10i4.969).
- YİĞİTPAŞA 2024  
Yiğitpaşa, D., New archaeological data from Claudiopolis in the light of stadion excavations, *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi* 14/4, 1321-1339. DOI: [10.48146/odusobiad.1313024](https://doi.org/10.48146/odusobiad.1313024).
- ZAMPIERIN *et alii* 2024  
Zampierin, D./Moita, P./Lischi, S./van Aerde, M./Barrulas, P./Mirao, J., A multi-analytical approach applied to pottery from Oman as a key to understanding ancient Indian Ocean maritime trade, *Archaeometry* 66/5, 967-1015. DOI: [10.1111/arcm.12949](https://doi.org/10.1111/arcm.12949).
- ZANIER *et alii* 2024  
Zampierin, D./Moita, P./Lischi, S./van Aerde, M./Barrulas, P./Mirao, J., A multi-analytical approach applied to pottery from Oman as a key to understanding ancient Indian Ocean maritime trade, *Archaeometry* 66/5, 967-1015. DOI: [10.1111/arcm.12949](https://doi.org/10.1111/arcm.12949).