

<https://doi.org/10.1038/s41538-025-00665-8>

# Residues on unglazed Roman pottery: extraction of organic spectral signatures via ATR-FTIR and machine learning

Check for updates

S. Giampaolo<sup>1</sup>, C. Scatigno<sup>1</sup> ✉, L. Migliorati<sup>2</sup>, T. Sgrulloni<sup>3</sup> & G. Festa<sup>1</sup>

The study of organic residues on archaeological objects offers invaluable insights into past human behaviours, including diet, trade, storage, and ritual practices. Here, we present an innovative analytical approach that combines attenuated total reflectance Fourier transform infra-red spectroscopy (ATR-FTIR) with machine learning techniques to extract spectral signatures for identifying organic residues within complex inorganic matrix, primarily depositional materials in archaeological contexts. The methodology was applied to the study of six unglazed ceramic artefacts coming from a sealed context of a *taberna* looking onto the *Forum* area of the archaeological site of *Peltuinum* Roman city, central Italy. The probable commercial nature of the *taberna* and the porous, unglazed pottery supported the high probability of organic residue preservation. The analysis successfully identified lipidic residues in two objects: a few African type D sealed earth plate fragments with black staining and a few common pottery olla fragments. These residues were notable for the absence of proteins and polysaccharides. Infra-red spectroscopy together with machine learning extracts organic spectral features at 2921, 2853, 1737  $\text{cm}^{-1}$ , indicative of resins, suggesting its use in food preparation or ritual activity. Furthermore, additional markers at 2951, 2918, 2852, 1701, 1558, 1462, 1247, and 1160  $\text{cm}^{-1}$  were attributed to a degraded oil, likely associated with ancient culinary or ritual practices. The findings underscore the value of a non-destructive methodology, enhanced by machine learning, for the detection of organic residues in complex archaeological contexts, especially where molecular signals are masked by interference from archaeological soils. This advanced approach provides a reproducible and chemical-free strategy for the detection of organic spectral features in ancient pottery, significantly enhancing the reliability and efficiency of archaeological organic residue analysis.

Pottery is one of the most significant classes in archaeological findings, offering crucial insights into daily practices, social organisation of past populations<sup>1–3</sup>. As containers, they stored traces of their use, trapping residues of content materials which underwent to degradation process. The analysis of residue traces allows for the identification of organic compounds that may serve as markers of past cooking or storage activities, thus helping to distinguish between functional and decorative vessels. Specifically, vessels may retain organic matter in three distinct forms: (1) as loose or semi-aged residues of the original contents that remain inside the vessel; (2) as visible surface deposits or encrustations on the interior and exterior surfaces; and (3) as organic compounds that have been absorbed into the ceramic matrix

through prolonged contact with foodstuffs or liquids, and are not detectable through visual inspection<sup>4</sup>. Therefore, the analytical investigation of residues on vessels used for storage, cooking, and related functions plays a key role in reconstructing ancient dietary practices, ritual behaviours, and trade systems<sup>5,6</sup>.

Organic residue analysis (ORA) encompasses a range of analytical techniques aimed at the identification and characterisation of archaeological organic spectral features that serve as indicators for tracing back the former presence of high-molecular-weight compounds such as lipids, proteins, carbohydrates<sup>7</sup>. Lipids, in particular, have been a focus of sustained attention because their relative chemical stability and hydrophobic nature often

<sup>1</sup>CREF—Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi, Rome, Italy. <sup>2</sup>Sapienza, University of Rome, Rome, Italy. <sup>3</sup>PhD Independent Researcher, Rome, Italy. ✉e-mail: [claudia.scatigno@cref.it](mailto:claudia.scatigno@cref.it)

contribute to their excellent preservation in archaeological contexts<sup>8</sup>. However, recent research has shown that proteins can also survive over extended timescales under favourable conditions. Fats and oils were widely used in antiquity, both in food preparation and in the production of unguents, often for ritual or ceremonial purposes.

Lipids undergo physicochemical alterations that may compromise the preservation of their original molecular signatures. Notably, the formation of calcium fatty acid salts, promoted by calcium carbonate, and common degradation processes such as oxidative cleavage of aliphatic double bonds and ester hydrolysis, contribute to the generation of free fatty acids (FFAs)<sup>9–13</sup>. Key lipid organic spectral features include FFAs, monoacylglycerols (MAGs), triacylglycerols (TAGs), diacylglycerols (DAGs), and sterols<sup>9–13</sup>. Therefore, pinpointing specific organic molecules is crucial for reconstructing the foods and materials stored or processed in the containers.

The detection and characterisation of these materials is commonly carried out using analytical techniques including gas chromatography-mass spectrometry (GC-MS), nuclear magnetic resonance (NMR) and Fourier-Transform Infra-red Spectroscopy (FT-IR)<sup>14–18</sup>. Recent studies demonstrate that machine learning algorithms effectively uncover hidden patterns in complex, large-scale datasets through both supervised and unsupervised approaches<sup>19–24</sup>.

In this study, unsupervised machine learning together with Fourier transform infra-red spectroscopy (FTIR) was employed to detect traces of organic residues on a set of unglazed pottery fragments excavated from the Roman *forum* of *Peltuinum* (AQ, Italy)<sup>25</sup>, a non-permanent ancient settlement in the Central Apennines (province of L'Aquila). *Peltuinum* was founded in the mid-1st c. BC on a plateau, which, since pre-Roman times, was a resting place for sheep moving from Central Italy to Northern Puglia. The abandonment of *Peltuinum* is due to a sequence of strong earthquakes that hit the region from the mid-4th c. to the mid-5th c. AD. The final seismic event led to the spoiling of all structures and to the reusing of the area for agricultural purposes. The Roman city had an orthogonal layout with the major public buildings located in the most suitable area. The main temple faced one of the short sides of the rectangular square, the *forum*; on both long sides were public buildings with different functions. On the eastern side, a row of rooms of equal size housed *tabernae*, offices, workshops, etc. Despite agricultural works, small parts of the rooms remained undamaged. The excavations document many interesting situations. An intact collapse context inside a probable commercial place (room 2) offered the opportunity to verify the contents of some containers that remained crushed on the floor under the roll fall.

Given the heterogeneity of the spectroscopic dataset and the potential masking of organic residues by signals from archaeological soils, the identification of spectral bands is achieved through a computational extraction method, thereby avoiding solvent-based chemical procedures. This study aims to provide a preliminary investigation into the application of spectroscopic and chemometric techniques for the non-destructive analysis of archaeological residues. Figure 1 shows the archaeological site of *Peltuinum*.

## Results

The methodological details are provided in the “Materials and methods” section.

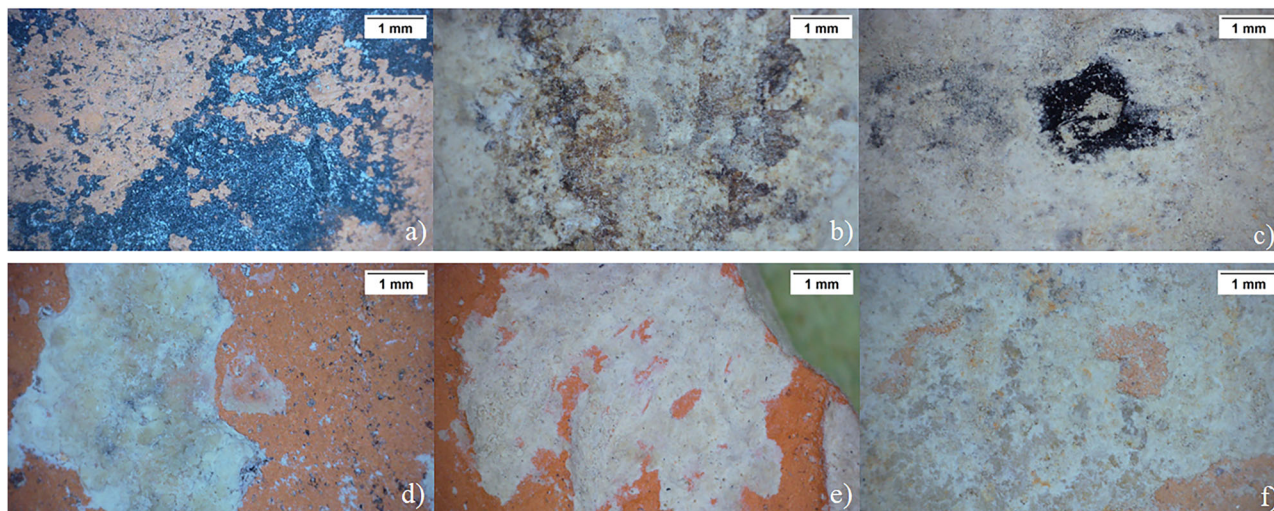
Microscopic pre-screening analysis revealed a variety of traces distinguished by differing morphologies and textures. Owing to the presence of concretions and residues on both the interior and exterior surfaces of the fragments, multiple internal and external areas were examined. Figure 2 shows six microscopic images corresponding to six areas of interest. While visible spots in black and white-yellow hues are evident, the heterogeneity in texture and morphology prevents the identification of consistent similarities among these spots across the examined areas. Figure 2a, c both show black spots; however, Fig. 2a exhibits a diffuse, smooth and flat appearance, suggesting the absorption of materials by the bulk of the object. In contrast, Fig. 2c shows a dense, thick formation resembling a surface crust. Similar distinctions are observed in the white-yellow staining. Figure 2d, e are representative of white-yellow crusts found on different portions of the same object, notably in Fig. 2d, where the crust exhibited a waxy consistency. In comparison, Fig. 2b, f reveals heterogeneous spots, consisting of both white deposits and darker areas in various shades of black, with a diffuse and layered structure.

Preliminary microscopic screening revealed the heterogeneous and multilayered nature of the observed traces, whose variability in morphology and texture does not allow any reliable classification based on images alone. Although two dominant colours, white-yellow and black, were observed, no consistent correlation with specific morphological features could be established. To overcome this limitation and investigate their molecular composition, ATR-FTIR spectroscopy was employed. Figure 3 shows the FTIR absorbance spectra acquired from the analysed residues, where three distinct ranges can be identified (Fig. 3a, b), corresponding to the presence of organic materials, carbonates and silicates. Analysis of the spectral data from all 72 measurement points (MPs) reveals that the chemical heterogeneity of the residues is reflected in the spectral features, which arise from the coexistence of mixed materials with overlapping vibrational signals. Furthermore, the slow degradation of the original materials over time has resulted in the formation of secondary compounds and the alteration of functional groups. A detailed examination of three representative spectra (Fig. 3b) reveals significant differences in the functional group and fingerprint regions. The spectral region between 3000 and 2000  $\text{cm}^{-1}$  is attributed to organic materials, with variations in shape and intensity relating to the vibrational bands of OH, CH<sub>2</sub> and CH<sub>3</sub><sup>26–28</sup>. It is interesting to note that at least one spectrum, R12, does not display these vibrations. In the fingerprint region, vibrations associated with silicates (1100–900  $\text{cm}^{-1}$ )<sup>29,30</sup> and carbonates (1500–1200  $\text{cm}^{-1}$ )<sup>31,32</sup> are clearly identifiable. However, the 1800–900  $\text{cm}^{-1}$  range shows significant variation in at least one sample, R8. This sample exhibits a markedly different texture and morphology under microscopic examination. Furthermore, compound assignment within this range is difficult due to overlapping spectral contributions.

To identify individual spectral regions without interference or peak overlap and, more importantly, to determine which compounds are specific



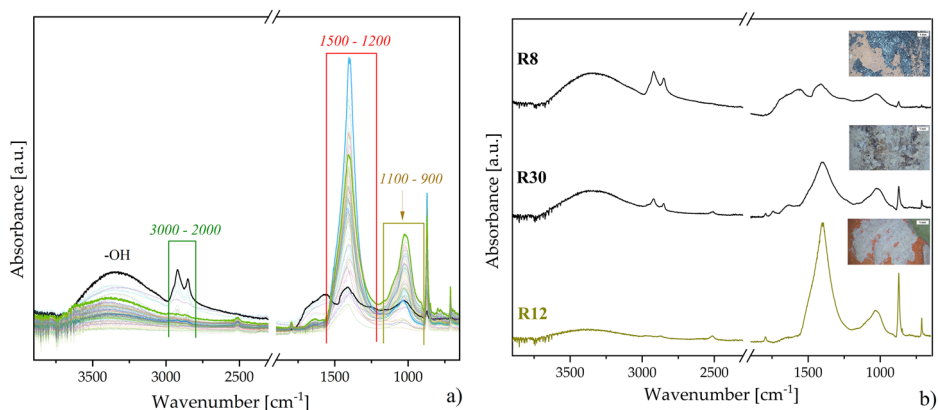
**Fig. 1 | Archaeological site of *Peltuinum*.** **a** Aerial view of *Peltuinum*. In the foreground, the main public area (oblique aerial photograph from east). **b** Forum schematic drawing, north top. Yellow: road network; asterisk: location of the analysed artefacts. **c** Pottery crushed under the collapsed roof.



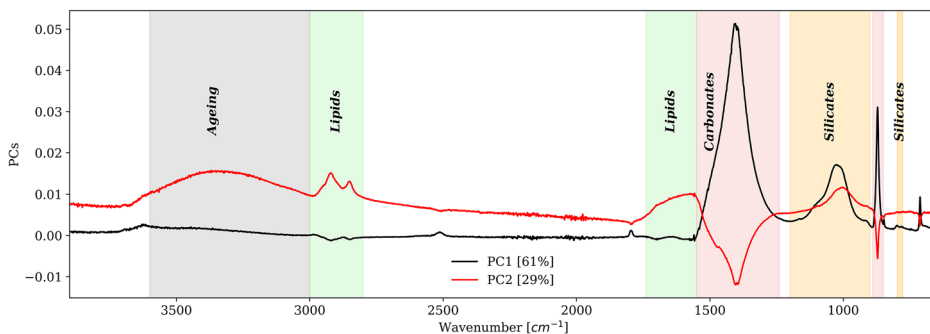
**Fig. 2 | Microscopic images of selected areas.** **a** Black staining R8, inner surface of a plate—PE 22 S.97\_US 1773\_3fr (B); **b** co-presence of black and white–yellow spots R30, outer part of the olla fragment—PE 22 S.97\_US 1773\_10fr (H); **c** black spots R27, inner part of the olla fragment—PE 22 S.97\_US 1773\_10fr (H); **d** yellow–white spots with waxy appearance R18, outer part of the plate fragment—PE 22 S.97\_US 1773\_3fr (C); **e** white–yellow crusts R12, external area of the plate fragment—PE 22 S.97\_US 1773\_3fr (B); and **f** co-presence of both colours R24, outer part of a different fragment of the olla—PE 22 S.97\_US 1773\_10fr (F).

**Fig. 3 | ATR-FTIR of collected spectra.**

**a** Vibrational bands of 72 collected samples, wavenumbers are on X-axis, ranging from 3900  $\text{cm}^{-1}$  to 650  $\text{cm}^{-1}$ . Spectra in bold emphasise recurring absorption bands. **b** FTIR spectra of three points, R8, R30, R12, selected for their different spectral profile and observed microscopic variations.



**Fig. 4 | Loading plot PC1 and PC2.** The different colour line of the absorption bands is due to the PC they belong to, black for PC1 [61%] and red for PC2 [30%]. The figure highlights both specific bands and spectral ranges of interest. Green shaded zones are ascribed to organic contributions of lipids; red shaded zones are the absorptions of carbonates, while yellow stripes are silicates. The grey-shaded region corresponds to the hydroxyl stretching band, which is indicative of ageing.



to each object or fragment, a principal component analysis (PCA) was computed. ATR-FTIR spectroscopic data were organised into a  $72 \times 13,486$  matrix as variables on the entire spectral range. PCA, consisting of three principal components (PCs), accounts for a total of 96% of the variance.

Figure 4 shows the loading plot for the first two PCs. Sixty-one percent of the data (PC1) indicates the presence of inorganic materials, specifically carbonates and silicates, which are associated with the archaeological context of silt-lime-rich substrates. It is also noteworthy that the loading highlights an overlap in the carbonate band, evidencing two bands at 1408 and 1394  $\text{cm}^{-1}$ , which are related to the asymmetric stretching mode of the

carbonate ion ( $\text{CO}_3^{2-}$ )<sup>31,32</sup>. The overlapping in the carbonate band is attributed to the coexistence of different polymorphs of calcium carbonate, such as dolomite and calcite<sup>31</sup>, and the presence of silicates<sup>33</sup>, consistent with the archaeological context.

Table 1 reports the relative wavenumber absorbances.

Lipid vibrations at 3000–2800  $\text{cm}^{-1}$  (Fig. 4 and Table 1) are negative, indicating their absence. Interestingly, the OH bending region is absent in the total variance of PC1, indicating that this range is not related to interstitial water from the surrounding soil. One-third of the data (PC2), on the other hand, shows characteristic lipid vibrations, as well as a silicate region

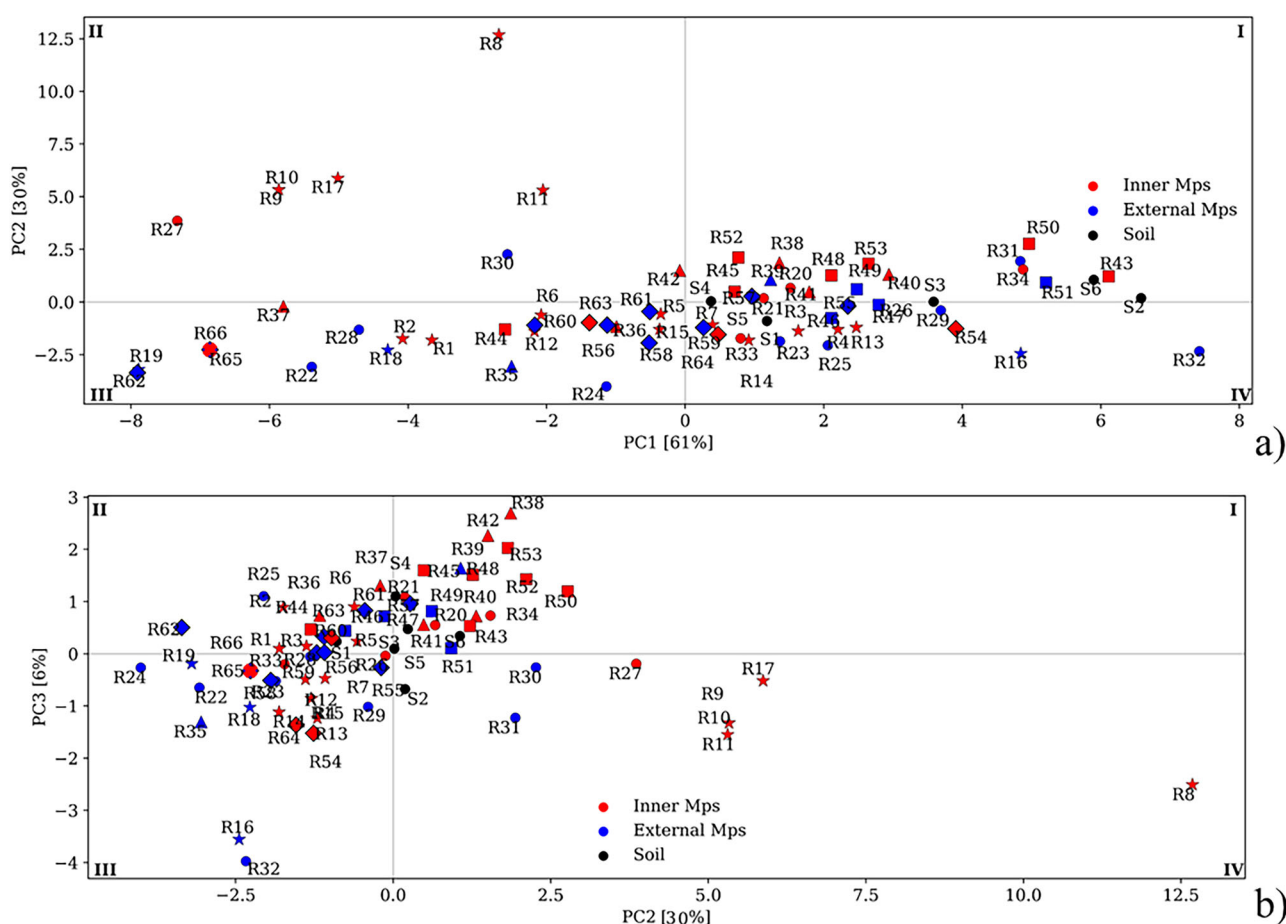
**Table 1 | Assignments of the IR absorption bands are evidenced by the loading plot of PC1 vs PC2, as illustrated in Fig. 4**

Wavenumber [cm <sup>-1</sup> ]	Functional group	Compound	Reference
3380-3330	νOH	Water/Lipids	26
2951	νCH <sub>3</sub>	Lipids	27
2920	νCH	Lipids	26-28
2852	νCH <sub>2</sub>	Lipids	26-28
2515	2νCO <sub>3</sub> <sup>2-</sup>	Carbonates	26,34
1795	2δCO <sub>3</sub> <sup>2-</sup>	Carbonates	26,34
1740	νC=O	Lipids	26,35
1701	νC=O	Lipids	27,36
1652	δOH	Water/lipids	34,51
1558	νCOO <sup>-</sup>	Lipids	27
1462	scissoring CH <sub>2</sub>	Lipids	51
1408-1394	νCO <sub>3</sub> <sup>2-</sup>	Carbonates	31,32
1247	δ out of plane CH <sub>2</sub>	Lipids	51
1160	νC-O	Lipids	28
1008-1040	νSiO	Silicates	29
778	SiO	Silicates	30
798	SiO	Silicates	30
876	δCO <sub>3</sub> <sup>2-</sup>	Carbonates	31
713	δCO <sub>3</sub> <sup>2-</sup>	Carbonates	31

associated with the bulk material of the pottery. Furthermore, the loading plot reveals the presence of different compounds, particularly at 1160 cm<sup>-1</sup>, 1701 cm<sup>-1</sup>, 1652 cm<sup>-1</sup>, 1740 cm<sup>-1</sup> (Table 1) and where triglyceride esters can be identified<sup>27,34,35</sup>, providing additional confirmation of the presence of organic materials. PC2 shows that the intensity of the OH stretching band at 3380-3330 cm<sup>-1</sup> is linked with the organic compounds. Indeed, the occurrence of the hydroxyl group can be attributed to both the O-H group in clay minerals<sup>26</sup> and to the presence of FFAs<sup>35,36</sup>. In fact, certain organic compounds related to lipids, such as oils, exhibit a broad band in the hydroxyl stretching range upon ageing<sup>11,37</sup>. This is due to the chemical reaction between the double bond of the aliphatic chain (C=C) and water, which may lead to the formation of by-products such as alcohol and FFAs through partial hydrolysis. In addition, absorption related to the carboxylate group of fatty acids is evidenced by the absorption bands at 1558 cm<sup>-1</sup>, confirming the presence of lipid-based organic materials in these MPs.

Signals attribution to organic resin-oil-based materials and spectral regions of interest were compared with modern resin and oil standards (see Supplementary Fig. 7 and Supplementary Table 3). The standard spectra are pure, undegraded materials, while the loading profiles along PC2 and PC3 describe the features of chemically altered lipidic residues<sup>26-28</sup>. Moreover, this comparison underscores the discriminative power of the technique, clearly showing how the spectral profiles of pure compounds differ from those of degraded archaeological materials.

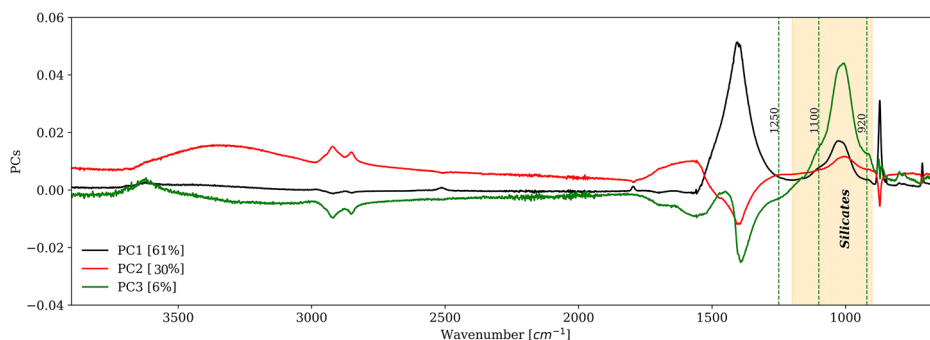
Figure 5 shows the score plot, where the MPs are distributed according to their multivariate profiles, reflecting similarities and differences associated with the variables highlighted in the corresponding loading plot previously discussed in Fig. 4. Approximately 60% of the MPs, including



**Fig. 5 | Score plot of all collected measurements.** The gathered data are presented with two different colours based on how the points have been sampled, more specifically, red dots stand for inner sampling and blue dots for external sampling. Soil measurements are signed with black dots. Soil samples collected from the

containment bags of the objects were labelled as “S” (for more details, please see Supplementary Materials). **a** Score plot of PC1 and PC2; **b** score plot of PC2 and PC3. To better highlight the attribution of individual MPs to specific ceramic objects, additional markers were used, see Supplementary Table 2.

**Fig. 6** | Loading plot of PC1, PC2 and PC3. The black curve is PC1, the red curve is PC2, and the green one is PC3, the absorption IR bands are variables of the loading plot.



both interior and exterior surfaces, show a common molecular fingerprint, suggesting a similar composition of the deposited material, likely resulting from long-term interaction with the original depositional environment. These points are distributed primarily along PC1, which corresponds to the mineralogical gradient of the soil, particularly in terms of carbonate content. Within this main group, a certain degree of dispersion is evident, likely due to micro-variations in the depositional context or degradation dynamics. However, two additional and distinct groups are detected: one still connected to the original depositional environment context but with a slightly modified profile, and another showing a divergent distribution dominated by PC2, indicative of a different compositional phase. These findings highlight the complexity of residue formation and the value of PCA in distinguishing between overlapping yet compositionally distinct signatures.

In the score plot (Fig. 5a), the MPs are mainly distributed along PC1, with low variance along PC2. The signal similarity along PC1, where both the MPs and soil samples cluster, is most likely attributable to the presence of common soil-related contaminants, such as carbonates and silicates. These are expected given the original depositional environment of the artefacts and reflect shared inorganic components rather than molecular homogeneity in organic residues (spectral differences described by PC2 and PC3). Therefore, PC1 suggests an overall spectral homogeneity due to a shared inorganic background. By observing the distribution of the MPs along PC1, the concentration of calcium carbonate increases from left to right, indicating a higher abundance of inorganic compounds associated with the geological matrix of the excavation site, as confirmed by the presence of soil, marked with the letter 'S' on the plot. Indeed, reference soil samples, S1, S2, S3, S4, S5, and S6, all of which fall within the region of high positive variance of the PC1 axis.

In contrast, PC2 is predominantly associated with lipid-related absorption bands, as shown by the loading plot shown in Fig. 4. Consequently, MPs projecting toward positive values along the PC2 axis are primarily characterised by the presence of organic compounds, likely due to lipid preservation. Along PC2, five MPs (R8, R9, R10, R11, and R17) exhibit notably distinct variance, forming a cluster in the second quadrant. These points were collected from different fragments of the same object, a plate (see Supplementary Fig. 1), in areas where black staining had been previously observed during microscopic pre-screening (Fig. 2a). R9 and R10 overlap in the plot, as samplings were carried out at the same location to investigate potential residue stratigraphy, which was ultimately not observed. Within this cluster, lipid-related features appear consistent across the MPs, except for R8, which shows significantly higher variance along PC2. FTIR spectral data are indicative of an organic residue, with the absence of protein-related bands (Amide I, II, and III)<sup>38</sup>, potentially pointing to a vegetal origin and a partial hydrolysis of triglycerides. Additional MPs, R1 and R2 (III quadrant), and R3 and R4 (IV quadrant), although sampled from adjacent areas within the interior surface of the same ceramic vessel, display differences in their spectral profiles. These variations are likely attributable to differing amounts of soil-related contaminants, which appear to increase progressively from left to right along PC1. Differences may also relate to the depositional environment exposure across the sampled areas of

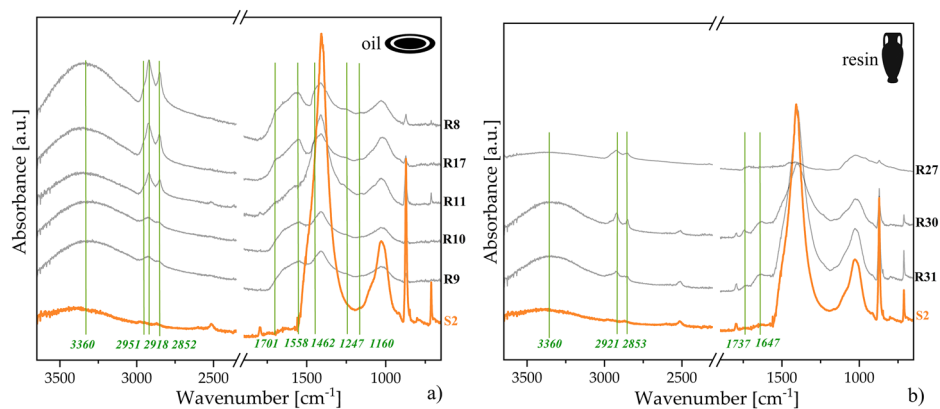
the same fragment. The clustering of R9, R10, and R17 further suggests compositional similarity of the organic residues, corroborated by the presence of black stains (see Supplementary Fig. 1). This supports the interpretation of residue homogeneity within the vessel, as R8, R9, R10, and R11, sampled from one fragment, share comparable spectral features with R17, obtained from a different fragment of the same plate. These groupings were consistently observed across replicates, thereby confirming the reliability and reproducibility of the analytical approach.

As a validation of the robustness of the model and its total cumulative variance of 97%, score plots (PC1 vs PC2 and PC2 vs PC3) with 95% confidence ellipses for each sample group were computed (see Supplementary Fig. 8).

Figure 5 shows PC2 and PC3 variance, where the influence of carbonates and silicates is reduced. Observing the residual variance (6%) is possible to observe another notable group that consists of three MPs, R29, R30, and R27, where organic traces are found both externally and internally. These MPs are collected from a common pottery olla fragment (see Supplementary Fig. 2). The FTIR spectra indicate the presence of specific absorbance bands at 2920, 2852, 1740  $\text{cm}^{-1}$  related to other vegetable compounds such as resins PC3 is associated with the shape of the silicate absorption band in the range 1008–1040  $\text{cm}^{-1}$  and with the doublet absorption bands of quartz at 778  $\text{cm}^{-1}$ , 798  $\text{cm}^{-1}$ . Moreover, in contrast to the PC1 and PC2 curves, the PC3 loading shows two shoulder bands in the silicate range, at 920  $\text{cm}^{-1}$  and 1100  $\text{cm}^{-1}$ . Silicates are indeed associated with the excavation soil (see Fig. 6). The shoulder bands can be attributed to the presence of clay minerals<sup>39,40</sup> and the additional presence of quartz underscores the dominant influence of the geological matrix also in the PC3 curve.

Figure 7 shows the spectra in which the PCA identifies organic traces in the MPs located on the plate PE 22 S.97\_US 1773\_3fr (Fig. 7a) and the MPs in the olla fragment H, PE 22 S.97\_US 1773\_10fr (Fig. 7b). Figure 7a shows that a vegetable resin is present on the internal and external surfaces of the olla, as indicated by the absorbance bands at 2921  $\text{cm}^{-1}$ , 2853  $\text{cm}^{-1}$  and 1737  $\text{cm}^{-1}$ . The MPs presented in Fig. 7b, R30, R31, and R27, are part of a broader group of MPs collected from the external surface of olla fragment H (see Supplementary Fig. 4). This group includes MPs R28, R29, and R32, which, despite being sampled in close spatial proximity to the others, did not exhibit evidence of organic residues in the PCA score plot. Spectral analysis of R28, R29, and R32 reveals that the absorbance profiles are dominated by intense signals associated only with silicate and carbonate phases. These findings highlight compositional differences through PCA even within sampling areas that are macroscopically uniform. Moreover, this outcome describes the spectral complexity of the dataset, in which inorganic matrix components can mask low-intensity bands indicative of organic residues. By contrast, Fig. 7b shows spectral features that are consistent with an oil-based compound found inside the plate, with characteristic absorbance bands at 2951  $\text{cm}^{-1}$ , 2918  $\text{cm}^{-1}$ , 2852  $\text{cm}^{-1}$ , 1701  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$ , 1462  $\text{cm}^{-1}$ , 1247  $\text{cm}^{-1}$  and 1160  $\text{cm}^{-1}$ <sup>134</sup>. These chemical differences highlight the distinct nature of the residues found in the two objects. Both figures exhibit hydroxyl-related bands at 3360  $\text{cm}^{-1}$  and 1647  $\text{cm}^{-1}$ , which are

**Fig. 7 | ATR-FTIR spectra of measurement points with organic traces.** **a** Spectra of MPs n. R8, R17, R11, R10, R9, collected in the internal surfaces of two fragments, B and C, of the plate PE 22 S.97\_US 1773\_3fr, absorption bands of lipid are highlighted by green lines. For comparative purposes, the spectrum of S2 MP, representative of the archaeological soil, is also shown. **b** Spectra of MPs: R30, R31, gathered from the external surface and R27 sampled on the internal surface of the fragment H of the olla PE 22 S.97\_US 1773\_10fr, absorption bands of lipid are highlighted by green lines. The spectrum of MP S2, corresponding to the archaeological soil, is included for comparison.



characteristic of degradation products such as alcohols or FFAs. Notably, the MPs identified as oil-based exhibit a larger number of organic-related absorbance bands, which also appear with significantly greater intensity than those observed in resinous residues. This combination contributes more substantially to the overall variance described by PC2 in the PCA model, resulting in the clustering of oil-associated points at higher PC2 values in the score plot.

## Discussion

This study demonstrates the efficacy of an integrated analytical approach combining ATR-FTIR spectroscopy with machine learning techniques to identify organic residues within complex inorganic archaeological matrices. The molecular analysis of pottery fragments revealed a heterogeneous composition dominated by silicates and carbonates, which makes the detection of organic spectral features challenging due to overlapping spectral features. Nonetheless, specific spectral signatures in the functional group region, supported by preliminary microscopic observations, indicate the preservation of organic compounds. The application of unsupervised methods, particularly PCA, facilitated the reduction of spectral complexity and the identification of key absorption bands characteristic of both inorganic and organic residues. Notably, residues containing FFAs associated with degradation processes were identified on several MPs, suggesting functional usage related to food storage, consumption, or ritual practices. Inorganic residues were characterised by absorption bands at  $2515\text{ cm}^{-1}$ ,  $1795\text{ cm}^{-1}$ ,  $1408\text{--}1394\text{ cm}^{-1}$ ,  $876\text{ cm}^{-1}$ , and  $713\text{ cm}^{-1}$ , associated with carbonates, and at  $1100\text{ cm}^{-1}$ ,  $1008\text{--}1040\text{ cm}^{-1}$ , and  $922\text{ cm}^{-1}$ , attributed to silicates. Distinctive spectral markers for organic residues were also identified at  $3380\text{--}3330\text{ cm}^{-1}$ ,  $2951\text{ cm}^{-1}$ ,  $2920\text{ cm}^{-1}$ ,  $2851\text{ cm}^{-1}$ ,  $1701\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ , and  $1160\text{ cm}^{-1}$ . These features may indicate the presence of lipids, while the absence of signals typically associated with polysaccharides and proteins could suggest that the organic residues are predominantly of plant origin, such as oils or resins. Furthermore, the hydroxyl stretching band within the  $3380\text{--}3330\text{ cm}^{-1}$  range serves as a marker for degraded material. Five MPs (R8, R9, R10, R11, and R17), located on the African type D sealed earth plate PE 22 S.97\_US 1773\_3fr were identified as black stains associated with degraded residues containing FFAs, which suggest ageing processes of organic compounds originally stored within the artefact. The spatial distribution and chemical nature of the residues suggest a functional use of the object, linked to food storage, consumption, or ritualistic activities. Organic residues were also detected in a second artefact, a common pottery olla (PE 22 S.97\_US 1773\_10fr) on R27, R9, and R30 MPs. Notably, residues were found on both the internal and external surfaces of the vessel, suggesting a functional use of the object. The presence of the same molecular signature in the external MP supports the hypothesis of an outward migration of the organic content, possibly resulting from culinary or ritual practices, or alternatively, the residues could be drips from pouring the liquid into a different container. These findings underscore the potential of

combining ATR-FTIR spectroscopy with machine learning as a non-destructive and effective strategy for reconstructing past human activities, offering valuable insights into the functional and cultural significance of ceramic artefacts within their archaeological context. As widely reported in the literature, the combined use of PCA and ATR-FTIR spectroscopy is an effective strategy for exploring similarities and compositional differences among materials. This approach allows for the identification of discrimination criteria based on spectral features, enabling unsupervised clustering without relying on predefined classification models<sup>41–43</sup>. Building on this established framework, the present study has applied this combined methodology to the individuation of functional objects in a dataset. The results of this study can be meaningfully compared to those from other Roman sites. Notably, our findings are consistent with widespread archaeological evidence indicating that amphorae were commonly employed across the Roman Empire for the transport and trade of consumable goods such as oil<sup>44,45</sup>. Moreover, the detection of resinous substances in the olla may be associated with the widespread practice of internal sealing or resin coating of amphorae, a technique well documented in Roman contexts, especially when the olla are used to trade wine<sup>46,47</sup>. This suggests that the vessel PE 22 S.97\_US 1773\_10fr might have been used for transporting liquid goods. These findings underscore the potential of integrating spectroscopic techniques with machine learning for the study of organic residues, laying the groundwork for future in-depth investigations. While the results demonstrate the potential of the non-destructive and new approach in providing insights into the material characteristics and functional use of the objects, it is important to acknowledge the limitations inherent in the absence of direct chemical validation. In this regard, complementary techniques such as GC-MS could offer more definitive molecular-level identification; however, their inherently destructive nature poses a significant constraint, particularly when working with rare or valuable archaeological materials.

## Methods

The six objects under examination were discovered in the Roman *forum* area of the Italian archaeological site of *Peltuinum* (AQ, Italy) and encompassed different pottery classes and shapes, including an African type D sealed earth plate decorated with a wheel, a common pottery olla, and amphorae decorated with engravings. The categorisation of the analysed objects was based on their archaeological denominations. The object set comprised six ceramic artefacts, including one fragmented plate (see Supplementary Fig. 1), four fragmented amphorae (see Supplementary Figs. 2, 4, and 5 and Supplementary Tables 1 and 2), one of which was a decorated amphora (see Supplementary Figs. 3 and Supplementary Tables 1 and 2), and one unique intact decorated amphora (see Supplementary Fig. 6 and Supplementary Tables 1 and 2).

The recovered objects were found within the same stratigraphic layer associated with the collapse of the roof of a tabula. From a conservation

perspective, the artefacts were effectively “sealed” by the overlying layer of collapsed roofing material. Differences in the morphology of the deposits were observed during both macroscopic and microscopic investigation of the samples. It was thus possible to distinguish between visible surface deposits (white and yellowish, of carbonate nature) and residues absorbed within the porous ceramic matrix. The latter were sampled primarily because they could be easily detached as compacted aggregates within the ceramic fabric. In contrast, those present as thin blackish films were scraped from the surface. Both types of residues yielded spectra exhibiting the same spectral features in the lipid-associated regions. No chemical pretreatment was applied to the samples prior to analysis, to preserve the original residue composition.

A total of 66 measurement points were collected from a total of 37 archaeological fragments belonging to all the 6 different objects to investigate the presence of inorganic and organic residues.

In addition, 6 samples of the archaeological soil were collected, one corresponding to each container bag of the objects.

### Optical microscopic observation

Microscopic inspection was executed with a Jiusion HD USB microscope with a maximum magnification of 1000× and a resolution of 1920 × 1082p.

Therefore, some areas were washed with distilled water to find possible layering beneath the calcareous-siliceous crusts.

### FTIR-ATR spectroscopic analysis

Collected samples were subsequently analysed with a Fourier Transform Infra-red (FTIR) spectrometer in attenuated total reflectance (ATR) mode. FTIR-ATR spectra were recorded with a Nicolet iS5 spectrometer (Thermo Fisher Scientific, Waltham, MA, USA), equipped with a DTGS detector and a KBr beam splitter. In ATR mode, the instrument was outfitted with an iD5 ATR accessory featuring a diamond laminate crystal. Measurements were performed in the 3900–650 cm<sup>-1</sup> range, with 128 scans and a resolution of 2 cm<sup>-1</sup>, resulting in a total of 72 FTIR spectra.

No normalisation procedures were applied to the entire dataset (mathematical correction, i.e. Kubelka–Munk transformations)<sup>48–50</sup>. Initially, all the spectra were visualised using the OMNIC and OriginPro® 2021 (OriginLab, 2023) software programs. The data were not pre-processed; normalisation and window selection were not applied. This allowed us to assess the minimal contribution of compounds in samples with different mixtures and identify benchmark absorption bands without the imposition of preconceived spectral region selections.

### Unsupervised machine learning analysis

PCA is an unsupervised statistical technique used to reduce the dimensionality of a dataset by transforming correlated variables into a new set of linearly uncorrelated variables, known as PCs. These components are linear combinations of the original features, ordered so that the first PC captures the maximum variance in the data, followed by the second PC, which captures the next highest variance, and so on.

In this study, mean-centred PCA was applied to experimental data acquired via ATR-FTIR spectroscopy, with the objective of identifying potential organic spectral features of organic residues and evaluating the contribution of soil dust to the chemical composition of the samples. The analysis was conducted on the full spectral range, without the application of energy window selection<sup>48–50</sup>. PCA results are typically visualised through score plots and loading plots. The score plot displays the projection of samples onto the new coordinate system defined by the PCs, highlighting inter-sample variance and grouping patterns. In contrast, the loading plot reveals how much each original variable contributes to each PC, effectively pinpointing which wavenumbers drive the variance observed among samples.

Since constant or systematic signals, such as baseline offsets, signals from the ceramic matrix, or consistent background soil contributions, exhibit low variance, they are naturally down-weighted or excluded from the PCA space. In this sense, PCA acts as a mathematical filter, emphasising

directions of maximum variance while ignoring invariant or low-variance features. Therefore, derivative preprocessing becomes unnecessary when the goal is to assess inter-sample differences in the presence of consistent background signals. PCA, through its loading structure, enables a mathematical separation of variance sources within the dataset. This makes it possible to isolate contributions from relatively fixed inorganic components (e.g. depositional environment) from more variable signals, such as those related to organic residues. Supporting the robustness of the PCA model, 95% confidence ellipses were included for each sample group (Inner MPs, External MPs, and Soil) to illustrate how the different groups are distributed within the multidimensional space defined by the PCs. These ellipses represent the multivariate spread of each group and allow a visual assessment of group separation, overlap, and internal variability. Specifically, the plots help highlight whether the sample types occupy distinct regions of the PCA space, suggesting differences in their underlying features, or if significant overlap occurs, implying similarity or shared characteristics.

The dataset, structured as a 72 × 13,486 matrix (samples × wavenumbers), was mean-centred by subtracting the mean of each variable. Since all variables shared the same scale, no further scaling was performed. PCA was implemented using singular value decomposition (SVD) via a custom script executed in Google Colab (version 6.5.1, Jupyter Notebook environment). The first three PCs were retained, explaining a cumulative 96% of the total variance (PC1: 61%, PC2: 30%, PC3: 6%).

### Data availability

Data are available upon request from the corresponding author.

Received: 1 August 2025; Accepted: 1 December 2025;

Published online: 17 December 2025

### References

- Sillar, B. & Tite, M. S. The challenge of ‘Technological choices’ for materials science approaches in archaeology. *Archaeometry* **42**, 2–20 (2020).
- Barker, D. & Majewski, T. Ceramic studies in historical archaeology. In: *The Cambridge Companion to Historical Archaeology* 205–232 (Cambridge University, 2006).
- Pikirayi, I. Ceramics and group identities: Towards a social archaeology in southern African Iron Age ceramic studies. *J. Soc. Archaeol.* **7**, 286–301 (2007).
- Rosiak, A., Kałużna-Czaplińska, J. & Gątarek, P. Analytical interpretation of organic residues from ceramics as a source of knowledge about our ancestors. *Crit. Rev. Anal. Chem.* **50**, 189–195 (2019).
- Germinario, G., De Luca, G., Messina, T., Scerra, S., De Benedetto, G. E. & Miliani, C. Exploring pottery function and cooking practices in Bronze Age Sicily: the results of high-resolution GC-MS of organic residues. *Environ. Archaeol.* <https://doi.org/10.1080/14614103.2024.2379639> (2024).
- Pecci, A., Reynolds, P., Mileto, S., Vargas Girón, J. M. & Bernal-Casasola, D. Production and transport of goods in the Roman period: residue analysis and wine derivatives in late republican Baetican ovoid amphorae. *Environ. Archaeol.* **30**, 104–116 (2021).
- Hamman, S. & Cramp, L. J. E. Towards the detection of dietary cereal processing through absorbed lipid biomarkers in archaeological pottery. *J. Archaeol. Sci.* **93**, 74–81 (2018).
- Evershed, R. P. Organic residue analysis in archaeology: the archaeological biomarker revolution. *Archaeometry* **50**, 895–924 (2008).
- Hamman, S., Scurr, D. J., Alexander, M. R. & Cramp, L. J. E. Mechanisms of lipid preservation in archaeological clay ceramics revealed by mass spectrometry imaging. *PNAS* **117**, 14688–14693 (2020).
- Cayme, J. M. C., Vahur, S., Teearu, A., Oras, E. & Leito, I. The impact of mineral composition on the yield and preservation of selected fatty acids in replicate archaeological ceramics. *J. Chem. Metrol.* **18**, 95–113 (2024).

11. Irto, A., Micalizzi, G., Bretti, C., Chiaia, V., Mondello, L. & Cardiano, P. Lipids in archaeological pottery: a review on their sampling and extraction techniques. *Molecules* **27**, 3451 (2022).
12. Spiteri, C., Craig, O. E. & Heron, C. Characteristic surviving residues from archaeological ceramics: a biomolecular approach. *Archaeological Ceramics: A Review of Current Research* (BAR International Series, 2011).
13. Mahesar, S. A., Sherazi, S. T. H., Khaskheli, A. R., Kandhero, A. A. & Uddin, S. Analytical approaches for the assessment of free fatty acids in oils and fats. *Anal. Methods* **6**, 4956 (2014).
14. Polla, S. & Springer, A. Organic residues analysis (ORA) in archaeology. In *Handbook of Cultural Heritage Analysis* 1075–1119 (Springer International Publishing, 2022).
15. Colombini, M. P., Modugno, F. Organic Materials in Art and Archaeology. John Wiley and Sons, Ltd: Chichester, 2009). 3–36.
16. Bonaduce, I., Ribechini, E., Modugno, F. & Colombini, M.P. Analytical approaches based on gas chromatography mass spectrometry (GC/MS) to study organic materials in artworks and archaeological objects. In: *Analytical Chemistry for Cultural Heritage. Topics in Current Chemistry Collections* (ed. Mazzeo, R.) (Springer, 2017).
17. Tanasi, D., Greco, E., Di Tullio, V., Capitani, D., Gulli, D. & Ciliberto, E. <sup>1</sup>H-<sup>1</sup>H NMR 2D-TOCSY, ATR FT-IR and SEM-EDX for the identification of organic residues on Sicilian prehistoric pottery. *Microchem. J.* **135**, 140–147 (2017).
18. Correa-Ascencio, M. & Evershed, R. P. High throughput screening of organic residues in archaeological potsherds using direct acidified methanol extraction. *Anal. Methods* **6**, 1330–1340 (2014).
19. Towarek, A., Halics, L., Matwin, S. & Wagner, B. Machine learning in analytical chemistry for cultural heritage: a comprehensive review. *J. Cult. Herit.* **70**, 64–70 (2024).
20. Anglisano, A., Casas, L., Queralt, I. & Di Febo, R. Supervised machine learning algorithms to predict provenance of archaeological pottery fragments. *Sustainability* **14**, 11214 (2022).
21. Ruan, F., Hou, L., Zhang, T. & Li, H. A modified backward elimination approach for the rapid classification of Chinese ceramics using laser-induced breakdown spectroscopy and chemometrics. *J. Anal. Spectrom.* **35**, 518–525 (2020).
22. Ruan, F., Hou, L., Zhang, T. & Li, H. A novel hybrid filter/wrapper method for feature selection in archaeological ceramics classification by laser-induced breakdown spectroscopy. *Analyst* **146**, 1023–1031 (2022).
23. Rosiak, A., Józefowska, A., Sekulska-Nalewajko, J., Goclawski, J. & Kałużna-Czaplińska, J. Funerary vs. domestic vessels from the Hallstatt period. A study on ceramic vases from the Milejowice settlement and the Domasław cemetery. *Sci. Rep.* **14**, 19942 (2024).
24. Fiorucci, M. et al. Machine learning for cultural heritage: a survey. *Pattern Recognit. Lett.* **133**, 102–108 (2020).
25. Migliorati, L. & Buonocore, M. Peltuinum: nuovi dati dal foro: Appendice II. Analisi fisico-chimiche di intonaci e affreschi. *Atti della Pontificia Accad. Romana di Archeologia. Rendiconti* **94**, 355–357 (2021).
26. Tarquini, G., Nunziante Cesaro, S. & Campanella, L. Identification of oil residues in Roman amphorae (Monte Testaccio, Rome): a comparative FTIR spectroscopic study of archaeological and artificially aged samples. *Talanta* **118**, 195–200 (2014).
27. Filopoulou, A., Vlachou, S. & Boyatzis, S. C. Fatty acids and their metal salts: a review of their infrared spectra in light of their presence in cultural heritage. *Molecules* **26**, 6005 (2021).
28. Lettieri, M., Giannotta, M.T., Investigations by FT-IR spectroscopy on residues in pottery cosmetic vases from archaeological sites in the Mediterranean Basin. *Int. J. Exp. Spectrosc. Tech.* **2**, 1–10 (2009).
29. Yan, B., Liu, S., Chastain, M. L., Yang, S. & Chen, J. A new FTIR method for estimating the firing temperature of ceramic bronze-casting moulds from early China. *Sci. Rep.* **11**, 3316 (2021).
30. Cantisani, E., Cavalieri, M., Lofrumento, C., Pecchioni, E. & Ricci, M. Ceramic findings from the archaeological site at Aiano-Torraccia di Chiusi (Siena, Italy): a multi-analytical approach. *Archaeol. Anthropol. Sci.* **4**, 29–46 (2012).
31. Kim, Y., Caumon, M. C., Barres, O., Sall, A. & Cauzid, J. Identification and composition of carbonate minerals of the calcite structure by Raman and infrared spectroscopies using portable devices. *Spectrochim. Acta, Part A.* **261**, 119980. <https://doi.org/10.1016/j.saa.2021.119980> (2021).
32. Gunasekaran, S., Anbalagan, G. & Pandi, S. Raman and infrared spectra of carbonates of calcite structure. *J. Raman Spectrosc.* **37**, 892–899 (2006).
33. Bálintová, M., Eštoková, A., Sicakova, A., Holub, M. & Singovszka, E. Analysis of building stone of the medieval historical building. *Adv. Mater. Res.* **897**, 305–308 (2014).
34. Lettieri, M. Infrared spectroscopic characterization of residues on archaeological pottery through different spectra acquisition modes. *Vib. Spectrosc.* **76**, 48–54 (2015).
35. Brambilla, L. et al. Characterization of fresh and aged natural ingredients used in historical ointments by molecular spectroscopic techniques: IR, Raman and fluorescence. *Anal. Bioanal. Chem.* **401**, 1827–1837 (2011).
36. Sacchi, R. et al. Olive oil from the 79 A.D. Vesuvius eruption stored at the Naples National Archaeological Museum (Italy). *npj Sci. Food* **4**, 19 (2020).
37. Regert, M., Bland, H. A., Dudd, S. N., Van Bergen, P. F. & Evershed, R. P. Free and bound fatty acid oxidation products in archaeological ceramic vessels. *Proc. R. Soc. Lond. B* **265**, 2027–2032 (1998).
38. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta Rev. Bioenerg.* **1797**, 1073–1101 (2007).
39. Madejov, J. & Komadel, P. Baseline studies of the clay minerals society source clay: infrared methods. *Clays Clay Miner.* **49**, 410–432 (2001).
40. Canevali, C. et al. A multi-analytical approach for the characterization of powders from the Pompeii archaeological site. *Anal. Bioanal. Chem.* **401**, 1801–1814 (2011).
41. Daher, C., Bellot-Gurlet, L., Le Ho, A. S., Paris, C. & Regert, M. Advanced discriminating criteria for natural organic substances of cultural heritage interest: spectral decomposition and multivariate analyses of FT-Raman and FT-IR signatures. *Talanta* **115**, 540–547 (2013).
42. De Palaminy, L., Daher, C. & Moulherat, C. Development of a non-destructive methodology using ATR-FTIR and chemometrics to discriminate wild silk species in heritage collections. *Spectrochim. Acta A* **270**, 120788 (2022).
43. Geminiani, L., Campione, F. P., Corti, C., Giussani, B., Gorla, G., Luraschi, M., Recchia, S. & Rampazzi, L. Historical silks: a novel method to evaluate their condition with ATR-FTIR spectroscopy and principal component analysis. *J. Cult. Herit.* **67**, 9–22 (2024).
44. Colombini, M. P., Giachib, G., Modugno, F. & Ribechini, E. Characterisation of organic residues in pottery vessels of the Roman age from Antinoe (Egypt). *Microchem. J.* **79**, 83–90 (2005).
45. Baniou, T. et al. Tracing culinary practices in the western provinces of the Roman empire using organic residue analysis. *Archaeol. Anthropol. Sci.* **16**, 103 (2024).
46. Tanasi, D., Greco, E., Pisciotta, F., Hassam S. Chemical characterization of organic residues on late Roman amphorae from shipwrecks off the coast of Marsala (Trapani, Italy). *J. Archaeol. Sci. Rep.* **40**, 103241 (2021).
47. Manhita, A., Martins, S., Gomes da Silva, M. et al. Transporting olive oil in Roman times: chromatographic analysis of dressel 20 amphorae from Pax Julia Civitas, Lusitania. *Chromatographia* **83**, 1055–1064 (2020).
48. Teodonio, L., Scatigno, C., Missori, M. & Festa, G. Late middle ages watermarked italian paper: a machine learning spatial-temporal approach. *J. Cult. Herit.* **57**, 53–59 (2022).
49. Scatigno, C. & Festa, G. FTIR coupled with machine learning to unveil spectroscopic benchmarks in the Italian EVOO. *Int. J. Food Sci. Technol.* **57**, 4156–4162 (2022).

50. Scatigno, C., Teodonio, L., Di Rocco, E. & Festa, G. Spectroscopic benchmarks by machine learning as discriminant analysis for unconventional Italian pictorialism photography. *Polymers* **16**, 1850 (2024).
51. Safar, M., Bertrand, D., Robert, P., Devaux, M. F. & Genot, C. Characterization of edible oils, butters and margarines by Fourier transform infrared spectroscopy with attenuated total reflectance. *J. Am. Oil Chem. Soc.* **71**, 371–377 (1994).

### Acknowledgements

This work was supported by the CREF—Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi, founded in the framework of the Physics for Cultural Heritage Laboratory.

### Author contributions

S.G.: measurements, data curation, formal analysis, and original draft preparation. C.S.: conceptualisation, methodology, investigation, measurements, software, validation, visualisation, script development, data curation, formal analysis, original draft preparation, and writing—reviewing and editing. L.M.: archaeological context, pottery materials. T.S.: archaeological context and pottery materials. G.F.: conceptualisation, methodology, investigation, project administration, supervision, measurements, writing—reviewing and editing.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41538-025-00665-8>.

**Correspondence** and requests for materials should be addressed to C. Scatigno.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025