Analytical and Bioanalytical Chemistry (2021) 413:6535–6550
https://doi.org/10.1007/s00216-021-03617-9

RESEARCH PAPER

Quantitative mineralogical analysis of clay-containing materials using ATR-FT-IR spectroscopy with PLS method

Signe Vahur[1] · Lisett Kiudorv[1] · Peeter Somelar[2] · Jan-Michael Cayme[1] · Mark Dennis Chico Retrato[1] · Rady Jazmin Remigio[1] · Varun Sharma[1] · Ester Oras[1,3] · Ivo Leito[1]

Received: 8 June 2021 / Revised: 21 July 2021 / Accepted: 10 August 2021 / Published online: 20 August 2021

Abstract

A universal method for quantitative analysis of clay components, in terms of mineral composition, using ATR-FT-IR spectroscopy (in the mid-IR and far-IR regions) combined with partial least squares (PLS) regression technique (ATR-FT-IR-PLS) is reported. For the PLS method development, altogether 222 samples covering natural clay sources and various archaeological/cultural heritage artefacts were used as calibration and validation standards. This is the largest calibration set used for creating an ATR-FT-IR-PLS method for clay minerals. The quantitative compositions of these standards containing combinations of altogether 29 minerals for the PLS method were determined using XRD analysis. The developed ATR-FT-IR with PLS method is quick and easy to use, and enables analysis of very small sample amounts (down to a few mg). This is very important when working with samples from archaeological and cultural heritage objects. The developed ATR-FT-IR-PLS method enables quantifying the contents of 7 main classes of minerals in different clays with a root mean square error of prediction (RMSEP) ranging from 0.9 to 5.1 g/100g. This means that in some cases, depending also on the content of the mineral in the sample, the accuracy is at a semiquantitative level. This quantitative method was successfully applied to 11 cultural heritage case-study samples.

Keywords Clays · ATR-FT-IR · PLS · XRD · Quantitative analysis · Minerals

Introduction

Clay-containing materials are widely used for different purposes — from daily utensils, building materials and artistic applications to the production of cosmetics and pharmaceuticals. Knowledge of the quantitative chemical composition of clay-based materials is valuable in many fields, such as cultural heritage and archaeology [1–7], geology [8–13], engineering [14], pharmaceutics [15] and cosmetics [14].

✉ Signe Vahur
signe.vahur@ut.ee

[1] Institute of Chemistry, University of Tartu, Ravila 14A, 50411 Tartu, Estonia

[2] Department of Geology, University of Tartu, Ravila 14A, 50411 Tartu, Estonia

[3] Department of Archaeology, University of Tartu, Jakobi 2, 51014 Tartu, Estonia

This study focuses on the quantitative investigation of clay-containing materials from archaeological, cultural heritage and art objects (such as pottery, technical ceramics, sculptures and construction materials). Analysis of these clay-based objects can provide valuable information about preparation technologies from the past, the geographical, including the geological, origin of raw materials, and quality of the clay, with broader applications for authenticity studies. In industrial settings, quantitative data can be used for estimating the quality and purity of objects.

Clays and clay minerals are widely distributed and can be found in soils, natural sediments and rocks [8, 16, 17]. Natural clay contains fine-grained (i.e. < 2.0 μm) clay minerals (kaolinite, illites, smectite, chlorite, etc.), including other finely divided non-clay compounds like quartz, feldspars and also metal oxides/hydroxides [8, 18].

Clay is one of the earliest raw materials used in manufacturing ceramics, pottery and building materials [19]. Ceramics are made by mixing clay with water and various temper additives, shaped to the desired form, dried and lastly fired at high temperatures (i.e. 900 to 1600 °C). The same

general process is also utilized for making pottery and bricks. The main difference is the type of clay material used and tempers added, which largely depend on manufacturing traditions and availability of different materials in the surrounding sediments, and the temperature reached during firing (usually below 1000 °C in historical samples). To improve the quality of ceramic vessels, tempers such as crushed rocks, stones, shells, different organic matter and grog (ground potsherds) are also added [4, 20–22]. Sand is combined with clay in some old brick preparations to reduce shrinkage. During firing, the clay materials go through significant mineralogical changes. The main mineral assemblages formed after firing depend on the original clay mineral type, the presence of other minerals, firing temperature and environment (i.e. reducing or oxidizing atmosphere) (see in the Supplementary information (SI) Fig. S1) [23–29].

The clay-containing materials or sediments are very complex mixtures, and investigation of their composition is challenging. For determining clay's quantitative mineralogical composition, X-ray diffraction (XRD) is the de facto standard method due to the minerals' regular and repeating crystal structure [30, 31]. Even though XRD is widely applied, it requires a rather large sample amount. Depending on sample nature, the minimum amounts for quantitative analysis, based on our experience, start from 0.5 to 1 g. It is possible to measure smaller sample quantities, but in both cases, the homogeneity and representativity of the sample must be ensured. This is problematic for archaeological and cultural heritage materials where only a minimal sample amount is usually available. Also, XRD has time-consuming sample preparation followed by interpretation of complex diffractograms with specific software that can be handled only by skilled specialists [32]. Furthermore, XRD cannot identify the amorphous phase, which is often present in natural samples as organic material or in ceramics as glass phase due to vitrification.

Because of these limitations in the XRD, other approaches for quantitative mineralogical analysis are sought. Fourier transform infrared (FT-IR) spectroscopy with attenuated total reflection (ATR) is a widely used technique for (mostly qualitative, less quantitative) analysis of cultural heritage and archaeological materials [33–40]. ATR-FT-IR spectroscopy (ATR-FT-IR) is a contact method that enables analysis of different types and states of samples (e.g. liquids, pastes, polymer films, fibres, solids and also powders) [41–44]. ATR-FT-IR has many advantages for clay analysis: it is a rapid, simple and relatively inexpensive technique that enables analysis with little or no sample preparation. It can work with very small samples (down to a few mg) and also non-destructively directly on the object. Furthermore, FT-IR offers highly characteristic information about clay sample composition and its mineral structure [45, 46]. This way it enables assigning samples to mineral families. More specifically, it is possible to determine isomorphic substitution and to discriminate between different forms of structural water and/or constitutional hydroxyl groups [45]. Compared to XRD, ATR-FT-IR can work with amorphous (non-crystalline), semi-crystalline and crystalline components.

Clays are complex mixtures, and in the ATR-FT-IR spectrum, the bands of some of their silicate minerals usually overlap strongly, thus making quantitative analysis challenging [47]. In this case, chemometric tools can be used. Partial least squares (PLS) is able to quantify a large number of components, whose bands can overlap severely in IR spectra, and spectral baselines can be variable, as long as there are at least minor differences in the ATR-FT-IR spectra of the components [48]. PLS is also good for being tolerant to noise in IR spectra, zero concentrations of components in some samples and some missing data in the model [49, 50]. So, coupled with PLS or similar chemometric methods, ATR-FT-IR is a useful technique for quantifying components from mixtures if a sufficient number of suitable calibration samples is available. Some quantitative studies of minerals from rocks (including shales) and soils have been published using the PLS regression method with ATR-FT-IR [42, 51–54]. However, mostly these reports are geology-related. Often, PLS combined with diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) has been used. For quantitative analysis, DRIFTS, compared to ATR-FT-IR, usually requires a somewhat larger sample amount (approx. 5–20 mg) and sample pre-treatment (e.g. preparation of KBr pellets) for more accurate results [32, 55–57]. Furthermore, for these investigations, the mid-IR region (4000–400 cm^{-1}) has commonly been used. At the same time, clay components have characteristic absorption bands also in the far-IR region (below 400 cm^{-1}), which are beneficial for the quantitative analysis [36, 41, 42]. To the best of our knowledge, there are no previous ATR-FT-IR-PLS–based complex quantitative investigations of main clay components in archaeological and cultural heritage samples without specific sample preparation and using an expanded wavenumber range (4000–225 cm^{-1}).

In this study, a quantitative method for the determination of the mineralogical composition of clays in different archaeological and cultural heritage samples using ATR-FT-IR (in the mid-IR and far-IR regions) with PLS quantification is presented. The creation of the ATR-FT-IR-PLS model is described step by step. Altogether, 222 calibration and validation standards covering natural clay sources and various samples (including fired samples in different temperatures) from archaeological/cultural heritage artefacts are used (compositions determined with XRD). This is the largest calibration set used for creating an ATR-FT-IR-PLS method for clay minerals. The developed ATR-FT-IR-PLS method was applied to the 11 case-study samples to evaluate the efficiency and accuracy of this approach.

Materials and methods

Reference standards for classification and quantitative analysis

Clays are complex mixtures that can consist of a number of different components. For the quantitative analysis, it is impossible to find commercial reference standards (reference samples) with varying concentrations, where all the clay components have been characterised on a quantitative basis. In this work, we used real-life archaeological, cultural heritage, geological and other clay materials as calibration and validation standards for building the PLS model.

For the classification method, 28 individual monomineralic samples (standards) were used. Minerals for the classification were obtained from the collection of the University of Tartu Department of Geology.

For creating the quantitative PLS model, initially, 236 reference samples covering natural clay sources and various archaeological/cultural heritage objects were considered as reference standards. For the quantitative analysis, reference standards of different types, ages and origin were selected. The initial set contains 127 geological samples (natural sediments, soils and moraines) [9–11]; 83 archaeological and cultural heritage samples (pottery and ceramic fragments, bricks); 11 commercial plasters, pigments, clays, including 15 clay samples from a private collection. SI Table S1 lists all the different reference standards and their origin. In the process of creating the PLS model for quantification of clay components, out of the 236 initial reference standards, 14 were excluded (see section "Creating a PLS model for quantitative analysis of clay components" for the reasons). The remaining 222 were used as calibration and validation standards for creating the final PLS model (more described in section "Creating a PLS model for quantitative analysis of clay components").

To obtain homogenous powders for XRD and ATR-FT-IR analysis, all the reference standards (amount ranged from 0.5 to more than 1g) were ground using ball mill (standards of geological origin) or agate mortar and pestle (standards of archaeological and cultural heritage origin).

Case-study samples

For the practical application, different archaeological and cultural heritage objects with variable clay components were analysed. Table 1 presents information about the analysed objects and the sample pieces taken from them. The photos of the artefacts and their samples are presented in the SI in Tables S2–S5 and Figs. S2 and S3.

For the ATR-FT-IR analysis from the archaeological objects (Egyptian clay vessel and potsherds from Narva Joaorg), it was possible to use only very small sample pieces (a few mg), and grinding was not carried out. Only samples from Tartu Cathedral and Tartu St John's church were large enough (about 1 to 2.8 g) to be ground using agate mortar and pestle for obtaining a homogeneous mixture. For the comparison for Tartu Cathedral and Tartu St John's Church samples (samples 7 to 11), XRD analysis was also carried out.

Table. 1 Information on the analysed case-study samples

Object	Samples	Age
Archaeological Egyptian clay vessel with bird mummy (KMM A 71:1; Otto Friedrich von Richter's Egyptian artefacts collection in University of Tartu Art Museum [58])	**Sample 1:** Small piece from the vessel	Mid 1st millennium cal BC
Archaeological clay vessels from different periods from the Narva Joaorg, Estonia (Tallinn University Archaeological Research Collections)	**Sample 2:** Fragment of the Narva Ware pottery (AI 4264: 870)	6th – 4th millennium cal BC
	Sample 3: Fragment of pottery production raw material (clay coil) of probably local origin (AI 4264: 847)	6th – 4th millennium cal BC
	Sample 4: Fragment of Corded Ware pottery (AI 4264: 214)	3rd millennium cal BC
	Sample 5: Fragment of textile ceramics (AI 4264: 216)	ca 1st millennium cal BC
	Sample 6: Fragment of Late Iron Age /Early Medieval wheel thrown pottery (AI 4264: 720)	ca 12th to 13th century AD
Bricks from Tartu Cathedral, Estonia (University of Tartu Museum collection)	**Sample 7:** Round brick from the upper part of the ribs of the pillar	14th to 15th century AD
	Sample 8: Rectangular brick from the floor	14th to 15th century AD
Sample pieces from Tartu St John's Church, Estonia (collection of fragments collected by the conservator Eve Alttoa)	**Sample 9:** Terracotta head sculpture	14th century AD
	Sample 10: Profile stone in the interior	14th century AD
	Sample 11: Fragment from brick	14th century AD

Measurements and processing of ATR-FT-IR spectra

All the ATR-FT-IR spectra were recorded using Thermo Scientific Nicolet 6700 FT-IR spectrometer with Smart Orbit diamond micro-ATR accessory. The FT-IR spectrometer has DLaTGS detector, CsI beamsplitter and Vectra Aluminium interferometer. The spectrometer is protected from atmospheric moisture by the constant flow of dry air.

For recording of an ATR-FT-IR spectrum, the powdered reference standard or piece of case-study sample was placed on the ATR crystal and pressed against the ATR crystal by a pressure applicator. The active sampling area of the micro-ATR crystal has a diameter of 1.5 mm. This small area is suitable for measuring powders as it enables applying high pressure for obtaining sufficient contact between the sample and the crystal, which is of key importance obtaining qualitative ATR-FT-IR spectra of powdered samples [44].

The FT-IR operating parameters were similar to the ones used earlier [36]: the resolution of 4 cm^{-1}, number of scans were 128 and 256, wavenumber range at 4000–225 cm^{-1}, zero filling factor was 0 and apodization window was Happ-Genzel. Thermo Electron's OMNIC 9 software was used to record and process the IR spectra.

To obtain homogeneous results, for every individual monomineralic samples 2 to 4, clay-related reference samples 2 to 4 and case study samples 1 to 3 ATR-FT-IR spectra were recorded. After that, IR spectra were averaged, and from the monomineralic samples (28), clay-related reference samples (236) and case-study samples (11), statistical average ATR-FT-IR spectra were obtained.

For all the recorded ATR-FT-IR spectra, atmospheric suppression correction (removes CO_2 and humidity absorptions from spectra) was applied (for that Thermo Fisher Scientific Inc. OMNIC Specta 2.0 was used). For some noisier IR spectra, the smoothing function of the OMNIC software (with the Savitzky-Golay algorithm) was used.

Analysis with X-ray diffraction (XRD)

XRD was used as the reference method for assigning the compositions of the reference standards, as it is the most reliable tool available for the quantitative determination of the mineralogical composition of materials. The XRD results of the mineralogical compositions of the reference standards used for calibration of the quantitative PLS model are shown in the SI in Table S6.

All the reference standards were measured in a similar way. Randomly oriented powdered standards (mostly over 1g) were measured on a Bruker D8 diffractometer using CuKα radiation and LynxEye positive sensitive detector in 3–70 2θ region. The quantitative mineral composition of the standards was interpreted and modelled using a Rietveld algorithm-based code Siroquant – 3 [59].

In this investigation, the XRD data of different geological samples that were previously published were also used to create the PLS model. XRD data of Estonian geological samples from Padu-Aru, Varja, Nudi and Lohu were obtained, together with XRD information on the crystallographic composition, from Liivamägi et al. (2014) [11] and Liivamägi et al. (2015) [10]. Some samples of African origin were also included, and their XRD data were obtained from Somelar et al. (2018) [9]. The rest of the reference standards were measured as part of this work. SI Table S1 lists all the reference standards.

Classification and quantitative analysis of clay components

Quantitative analysis of clay components was carried out using the PLS method. Classification was performed using discriminant analysis (DA) that is based on principal component analysis (PCA). For these chemometric techniques, Thermo Scientific TQ Analyst™ Professional Edition 9.0 software was used.

For creating DA classification and PLS models, ATR-FT-IR spectra of different reference standards were used. Even though all the reference standards were ground to homogenous powders and the measurements of the ATR-FT-IR spectra were made in a similar way, it is still difficult or impossible to ensure that the entry depth of beam is the same for every standard. In ATR-FT-IR, the depth of penetration of the IR radiation, and therefore the effective path length, depends on refractive index of standard sample, contact between crystal and sample, wavelength and also temperature [60]. So, in these spectra, effective path length in the samples can be different, but it is unknown how much it varies. To compensate for these differences, standard normal variate (SNV) correction was used for PLS and discriminant analysis. This correction algorithm improves the results if there is an interference that affects the signal proportionally to its magnitude (multiplicative influence). This can be due to the scattering of radiation or some other effect that is proportional to path length [48]. The SNV compensates for variations in sample thickness that can be due to particle size and scattering and also normalizes IR spectra to minimize the effects of scaling or offsets. The SNV correction eliminates the effects of scattering by normalizing the IR spectra of the reference standards individually so that chemical differences can be detected [48]. SNV is more often used for near-infrared (NIR) spectral data pre-processing before chemometric analysis [61]. However, our previous work [62] demonstrates that SNV can also be successfully used in the mid-IR region, and this work expands its usability in the far-IR region.

Creating DA classification model for individual minerals

DA classification enables estimating the compound classes based on similarities of minerals' spectral features. This classification method compares the IR spectrum of an analysed unknown sample with the spectra of classes of known standard materials.

For the DA classification, 28 averaged ATR-FT-IR spectra of individual monomineralic samples were inserted into the TQ Analyst software, and wavenumber ranges of 3720–2900 cm^{-1} and 1810–230 cm^{-1} were selected.

Creating a PLS model for quantification of clay components

Initially, 236 averaged ATR-FT-IR spectra of clay-related reference standards with known compositions were used for creating the PLS model. The contents of the mineral components (g/100g) of the reference standards were found using XRD analysis, and these were inserted into the TQ Analyst software alongside recorded ATR-FT-IR spectra.

Based on the appearance of the absorption bands of the clay components (see Fig. 1, and in the SI Fig. S4) for the PLS analysis, the following wavenumber ranges were used:

- 3720–3000 cm^{-1} (kaolinite, halloysite, chlorite, illite-smectite-mica and hydroxides have characteristic absorption bands in that range).
- 1710–275 cm^{-1} (all the clay components, including carbonates and oxides have characteristic bands in this range).

The software was allowed to choose the required number of factors, ranging from 8 to 10 for different minerals.

The PLS method was evaluated using the following parameters: root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), squared correlation coefficient (R^2), root mean square error of cross-validation (RMSECV) and modified Willmott performance index [63].

RMSEC, RMSEP and RMSECV describe the accuracy of the method, which means how similar the concentration found using the developed method is to the concentration found using the referents method and is found based on the following equation (1):

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2\right)} \quad (1)$$

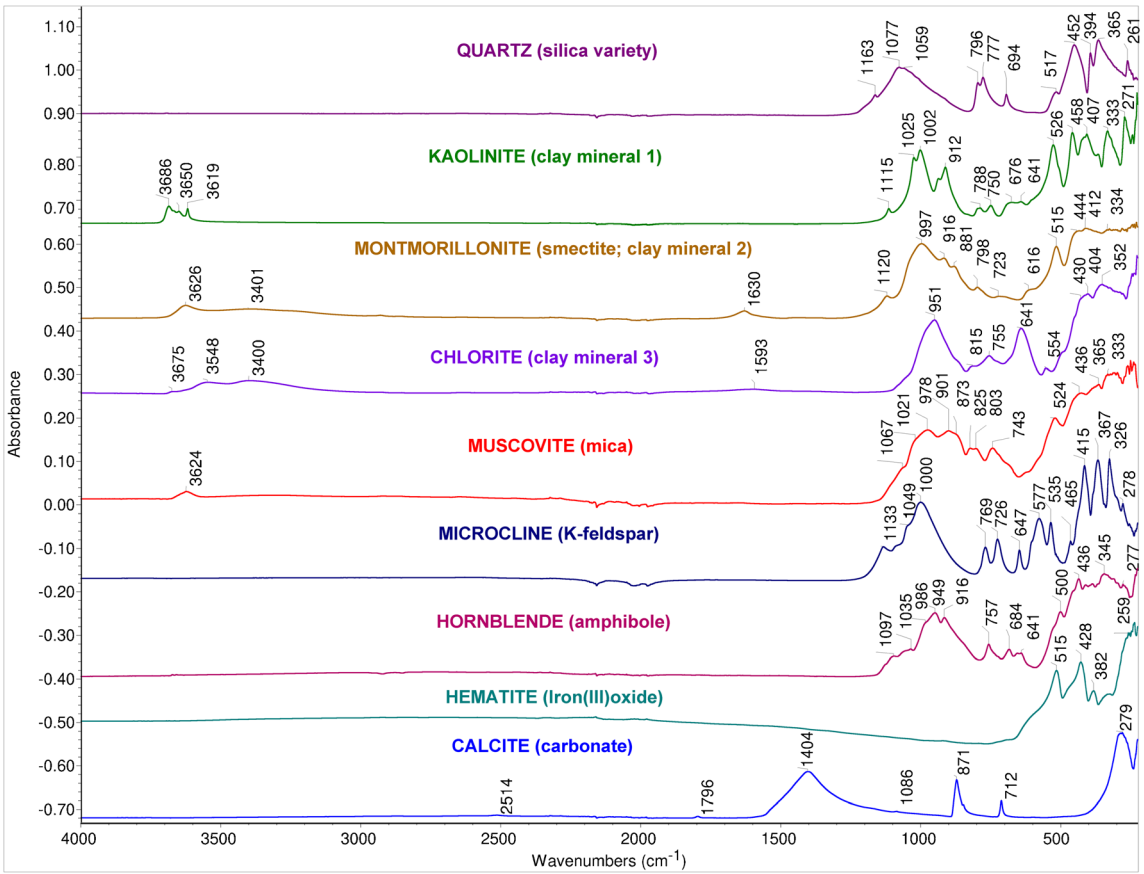

Fig. 1 Representative ATR-FT-IR spectra of individual minerals from each class

In this equation, n stands for the number of either calibration standards (by RMSEC), validation standards (RMSEP) or an overall number of standards (RMSECV) of the method. The i represents the number of a reference standard, and $\hat{y}_i$ the best estimate of the unknown value, which in this case is the concentration of the respective component in standard i found by XRD analysis. The y_i is the concentration of the reference standard i found using the PLS method. In the case of RMSECV, the y_i of every reference standard is found using all remaining standards as calibration standards. R^2 describes the correlation between measured signal and component concentrations (found using XRD). TQ Analyst calculates R^2 for calibration as well as validation standards.

The performance index shows the PLS method's ability to predict the component concentrations in the validation standards. Performance index enables easy comparison between different methods [63].

For monitoring the correlation between RMSEC, RMSEP and the number of factors, the PRESS function was used. Also, TQ Analyst Spectrum Outlier function for spotting the deviating spectra was used. The function finds the most different data point and evaluates the significance of deviations using the Chauvenet's test.

Results and discussion

Classification of the individual monomineralic samples

The first step in the quantitative method development was to evaluate ATR-FT-IR spectra of individual minerals commonly present in clays and observe how they can be differentiated and classified based on their spectral features. For that, DA classification was used. Based on the similarities of structures and observation of ATR-FT-IR spectra, initially, 28 individual monomineralic samples were grouped into seven classes (see Table 2). These classes were thereafter used for creating the DA classification model. ATR-FT-IR spectra of minerals from each class are presented in Fig. 1, and their complete interpretation can be found in the SI in Table S7.

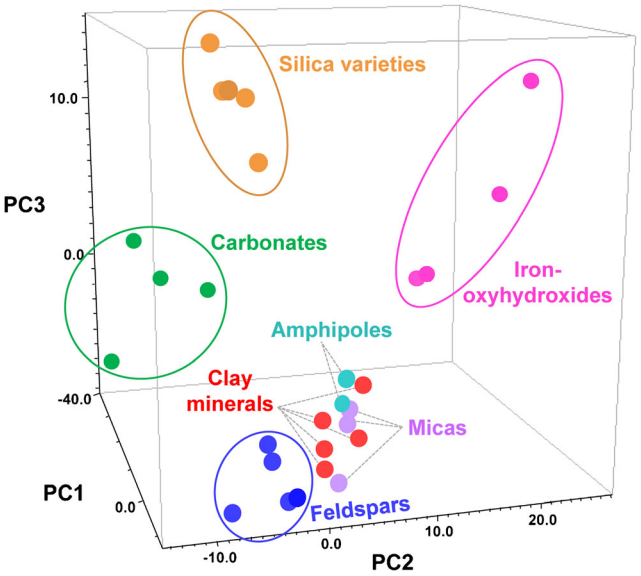

Fig. 2 Classification results of the individual minerals presented by the three principal components

The 3D PCA graph with the classification results of the individual minerals is presented in Fig. 2. Altogether, 10 principal components (PCs) were obtained for discriminant analysis, and the variances (or variabilities) described by the first three principal components were PC1 36.1%, PC2 18.6% and PC3 11.9%, leading to the cumulative variance of only 66.6%. However, the cumulative variability described by all 10 PCs is 95.3 %, which is sufficient for our purposes.

The PCA graph in Fig. 2 shows that silica varieties, feldspars, carbonates and iron-oxyhydroxides are distributed into distinctly separate groups. However, the clay minerals, micas and amphiboles are grouped as one group. The structure of amphiboles compared to micas and clay minerals is slightly different. Amphiboles are inosilicates with a double-chain structure, and micas and clay minerals are phyllosilicates, but both groups contain a hydroxyl (OH) group.

Ten PCs were involved in the DA classification method. The results indicate that the minerals under question are

Table. 2 The mineral classes used for the discriminant analysis classification

Classes	Analysed minerals
Silica varieties	Chalcedony, quartz, milky quartz, rock crystal, opal
Clay minerals	Illite-smectite, smectite, kaolinite, chlorite, glauconite
Feldspars	Albite, labradorite, microcline, sanidine, plagioclase
Micas	Biotite, phlogopite, muscovite
Amphiboles	Hornblende, amphibole
Iron-oxyhydroxides	Fe_2O_3, hematite, goethite, limonite
Carbonates	Calcite, aragonite, dolomite, siderite

different enough, and there are characteristic vibrations in the ATR-FT-IR spectra that the PLS algorithm can use for quantification. Also, these results suggest that for creating a reliable quantitative method, minerals grouped together (feldspars, clay minerals and micas, carbonates, etc.) should be handled as a group also in the PLS model.

Creating a PLS model for quantitative analysis of clay components

For creating an accurate ATR-FT-IR-PLS method, numerous calibration standards are needed. These must contain all the main components that are expected in the unknown sample, and their concentrations in the reference standards should vary independently. In this work, ATR-FT-IR spectra of 236 clay-related reference standards, with known contents of individual crystalline phases (components) distinguished with XRD analysis, were used. The XRD analysis detected altogether 36 crystalline phases from these standard samples. From those 36, the following 29 minerals (or mixtures of minerals) were explicitly used for creating the PLS model: quartz, cristobalite, K-feldspar, plagioclase, kaolinite, halloysite, illite, smectite, illite-smectite, mica, vermiculite, chlorite, chlorite-smectite, calcite, aragonite, dolomite, siderite, ankerite, hematite, goethite, amphibole, hydroxides (gibbsite, portlandite, brucite), apatite, Ti-oxides (anatase/rutile), pyroxene, Al-silicates (mullite/sillimanite), spinel. These mineral phases were chosen on the basis of two criteria: (1) the phase was detected in the composition of more than three reference standards, and (2) its content in at least one standard was above 1.5%. As a result, garnet, titanite, akermanite, pyrite and weddellite were left out of the model.

Examining the XRD results revealed that even the XRD analysis, which is commonly used for the quantification of a clay composition, is not always able to distinguish all the different components. This happens, for example, when the mineral phase with low content is overlapping with dominant mineral phases or if overlapping minerals are with similar structure and composition like micas, illite and illite-smectite.

Before creating the PLS model, the contents of the above-mentioned 29 individual crystalline phases of the reference standards were grouped based on the XRD results: minerals with similar structures and identical or similar chemical compositions were grouped into one group. For the grouping of the minerals, the results obtained by comparison of the ATR-FT-IR spectra of individual minerals, including their classification (see section "Classification of the individual monomineralic samples") and also information obtained from the literature [18, 64], were taken into consideration. The minerals in the reference standards were divided into the following groups:

1. **Silica varieties**: *quartz and cristobalite (SiO_2 polymorphs).* They have identical chemical compositions but different crystal structures.
2. **Feldspars**: *plagioclase* (Ca, $NaAlSi_2O_8$) and *K-feldspar* ($KAlSi_3O_8$). Feldspars are a group of tectosilicates ($KAlSi_3O_8$ – $NaAlSi_3O_8$ – $CaAl_2Si_2O_8$) with very similar structure, chemical composition and physical properties. Feldspars are the most abundant minerals in Earth crust and are common in igneous, metamorphic and sedimentary rocks.
3. **Clay minerals, group 1**: *kaolinite* and *halloysite* (both $[Al_2Si_2O_5(OH)_4]$). These minerals have a similar chemical composition and crystal structure. The difference is in water content; in the halloysite, the unit layers are separated by a monolayer of water molecules.
4. **Clay minerals, group 2**: *micas, illite, illite-smectite, smectite* and *vermiculite.* The discrimination of those minerals using XRD is difficult due to overlapping peaks in diffractograms. The overlapping is caused by very similar chemical compositions and structures. To distinguish those minerals, the separation and analyses of clay minerals/clay fractions are needed.
5. **Clay minerals, group 3**: *chlorite* and *chlorite-smectite.* Unlike other clay minerals, the interlayer space of chlorite group minerals is filled by the brucite-like layer.
6. **Carbonates**: *calcite* ($CaCO_3$), *aragonite* ($CaCO_3$), *dolomite* ($Ca,Mg(CO)_2$), *siderite* ($FeCO_3$) and *ankerite* ($Ca(Fe,Mg,Mn)(CO_3)_2$). These minerals differ in structure, and some of them also in chemical composition.
7. **Iron-oxyhydroxides**: *hematite* (Fe_2O_3) and *goethite* [FeO(OH)].

The contents of the grouped components in the clays-related reference standards and their recorded ATR-FT-IR spectra were inserted into the TQ Analyst software as calibration or validation standards. The contents of the components not belonging to any of the above groups, such as amphibole, pyroxene, Al-silicates, spinel, apatite, hydroxides and Ti-oxides, were also added into the TQ Analyst software. These minerals were used in the PLS model as support components in order to improve its predictive capability for the main seven groups. With the XRD analysis, these components were not detected in all the standard samples, and the variability of their concentrations in samples was not high enough. For these reasons, it was not expected that the PLS model would reliably evaluate their contents in unknown samples. However, including them in the PLS model will enable quantifying the main 7 component classes in different reference standards more accurately (see Table 3).

The PLS method was developed starting with the above-explained groups of reference standards. These standards occasionally contained also amorphous components that XRD analysis is not able to identify or other components that are not

Table. 3 Results of the ATR-FT-IR-PLS analysis of the main clay components

Component	Calibration		Validation		Leave-one-out cross-validation		Performance index (%)	Number of factors used
	R^2	RMSEC (g/100g)	R^2	RMSEP (g/100g)	R^2	RMSECV (g/100g)		
Silica varieties	0.9585	4.9	0.9816	3.6	0.9489	5.4	82.0	8
Feldspars	0.9722	5.0	0.9875	4.0	0.9636	5.8	85.0	10
Clay minerals, group 1	0.9840	2.9	0.9897	2.6	0.9785	3.3	86.8	9
Clay minerals, group 2	0.9029	6.1	0.9371	5.1	0.8563	7.4	79.7	10
Clay minerals, group 3	0.9142	2.4	0.9493	1.8	0.8754	2.8	83.5	10
Carbonates	0.9944	1.7	0.9993	0.9	0.9894	2.4	96.4	10
Iron-oxyhydroxides	0.9049	2.7	0.9570	1.9	0.8627	3.2	80.1	10

commonly found in natural clays. It was therefore expected that the set of reference standards contained samples whose compositions are not entirely correctly found using XRD or are not typical to most natural clays. To discover such samples, all the 236 reference standards were critically evaluated. At first, the TQ Analyst Spectrum Outlier diagnostic was used, which indicated that six reference standards failed the Chauvenet's test and were considered outliers. Secondly, it was observed how well the standard sample data points were distributed in the calculated vs actual plots and the % difference plots (see in the SI Fig. S5). Calculated vs actual plot compares the concentration value calculated with the PLS model to the reference concentration value (found using XRD) for each component in the reference standards. The % difference plot shows the differences between the calculated and the actual concentration values relative to the actual values. Using these plots, we were able to detect eight outliers that were farther from the others in the calibration plots.

Thus, altogether 14 outliers were detected. All of them were somewhat different in their composition from the other reference standards (see in the SI Table S6). Observing the XRD results of these standards, it is not surprising that these were ruled out. From the 14 outlier standard samples, 8 (3, 12, 13, 14, 98, 183, 218, 230) contain essentially just one main component with very high concentration (between 65 and 90 g/100g), and the other 1 to 3 components have low concentrations. Three reference standards (136, 142, 144) contain a higher concentration of Al-silicates (25 to 45 g/100g) that is however in the model as a support component and is not present in most reference standards, also a quite high content of quartz (40 to 59 g/100g) and 1 or 2 components with lower concentrations. The remaining three reference standards (128, 160, 192) were left out because, under closer inspection, doubts arose about their homogeneity and, consequently, whether their IR spectra and XRD results refer to the same compositions. Based on these considerations, the ATR-FT-IR spectra of these 14 reference standards were excluded from the PLS model.

All in all, out of the initially considered 236 reference standards, 14 were excluded, as described above. From the remaining 222 reference standards, 202 were defined as calibration standards (training set) (see in the SI Table S6) and 20 were defined as validation standards (validation set). The main criteria for choosing validation standards were the presence and content ranges of the different components in the clay reference standards. Whenever possible, reference standards with extreme concentrations (i.e. standards in which the concentration of one of the components was higher or lower than in all remaining reference standards) were avoided. Also, each of the seven mineral groups of reference standards were observed separately to make sure that enough calibration and validation standards would be available for the quantification of the respective group of minerals. By the end of the process, 20 were defined as validation standards (see Table 4).

The resulting PLS model can quantify the main seven classes of minerals in clay-related samples. Table 3 presents R^2 for calibration and validation standards and RMSEC and RMSEP values, including performance index and number of factors results for each main clay components. The overall performance index of the PLS model (7 classes of main minerals including support components) is 82.6%. The average performance index of 7 main classes of minerals in clays is 84.8%.

For the seven groups of main components, the calibration R^2 varied between 0.90 and 0.99 and validation R^2 between 0.94 and 0.99. Table 3 indicates that RMSEC values of main components varied between 1.7 and 6.1 g/100g and RMSEP values between 0.9 and 5.1 g/100g. One would expect that RMSEP values are higher than RMSEC values. The reason for the opposite situation is that (1) the selection criteria for the validation standards in terms of composition were much more stringent, while for the calibration sets, much more relaxed selection criteria were applied (see above) and (2) the fact that

Table. 4 Comparison of actual (XRD values) and calculated (ATR-FT-IR with PLS) validation results using developed PLS method

Sample no.	**Silica varieties** (g/100g)			**Feldspars** (g/100g)			**Clay minerals 1** (g/100g)			**Clay minerals 2** (g/100g)			**Clay minerals 3** (g/100g)			**Carbonates** (g/100g)			**Iron-oxyhydroxides** (g/100g)		
	Actual	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**	**Actual**	**Calc.**	**Diff.**
10	22.3	27.0	**4.7**	8.6	1.7	**−6.9**	40.8	38.8	**−2**	27.4	30.2	**2.8**	0	−1.1	**−1.1**	0	−0.4	**−0.4**	0	0	**0**
20	4.3	5.2	**0.9**	1.9	−2.7	**−4.6**	63.6	67.8	**4.2**	0	2.7	**2.7**	0	−0.3	**−0.3**	0	−1.7	**−1.7**	28.6	26.2	**−2.4**
35	15.0	12.8	**−2.2**	6.5	12.5	**6.0**	9.3	6.3	**−3**	21.2	26.9	**5.7**	17.7	21.4	**3.7**	1.3	−0.3	**−1.6**	2.1	6.1	**4**
44	43	40.4	**−2.6**	1.0	2.0	**1**	33.2	36.2	**3**	15.8	9.4	**−6.4**	0	−0.3	**−0.3**	0	0.6	**0.6**	7.0	8.6	**1.6**
73	43.3	42.6	**−0.7**	11.6	8.0	**−3.6**	28.3	27.8	**−0.5**	12.4	16.1	**3.7**	0	0	**0**	1.8	2.1	**0.3**	1.0	0.1	**−0.9**
82	25.6	23.1	**−2.5**	19.4	23.0	**4.5**	10.8	6.3	**−4.5**	34.0	27.9	**−6.1**	5.1	8.8	**3.7**	0.5	0.7	**0.2**	0	2.2	**2.2**
105	3.6	4.3	**0.7**	68.9	69.8	**0.9**	19.8	17.6	**−2.2**	0	−0.4	**−0.4**	0	1.1	**1.1**	0	0	**0**	3.1	3.5	**0.4**
106	1.6	1.8	**−0.2**	74.9	73.4	**−1.5**	16.3	21.2	**4.9**	0	−8.2	**−8.2**	0	−1.0	**−1**	0	0.4	**0.4**	2.5	5.3	**2.8**
122	2.0	−1.5	**−3.5**	83.6	84.7	**1.1**	9.3	13.3	**4**	0	−5.9	**−5.9**	0	0.7	**0.7**	0	0.2	**0.2**	0.8	3.1	**2.3**
130	55.8	48.5	**−7.3**	25.4	29.5	**4.1**	0	−2.1	**−2.1**	11.9	12.5	**0.6**	0	1.2	**1.2**	1.5	2.0	**0.5**	3.8	2.9	**−0.9**
147	41.9	41.7	**−0.2**	23.1	25.3	**2.2**	0	3.0	**3**	34.3	30.7	**−3.6**	0.6	−0.4	**−1**	0	−0.9	**−0.9**	0	0	**0**
150	38.2	36.9	**−1.3**	39.7	41.5	**1.8**	0	1.1	**1.1**	19.4	23.9	**4.5**	0	−1.4	**−1.4**	0	−1.0	**−1.0**	0	0.7	**0.7**
156	34.0	36.3	**2.3**	47.5	52.0	**4.5**	0	1.9	**1.9**	15.0	12.4	**−2.6**	0	−1.2	**−1.2**	0	−1.7	**−1.7**	0	1.1	**1.1**
163	47.7	40.4	**−7.3**	36.7	37.7	**1**	0	0.4	**0.4**	12.3	20.4	**8.1**	0	−1.9	**−1.9**	0	−0.3	**−0.3**	0	0.9	**0.9**
182	18.0	19.0	**1**	25.0	21.5	**−3.5**	6.1	8.2	**2.1**	39.1	42.8	**3.7**	7.6	6.3	**−1.3**	3.2	3.9	**0.7**	0	−0.3	**−0.3**
195	20.7	22.4	**1.7**	59.3	53.2	**−6.1**	0	−0.3	**−0.3**	14.3	11.7	**−2.6**	0	−0.7	**−0.7**	0	0.3	**0.3**	0	0.5	**0.5**
204	18.0	22.4	**4.4**	25.4	27.8	**2.4**	0	0.4	**0.4**	51.7	42.5	**−9.2**	0	3.8	**3.8**	4.1	3.2	**−0.9**	0	1.0	**1**
216	2.3	−1.3	**−3.6**	0	−5.4	**−5.4**	0	0.8	**0.8**	5.6	12.6	**7.0**	0	−1.2	**−1.2**	88.6	87.0	**−1.6**	3.5	2.4	**−1.1**
226	40.7	36.3	**−4.4**	17.3	22.2	**4.9**	0	2.0	**2.0**	7.9	9.7	**1.8**	4.9	3.0	**−1.9**	29.2	28.2	**−0.4**	0	0.6	**0.6**
233	61.2	67.2	**6**	23.3	19.0	**−4.3**	0	−2.3	**−2.3**	3.5	0.9	**−2.6**	0	1.3	**1.3**	12.0	11.4	**−0.6**	0	−4.4	**−4.4**

Diff., difference between calculated and actual values (Diff. = Calculated - Actual); *Calc.*, calculated - the calculated concentration value is the value that is calculated by the PLS calibrated method for the selected component in each reference standard. See the text regarding negative values

in almost all calibration standards the content of at least some components was zero.

The fact that some found component concentrations are not within limits defined by RMSEP is not unexpected, since the probability of that is about 68% (assuming approximately normal distribution of deviations), and in addition to the uncertainty of ATR-FT-IR-PLS results, the uncertainties of XRD reference values also have an impact. The relative standard deviation of XRD Rietveld-based phase analysis of diffractograms using Siroquant-3 has been estimated 5 (for higher concentrations) to 10% (for lower concentrations), based on the investigation of mixtures of pure phases.

The performance and applicability of the developed PLS model can be evaluated by the results of validation standards. Table 4 shows how well a calibrated quantitative PLS method performs when quantifying the main clay components of the 20 validation standards. The performance index indicates how accurately the PLS method can quantify validation standards (see Table 3). In Table 4, the results of the validation standards show that carbonates have the most accurate results where differences between actual and calculated values vary between −1.7 and 0.7 g/100g. Also, the performance index of carbonates is the highest (96.4%) compared with the values of other components (see Table 3). This result is expected because the IR spectra of carbonates (see IR spectra in Fig. 1 and their complete interpretation in the SI in Table S7) are very different from other clay components. Comparing the results of three groups of clay minerals, group 1 (contains kaolinite and halloysite) has a better performance index (86.8%) than group 2 (contains summarized values of smectite, illite, illite-smectite, micas, vermiculite) and group 3 (chlorite and chlorite-smectite) which have performance indices 79.7% and 83.5%, respectively. The results presented in Table 4 indicate that the PLS method has quantified group 1 clay minerals better than group 2 clay minerals. As Fig. 1 shows, the ATR-FT-IR spectrum of kaolinite (a mineral in group 1) is different from that of illite and muscovite (clay minerals in group 2). Kaolinite has characteristic bands in the mid-IR ranges of 3686–3619 cm^{-1} and 1115–641 cm^{-1}, as well as in the lower wavenumber region at 526–271 cm^{-1}. IR spectra of illite and muscovite have broader and less characteristic bands. Clay minerals in group 2 contain components with very similar structure and chemical composition, and even XRD had difficulties in discriminating between them. Due to the similarities in the IR spectra, their absorption bands overlap extensively, and the PLS model has difficulties differentiating them. However, some uncertainty comes also from the XRD analysis, i.e. from the "actual" compositions. Figure 1 indicates that the IR spectra of chlorite (group 3 clay mineral) are somewhat different (there are characteristic bands in the region of 3700 to 3300 cm^{-1} and 1000–350 cm^{-1}) from other clay minerals (groups 1 and 2) and the PLS model was able to quantify these components.

Table 4 reveals that some of the quantified components are absent in some of the validation standards. For these standards, ATR-FT-IR-PLS typically finds concentrations close to zero, which are quite reasonable. Importantly, when trying to detect components at zero concentration level and if the results are approximately symmetrically distributed (which is a reasonable assumption), then roughly half of the results should be positive, and half should be negative. Replacing negative results with zeroes (which is sometimes done) would introduce bias to the results. Right now, the average result of the 48 results where a component missing in the sample was quantified as −0.4 g/100g, indicating that positive and negative results essentially cancel and thus the bias is negligible. For this reason also, negative values are listed in the table.

Clays also contain some amount of non-clay minerals, such as quartz, feldspars (K-feldspar, plagioclase) and iron-oxyhydroxides (hematite and goethite). Figure 1 reveals that quartz, microcline (K-feldspar) and hematite all have absorption bands in the range of 1170 to 250 cm^{-1}. This is the same range where all the clay components have absorption bands, causing considerable overlapping. PLS can quantify overlapping minerals as long as there are at least small differences in the spectra [36]. Looking closer, it can be seen that quartz has characteristic bands in the range of 796 to 261 cm^{-1}, microcline in the range of 769 to 278 cm^{-1} and for hematite could be useful bands at 428 and 382 cm^{-1}. Based on their spectral features, PLS was able to adequately quantify these components (see Table 3 and Table 4).

In addition to the validation with the pre-defined validation standards, the leave-one-out cross-validation diagnostic in the TQ Analyst software was used to monitor how well the PLS method performs by quantifying each calibration standard out of 222 as if it were a validation standard (and leaving it out of the calibration set). For the seven groups of main components, the R^2 varied between 0.86 and 0.99 and root mean square error of cross-validation (RMSECV) values varied between 2.4 and 7.4 g/100g (see Table 3).

Considering that for the calibration and validation standards real-life archaeological and geological reference standard samples were used, the obtained results are satisfactory to good. Quantitative determination of some components is possible while others are determined at a semiquantitative level.

Application of the developed method for analysis of case-study samples

The developed ATR-FT-IR-PLS method was applied for quantitative analysis of a selection of archaeological and cultural heritage samples (see Table 1) to evaluate its capabilities. All the ATR-FT-IR spectra of case-study samples were inserted into TQ Analyst software, and the PLS model calculated the quantitative contents of these samples. Table 5

Table. 5 Quantitative contents of clay components in the samples using the ATR-FT-IR-PLS method[a]

Samples	Content (g/100g) of the component							Total quantified (g/100g)
	Silica varieties	Feldspars	Clay minerals 1	Clay minerals 2	Clay minerals 3[b]	Carbonates	Iron-oxy-hydroxides	
Sample 1. Egyptian vessel	23.0±3.6	26.5±4.0	1.0±2.6	21.1±5.1	1.5±1.8	1.1±0.9	8.4±1.9	82.6±8.3
Sample 2. Narva Ware pottery from Narva Joaorg	26.8±3.6	52.3±4.0	6.5±2.6	11.2±5.1	−3.0±1.8	1.1±0.9	10.4±1.9	105.3±8.3
Sample 3. Clay coil from Narva Joaorg	26.7±3.6	43.8±4.0	2.4±2.6	20.2±5.1	1.0±1.8	1.2±0.9	7.5±1.9	102.8±8.3
Sample 4. Corded Ware pottery from Narva Joaorg	25.3±3.6	31.5±4.0	7.2±2.6	30.2±5.1	−0.4±1.8	0.6±0.9	9.1±1.9	103.5±8.3
Sample 5. Textile ceramics from Narva Joaorg	27.5±3.6	46.3±4.0	4.0±2.6	22.6±5.1	1.3±1.8	−1.1±0.9	1.8±1.9	102.4±8.3
Sample 6. Wheel thrown pottery from Narva Joaorg	35.8±3.6	40.4±4.0	1.6±2.6	20.8±5.1	−2.1±1.8	−1.2±0.9	0.4±1.9	95.7±8.3
Sample 7. Brick from Tartu Cathedral[c]	48.2±3.6	30.5±4.0	0.63±2.6	12.8±5.1	−2.2±1.8	0.7±0.9	9.1±1.9	99.7±8.3
Sample 8. Brick from Tartu Cathedral[c]	44.6±3.6	35.7±4.0	2.0±2.6	7.0±5.1	−2.3±1.8	0.8±0.9	12.3±1.9	100.1±8.3
Sample 9. Head sculpture from the St. John's church[c]	53.4±3.6	37.0±4.0	2.3±2.6	0.6±5.1	−1.7±1.8	0.7±0.9	8.1±1.9	100.4±8.3
Sample 10. Profile stone from the St. John's church[c]	38.4±3.6	36.5±4.0	3.1±2.6	12.2±5.1	−2.5±1.8	−0.2±0.9	13.7±1.9	101.2±8.3
Sample 11. Brick from the St. John's church[c]	51.2±3.6	33.0±4.0	0.1±2.6	6.3±5.1	−1.0±1.8	−0.1±0.9	5.6±1.9	95.1±8.3

[a] Uncertainties of component contents correspond to RMSEP values. The uncertainty of the total quantified amount was found as the square root of sums of squares of individual component RMSEP values.

[b] See section "Creating a PLS model for quantitative analysis of clay components" regarding negative values.

[c] See Table S8 in the SI for compositions determined using XRD.

presents all the results of the quantitative PLS analysis of 11 real-life samples.

The results presented in Table 5 demonstrate that the developed ATR-FT-IR-PLS method can quantify a group of minerals in the different kinds of clay samples, while in some cases the results are to be considered semiquantitative. It is instructive to look closer at the total quantified amounts listed in the table. These have been obtained as sums of the contents of the individual components. If all main components have been quantified, then, ideally, the sum should be exactly 100 g/100g. However, all individual amounts have uncertainties. They are quantified here as the RMSEP values, used as estimates of combined standard uncertainties (see the course in ref [65] for full details of uncertainty-related terminology). Consequently, the total quantified amount also has an uncertainty, found (again, at standard uncertainty level) as the square root of sums of squares of individual component RMSEP values (see Table 5). Looking at the values, it is remarkable how close most values are to 100%, demonstrating that in many cases, the approach performs better than the uncertainties imply. The differences of the total quantified amounts from 100 g/100g should be judged keeping in mind that standard uncertainty presents uncertainty at 68% coverage probability level and two times the standard uncertainty has coverage probability of 95%. One can see that all total quantified amounts are within the two times standard uncertainty region. If values are strongly below 100%, then the cause might also be that there is a non-negligible amount of unknown material that was not explicitly modelled and consequently was not quantified.

Because of the nature of the mathematical procedure, the absolute uncertainty estimates of the components' contents (estimated via the RMSEP values) are constant over the whole content range. This means that the relative uncertainties increase sharply when the content of the respective component decreases. This makes sense— the lower the content of the component, the more disturbing are the spectral imperfections, components not accounted for and the measurement uncertainties of XRD analysis. For this reason, caution should be exercised with low content values. Specifically, if the content of a component is not higher than two times the standard uncertainty then it is not possible to state with confidence that the component is present in the sample. For example, in the

majority of samples, it is not possible to state the presence of the first group of clay minerals (i.e. kaolinite).

The following shortly describes the results of each sample from archaeological and cultural heritage objects and the information obtained with the ATR-FT-IR-PLS method (see Table 5).

Quantitative analysis of Egyptian clay vessel with bird mummy

The photos of clay vessel with bird mummy and small analysed sample piece are presented in the SI in Table S2, with ATR-FT-IR spectrum provided in the SI Fig. S6.

The ATR-FT-IR-PLS method enabled quantifying most of the main components of the sample as a total of 82.6±8.3 g/100g. The rest of the content belong to components not explicitly modelled. PLS model quantified the contents of silica varieties (quartz and/or cristobalite), feldspars and group 2 clay minerals (smectite, illite, micas, illite-smectite, vermiculite) all over 20 g/100g. The red colour of the vessel was caused by hematite which was detected with the concentration of (8.4±1.9) g/100g. Based on these results, it is impossible to confirm the presence of kaolinite (clay minerals group 1) in the sample.

The composition of the bird mummy-urn is somewhat different from the rest of the case-study examples, all from Estonian archaeological sites (somewhat lower content of quartz, feldspars and absence of kaolinite) [66]. The results of vessel's clay quantitative composition seem to correlate nicely with its function — an urn in an Egyptian cemetery context. It was meant as a burial container rather than a vessel for eating and drinking, and hence the firing and making the clay vessel strong and heat- and water-resistant were not a priority.

Quantitative analysis of archaeological samples from Narva Joaorg, Estonia

Using the developed ATR-FT-IR-PLS method, five archaeological pottery fragments from the different occupation phases covering a total of 6 millennia from Narva Joaorg (NE-Estonia) were analysed and compared. See photos of the fragments of archaeological pottery and analysed small samples in the SI in Table S3 and ATR-FT-IR spectra in Fig. S7.

These results are rather significant, especially considering that in the Narva Joaorg area the local clay sources are expected to have been used for pottery manufacture [67]. The compositions (see Table 5) of samples 3 (raw material clay coil found in Mesolithic context), 5 (fragment of textile ceramics) and 6 (Iron Age wheel-thrown specimen) are quite similar. The only significant differences can be observed in the content of quartz (about 36 g/100g) for sample 6 and the higher content of Fe-oxyhydroxides (7.5±1.9 g/100 g) for sample 3. Based on these quantitative results, samples 2 (fragment of the Narva Ware pottery) and 4 (fragment of Corded Ware pottery) differ more from the others. In sample 2, the contents of feldspars are about 52 g/100g, clay minerals 2 is significantly lower (11.2±5.1 g/100g) and Fe-oxyhydroxides (hematite) is quite high (about 10 g/100g). However, sample 4 has considerably higher group 2 clay minerals (smectite, illite, micas, illite-smectite), lower feldspar concentration (about 30.2 g/100g and 31.5 g/100g, respectively) and significantly high Fe-oxyhydroxides content (about 9 g/100g). Both samples 2 and 4 have higher contents of clay minerals 1 (kaolinite), around 7 g/100g, compared to other samples. Based on that, it can be supposed that samples 2 and 4 were fired at a lower temperature, not higher than 550 °C.

The different clay content of Narva Ware sherd (sample 2) is somewhat unexpected as it ought to resemble the raw material clay coil (sample 3). The difference may be due to the inhomogeneity of the analysed samples. However, the difference might also relate to the exploitation of different natural clay sources in the early pottery production phases compared to later periods (samples 3, 5, 6), potentially also different tempers added to the initial clay mixture or indeed potential temporal difference of these two artefacts. The considerably deviating composition of Corded Ware sherd (sample 4) in this context might, however, indicate that this vessel was not locally produced but rather imported. Indeed, the Corded Ware phenomenon as such in Europe has been widely related to human migrations (supported by aDNA studies [68]) and dispersal of several novel cultural traits from subsistence (arrival of domesticated species) to burial customs and artefacts (inc. pottery) [69].

Quantitative analysis of samples from the Tartu Cathedral, Estonia

Table S4 in the SI presents the photos of the analysed brick samples from Tartu Cathedral, and Fig. S8 shows their ATR-FT-IR spectra. For the comparison, XRD analyses for the brick samples were also carried out, and these data are presented in Table S8 in the SI.

The ATR-FTIR-PLS method enabled quantifying some of the main components of the two samples as a total of about (100±8) g/100g. Results of the quantitative analysis indicate that the composition of the bricks originating from the pillar and floor (samples 7 and 8) is expectedly quite similar. For both samples, the PLS model quantified the contents of silica varieties between (44.6±3.6) and (48.2±3.6) g/100g and feldspars between (30.5±4.0) and (35.7±4.1) g/100g. These values (especially quartz) are quite similar also with the XRD quantitative results. However, the ATR-FT-IR-PLS method could not adequately quantify clay minerals of groups 1 and 2 in these samples. When visually observing the powdered samples of the bricks (see in the ESI, Table S8), then

sample 8 is more reddish than sample 7. The quantitative results support that PLS model quantified contents of iron-oxyhydroxides (in this case hematite that gives red colour) as (12.3±1.9) g/100g and (9.1±1.9) g/100g, respectively. XRD results also support these results. The results show that despite the overlapping of the bands and some noisiness in the IR spectra in the region of 600–225 cm^{-1}, our ATR-FT-IR-PLS method can also quantify iron-oxyhydroxides (e.g. hematite) with reasonable accuracy.

Based on the results with ATR-FT-IR-PLS and XRD methods, it is possible to conclude that the firing temperatures were below 900–950°C (micas are still present) and the initial source of the studied Tartu Cathedral bricks is probably the same. Since brick kilns already existed during the Medieval Period in Tartu, it is possible that the clay used in manufacturing of the bricks was of local origin [70].

Quantitative analysis of samples from the St. John's Church in Tartu, Estonia

Table S5 in the SI presents the photos of the analysed samples, and Fig. S9 their ATR-FT-IR spectra. Also, for these samples, XRD analyses were carried out, and obtained data are presented in Table S8 in the SI.

The developed ATR-FT-IR-PLS method enabled quantifying some of the main components of the three samples as a total of between (95.1±8.3) and (101.2±8.3) g/100g. PLS results show that the compositions of samples 9 (terracotta head sculpture) and 11 (brick) are more similar to each other compared to sample 10 (profile stone). Both samples 9 and 11 contain over 50 g/100g silica varieties (quartz), over 32 g/100g feldspars and under 6 g/100g of clay minerals and/or micas. There is a small difference in the composition of Fe-oxyhydroxides (reddish colour indicates the presence of hematite) — sample 9 has the content of (8.1±1.9) g/100g, and sample 11 has (5.6±1.9) g/100g. XRD results also support some of these results, indicating the same material source for those objects. XRD results also show that the ATR-FT-IR-PLS model quantified clay minerals of groups 1 and 2 incorrectly, and in this case, XRD quantified them correctly. However, in sample 10, the ATR-FT-IR-PLS quantified mineral content differs dramatically from XRD results — the contents of silica varieties and also feldspars both are around 36 to 38 g/100g, groups 1 and 2 clay minerals between 3 and 12 g/100g and Fe-oxyhydroxides (hematite) 13.7±1.9 g/100g. The dramatic difference is caused by the presence of an amorphous phase (glass) in sample 10. It is impossible to quantify the amorphous glass directly using the XRD. In this case, the XRD overestimates other mineral phases at the expense of the amorphous phase. The presence of the amorphous phase, as well as the detection of spinel with XRD, indicates that the firing temperature for the profile stone (sample 10) may have reached 900 to 950 °C.

In general, comparing the quantitative results of the samples from St. John's Church and Tartu Cathedral, we can conclude that these are quite similar, and supposedly the clay and other raw materials could be from the same location and of the same origin.

Conclusions

In this study, a quantitative method for determining different minerals in clay samples using ATR-FTIR spectroscopy in the range of 3720 to 275 cm^{-1} with PLS data analysis is presented.

For creating the ATR-FT-IR-PLS method, 222 real-life archaeological, geological and commercial clay-related reference standard samples (the concentrations of components were found by XRD analysis) were selected after a careful evaluation for calibration and validation. These reference standards are unique, and this is the first time that these kinds of samples have been used for building the chemometric PLS method for the analysis of clays using FT-IR. As a result, the developed ATR-FT-IR-PLS method can quantify seven main classes of minerals (silica varieties, feldspars, three groups of clay minerals, carbonates, iron-oxyhydroxides) in different sizes and amounts of clay samples. The performance of the developed quantitative method is characterised as follows: the average performance index of 7 main classes of minerals is 84.8 %, R^2 of the calibration and validation standards is >0.90, RMSEC values of main components varied between 1.7 and 6.1 g/100g and RMSEP values between 0.9 and 5.1 g/100g. We rate these figures of merit as satisfactory to good.

The ATR-FT-IR-PLS method was applied to the 11 (mostly very small) case-study samples from archaeological and cultural heritage objects, covering a large timespan and different geographical origins. The results indicate that the developed method can quantify tiny samples even if only one IR spectrum is available (however, higher uncertainty is expected, often leading to semiquantitative results). In general, the results of all the case-study samples show that the ATR-FT-IR-PLS method quantified quite well the higher contents of minerals like silica varieties (e.g. quartz), feldspars, for some samples also clay minerals 2 (illite, illite-smectite, micas) and surprisingly well also iron-oxyhydroxides. However, the quantified lower content values from the sample should be evaluated critically. Specifically, if the content of a mineral is not higher than two times the standard uncertainty, then it is not possible confidently to say that the component is present in the sample. In order to get adequate results, sample homogeneity is a key aspect. For the ATR-FT-IR-PLS model, the average spectrum from at least three IR spectra of the sample (if the sample size is large enough) is recommended.

Our results show that ATR-FT-IR with the PLS quantification method is a suitable alternative to the XRD if high

accuracy is not needed (in some cases results are to be considered semiquantitative) and especially when the available sample amounts are very low. The developed method is quick and simple and does not have time-consuming sample preparation. However, the accuracy and ability to distinguish individual minerals of similar composition and/or structure are clearly lower in the case of ATR-FT-IR-PLS than in the case of XRD.

The developed quantitative ATR-FT-IR-PLS method can improve by adding new calibration standards into the model. Also, this serves as a baseline model for future research on clay components analysis using ATR-FT-IR spectra, particularly relevant for the analysis of heritage objects where small sample quantities are an essential aspect of the analysis.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00216-021-03617-9.

Acknowledgements This work was carried out using the instrumentation of the Estonian Center of Analytical Chemistry (www.akki.ee). We thank Jaan Aruväli from Department of Geology at University of Tartu for assistance with XRD analysis. For help in sample preparation and ATR-FT-IR measurement, the authors would like to thank Jaak Veske and Kristina Kundla. We would like to show our gratitude to Dr. Jaanika Anderson, Dr. Ivari Kaljurand, Dr. Heili Kasuk, Dr. Uwe Sperling, Dr. Andres Tvauri and Arvi Haak for providing us with reference samples. The authors gratefully acknowledge Tiina Vint from University of Tartu Museum for the brick samples of Tartu Cathedral and conservator Eve Alttoa for the samples of St. John's Church in Tartu and their valuable advice.

Author contribution Signe Vahur: conceptualization, methodology, investigation, formal analysis, validation, visualization, supervision, writing (original draft), writing (review and editing), funding acquisition. Lisett Kiudorv: investigation, formal analysis, visualization, writing (original draft). Peeter Somelar: formal analysis, writing (review and editing), resources, funding acquisition. Jan-Michael Cayme: investigation, writing (review and editing). Mark Dennis Chico Retrato: investigation. Rady Jazmin Remigio: investigation. Varun Sharma: investigation. Ester Oras: writing (review and editing), resources, funding acquisition. Ivo Leito: conceptualization, supervision, writing (review and editing), resources, funding acquisition

Funding This work was supported by Estonian Research Council grant (PUT1521, PUT1511, MOBERC14, PRG1198, PSG492) and by the EU through the European Regional Development Fund (TK141 "Advanced materials and high-technology devices for energy recuperation systems").

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Bitossi G, Giorgi R, Mauro M, Salvadori B, Dei L. Spectroscopic techniques in cultural heritage conservation: a survey. Appl Spectrosc Rev. 2005;40:187–228.
2. Goffer Z. Archaeological chemistry. 2nd ed. Hoboken, N.J: Wiley-Interscience; 2007.
3. Santacreu DA. Materiality, techniques and society in pottery production: the technological study of archaeological ceramics through paste analysis. Warsaw, Berlin: De Gruyter Open Ltd; 2015.
4. Holmqvist E, Larsson ÅM, Kriiska A, Palonen V, Pesonen P, Mizohata K, Kouki P, Räisänen J. Tracing grog and pots to reveal Neolithic Corded Ware Culture contacts in the Baltic Sea region (SEM-EDS, PIXE). J Archaeol Sci. 2018;91:77–91.
5. Darchuk L, Rotondo GG, Swaenen M, Worobiec A, Tsybrii Z, Makarovska Y, Van Grieken R. Composition of prehistoric rock-painting pigments from Egypt (Gilf Kébir area). Spectrochim Acta A Mol Biomol Spectrosc. 2011;83:34–8.
6. Aquilia E, Barone G, Mazzoleni P, Ingoglia C. Petrographic and chemical characterisation of fine ware from three Archaic and Hellenistic kilns in Gela. Sicily J Cult Herit. 2012;13:442–7.
7. Hradil D, Grygar T, Hradilová J, Bezdička P. (2003) Clay and iron oxide pigments in the history of painting. Appl Clay Sci. 2003;22:223–36.
8. Velde B. Origin and mineralogy of clays. Berlin, New York: Springer; 1995.
9. Somelar P, Vahur S, Hamilton TS, Mahaney WC, Barendregt RW, Costa P. Sand coatings in paleosols: evidence of weathering across the Plio-Pleistocene boundary to modern times on Mt. Kenya. Geomorphology. 2018;317:91–106.
10. Liivamägi S, Somelar P, Vircava I, Mahaney WC, Kirs J, Kirsimäe K. Petrology, mineralogy and geochemical climofunctions of the Neoproterozoic Baltic paleosol. Precambrian Res. 2015;256:170–88.
11. Liivamägi S, Somelar P, Mahaney WC, Kirs J, Vircava I, Kirsimäe K. Late Neoproterozoic Baltic paleosol: intense weathering at high latitude? Geology. 2014;42:323–6.
12. Türkmenoğlu AG, Yavuz-Işık N. Mineralogy, chemistry and potential utilization of clays from coal deposits in the Kütahya province, Western Turkey. Appl Clay Sci. 2008;42:63–73.
13. Štyriaková I, Mockovčiaková A, Štyriak I, Kraus I, Uhlík P, Madejová J, Orolínová Z. Bioleaching of clays and iron oxide coatings from quartz sands. Appl Clay Sci. 2012;61:1–7.
14. Savic I, Stojiljkovic S, Savic I, Gajic D. Industrial application of clays and clay minerals. In: Wesley LR, editor. Clays Clay Miner Geol Orig Mech Prop Ind Appl. Hauppauge, New York: Nova Science Publishers, Inc; 2014. p. 379–402.
15. Aguzzi C, Cerezo P, Viseras C, Caramella C. Use of clays as drug delivery systems: possibilities and limitations. Appl Clay Sci. 2007;36:22–36.
16. Sposito G, Skipper NT, Sutton R, Park S, Soper AK, Greathouse JA. Surface geochemistry of the clay minerals. Proc Natl Acad Sci. 1999;96:3358–64.
17. Mana SCA, Hanafiah MM, Chowdhury AJK. Environmental characteristics of clay and clay-based minerals. Geol Ecol Landsc. 2017;1:155–61.
18. Wilson MJ. A handbook of determinative methods in clay mineralogy. Glasgow: Blackie [u.a.]; 1987.
19. Mukherjee S, Ghosh B. The science of clays: applications in industry, engineering and environment. Dordrecht: Springer; 2013.
20. Kriiska A. (1996) The Neolithic pottery manufacturing technique of the lower course of the Narva River. In: Hackens T, Hicks S, Lang V, Miller U, Saarse L, editors. Coastal Est Recent Adv Environ Cult Hist. Strasbourg: 1996 Rixensart PACT; 1996.
21. Feathers JK. Explaining shell-tempered pottery in prehistoric eastern North America. J Archaeol Method Theory. 2006;13:89–133.
22. Moskal-del Hoyo M, Rauba-Bukowska A, Lityńska-Zając M, Mueller-Bieniek A, Czekaj-Zastawny A. Plant materials used as temper in the oldest Neolithic pottery from south-eastern Poland. Veg Hist Archaeobotany. 2017;26:329–44.
23. Daghmehchi M, Rathossi C, Omrani H, Emami M, Rahbar M. Mineralogical and thermal analyses of the Hellenistic ceramics from Laodicea Temple, Iran. Appl Clay Sci. 2018;162:146–54.

24. Castellanos O, Rios C, Ramos M, Plaza E. A comparative study of mineralogical transformations in fired clays from the Laboyos Valley, Upper Magdalena basin (Colombia). Boletin Geol. 2012;34:43–55.
25. El Ouahabi M, Daoudi L, Hatert F, Fagel N. Modified mineral phases during clay ceramic firing. Clay Clay Miner. 2015;63: 404–13.
26. Moropoulou A, Bakolas A, Bisbikou K. Thermal analysis as a method of characterizing ancient ceramic technologies. Thermochim Acta. 1995;269–270:743–53.
27. Liu H, Chen T, Zou X, Qing C, Frost RL. Thermal treatment of natural goethite: thermal transformation and physical properties. Thermochim Acta. 2013;568:115–21.
28. Maggetti M, Neururer C, Ramseyer D. Temperature evolution inside a pot during experimental surface (bonfire) firing. Appl Clay Sci. 2011;53:500–8.
29. Maggetti M. Phase analysis and its significance for technology and origin. In: Olin JS, Franklin AD, editors. Archaeol Ceram. Washington: Smithsonian Institution Press; 1982. p. 121–33.
30. Bish DL, Howard SA. Quantitative phase analysis using the Rietveld method. J Appl Crystallogr. 1988;21:86–91.
31. Zhou X, Liu D, Bu H, Deng L, Liu H, Yuan P, Du P, Song H. XRD-based quantitative analysis of clay minerals using reference intensity ratios, mineral intensity factors, Rietveld, and full pattern summation methods: a critical review. Solid Earth Sci. 2018;3:16–29.
32. Hahn A, Vogel H, Andó S, Garzanti E, Kuhn G, Lantzsch H, Schüürman J, Vogt C, Zabel M. Using Fourier transform infrared spectroscopy to determine mineral phases in sediments. Sediment Geol. 2018;375:27–35.
33. Thickett D, Pretzel B. FTIR surface analysis for conservation. Herit Sci. 2020;8:5.
34. Prati S, Sciutto G, Bonacini I, Mazzeo R. New frontiers in application of FTIR microscopy for characterization of cultural heritage materials. Top Curr Chem. 2016;374:26.
35. Madariaga JM. Analytical chemistry in the field of cultural heritage. Anal Methods. 2015;7:4848–76.
36. Vahur S, Eero L, Lehtaru J, Virro K, Leito I. Quantitative non-destructive analysis of paper fillers using ATR-FT-IR spectroscopy with PLS method. Anal Bioanal Chem. 2019;411:5127–38.
37. Hayes PA, Vahur S, Leito I. ATR-FTIR spectroscopy and quantitative multivariate analysis of paints and coating materials. Spectrochim Acta A Mol Biomol Spectrosc. 2014;133:207–13.
38. Wiesinger R, Pagnin L, Anghelone M, Moretto LM, Orsega EF, Schreiner M. Pigment and binder concentrations in modern paint samples determined by IR and Raman spectroscopy. Angew Chem Int Ed. 2018;57:7401–7.
39. Jordan MM, Jorda J, Pardo F, Montero MA. Mineralogical analysis of historical mortars by FTIR. Materials. 2019;12:55.
40. Anderson E, Almond MJ, Matthews W. Analysis of wall plasters and natural sediments from the Neolithic town of Çatalhöyük (Turkey) by a range of analytical techniques. Spectrochim Acta A Mol Biomol Spectrosc. 2014;133:326–34.
41. Vahur S, Teearu A, Peets P, Joosu L, Leito I. ATR-FT-IR spectral collection of conservation materials in the extended region of 4000-80 cm^{-1}. Anal Bioanal Chem. 2016;408:3373–9.
42. Heath C, Pejcic B, Delle Piane C, Esteban L. Development of far-infrared attenuated total reflectance spectroscopy for the mineralogical analysis of shales. Fuel. 2016;182:771–9.
43. Kontopoulos I, Presslee S, Penkman K, Collins MJ. Preparation of bone powder for FTIR-ATR analysis: the particle size effect. Vib Spectrosc. 2018;99:167–77.
44. Smith BC. Fundamentals of Fourier transform infrared spectroscopy. 2nd ed. Boca Raton, etc: CRC Press; 2011.
45. Madejová J, Komadel P. Baseline studies of the Clay Minerals Society Source Clays: infrared methods. Clay Clay Miner. 2001;49:410–32.
46. Madejová J. FTIR techniques in clay mineral studies. Vib Spectrosc. 2003;31:1–10.
47. Doménech-Carbó A, Bosch-Reig F, Montoya N. ATR-FTIR and XRD quantification of solid mixtures using the asymptotic constant ratio (ACR) methods. Application to geological samples of sodium and potassium feldspars. Spectrochim Acta A Mol Biomol Spectrosc. 2020;236:118328.
48. TQ Analyst Software User Guide. Madison: Thermo Fisher Scientific Inc; 2008.
49. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst. 2001;58:109–30.
50. Garson GD. Partial least squares: regression and structural equation models. Asheboro, NC, USA: Statistical Publishing Associates; 2016.
51. Ge Y, Thomasson JA, Morgan CLS. Mid-infrared attenuated total reflectance spectroscopy for soil carbon and particle size determination. Geoderma. 2014;213:57–63.
52. Mueller CM, Pejcic B, Esteban L, Delle Piane C, Raven M, Mizaikoff B. Infrared attenuated total reflectance spectroscopy: an innovative strategy for analyzing mineral components in energy relevant systems. Sci Rep London: Nature Publishing Group. 2014;4:6764.
53. Mroczkowska-Szerszeń M, Orzechowski M. Infrared spectroscopy methods in reservoir rocks analysis – semiquantitative approach for carbonate rocks. Nafta-Gaz. 2018;74:802–12.
54. Ma F, Du CW, Zhou JM, Shen YZ. Investigation of soil properties using different techniques of mid-infrared spectroscopy. Eur J Soil Sci. 2019;70:96–106.
55. Ritz M, Vaculíková L, Plevová E, Matýsek D. Determination of chlorite, muscovite, albite and quartz in claystones and clay shales by infrared spectroscopy and partial least-squares regression. Acta Geodyn Geomater. 2012;9:511–20.
56. Ritz M, Vaculíková L, Plevová E, Matýsek D, Mališ J. Determination of the predominant minerals in sedimentary rocks by chemometric analysis of infrared spectra. Clay Clay Miner. 2012;60:655–65.
57. Vercruysse K, Grabowski RC. Using source-specific models to test the impact of sediment source classification on sediment fingerprinting. Hydrol Process. 2018;32:3402–15.
58. Oras E, Anderson J, Tõrv M, Vahur S, Rammo R, Remmer S, Mölder M, Malve M, Saag L, Saage R, Teearu-Ojakäär A, Peets P, Tambets K, Metspalu M, Lees DC, Barclay MVL, Hall MJR, Ikram S, Piombino-Mascali D. Multidisciplinary investigation of two Egyptian child mummies curated at the University of Tartu Art Museum, Estonia (Late/Graeco-Roman Periods). PLoS One. 2020;15(1):e0227446.
59. Taylor JC. Computer programs for standardless quantitative analysis of minerals using the full powder diffraction profile. Powder Diffract. 1991;6:2–9.
60. Peets P, Leito I, Pelt J, Vahur S. Identification and classification of textile fibres using ATR-FT-IR spectroscopy with chemometric methods. Spectrochim Acta A Mol Biomol Spectrosc. 2017;173: 175–81.
61. Rinnan Å, van den Berg F, Engelsen SB. Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal Chem. 2009;28:1201–22.
62. Peets P, Kaupmees K, Vahur S, Leito I. Reflectance FT-IR spectroscopy as a viable option for textile fiber identification. Herit Sci. 2019;7:93.
63. Willmott CJ, Robeson SM, Matsuura K. A refined index of model performance. Int J Climatol. 2012;32:2088–94.
64. Stuart B. Infrared spectroscopy: fundamentals and applications. Chichester, West Sussex, England. J. Wiley: Hoboken, NJ; 2004.
65. Leito I, Helm I, Jalukse L. Using MOOCs for teaching analytical chemistry: experience at University of Tartu. Anal Bioanal Chem. 2015;407:1277–81.

66. Kalm V, Kriiska A, Aruväli J. Mineralogical analysis applied in provenance studies of Estonian Neolithic pottery. Proc Est Acad Sci Geol. Tallinn, Estonia: Estonian Academy Publishers; 1997. p. 17–33.
67. Kriiska A, Oras E, Lõugas L, Meadows J, Lucquin A, Craig OE. Late Mesolithic Narva stage in Estonia: pottery, settlement types and chronology. Est J Archaeol. 2017;21:52.
68. Saag L, Varul L, Scheib CL, Stenderup J, Allentoft ME, Saag L, Pagani L, Reidla M, Tambets K, Metspalu E, Kriiska A, Willerslev E, Kivisild T, Metspalu M. Extensive farming in Estonia started through a sex-biased migration from the steppe. Curr Biol. 2017;27:2185–2193.e6.
69. Kriiska A. Noorem kiviaeg (3900–1750 eKr). In: Lang V, Ratas J, editors. Eesti ajalugu I. Eesti esiaeg. Tartu Ülikooli Ajaloo- ja Arheoloogia Instituut: Tartu, Estonia; 2020.
70. Bernotas R. Brick-making in medieval Livonia - the Estonian example. Est J Archaeol. 2013;17:139.