Geophysical Journal International

REPRODUCIBILITY OF ARCHAEOINTENSITY DETERMINATIONS WITH A MULTIMETHOD APPROACH ON ARCHAEOLOGICAL MATERIAL REPRODUCTIONS

Journal:	*Geophysical Journal International*
Manuscript ID	GJI-19-0260.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Calvo-Rathert, Manuel; Universidad de Burgos, Departamento de Física Morales-Contreras, Juan; Universidad Nacional Autonoma de Mexico, Laboratorio Interinstitucional de Magnetismo Natural Carrancho, Ángel; Universidad de Burgos, Departamento de Historia, Geografía y Comunicación Camps, Pierre; Universite de Montpellier, Géosciences Montpellier Goguitchaichvili, Avto; Universidad Nacional Autonoma de Mexico, Laboratorio Interinstitucional de Magnetismo Natural Hill, Mimi; University of Liverpool School of Environmental Sciences
Keywords:	Archaeomagnetism < GEOMAGNETISM and ELECTROMAGNETISM, Palaeointensity < GEOMAGNETISM and ELECTROMAGNETISM, Magnetic mineralogy and petrology < GEOMAGNETISM and ELECTROMAGNETISM

SCHOLARONE™
Manuscripts

REPRODUCIBILITY OF ARCHAEOINTENSITY DETERMINATIONS WITH A MULTIMETHOD APPROACH ON ARCHAEOLOGICAL MATERIAL REPRODUCTIONS

Manuel Calvo-Rathert[1], Juan Morales Contreras[2], Ángel Carrancho[3], Pierre Camps[4], Avto Goguitchaichvili[2], Mimi J. Hill[5]

[1]Departamento de Física, EPS, Universidad de Burgos, Av. Cantabria s/n, 09006 Burgos, Spain
[2]Laboratorio Interinstitucional de Magnetismo Natural, UNAM, Campus Morelia, 58098, Mexico
[3]Departamento de Historia, Geografía y Comunicación, Universidad de Burgos, 09001 Burgos, Spain
[4]Géosciences Montpellier, CNRS and Univ. Montpellier, 34095 Montpellier Cedex 5, France.
[5]Geomagnetism Laboratory, School of Environmental Sciences, University of Liverpool, UK

Abstract

Archaeointensity determinations on burnt archaeological material are complex and reliable data scarce, although this kind of material can be of great interest in archaeological investigations. With the goal of analysing the reliability of archaeointensity determinations, an interlaboratory comparison study has been performed combining different experimental protocols on present-day reproductions of Pre-Columbian Mesoamerican archaeological artifacts and two brick samples. Samples were baked in an original kiln from an artisan workshop in western Mexico. The ambient magnetic field at the site during the experiment was measured and continuous temperature data were recorded at four different positions in the kiln during the heating-cooling procedure.

Archaeointensity determinations were carried out with four different methods at four different palaeomagnetic laboratories: Thellier-Coe (Burgos, Spain), microwave (Liverpool, U.K.), multispecimen (Morelia, Mexico) and multispecimen with the extended protocols for fraction and domain-state correction (Montpellier, France). 26 conventional resistive heating determinations with the Thellier-Coe protocol yielded a 100% success rate, while 7 out of 8 microwave-heating determinations with the Thellier-Coe protocol also provided successful results. Also, two multispecimen determinations performed with both multispecimen methods provided statistically reliable results. In all cases, a good agreement between the determined archaeointensities and the ambient field at the production site could be observed.

Highly reversible magnetisation-versus-temperature curves yielded slightly Al, Mg or Ti-substituted magnetite as the main ferromagnetic (*s.l.*) phase. In addition, in several samples, a thermally stable low Curie-temperature phase displaying a high coercivity behaviour could be observed in thermomagnetic curves and by thermal demagnetisation of saturation isothermal remanent magnetisation (SIRM). This phase is interpreted as ε-Fe_2O_3. To our knowledge, its occurrence has never been reported through the experimental recreation of burnt archaeological materials. No correlation could be observed between the proxies of domain-state behaviour and deviation of palaeointensity determinations from the expected result.

Results obtained on clay samples heated in this type of ancient kiln can be considered a good source for determining the geomagnetic field strength variation in the past. Matching palaeointensity results obtained with different methods based on different principles can be taken as a quality criterion for result reliability and consistency.

1. Introduction

Heated archaeological material is an important source of information about geomagnetic secular variation beyond the historical record as it can register a thermoremanent magnetisation (TRM) parallel to the direction and proportional to the strength of the ambient magnetic field, which is usually the Earth's magnetic field at the time of its last heating/cooling. Although many artefacts like potsherds, bricks or tiles may have been fired in an unknown position, these archives of the ancient field nevertheless retain information on its intensity that can be retrieved by means of different experimental techniques.

However, the determination of the palaeointensity is experimentally much more difficult than the determination of the palaeofield vector direction. Several different methods have been proposed so far, but those based on the original Thellier method (Thellier and Thellier, 1959) are considered the most reliable ones, as they are based on a rigorous physical background. In Thellier-type experiments several requirements have to be fulfilled in order to be able to provide a reliable palaeointensity determination: (i) Remanence must be a TRM; (ii) Samples must obey the Thellier laws of reciprocity, independence and additivity of partial TRMs (pTRMs) (Thellier and Thellier, 1959), a condition which is fulfilled by non-interacting single-domain (SD), but not multi-domain (MD) particles (e.g., Dunlop, 2011, and references therein); (iii) Sample remanence must be stable. During heating, irreversible chemical/mineralogical or physical changes (e.g., Kosterov and Prévot, 1998) can affect magnetic phases, resulting in spurious palaeointensity estimates. Therefore, the failure rate of palaeointensity experiments can often be large and, in addition, the scatter observed in palaeointensity (or archaeointensty) results is much higher than in directional results, which is often related to the fact that incorrect determinations are not detected because they pass through the selection filters (e.g., Calvo et al., 2002).

Some different methods involving different protocols and different physical types of heating have been proposed to avoid or lessen problems related to the presence of MD grains or chemical/mineralogical alterations in specimens subjected to palaeointensity experiments. The so-called microwave method is a Thellier-type protocol in which the laboratory heatings involve electromagnetic waves and heat transfer. The main difference with resistive-heating lies in the fact that when a sample is subjected to microwave demagnetisation, most of the energy is absorbed by the magnetic system, the bulk sample not being heated significantly. In addition, microwave application takes place only for seconds (usually 5-10s) as opposed to much longer times in conventional heating. For these reasons, the probability of alteration during palaeointensity experiments can be reduced (e.g., Hill and Shaw, 1999).

Dekkers and Böhnel (2006) developed the multispecimen (MS) protocols from a theoretical model proposed by Biggin and Poidras (2006), in which a pTRM is imparted to a sub-specimen taken from a sample in a direction parallel to NRM at a specific temperature and a chosen field. Subsequently the experiment is repeated at the same temperature but at different fields on other sub-specimens of the same sample. With this method, palaeointensity should be independent of domain structure, as it would eliminate magnetic history effects. Alteration would also be reduced, because specimens are heated only once at temperatures below those producing significant alterations. Fabian and Leonhardt (2010), however, questioned the Biggin and Poidras model, claiming that this method might produce systematic palaeointensity overestimates on samples containing MD grains. This has been the case for lavas containing a significant MD fraction as reported by Michalk et al. (2008; 2010) and Calvo-Rathert et al. (2016). From new theoretical inferences Fabian and Leonhardt (2010) included some correction steps in the MS measurement protocol to avoid this palaeointensity overestimation.

In order to successfully retrieve an accurate archaeointensity result, it is necessary that the method applied for palaeointensity determination has not produced physical, chemical or mineralogical alterations inadvertedly yielding incorrect archaeointensity results. To analyse the reliability of archaeointensity determinations, it is of interest to perform archaeointensity experiments under controlled conditions and combining different experimental protocols. We promote the implementation of such an approach with palaeointensity experiments performed on archaeological baked clays specifically manufactured for the experiment that acquired a remanent magnetisation in a known field. Additional rock-magnetic data can also provide useful information and constraints regarding the success or failure of the archaeointensity experiments.

Specific studies which aim to relate the accuracy, quality and reliability of palaeointensity determinations obtained from materials of archaeological interest to the characteristics of the applied experimental procedures are nonetheless still rather scarce, especially if methods other than Thellier-type ones are considered.

An archaeomagnetic quality control test was conducted by Catanzariti et al. (2008) in a partially heated brick kiln from 1959. Using the classical Thellier palaeointensity method (Thellier and Thellier, 1959) they obtained results consistent with the known field value. Morales et al. (2011) studied rock-magnetic properties and the palaeointensity of in situ manufactured ceramic and bricks with the Thellier-Coe method (Coe, 1967) and with a TRIAXE magnetometer (Le Goff and Gallet, 2004), observing a good agreement with the field at the manufacturing site. However, they also point out the significant scatter which can be observed in archaeointensity determinations even from pieces fired together in the same oven, depending on their position. Nakajima et al. (1974) reconstructed a kiln imitating a seventh century one to measure palaeomagnetic directions from baked clay samples taken from the kiln. Yamamoto et al. (2015) performed a palaeointensity study on these samples applying the Tsunakawa-Shaw method with anisotropy correction but no cooling-rate correction. They obtained results consistent with the in situ geomagnetic field on kiln floor samples, but not on samples at a 20 cm level above, apparently due to the acquisition of only partial TRMs. The MS method has only been tested in a few studies on archaeological materials. Carrancho et al. (2014) performed a rock-magnetic and archaeointensity study on clasts of different lithologies (chert, quartzite, limestone, sandstone and obsidian) heated under controlled field and temperature conditions to estimate the feasibility of these raw materials, which are commonly found in prehistoric archaeological sites for archaeomagnetic purposes. Application of the MS palaeointensity technique was successfully applied to obsidian and sandstone specimens yielding a field estimation in agreement with the expected one. Schnepp et al. (2016) performed an archaeomagnetic and rock magnetic investigation on an experimental pottery kiln, carrying out archaeointensity experiments with both the Thellier–Coe and the MS domain-state corrected method. In both cases, accurate intensity estimations within their standard deviations were obtained. In a paleointensity study performed by Calvo-Rathert et al. (2016) on historical lava flows from the island of Lanzarote (Canary Islands, Spain) with the Thellier-Coe and the MS method, expected values or moderately lower ones were obtained with the former method, but a large deviation from the expected result in one case with the latter one. The microwave method was applied on samples of archaeological interest magnetised in a known field together with the Thellier-Coe method by Calvo-Rathert et al. (2012) in an experiment devised to reproduce the prehistoric use of fire on a clayish soil substratum. Results were in reasonable agreement with the expected field value.

It is also interesting to note that, since different palaeointensity determination methods are based on different experimental procedures, which depend on the attainment of different energy equilibrium states related to temperature, applied field and demagnetising field at all heating steps. Hence, consistency of results obtained with procedures relying on distinct physical principles can be considered a way to strengthen the reliability of palaeointensity determinations (e.g. Böhnel et al., 2009; De Groot et al., 2013, 2015; Enterpinar et al., 2016; Monster et al., 2015; Calvo-Rathert et al., 2016). Accordingly, a multimethod palaeointensity study on archaeological material heated and magnetised under controlled conditions is of interest for future archaeointensity determinations following a similar approach.

In this study we used samples from ceramic materials manufactured and baked in an original style open kiln by an artisan workshop from the town of Zinapécuaro (Michoacán, Mexico). The workshop was founded in 1815 and uses most of the local ancestral manufacturing procedures. In fact, it is authorised by the National Institute of Anthropology and History of Mexico (INAH) to produce reproductions of local archaeological items. A preliminary rock-magnetic and synthetic archaeointensity study had been already performed by the same research group on *in situ* manufactured ceramic and bricks (Morales et al., 2011). During that experiment, a single thermocouple had been placed in the middle of the cavity to monitor heating temperatures in the kiln. In this new and improved version of the experiment four thermocouples were placed in the same furnace to simultaneously record the temperature at different positions. In addition to reproductions of archaeological samples, two bricks previously manufactured and baked at another place were introduced into the furnace and exposed to the same heating procedure. The latter were thus subjected to a second new heating and acquisition of TRM.

In the original experiment, samples were only subjected to the Thellier-Coe method (Thellier and Thellier, 1959; Coe, 1967) and to an alternative palaeointensity experiment with a TRIAXE magnetometer (Le Goff and Gallet, 2004). In the new approach, four archaeointensity determination experiments were carried out independently on specimens from the same samples at four different palaeomagnetic laboratories: At the palaeomagnetic laboratory of the University of Burgos (Spain) a Thellier-type double heating experiment (Thellier and Thellier, 1959) as modified by Coe (1967) was performed, while a microwave archaeointensity determination with the Thellier-Coe protocol was carried out at the palaeomagnetic laboratory of the University of Liverpool (UK). In addition, archaeointensity experiments with the MS method were performed on two single selected samples of the manufactured set both at the palaeomagnetic laboratory of UNAM in Morelia (Mexico) and the palaeomagnetic laboratory of Géosciences Montpellier (France). At UNAM, the original MS method as proposed by Dekkers and Böhnel (2006) was used by means of a resistive-heating furnace, whilst in Montpellier the extended MS method including protocols for fraction and domain-state correction (Fabian and Leonhardt, 2010) was applied by means of an infrared-heating furnace. The heat transfer to the sample is achieved by means of two different physical process, convection plus radiation at high temperatures in the former, mainly by radiation with a small part of conduction in the latter. This kind of interlaboratory comparison is an advantageous way of assuring quality control among the different participating laboratories, allowing them to detect problems or deficiencies in their applied methodology, because despite comparing different methods, the same results should be obtained in all cases. The fact that the external conditions (magnetic field strength, temperature, duration of heating and cooling) giving rise to the analysed signal (remanent magnetisation) were known, allows to estimate the precision and reliability of palaeointensity determinations obtained with different protocols and experimental setups.

2 Experimental setups

The kiln used to bake the archaeological reproductions is shown in Fig. 1a. It is a circular, 100 cm wide open clay structure made up of two chambers, the lower 20 cm-high burning cavity and the upper 60 cm high open baking compartment. The kiln floor is built of clay blocks and covered with potsherds coming from broken or defective pieces.

Heating of the samples in the kiln was carried out in 2010. Once modelled and sun-dried for several hours, the raw pieces of ceramic were placed into the baking chamber. These pieces included vessels (sample L), flowerpots (samples M) and zoomorphic vessels (sample N). In addition, two bricks which had been previously baked in 2010 in the artisan workshop were also included in the experiment. One (sample LQ) was put into the oven and another one (sample LN) was not heated in the kiln but subjected to paleointensity experiments for comparison. Four thermocouples were placed at different positions in the baking compartment of the oven (Fig. 1b). Thermocouple T1 was placed in the middle of the kiln, near the bottom of the baking cavity. Thermocouples T2 and T4 were positioned nearer to the oven's rim at different heights (T2 at 16 cm from the bottom and T4 at 16 cm from the top). Thermocouple T3 was placed near the centre of the oven, but near its top. This latter thermocouple cannot be seen in its final position in Fig. 1b, as it was placed on a horizontal clay disk which partially covered other pieces in the oven and is not shown in the figure for the sake of clarity. Temperature was first increased up to approximately 100 °C and maintained at that value for approximately one hour to eliminate the remaining water in the clay. Subsequently, during the next four hours, the temperature of the oven was augmented until a maximum temperature above 700°C was reached in the middle of the kiln, near the bottom (thermocouple 1) and temperatures near or above 650°C in other parts of the oven (Fig. 2). Unfortunately, one of the thermocouples (T2) stopped working after approximately 150 minutes of heating. Finally, the oven cooled down naturally over approximately three hours. The maximum temperature reached in the lower central part of the kiln thus exceeded the Curie temperature (T_c) of hematite, however in other parts of the kiln this temperature is nearly, but not completely reached. The Curie-temperature of magnetite, on the other hand, seems to be exceeded in all parts of the kiln. It is interesting to note that during regular heating procedures no temperature measurements are performed in the kiln, and the temperatures believed by the artisans to be reached were much higher than the actual ones.

The field strength at the experiment site was measured with a MEDA µMAG-01N Fluxgate Magnetometer in 2011, one year after the experiment, obtaining an averaged value of 40.5 ± 0.5 µT (Tab. 1). This value is consistent with the data retrieved from the Coeneo magnetic observatory in 2011, which is located 100 km west from the site. Using model IGRF12 (Thébault et al., 2015) for calculation of the Earth's magnetic field intensity at the same location in 2010 and 2011 yields a difference of 0.22%. Direct field measurement inside the kiln yielded 40.7 μT in the upper and 39.9 μT in the lower centre of the kiln (Tab. 1). It is therefore concluded that no significant magnetic anomaly is observed at the experiment site.

3. Rock-magnetic properties of the samples

Rock-magnetic experiments have been performed to obtain knowledge about the magnetic properties of the studied bricks and ceramics as well as of the clay used for preparing the ceramic paste. This information allows the magnetic characterisation of the studied materials by determining the nature of their remanence carriers. It is also useful to gain insight regarding their thermal stability and grain size, as this can be used as a criterion to appraise the suitability of the studied samples for palaeointensity determinations. Experiments carried out include the measurement of strong-field (38 mT) magnetisation versus temperature (M_S-T) curves, the determination of hysteresis parameters and the recording of isothermal remanent magnetisation (IRM) acquisition curves. All were carried out at the palaeomagnetic laboratory of the University of Burgos (Spain) with a Variable Field Translation Balance (VFTB) on whole-rock powdered samples from all archaeological reproductions and brick samples used for archaeointensity experiments, as well as on a specimen of the original clay mixture for the archaeological reproductions. Artisans usually use two or three clay varieties obtained from different sites located within 3 to 8 km from the town (Rojas-Navarrete, 1995) to prepare the ceramic paste. At the workshop, the different clay varieties are dried in the sun and subsequently pulverised and sieved. Finally, they are mixed in different proportions and water is added until a homogeneous paste with the desired characteristics is obtained.

The measurement sequence performed by the balance was the following: (i) IRM acquisition, (ii) hysteresis curve, (iii) back-field and (iv) strong-field magnetisation versus temperature (M_S-T) curve. In stepwise IRM acquisition a maximum field of approximately 1T was applied. Hysteresis parameters were determined from hysteresis and backfield curves after correction for the dia- and paramagnetic contribution. Thermomagnetic M_S-T curves were recorded heating samples in air up to 600 or 700°C and cooling them down to room temperature with heating/cooling rates of 20 or 30°C/min. Before starting the thermomagnetic curve record, the sample is subjected to a 1T field, acquiring a (near) saturation magnetisation. Data were analysed with the RockMagAnalyzer 1.0 software (Leonhardt, 2006).

Curie temperatures (T_C) were determined from M_S-T curves with the two-tangent method (Grommé et al., 1969). The M_S-T curve of the original clay sample (Fig. 3a) displays a basically paramagnetic behaviour. It is interesting to note that this sample shows a rather high degree of thermomagnetic reversibility. All baked archaeological reproductions were made from the same material but were positioned at different places in the kiln, thus experiencing different maximum temperatures. Nevertheless, all display a very similar thermomagnetic behaviour, with Curie temperatures between 520 and 560°C and showing a high degree of reversibility (Fig. 4a, 4b). This phase can be interpreted as slightly Al, Mg or Ti-substituted magnetite. Sample NLE (Fig. 4c) shows less reversibility, which might be explained by the fact that it is heated to a higher temperature than most other samples (700°C instead of 600°C). As samples have already been heated to similar temperatures for several hours during the experimental heating in the kiln, either they still have not reached thermo-chemical equilibrium or oxygen and carbon-dioxide partial pressure might be different in the kiln and in the VFTB-furnace. The brick samples (Fig. 4b) also show a curve type very similar to ceramic samples (Fig. 4a). It is interesting to note that heating of the original clay sample during the thermomagnetic experiment (Fig. 3a) does not produce changes in its magnetic mineralogy, generating a similar composition to that of the archaeological pieces, which were obtained from the same clay material after heating in the kiln. In order to check if heating time would have a noticeable effect on the magnetic properties of the original clay material, it was heated during two hours in the laboratory furnace (still less time than the heating procedure in the kiln, but much longer than the approximately 30 minutes heating time in the VFTB-furnace). This procedure generated a near magnetite phase (Fig. 3b).

In several cases, including the brick samples, a tiny inflection can be detected in the heating and cooling curves from the thermomagnetic experiments between 100 and 250°C (Fig. 4a). Although some IRM acquisition curves show a strong predominance of low-coercivity phases, , in many other cases a strong coercivity phase can be observed (Fig. 5). Thus, although low-coercivity phases – probably the Al, Mg or Ti-substituted magnetite phase observed in thermomagnetic curves – can be recognized in all samples, a high coercivity phase is also present. The simultaneous observation of a thermally stable low Curie-temperature phase and high coercivity behaviour points to the presence of the phase observed in well-heated archaeological material and reported by McIntosh et al. (2007). This phase was termed by the authors HCSLT (high coercivity, thermally stable, low Curie temperature) phase and has been documented in several archaeological features from different parts of the world (López-Sánchez et al., 2017 and references therein). This mineral has been interpreted as epsilon iron oxide ε-Fe_2O_3 (e.g., Lee and Xu, 2018) by means of Confocal Raman Spectroscopy and rock-magnetic measurements by López-Sánchez et al. (2017).

In order to confirm the presence of this HCSLT phase, a supplementary experiment was performed: Specimens from all samples showing a high coercivity fraction and the previously mentioned thermally stable tiny low Curie-temperature phase were imparted an IRM in a strong 2T field along their z-axis. Subsequently, all these specimens were subjected to alternating-field (AF) demagnetisation up to 100mT, removing between 30 and 60% of the previously acquired IRM. Finally, the remaining remanence was stepwise thermally demagnetised. In all specimens a clear inflection can be recognised at 200-240°C (Fig.6a, b). In most cases 65 to 80% of the IRM remaining after the 100mT AF-demagnetisation has been removed at this temperature, and only less than 5% of this remanence remains in the samples after heating to 556 or 587°C. Brick samples, however, display a somewhat different behaviour (Fig. 6b). They also show a noticeable inflection at 200-240°C but losing only 35 to 50% of the IRM remaining after AF-demagnetisation. In addition, a significant part of the remanence is only removed at temperatures above 600°C. Thus, the presence of the HCSLT phase is confirmed in all these specimens, although in the brick samples it seems to coexist with another high-coercivity phase, apparently hematite.

Measurement of hysteresis and backfield curves allowed determination of hysteresis parameters such as M_S (saturation magnetization), M_{RS} (saturation remanence), B_C (coercivity) and B_{CR} (coercivity of remanence). When hysteresis parameter ratios are displayed in a Day-plot (Day et al., 1977) most show a PSD-like (pseudo-single-domain) behaviour (Fig. 7a) which can also be interpreted as due to a mixture of single-domain (SD) and multi-domain (MD) grains (Dunlop, 2002). Nevertheless, it should be borne in mind that interpretation of data plotted in a Day diagram in terms of domain state diagnosis might be highly ambiguous, because hysteresis parameter ratios may be affected by several conditions such as magnetic mineralogy, mineral stoichiometry, internal stress, magnetostatic interactions or magnetic particle mixtures, among others (Roberts et al., 2018). Nevertheless, in the following lines a qualitative interpretation is attempted, taking into account that the studied samples contain magnetic particle mixtures of different coercivity. Comparison with theoretical mixing curves for magnetite (Dunlop, 2002) shows that most samples lie in a field between SD-MD and superparamagnetic (SP)-SD mixing curves. However, as suggested by thermomagnetic curves and the IRM demagnetisation experiment described above, at least a part of the analysed samples contain a mixture of low-coercivity and high coercivity minerals. The shift in the day-plot of some of the samples towards higher B_{CR}/ B_C ratios and intermediate M_{RS}/ M_S values might be explained with this mixing. In this mixture, B_C would be largely controlled by the low-coercivity component, while B_{CR} would be controlled by the high-coercivity component, yielding a higher B_{CR}/ B_C ratio (Wasilewski, 1973; Roberts et al., 1995). The M_{RS}/ M_S ratio, on the other hand, obeys the following relationship: M_{RS}/ M_S (low coercivity composition) < M_{RS}/ M_S (mixture) < M_{RS}/ M_S (high coercivity composition) (Wasilewski, 1973). In fact, specimens from brick samples LQ and LN, which have shown to contain the HCSLT phase, display the most pronounced shift from the SD-MD mixing curve in the Day plot (Fig. 7a). The SD-MD mixing-curve sector nearest to the Zinapécuaro samples yields a relative MD content in the mixture varying between approximately 40% and 80%.

Assemblages of various magnetic components with different mineralogy or grain size may result in specific shapes of hysteresis loops [e.g., Roberts et al., 1995; Muttoni, 1995, Tauxe et al., 1996], which can be quantified by shape parameter $\sigma_{HYS.}$ Wasp-waisted loops have $\sigma_{HYS} > 0$ and pot-bellied loops $\sigma_{HYS} < 0$. In the present study, σ_{HYS} showed a large scatter, with most specimens displaying wasp-waisted loops (Fig. 7b) with positive σ_{HYS} values varying between 0 and 1.2, reflecting the mixture of low and high coercivity magnetic components observed in these samples.

4 Archaeointensity experiments and results

4.1 Thellier-Coe method (TC)

Archaeointensity determinations by means of the Thellier type double heating method (Thellier and Thellier, 1959) as modified by Coe (1967) were carried out at the University of Burgos (Spain). The experiments were performed on 19 unoriented small cylindrical specimens (0.9 cm diameter and 1 to 2 cm length) taken from different artisanal pieces (vessel L, flowerpot M, zoomorphic vessel N) and from two bricks (LQ and LN). In addition, 7 flowerpot specimens (R4) baked in 2009, in the preceding synthetic archaeointensity experiment mentioned above (Morales et al., 2011) were also included in the palaeointensity determination experiment. They were baked in the same kiln and the measured field values at the site (41.0 ± 0.5) μT and inside the kiln (40.3 ± 0.5) μT show an excellent agreement with those measured for the present study. Inclusion of these seven specimens R4 allows an interlaboratory comparison of the Thellier-Coe experiments performed in the Morelia laboratory by Morales et al. (2011), and the Thellier-Coe results obtained in the present study in the Burgos laboratory. All samples were subjected to heating and cooling cycles in an ASC TD-48 palaeointensity oven under argon atmosphere for preventing (or at least minimising) oxidation. After reaching the peak temperature, this maximum temperature was kept constant for about 10 minutes and subsequently the oven was turned off and the samples cooled down naturally over several hours, depending on the heating temperature. In-field steps were performed leaving the laboratory field switched on during the whole cycle. The palaeointensity determination was carried out in 11 temperature steps between room temperature and 581°C, a temperature at which the natural remanent magnetisation (NRM) left of most of the specimens was less than 3%. The temperature reproducibility between heating runs to the same temperature was within 2°C. The laboratory field intensity was set to 40 μT (chosen to fit the expected palaeointensity value) and it was held at a precision better than 0.1μT. During the experiment, several control heating cycles were performed: Six pTRM-checks (Coe, 1967) and six PTRM tail-checks (Riisager and Riisager, 2001). Remanence was measured with a 2G cryogenic magnetometer. Data obtained were interpreted with the ThellierTool4.0 software (Leonhardt et al., 2004) to determine archaeointensity results.

The reliability of the archaeointensity results depends on different factors regarding the quality of experimental conditions, the occurrence of chemical and/or mineralogical alterations and the presence of a remanence fraction carried by MD grains. Different parameters and reliability criteria have been proposed to assess and quantify the degree of reliability of palaeointensity determinations (e.g., Selkin and Tauxe, 2000; Kissel and Laj, 2004; Biggin et al., 2007; Paterson et al., 2014). However, as opposed to standard palaeomagnetic studies, no particular parameter and criteria set is customary applied, although they do not vary markedly among different palaeointensity studies. Moreover, criteria that are better at excluding inaccurate results may be not so effective at including accurate results and vice versa. Accordingly, Paterson et al. (2014) proposed some modifications to widely used criteria sets to increase the acceptance of accurate determinations.

For the Thellier-Coe experiment performed in the present study, we chose the sets of criteria included in the Thellier-Tool (version 4.22) software (Leonhardt et al., 2004) with the modifications proposed by Paterson et al., (2014) (Tab. 2). These criteria comprise two quality levels, A and B, of different stringency. As in the present study the archaeointensity results were obtained from remanence acquired in a known field, it is of interest to relate the quality level A or B assigned to each palaeointensity determination to the amount of deviation from the expected intensity values in order to evaluate the accuracy of the results and the quality of the determinations.

Application of the reliability criteria yields successful palaeointensity determinations in all 26 analysed specimens (Fig. 8, Tab. 3); 20 (76.9%) fulfil all class A criteria and 6 (23.1%) only class B criteria. Type A specimens yield a mean palaeointensity $F_{TC(A)}$ = (43.0 ± 5.2) µT while type B specimens display a higher mean value $F_{TC(B)}$ = (47.5 ± 4.9) µT (Tab. 4). The mean result F_{TC} = (44.0 ± 5.4) µT obtained for all 26 samples (Tab. 4) agrees within the error bars with the field value at the experiment site, which varies between 40 and 41µT depending on the position in the kiln (Tab. 1). If specimens heated 2009 and 2010 are considered separately, a significantly higher mean intensity F_{2009} = (48.0 ± 5.1) µT is obtained for the specimens fired in the 2009 experiment than in those baked in 2010, which yield an intensity F_{2010} = (42.5 ± 4.8) µT (or F_{2010} = (43.1 ± 4.8) µT if brick samples LN, which were heated in 2010, but not in the present experiment are excluded from the mean) . The mean raw archaeointensity (averaged over all specimens) obtained by Morales et al. (2011) on pieces baked during the same experiment in 2009 yields F = (38.4 ± 4.5) µT, which is also significantly lower than the mean 2009 results from the present study.

Archaeological materials such as ceramics or bricks are often characterised by a strong magnetic anisotropy (e.g., Aitken et al., 1981). As the strength of the laboratory acquired pTRMs depends of the direction along which the laboratory field is applied, a significant error in archaeointensity determination may occur unless the field is applied in the same direction as the ancient original field. For this reason, archaeointensity measurements were corrected for magnetic anisotropy by determining the anisotropy of TRM tensor (ATRM). These measurements were performed at the palaeomagnetic laboratory of Géosciences Montpellier, according to their standard procedure (Fanjat et al., 2013). ATRM measurements were carried out after completion of the palaeointensity experiments in Burgos by inducing a pTRM (550°C to room temperature) in six sample directions (i.e. +x, +y, +z, -x, -y, -z). Zero-field thermal demagnetizations at 580°C before each pTRM were used as a baseline. As the studied specimens had not experienced significant alteration during the palaeointensity experiment, performing the ATRM measurements at the end of the experiment at this temperature should not introduce a significant inaccuracy in the calculation of the anisotropy correction factor. The values of the latter are shown on Tab. 3 together with the corrected archaeointensity values. All archaeointensity values were corrected for the ATRM following Veitch et al.'s (1984) method with a Matlab® code developed in Montpellier, which is provided as supplementary material in Tema et al. (2015). In seven cases, the small specimens used in the palaeointensity experiments deteriorated during the ATRM measurement, and no anisotropy correction factor could be obtained. In such cases, the anisotropy factor was calculated from the mean values of the anisotropy factors of other specimens of the same archaeological artefact (specimens L1-1, L1H), from the mean values of the anisotropy factors of other brick specimens (LNE, LNI,) or from the mean of all archaeological artefact specimens (N1A, N1D, N1F). As can be recognised in Tab. 3, brick specimens show a lower degree of anisotropy than pottery specimens. This observation is in accordance with results from Jordanova et al. (1995) and Kovacheva et al. (1996), showing a lower effect of remanence anisotropy on palaeointensity determination on brick or tile samples than on pottery. After correction, type A specimens yield a mean corrected palaeointensity $F_{TC(A)CORR}$ = (38.3 ± 3.6) µT and in type B specimens, $F_{TC(B)CORR}$ = (37.9 ± 2.8) µT (Tab. 4). Both are indistinguishable, and the mean result obtained for all 26 samples $F_{TC\text{-}CORR}$ = (38.2 ± 3.6) µT (Tab. 4) agrees within the error bars with the field value at the experiment site. If considered separately, anisotropy-corrected archaeointensity results of specimens heated in 2009 and 2010 show an excellent agreement, as $F_{2009(CORR)}$ = (38.2 ± 3.2) µT and $F_{2010(CORR)}$ = (38.2 ± 3.8) µT (Tab. 3). No difference can be observed between the anisotropy corrected 2010 results with or without brick specimens LN, which were not heated in the present experiment (without LN, $F_{2010(CORR)}$ = 38.3 ± 4.0 µT). The mean 2009 anisotropy-corrected archaeointensity result from the present study is, however, higher than the mean anisotropy-corrected archaeointensity result (averaged over all specimens) obtained by Morales et al. (2011) on samples baked during the 2009 experiment, which yields F = (35.6 ± 3.1) µT. Application of anisotropy correction to the studied samples moderately diminishes the scatter of archaeointensity results. While the standard deviation to mean archaeointensity ratio yields values between 10.3 and 12.3% for all non-corrected means shown in Tab. 4, the same ratios in the case of anisotropy corrected values are reduced to 7.3 to 9.4%.

In the present study, no cooling rate correction (e.g., McClelland-Brown, 1984) needed to be applied, because during the Thellier-Coe palaeointensity experiments samples were left to cool down naturally over several hours, with a similar duration than the heating procedure during remanence acquisition in the kiln that was directly measured. Comparison of the mean 2009 anisotropy-corrected archaeointensity result from the present study with the anisotropy and cooling rate corrected mean archaeointensity averaged over all specimens F = (38.9 ± 3.6) μT obtained by Morales et al. (2011) shows an excellent agreement. This result confirms that samples can be left cooling down naturally over several hours to avoid extra measurements for the cooling-rate correction when original cooling times of a similar order of magnitude are involved.

All samples were fired in the same field, most at the same time in 2010 and one sample in 2009, which from an archaeological point of view is basically the same time. Therefore, calculation of the mean corrected palaeointensity has been performed so far averaging over all specimens. In a standard archaeointensity study, however, results would be averaged for each ceramic piece or brick, and then a mean palaeointensity could be calculated for all these pieces if they were considered to belong to the same time unit. In such case, the mean intensity obtained from six pieces (two bricks, one flowerpot from 2009 and one from 2010, one vessel and one zoomorphic pot) with the same weight would yield a mean result $F_{TC\text{-}CORR}$ = (38.0 ± 3.7) μT (Tab. 4) which agrees with the value obtained when averaging over all specimens. A slightly smaller value $F_{TC\text{-}CORR\text{-}W}$ = (37.1 ± 2.6) is obtained if a weighted mean of the six pieces is calculated by means of the inverse square of the standard error as the weight of the individual data (Kono et al., 1986) with a weighted standard deviation of the six paleointensity estimates (Heckert and Filliben, 2003).

4.2 Microwave method (MW)

Additional archaeointensity determinations were carried out at the palaeomagnetic laboratory of the University of Liverpool using integrated SQUID magnetometer and 14 GHz microwave systems, MWS. Both the older horizontally aligned system (Betty) and the newer vertically aligned system (Tristan) were used (see e.g. Böhnel et al., 2003; Stark et al. 2010). Mini core samples (5mm diameter by 1-3mm length) were drilled from ceramic pieces (vessel L, flowerpot M and zoomorphic vessel -N fired in 2010 and flowerpot R4 fired in 2009) and brick samples (LQ and LN) to make a total of 32 specimens.

One specimen is mounted (via vacuum in Tristan or attached with ceramic glue when using Betty) into the MWS and moved via computer control between the resonant microwave cavity and the magnetometer. The resonant frequency of the cavity plus specimen is determined by monitoring the amount of power reflected when the frequency is swept at very low (0.1W) power. The maximum power the amplifier can deliver is 80W, so to generate greater microwave energy the length of exposure can be increased. Exposure time was typically between 5 and 10 seconds.

Firstly, all specimens underwent a microwave demagnetisation experiment. For this sample set microwave absorption was poor so that high powers and longer exposure times than often used were needed to demagnetise the specimens and even so, only 14 specimens could successfully be demagnetised. Six of these were deemed too weak to undergo a palaeointensity experiment, leaving 8 specimens for microwave palaeointensity determination.

The microwave intensity determinations followed the Coe (1967) protocol with repeated infield steps (pT_MRM-checks) to monitor possible magnetic mineralogical alteration during the experiments. In order to monitor multidomain behaviour repeated zero field steps (MD checks) were performed. During re/demagnetization a field of 41 µT was applied parallel to the direction of the natural remanent magnetization (NRM). All microwave experiments were conducted in air under atmospheric conditions.

For the microwave experiment, reliability criteria were selected regarding the quality of experimental conditions, the occurrence of chemical and/or mineralogical alterations and the presence of a remanence fraction carried by MD grains. Again, we chose the sets of criteria included in the Thellier-Tool (version 4.22) software (Leonhardt et al., 2004) with the modifications proposed by Paterson et al., (2014), although not all parameters used for Thellier-Coe experiments were used for the microwave experiments (Tab. 2). As for the Thellier-Coe experiments, the chosen criteria included two quality levels, A and B, of different stringency. Strict application of these criteria yielded successful palaeointensity determinations in only 5 of 8 analysed specimens (Tab. 2; Fig. 9a). Nevertheless, although sample N1A(i)_2 did not formally fulfil the fraction parameter criterion $f \geq 0.35$, yielding only $f = 0.34$, this difference is small, and due to the difficulties in fully demagnetising the sample. As all other parameters display acceptable values (Tab. 5), this determination has been considered successful. Two archaeointensity determinations, on the other hand, were initially rejected because samples experienced alteration during the experiment (Fig. 9b). However, sample M(i)_3 only fails marginally criterion δ(CK), and the high δ(CK) value originates in a rather anomalous check. As all other parameters display excellent values (Tab. 5), in a standard palaeointensity experiment this result would have probably been accepted, so that we have considered it a successful determination. On the other hand, although the archaeointensity result from sample N1E(i)_3 is consistent with the results obtained on the remaining seven samples, δ(CK) is rather high. In a standard study, without a previous knowledge of the field value to be obtained, this determination would be rejected. Thus, 7 determinations were considered successful, 3 fulfilling all class A criteria and 4 only class B criteria.

As could be observed in the Thellier-Coe determinations obtained with a resistive heating described in the previous section, type A specimens determined with the MW method yield a somewhat lower mean palaeointensity $F_{MW}(A) = (43.2 \pm 3.5)$ µT than type B specimens, which display a mean value $F_{MW}(B) = (46.7 \pm 4.9)$ µT. Both yield very similar estimates of palaeointensity as those obtained with the Thellier-Coe method. Error bars of both means overlap and the mean result obtained for all 7 samples F_{MW} = (44.9 ± 4.3) µT, agrees with the field value (40 - 41µT) at the experiment site within error. In the microwave determinations the field is applied in the same direction as the natural remanence, thus no anisotropy correction has to be applied in this case.

Poletti et al. (2013) demonstrated that for a set of Brazilian ceramics the differences in cooling rate between the MW and standard heating methods can result in overestimates of MW results of up to 25%. The higher average results obtained from the MW may be due to cooling rate.

4.3 Original multispecimen method without corrections (MSP-DB)

An absolute archaeointensity determination was performed on a clay-pot sample (M) at the palaeomagnetic laboratory of UNAM in Morelia (Mexico) using the original multispecimen method as proposed by Dekkers and Böhnel (2006) and including alignment correction. The clay-pot specimens were cut into 6 sub-specimens and pressed into salt pellets in order to obtain standard-dimension cylindrical palaeomagnetic specimens. The experiment was performed employing laboratory fields from 30 to 50 µT, with increments of 10 µT. Specimens were oriented in the heating chamber in such a way that the natural remanent magnetization (NRM) directions of each sub-specimen lay parallel to the furnace axis. The heating temperature was set at 450°C. This temperature appeared suitable to allow the selected samples to retain enough NRM and acquire enough pTRM to obtain reliable results. In addition, it was low enough to avoid thermochemical alterations on the specimens. Specimens were heated during 20 minutes in air. The relative differences between remanences after each archaeointensity step (remaining NRM + pTRM acquired in the laboratory) and NRMs of specimens were calculated and the corresponding results plotted; a least square fit was performed for the data and intersections with the horizontal axis (zero difference) were calculated for palaeointensity determinations. Special care was taken regarding the difference between NRM and applied pTRM directions, taking a maximum angle of 5° as a cut-off value. A good linear least square fit was obtained, with R^2 = 0.95, and the multispecimen determination was considered successful (Fig. 10). A palaeointensity value $F_{MSP\text{-}DB}$ = (40.6 ± 0.4) µT in full agreement with the field value at the experiment site was obtained.

4.4 Fraction and domain-state corrected multispecimen method (MSP-FC and MSP-DSC)

A second multispecimen archaeointensity experiment was performed on specimens taken from different potsherds of clay-pot samples M1 and M2 at the palaeomagnetic laboratory of Géosciences Montpellier (France) with a very fast-heating furnace by infra-red (FUReMAG patent #1256194). Two key points determine its characteristics. The first is to heat uniformly by radiation a single rock sample of a 10cm^3-standard volume as fast as the thermal conductivity of the sample allows. This feature theoretically allows to reduce chemical changes of the magnetic oxides during the heating. The second is to apply to the sample during the heating/cooling cycle a precise magnetic induction field, perfectly controlled in 3D with a measured precision on its direction of less than 1°. The extended method that includes protocols for fraction and domain-state correction (Fabian and Leonhardt, 2010) was used. A pTRM was imparted at a dwell step of 320°C on 11 samples with a different magnetising field for each sample chosen every ten µT between 10 and 80 µT. The dwell temperature was chosen so that a NRM fraction of about 50 percent was replaced by the laboratory induced pTRM. Note that in the Montpellier laboratory approach, the shape of the distribution of the bootstrapped palaeointensity estimates is a strong criterium to accept a paleointensity estimate. The distribution has to be unimodal and symmetric about the mean, approaching a normal distribution for an ideal case. If not, the only way to proceed is to add supplementary data. In the present case, only 11 samples were required in order to generate an empirical bootstrap confidence interval from a normal distribution of the bootstrapped palaeointensity estimates (Figure 11). For the fraction correction and domain state correction determination, we anchored the linear regression to the point (0, -1) since it represents a theoretical point: when a sample is cooled in zero field there is no pTRM acquisition. The 95% confidence interval on the palaeointensity determination is determined by bootstrapping the least-squared regression. The influential data are detected and discarded recursively from the regression analysis by means of the Cook's distance. A cut-off value is arbitrarily choosen at 3 times the mean value of the Cook's distances. Alpha parameter is arbitrarily chosen at 0.5 as recommended by Fabian and Leonhardt (2010). Archaeointensity determinations obtained in this experiment are shown in Tab. 6 and Figure 11. The determination results are displayed without any corrections (i.e., equivalent to the original method from Dekkers and Böhnel, 2006), as well as fraction-corrected (FC) and domain-state corrected (DSC). A very good linear least square fit was obtained in all cases, with $R^2 \geq 0.97$. In all cases, very similar archaeointensity values varying between $F_{MSP\text{-}DSC}$ = 37.8 µT and $F_{MSP\text{-}FC}$ = 38.7 µT were obtained (Tab. 6, Fig. 11). In fact, all 95% confidence intervals overlap, although formally only those corresponding to the fraction corrected determination agree with the field value at the experiment site (40 - 41 µT). The remaining determinations show slightly lower values.

5 Discussion

As described in the previous section, in all four laboratories a good agreement between the archaeointensities determined and the original magnetising field was observed. Despite the use of four different archaeointensity determination protocols based on different heating processes, in all cases precise and reliable results could be obtained that were correct within error bounds (Fig. 12).

Thellier-Coe determinations yielded a 100% success rate, although two quality levels, A (76.9% of the cases) and B (23.1% of the cases) were distinguished. However, non anisotropy-corrected values only matched field values at the study site if averaged over all specimens. If results are first averaged for each ceramic piece or brick, and then a mean palaeointensity of these pieces is calculated, as usually done in standard archaeointensity studies, agreement with the expected field is only observed after anisotropy correction. Calculation of a weighted mean yields a somewhat smaller value, due to the excellent agreement of archaeointensity results determined on specimens belonging to the two pieces showing the weakest palaeointensity. Thus, caution must be exercised choosing weighting criteria of archaeointensity determinations, not to artificially bias results. No cooling rate correction needed to be applied, because cooling time of the samples during paleointensity experiments was similar than the duration of sample heating in the kiln.

Microwave determinations with the Thellier-Coe protocol also yielded a high success rate, as 7 of 8 analysed samples provided successful determinations. It must be however noted that after performing an initial microwave demagnetization capacity experiment on 32 samples, microwave palaeointensity experiments could only be performed on 8 of them, because the remaining ones either could not be demagnetised or showed an NRM intensity too weak. In this case, also two quality levels were defined, with 50% of the determinations belonging to type A and 50% to type B. The mean result obtained was somewhat higher than the field value at the experiment site but showed agreement within error bounds (Fig. 12). As in this case the field was applied in the direction of NRM, no anisotropy correction had to be performed. The higher mean result obtained might be ascribed to the fast cooling rate of this method (Poletti et al., 2013). One sample did not provide reliable results due to alteration during the microwave experiment. A specimen from the same sample subjected to the Thellier-Coe procedure did, however, show less alteration. This result indicates that a significant amount of dielectric heating is occurring in the specimens due to them being poor microwave absorbers as evidenced by the need to use high power and longer exposure times to de(re)magnetise.

Multispecimen experiments were only performed on flowerpot samples with each of both methods used. The original uncorrected multispecimen MSP-DB method performed in Morelia laboratory with samples taken from a single flowerpot agreed best with the expected value (Fig. 12). The extended multispecimen method also supplied results near to the expected one, the most accurate ones with the fraction corrected results, with a 95% confidence interval including 37.1 μT $\leq$ FMSP-FC $\leq$ 40.4 μT (Fig. 12). A nearer value to the expected one obtained with the MSP-FC protocol than the one obtained with the MSP-DSC protocol could indicate that the alpha parameter taken at

0.5 overestimates in the present case the multi-domain effect. Specimens from flowerpot sample M, which were used in both multispecimen experiments, show a trend towards MD characteristics on the Day-plot (Fig. 7a), but determinations with both applied multispecimen methods yield correct results. strenghting the conclusion of a non-adequate alpha parameter value. We clearly show that the MSP-DSC protocol requires more developments to fix precisely the alpha parameter value. We propose that in a multi-protocol approach for palaeointensity determination, alpha parameter should be fixed between 0.1 and 0.2 for samples yielding linear Thellier plots, and around 0.5 for samples yielding concave-up Thellier plots.

Palaeointensity determinations are experimentally difficult, and the presence of MD-grains and/or the occurrence of irreversible chemical/mineralogical or physical changes during the experiments can produce failed or erroneous palaeointensity determinations. Thus, the failure rate of palaeointensity experiments can be large and, even worse, incorrect determinations may be taken as correct palaeointensity results. In the present study, almost all samples analysed in all four laboratories yielded reliable results regarding the applied quality criteria and palaeointensities obtained agreed well with the original magnetising field.

These successful results pose, however, some questions. Rock magnetic characteristics are frequently used as preselection criteria for samples to be used in palaeo- or archaeomagnetic determinations. In the present study, the original magnetising field was known, and there was no sense in carrying out a preselection of samples with better characteristics to provide reliable archaeointensity results. On the contrary, knowing the result to be obtained, rock-magnetic or other experimental characteristics related to correct or incorrect determinations may provide clues about the causes behind successful or failed determinations.

As shown in Fig. 7a, hysteresis parameter ratios display PSD behaviour. Most samples lie between SD-MD and SP-SD theoretical mixing curves for magnetite (Dunlop, 2002), with the SD-MD mixing-curve sector nearest to them yielding a relative MD content varying approximately between 40% and 80%. Based on an analysis of comprehensive rock-magnetic and paleointensity data, Paterson et al. (2017) quantified a stability trend in hysteresis data that characterises the bulk domain stability (BDS) of the magnetic carriers in a palaeomagnetic sample. In that study BDS is considered an approximate quantitative measure of the effective bulk domain state of an assemblage of magnetic carriers, irrespective of the specific mechanisms that may influence the sample's bulk domain state. It provides a relative stability measure, with larger values being related to more stable remanent carriers and lower, negative values to less stable remanence carriers. Figure 13 shows a comparison between BDS and inaccuracy of the Zinapécuaro paleointensity data obtained with the standard and the microwave procedures. The inaccuracy of a paleointensity result B_{anc} is estimated from its deviation D from the expected intensity B_{exp}, with D= $\ln(B_{anc}/B_{exp})$ Paterson et al. (2017). No correlation appears between paleointensity accuracy and BDS, probably since neither of both parameters displays large variations. Most BDS values lie between 0.2 and 0.4. (For comparison, Paterson et al. (2017) obtain BDS = -0.94 for a large 220μm grain and BDS = 0.79 for an idealized assemblage of Stoner-Wohlfarth particles). In addition, Paterson et al. (2017) suggest that when less than approximately 100 specimens are used, a significant relationship between both parameters may be missed. Nevertheless, Fig. 13 shows slightly lower BDS values for type-B paleointensity determinations (squares) than for type-A determinations (circles).

Specific characteristics of the remanence acquisition procedure may also be compared with the experimental quality of the determinations and the results obtained. During heating of the archaeological reproductions in the kiln, four thermocouples had been placed at different positions in the baking compartment of the oven (Fig. 1b), as described in section 2. As previously mentioned, the maximum temperature reached in the lower central part of the kiln exceeded 700°C, while in other parts of the oven somewhat lower maximum temperatures - still above 640°C - were reached. During most of the experiment all thermocouples recorded very similar temperatures. Thus, thermal conditions in the kiln were rather similar for all heated artefacts, independently of their position. Nevertheless, centrally placed objects (near thermocouple 1, Fig. 1b) experienced higher temperatures – between 30 and 80°C – during the 3^{rd} and 5^{th} hour than objects placed at other positions within the kiln. In the Thellier-Coe experiment no relation could be observed, however, between determination quality A or B and position in the furnace. Regarding the difference between original field value and actual archaeointensity, archaeointensities from specimens from sample R4, which was fired in the 2009 experiment, showed the largest discrepancies with the original field strength, which had the same value both in the 2009 and 2010 experiments.

6 Conclusions

Archaeointensity determinations have been performed on present-day reproductions of Pre-Columbian Mesoamerican archaeological ceramics and two bricks in four different palaeomagnetic laboratories by means of different archaeointensity determination protocols based on different heating processes: Thellier-Coe (Coe, 1967) with a resistive heating in Burgos (Spain), Thellier-Coe with microwave heating (Walton et al., 1992) in Liverpool (U.K.), uncorrected multispecimen method with a resistive heating (Dekkers and Böhnel, 2006) in Morelia (Mexico) and extended multispecimen method including protocols for fraction and domain-state correction (Fabian and Leonhardt, 2010) with an infrared heating in Montpellier (France). Reliable determinations and a good agreement between the magnetising field strength (40-41 μT) and the archaeointensities obtained was achieved in all participating laboratories and with all methods used (Fig. 12). Thus, this study demonstrates the potential use of this type of ancient kiln as a source for determining geomagnetic field strength variation in the past. Agreeing palaeointensity results obtained from methods relying on different principles can bestow consistency and reliability to these results, even if only a limited number of determinations is available, as with multispecimen determinations in the present study.

Thellier-Coe determinations on 26 samples were successful in all cases and after anisotropy correction (Fanjat et al., 2013) a mean archaeointensity value F_{TC}=(38.0 ± 3.7) µT was obtained. Results confirmed that no cooling rate correction was needed, because samples were left cooling during paleointensity experiments for a time of the same order of magnitude than the duration of sample heating in the kiln.

Microwave determinations with the Thellier-Coe method could be performed on 8 samples and 7 of them provided successful determinations yielding a mean archaeointensity result F_{MW} = (44.9 ± 4.3) µT. This higher average result (Fig. 12) may be explained by the fast cooling rate during the MW experiments. One microwave determination had to be rejected because of alteration occurred during the experiment. A sister specimen of the rejected one subjected to the standard Thellier-Coe protocol was not affected by significant alteration and provided a reliable determination. Although the microwave procedure is devised in such way as to reduce the probability of alteration, in this case dielectric heating is likely to have been significant due to the high power and longer exposure times needed to de(re)magnetise these samples.

Both MS methods were only applied to flowerpot sample M. The original uncorrected multispecimen MSP-DB method yielded an archaeointensity $F_{MSP\text{-}DB}$ = (40.6 ± 0.4) µT, showing the best agreement with the field value at the kiln of all four methods. The extended multispecimen method also supplied results near to the expected one, the most accurate ones with the fraction corrected results, with a 95% confidence interval including 37.1 µT ≤ $F_{MSP\text{-}FC}$ ≤ 40.4 µT. Specimens from flowerpot sample M, show a certain trend towards MD characteristics on the Day-plot (Fig. 7a), but determinations with both applied multispecimen methods yield correct results, as expected for the MSP-DSC, but not necessarily for the MSC-DB method.Thermomagnetic magnetisation-versus-temperature curves showed a highly reversible behaviour, the main ferromagnetic (*s.l.*) phase being carried by slightly Al, Mg or Ti-substituted magnetite. These characteristics are in agreement with the good archaeointensity results obtained. Hysteresis parameter ratios displayed in a Day-plot (Day et al., 1977), however, mostly showed PSD behaviour, which if interpreted as due to a SD and MD grain mixture, displayed a trend towards a relatively high MD content. This behaviour would not be favourable for reliable archaeointensity determinations. Nevertheless, no correlation appears between paleointensity inaccuracy and the BDS parameter proposed by Paterson et al. (2017), which can be considered an approximate quantitative measure of the effective bulk domain state of an assemblage of magnetic carriers, irrespective of the specific mechanisms that may influence the sample's bulk domain state. A certain relation might be, however, discerned between BDS values and quality of paleointensity determinations (types A or B). Regarding rock-magnetic experiments, thermomagnetic curves seem to be a more useful means of preselecting samples for paleointensity determination than analysis of hysteresis parameters.

For several samples, the presence of a thermally stable low Curie-temperature phase and a high coercivity behaviour could be deduced from thermomagnetic curves. This observation was confirmed by experiments in which a SIRM was thermally demagnetised. This behaviour points to the presence of the HCSLT phase which has been observed in well-heated archaeological material (McIntosh et al., 2007) and is interpreted as ε-Fe2O3 (López-Sánchez et al., 2017). To our knowledge, its occurrence has never been reported through the experimental recreation of burnt archaeological materials and the capacity of this mineral to accurately record a palaeointensity is unknown. If this mineral is frequently present in archaeological baked clays, specific studies have to be performed in the future.

During heating of the archaeological reproductions, thermocouples placed at different positions in oven recorded very similar temperatures for the duration of most of the experiment. However, a maximum temperature (>700°C) was reached in the lower central part of the kiln, while in other parts of the oven somewhat lower maximum temperatures (> 640°C) were recorded. Nevertheless, in the Thellier-Coe paleointensity experiment no relation could be observed between determination quality and position in the furnace.

Acknowledgements

This work was funded by projects BU0066U16 and BU235P18 (Junta de Castilla y León, Spain) and the European Regional Development Fund (ERDF). AG is grateful for financial support of CONACyT 252149 and UNAM-PAPIIT project 101717. The Géosciences Montpellier survey was supported by a grant from the CNRS-PNP. The FUReMAG rapid furnace construction was supported by the French National Agency for Research (ANR-12-BS06-0015). We wish to thank Simo Spassov and two anonymous reviewers as well as editor Eduard Petrowsky for their constructive and useful comments and suggestions.

References

Aitken, M.J., Alcock, P.A., Bussell, G.D., and Shaw, C.J., 1981. Archaeomagnetic determination of the past geomagnetic intensity using ancient ceramics; allowance for anisotropy. *Archaeometry*, 23, 53-64.

Biggin, A., and Poidras, T., 2006. First-order symmetry of weak field partial thermoremanence in multidomain ferromagnetic grains. 1. Experimental evidence and physical implications. *Earth Planet. Sci. Lett.*,245, 438-453. DOI: 10.1016j.epsl.2006.02.035.

Biggin, A., Perrin, M., and Dekkers, M.J., 2007. A reliable absolute palointensity determiation obtained from a non-ideal recorder. *Earth Planet. Sci. Lett.*, 257, 545-563.

Böhnel, H., Biggin, A.J., Walton, D., Shaw, J., and Share, J.A., 2003. Microwave palaeointensities from a recent Mexican lava flow, baked sediments and reheated pottery. *Earth Planet. Sci. Lett.*, 214, 221-236.

Böhnel, H.M., Dekkers, M.J., Delgado-Argote, L.A., and Gratton, M.N., 2009. Comparison between the microwave and the multispecimen partial difference pTRM palaeointensity methods, *Geophys. J. Int.*, 17(2), 383-394.

Calvo, M., Prévot, M., Perrin, M. and Riisager, J., 2002. Investigating the reasons for the failure of palaeointensity experiments: A study on historical lava flows from Mt. Etna. *Geophys. J. Int.*, 149, 44-63.

Calvo-Rathert, M., Carrancho, Á., Stark, F., Villalaín, J.J., Hill, M., 2012. Are burnt sediments reliable recorders of geomagnetic field strength*? Quaternary Research*, 77 326–330.

Calvo-Rathert, M., Morales-Contreras, J., Carrancho, Á. and Goguitchaichvili, A., 2016. A comparison of Thellier-type and multispecimen palaeointensity determinations on Pleistocene and historical lava flows from Lanzarote (Canary Islands, Spain). *Geochem. Geophys. Geosyst.*, 17, 3638-3654, doi: 10.1002/2016GC006396.

Carrancho, Á., Morales, J., Goguitchaichvili, A., Alonso, R. and Terradillos, M. 2014. Thermomagnetic monitoring of lithic clasts burned under controlled temperature and field conditions. Implications for archaeomagnetism. *Geofísica Internacional,* 53-4: 473-490.

Catanzariti, G., McIntosh, G., Gómez-Paccard, M., Ruiz-Martínez, V.C., Osete, M.L., Chauvin, A. and Team, T.A.S., 2008. Quality control of archaeomagnetic determination using a modern kiln with a complex NRM. *Phys. Chem. Earth*, 33, 427–437.

Coe, R., 1967. Paleointensities of the Earth's magnetic field determined from Tertiary and Quaternary rocks, *J. Geophys. Res.*, 72, 3247-3262.

Coe, R., Grommé, S. and Mankinen, E.A., 1978. Geomagnetic paleointensities from radiocarbon-dated lava flows on Hawaii and the question of the Pacific nondipol low. *J. Geophys. Res.*, 83, 1740-1756.

Day, R., Fuller, M. and Schmidt, V.A., 1977. Hysteresis properties of titanomagnetites: Grain-size and compositional dependence. *Phys. Earth Planet. Int.*, 13, 260-267

De Groot, L.V., Biggin, A.J., Dekkers, M.J., Langereis, C.G., and Herrero Bervera, E., 2013. Rapid regional perturbations to the recent global geomagnetic decay revealed by a new Hawaiian record, *Nat. Commun.*, 4, 2727, doi:10.1038/ncomms3727.

De Groot, L.V., Béguin, A., Kosters, M.E., van Rijsingen, M., Struijk, E.L.M., Biggin, A.J., Hurst, E.A., Langereis, C.G., and Herrero Dekkers, M.J.., 2015. High paleointensities for the Canary Islands constrain the Levant geomagnetic high, *Earth Planet Sci. Lett.*, 419, 154-167, doi: 10.1016/jepsl.2015.03.020.

Dekkers, M.J. and Böhnel, H.N., 2006. Reliable absolute palaeointensities independent of magnetic domain state. *Earth Planet. Sci. Lett.*, 284, 508-517.

Dunlop, D., 2002. Theory and application of the Day plot (Mrs/Ms versus Hcr/Hc) 1. Theoretical curves and tests using titanomagnetite data. *J. Geophys. Res.*, Vol. 107, No. B3, doi: 10.1029/2001JB000486.

Dunlop, D. 2011. Physical basis of the Thellier-Thellier and related palaeointensity methods. *Phys. Earth Planet. Int.* , 187 (3), 118-138. DOI: 10.1016/j.pepi.2011.03.006.

Enterpinar, P., Langereis, C.G., Biggin, A.J., de Groot, L.V., Kulakoğlu, F., Osmura, S., and Süel, A., 2016. Full vector archaeomagnetic records from Anatolia between 2400 and 1350 BCE: Implications for geomagnetic field models and the dating of fires in antquity. *Earth Planet. Sci. Lett.*, 434, 171-186.

Fabian, K. and Leonhardt, R., 2010. Multi-specimen absolute palaeointensity determination: An optimal protocol including pTRM normalization, domain-state correction and alteration test. *Earth Planet. Sci. Lett.*, 297, 84-94.

Fanjat, G., Camps, P., Alva-Valdivia, L.M., Sougrati, M.T., Cuevas-García, M., and Perrin, M., 2013. First archaeointensity determinations on Maya incense burners from Palenque temples, Mexico: New data to constrain the Mesoamerican secular variation curve. *Earth Planet. Sci. Lett.*, 363, 168-180.

Grommé, C.S., T.L. Wright and D.L. Peck , 1969. Magnetic properties and oxidation of iron-titanium oxide minerals in Alae and Makaopuhi lava lakes, Hawaii, *J. Geophys. Res.*, 74, 5277-5249.

Heckert, N.A. and Filliben, J.J., 2003. *NIST Handbook 148: Dataplot Reference Manual, Volume 2: Let Subcommands and Library Functions*. National Institute of Standards and Technology Handbook Series.

Hill, M. and Shaw, J.,1999. Palaeointensity results for historic lavas from Mt. Etna using microwave demagnetization/remagnetization in a modified Thellier-type experiment. *Geophys. J. International*, 139 (2), 583–590, https://doi.org/10.1046/j.1365-246x.1999.00980.x

Jordanova, N., Karloukovski, V. and Spataras, V. (1995). Magnetic anisotropy studies on Greek pottery and bricks. *Bulg. Geophys. J.*, 21(4), 49-58.

Kissel, C. and C. Laj, 2004. Improvements in procedure and palaeointensity selection criteria (PICRIT-03) for Thellier and Thellier determinations: application to Hawaiian basaltic long cores. *Phys. Earth Planet. Int.*, 147, 155-169.

Kono, M., Ueno, N. and Onuki, Y., 1986. Paleointensities of the geomagnetic field obtained from pre-Inca potsherds near Cajamarca, Northern Peru. *J. Geomag. Geolelectr.*, 1339-1348.

Kosterov, A. and M. Prévot, 1998. Possible mechanisms causing failure of Thellier palaeointensity experiments in some basalts, *Geophys. J. Int.*, 134, 554-572.

Kovacheva, M., Jordanova, N. and Karloukovski, V. (1996). Geomagnetic field variations as determined from Bulgarian archaeomagnetic data. Part II: The last 8000 years. *Surv.* Geophys, 19, 431-460.

Le Goff, M. and Gallet, Y., 2004. A new three-axis vibrating magnetometer for continuous high-temperature magnetization measurements. Earth Planet. Sci. Lett., 229, 31-43.

Lee, S. Y. and Xu, H.F., 2018. The role of ε-Fe2O3 nano-mineral and domains in enhancing magnetic coercivity: implications for the natural remanent magnetization. *Minerals*, 8, 97; doi:10.3390/min8030097.

Leonhardt, R., Hufenbecher, F., Heider, F. and Soffel, H., 2000. High absolute palaeointensity during a mid-Miocene excursion of the Earth’s magnetic field. *Earth Planet. Sci. Lett.*, 184, 1, 30, 141-154.

Leonhardt, R., Matzka, J. and Menor, A., 2003. Absolute paleointensities and paleodirections from Fernando de Noronha, Brazil. *Phys. Earth Planet. Int*., 139, 285-303.

Leonhardt, R., Heunemann, C. and Krása, D., 2004. Analyzing absolute palaeointensity determinations: Acceptance criteria and the software ThellierTool4.0.*Geochem. Geophys. Geosyst.*, Vol. 5, no. 12, doi.: 10.1029/2004GC000807.

Leonhardt, R., 2006. Analyzing rock magnetic measurements; The RockMagAnalyzer 1.0 software.*Computers and Geosciences*, 32, 1420-1431.

López-Sánchez, J., G. McIntosh, M. L. Osete, A. del Campo, J. J. Villalaín, L. Pérez, M. Kovacheva, and O. Rodríguez de la Fuente, 2017. Epsilon iron oxide: Origin of the high coercivity stable low Curie temperature magnetic phase found in heated archaeological materials. *Geochem. Geophys. Geosyst.*, 18, 2646–2656, doi:10.1002/2017GC006929.

McClelland-Brown, E., 1984. Experiments on TRM intensity dependence on cooling rate. *Geophys. Res. Lett.*, 11, 205-208.

McIntosh, G., M. Kovacheva, G. Catanzariti, M. L. Osete, and L. Casas, 2007. Widespread occurrence of a novel high coercivity, thermally stable, low unblocking temperature magnetic phase in heated archaeological material. *Geophys. Res. Lett.,* 34*, L21302,* doi: 10.1029/2007GL031168.

Michalk, D., Muxworthy, A.R., Böhnel, H., MacLennan, J., Nowaczyk, N.R., 2008. Evaluation of the multispecimen parallel differential pTRM method: a test on historical lavas from Iceland and Mexico. *Geophys. J. Int.*, 173, 409–420.

Michalk, D. M., Biggin, A. J., Knudsen, M. F., Böhnel, H. N., Nowaczyk, N., Ownby, S., López-Martínez, M., 2010. Application of the multispecimen palaeointensity method to Pleistocene lava flows from the Trans-Mexican Volcanic Belt. *Phys. Earth Planet. Int.*, 179, 3-4, 139—156.

Monster, M.W.L., de Groot, L.V., Biggin, A.J., and Dekkers, M.J., 2015. The performance of various palaeointensity techniques as a function of rock-magnetic behaviour - A case study for La Palma. *Phys. Earth Planet. Inter*., 242, 36-49.

Morales, J., Goguitchaichvili, A., Aguilar-Reyes, B., Pineda-Durán, M., Camps, P., Carvallo, C. and Calvo-Rathert, M., 2011. Are ceramics and bricks reliable absolute geomagnetic intensity carriers? *Phys. Earth Planet. Int.*, 187, 310-321.

Muttoni, G., 1995.Wasp-waisted hysteresis loops from a pyrrothite and magnetite-bearing remagnetized Triasssic limestone. *Geophys. Res. Lett., 22, 3167-3170.*

Nakajima, T., Torii, M., Natsuhara, N., Yaskawa, K., Takagi, M., Ikeguchi, K. and Kawai, N., 1974. Remanent magnetism of the reconstructed ancient kiln. *Rock Magn. Paleogeophy*., 2, 28–31.

Paterson, G.A., Tauxe, L., Biggin, A.J., Shaar, R., and Jonestrask, L.C., 2014. On improving the selection of Thellier-type palaeointensity data. *Geochem. Geophys. Geosyst.*, 15, 1180-1192, doi: 10.1002/2013GC005135.

Paterson, G.A., Muxworthy, A.R., Yamamoto, Y. and Pan, Y. (2017). Bulk magnetic domains stability controls paleointensity fidelity. *PNAS*, vol. 114 no. 50. 13120-13125, doi: https://doi.org/10.1073/pnas.1714047114

Poletti, W., Hartmann, G.A., Hill, M., Biggin, A.J., Trindade, R.I.F., 2013. The cooling‐rate effect on microwave archeointensity estimates. *Geophys. Res. Lett.,* https://doi.org/10.1002/grl.50762

Riisager, P. and Riisager, J., 2001. Detecting multidomain magnetic grains in Thellier palaeointensity experiments. *Phys. Earth Planet. Int.*, 125, 111-117.

Roberts, A.P., Cui, Y.L. and Verosub, K.L., 1995. Wasp-waisted hysteresis loops: Mineral magnetic characteristics and discrimination of components in mixed magnetic systems. *J. Geophys. Res.*, Vol. 100, No. B9, 17909-17924.

Roberts, A.P., Tauxe, L., Heslop, D., Zhao, X. and Jiang, Z.X., 2018. A critical appraisal of the “Day Diagram”. *J. Geophys.Res.: Solid Earth*, vol. 123, no. 4, pp. 2618-2644. doi: 10.1002/2017JB015247

Rojas-Navarrete, L.L., 1995. *Desarrollo de un vidriado sin plomo de baja temperatura para la alfarería tradicional mexicana*. M. Sc. Thesis, Universidad Autónoma Metropolitana, Unidad Iztapalapa.

Schnepp, E., Leonhardt, R., Korte, M. and Klett-Drechsel, J., 2016. Validity of archaeomagnetic field recording: an experimental pottery kiln at Coppengrave, Germany. *Geophys. J. Int*. 205, 622–635.

Selkin, P. and Tauxe, L., 2000. Long-term variations in palaeointensity. *Phil. Trans. R. Soc. Lond.*, 358, 1065-1088.

Stark, F., Cassidy, J., Hill, M.J., Shaw, J. and Sheppard, P., 2010. Establishing a first archaeointensity record for the SW Pacific. *Earth Planet. Sci. Lett.*, 298 (1–2), 113-124.

Tauxe, L., Mullender, T.A.T. and Pick T., 1996. Potbellies, wasp-waists, and superparamagnetsim in magnetic hysteresis. *J. Geophys. Res.*, Vol. 101, No. B1, 571-583.

Tema, E., Camps, P., Ferrara, E., Poidras, T., 2015. Directional results and absolute archaeointensity determination by the classical Thellier and the multi-specimen DSC protocols for two kilns excavated at Osterietta, Italy. *Stud. Geophys. Geod.*, 59, 554–577.

Thébault, E.,et al., 2015. International geomagnetic reference field: The 12th generation. *Earth Planets Space,* 67(1), 1–19. https://doi.org/10.1186/s40623-015-0228-9

Thellier, E. and O. Thellier, 1959. Sur l'intensité du champ magnétique terrestre dans le passé historique et géologique, *Ann. Geophys.*, 15, 285-376.

Veitch, J., Hedley, I. and Wagner, J.J., 1984. An investigation of the intensity of the geomagnetic field during roman times using magnetically anisotropic bricks and tiles. *Archeological Sciences, Geneve* 37: 359-373.

Wasilewski, P.J., 1973. Magnetic hysteresis in natural materials. Earth Planet. Sci. Lett., 20, 67-72.

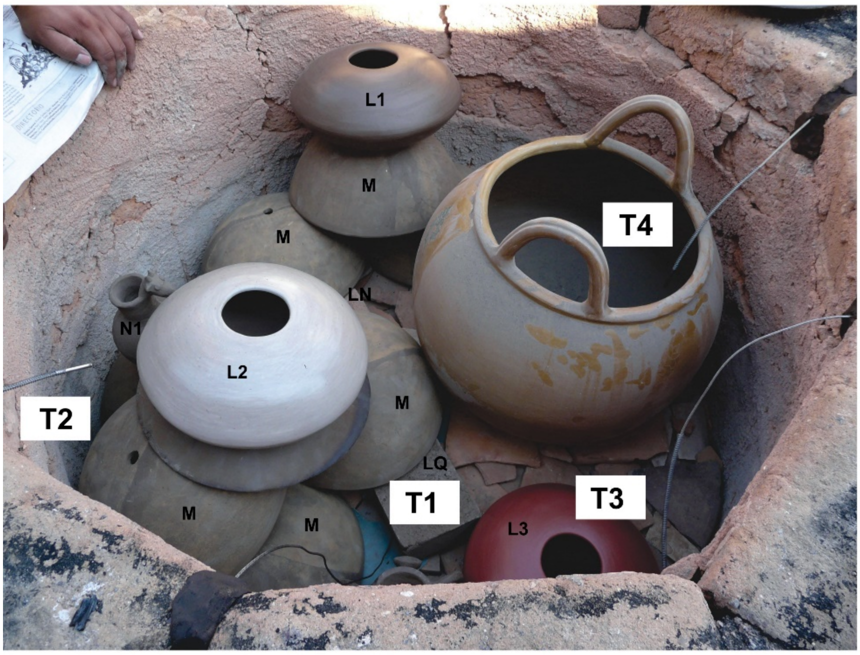
Walton, D., Shaw, J., Share, J. y Hakes, J., 1992. Microwave demagnetization, *J. Appl. Phys.*71, pp. 1549–1551.

Yamamoto, Y., Torii, M. and Natsuhara, N., 2015. Archeointensity study on baked clay samples taken from the reconstructed ancient kiln: implication for validity of the Tsunakawa-Shaw paleointensity method, *Earth Planets, Space*, 67, 63–76.

Figure 1. *Baking of ceramic reproductions*

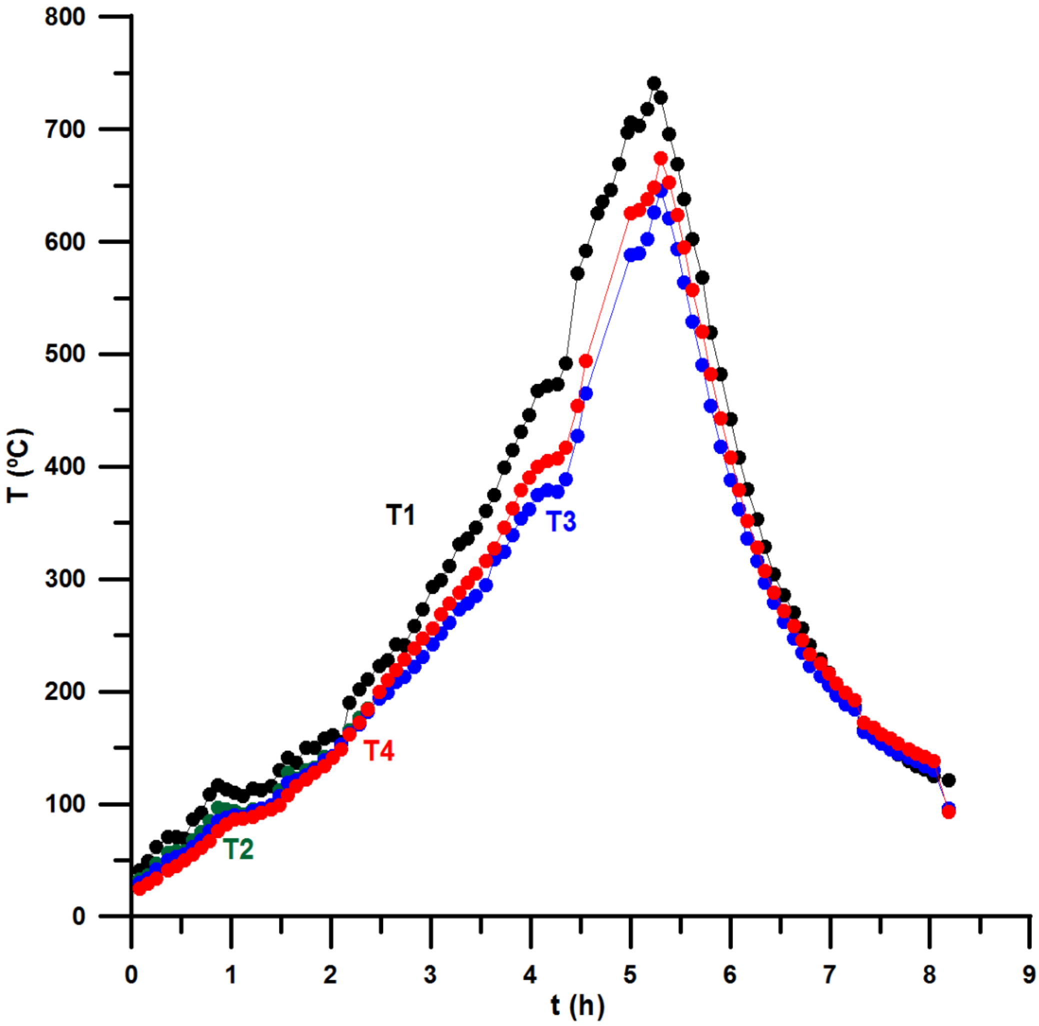
(a)

(b)



(a) Picture of the kiln during the heating procedure; *(b)* Baking compartment of the oven with archaeological artefacts and thermocouples T1 to T4. Names of archaeological pieces are indicated.

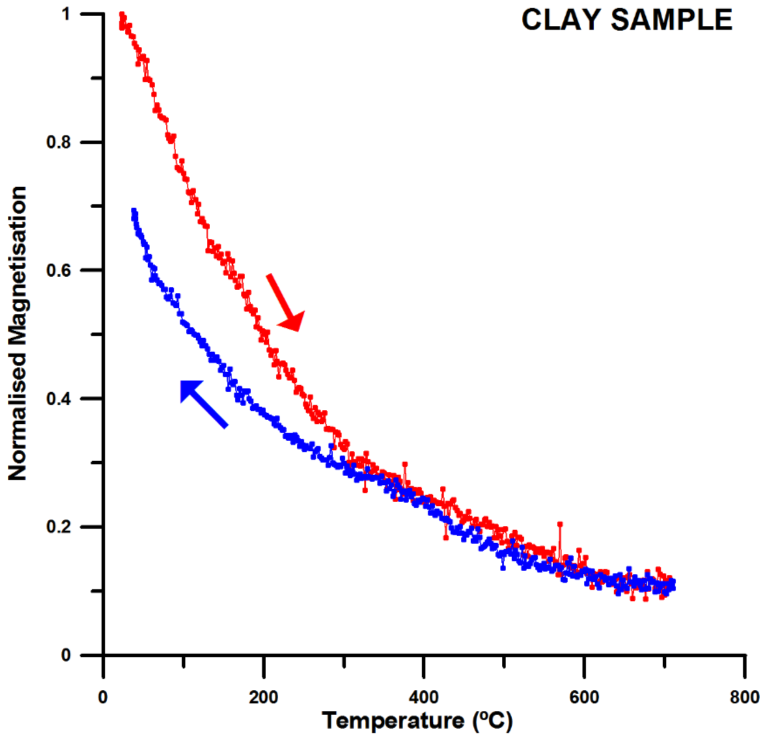
Figure 2. *Temperature variation in the kiln during the baking of ceramic reproductions.*



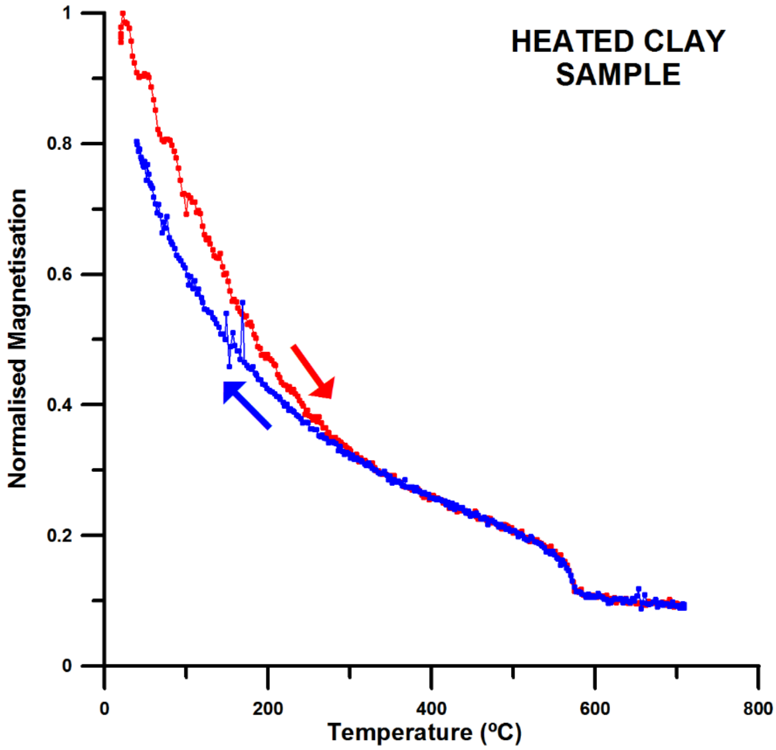
Thermocouples T1 (black), T2 (green), T3 (blue) and T4 (red) were placed at different positions in the oven (see figure 1b). T2 stopped working after approximately 150 minutes of heating.

Figure 3. *Thermomagnetic curves*

(a)

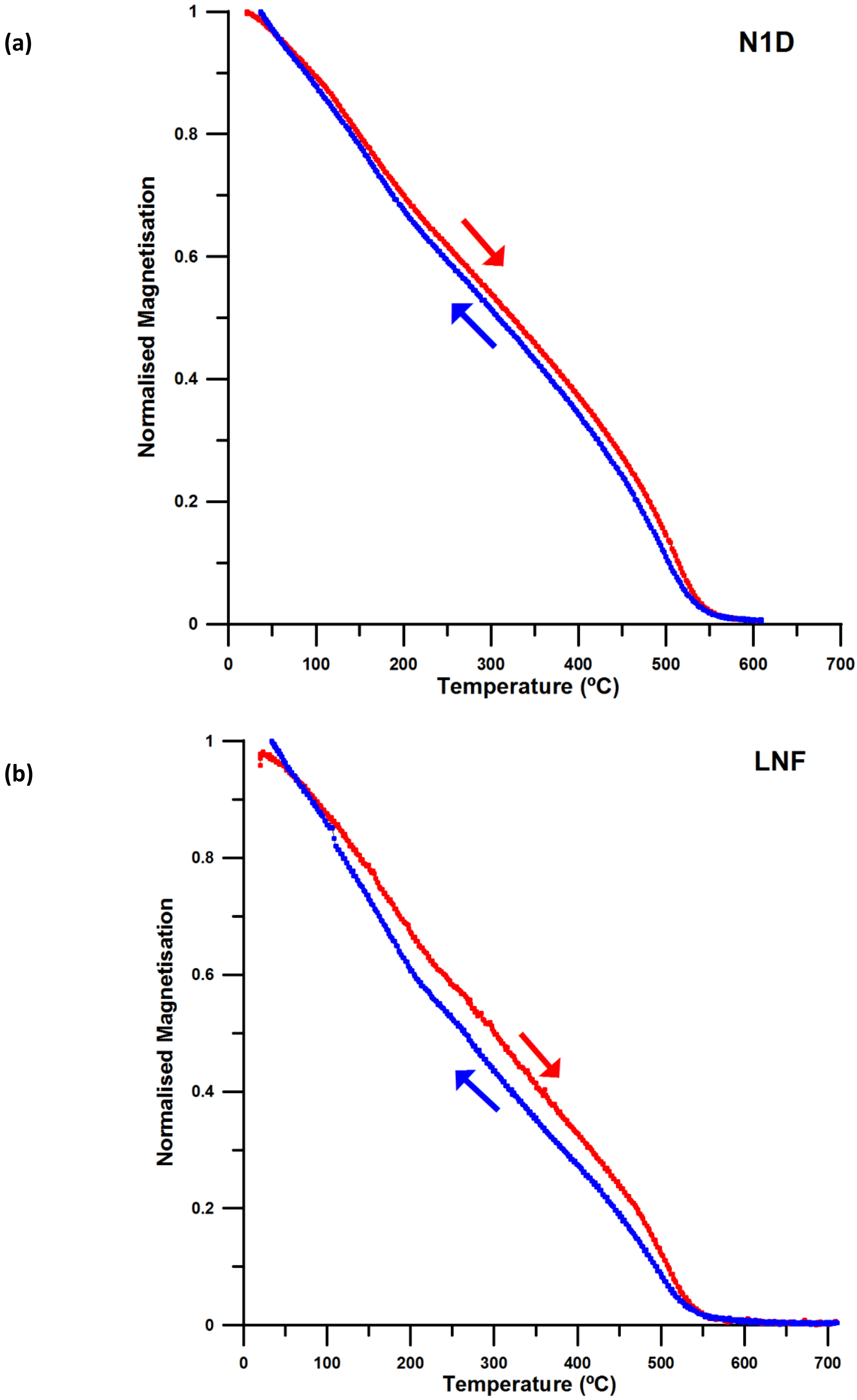


(b)

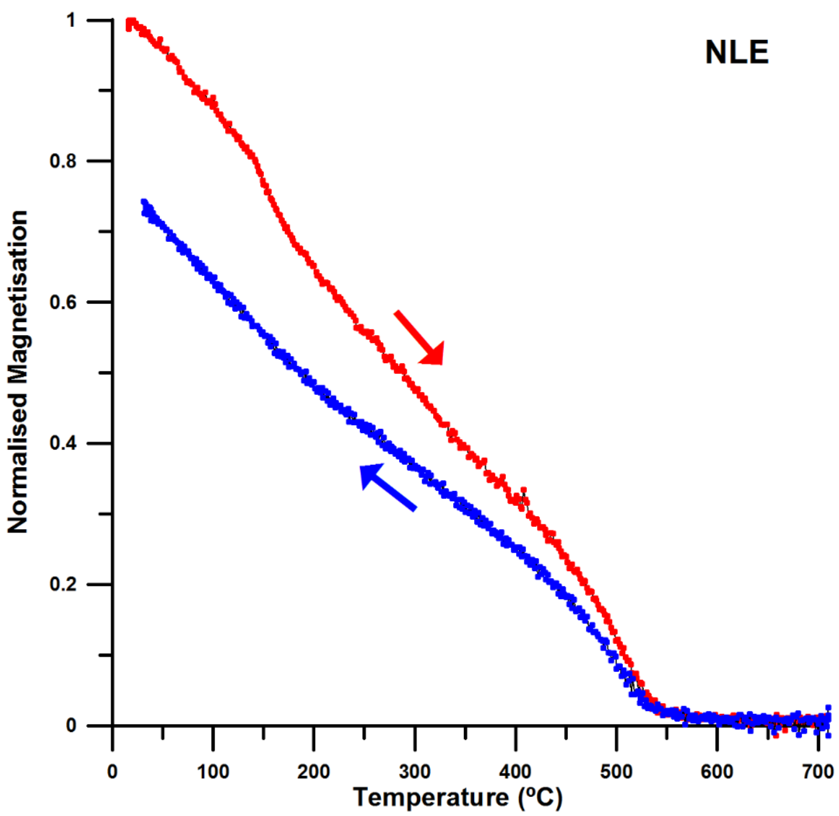


Magnetisation-vs-temperature curve of (a) sample of the clay raw material used to prepare the samples; (b) sample of the clay raw material after being heated for two hours in a furnace and left cooling down for several hours. Heating curve in red, cooling curve in blue.

Figure 4. *Thermomagnetic curves*

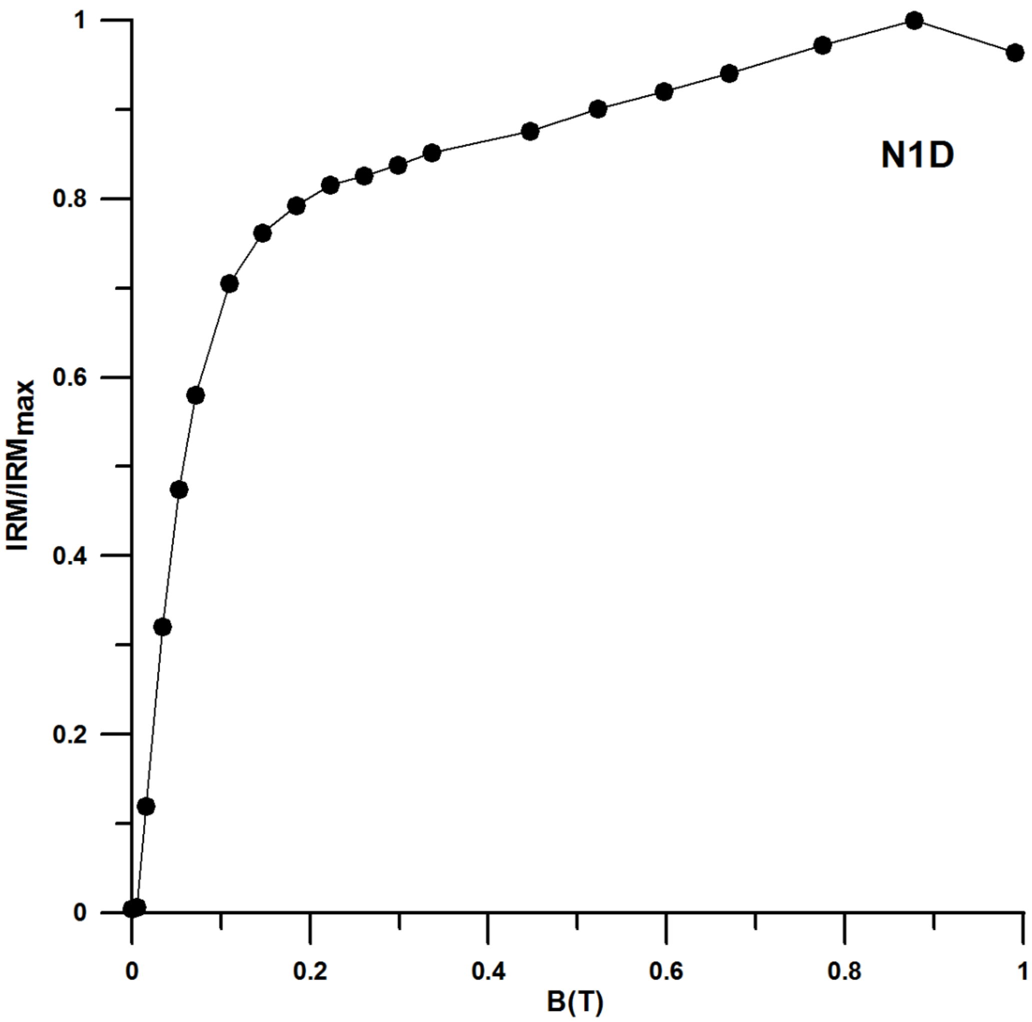


(c)



Magnetisation-vs-temperature curve of (a) zoomorphic vessel N1D; (b) brick sample LNF; (c) zoomorphic vessel NLE. Heating curve in red, cooling curve in blue.

Figure 5. *IRM acquisition curve*

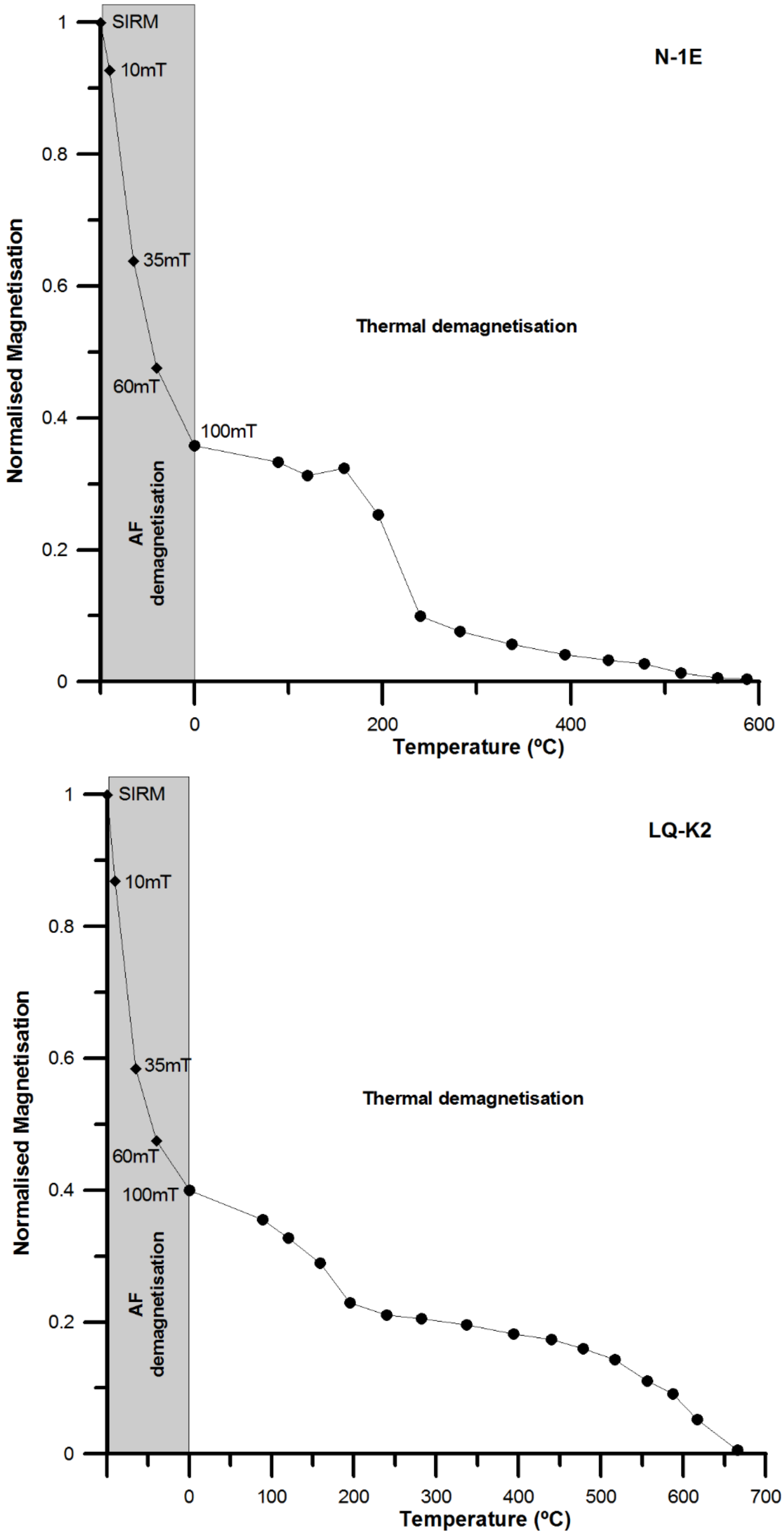


Isothermal remanence acquisition curve of zoomorphic vessel sample N1D.

Figure 6. *Identification of a HCSLT phase*

(a)

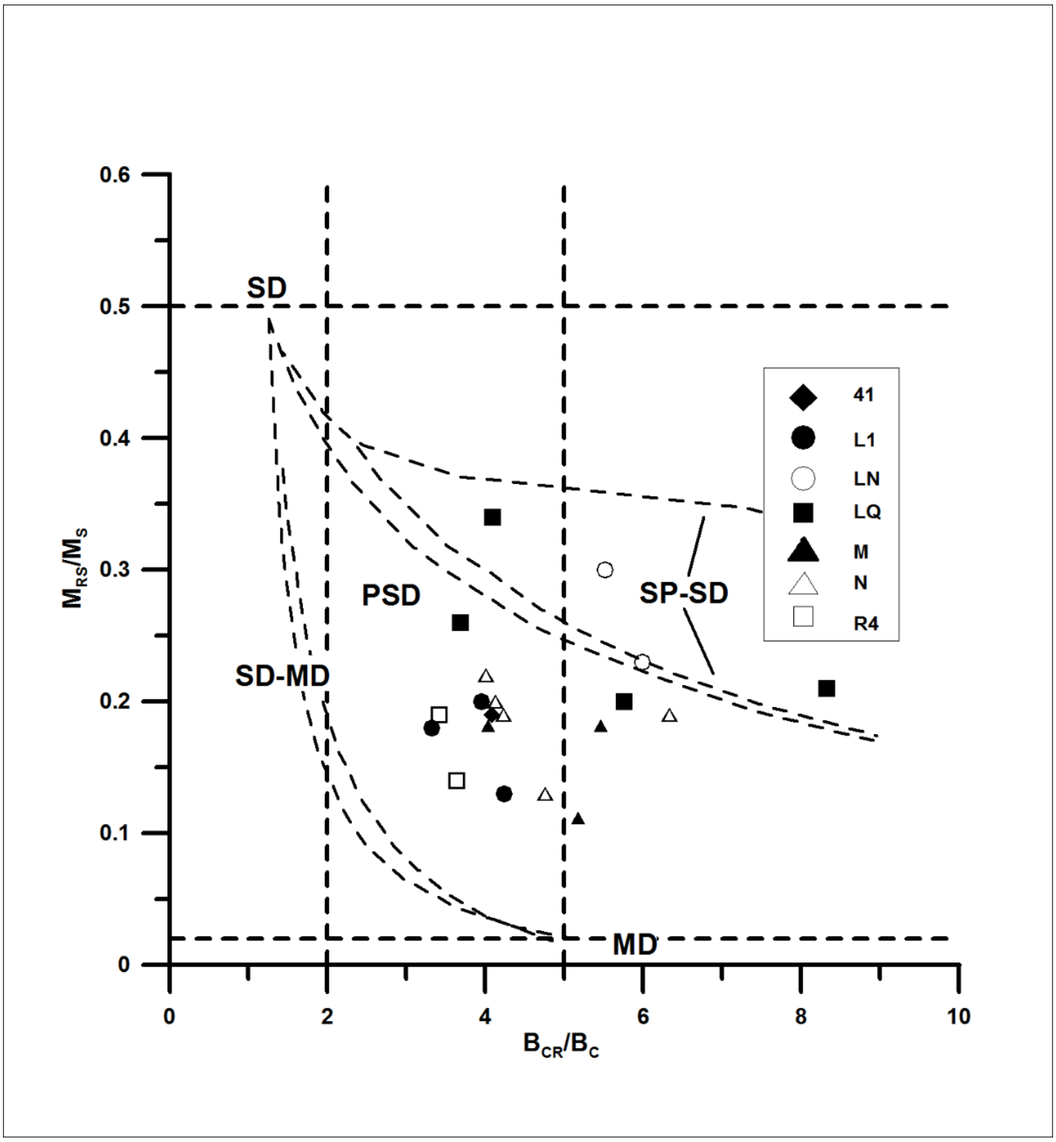
(b)



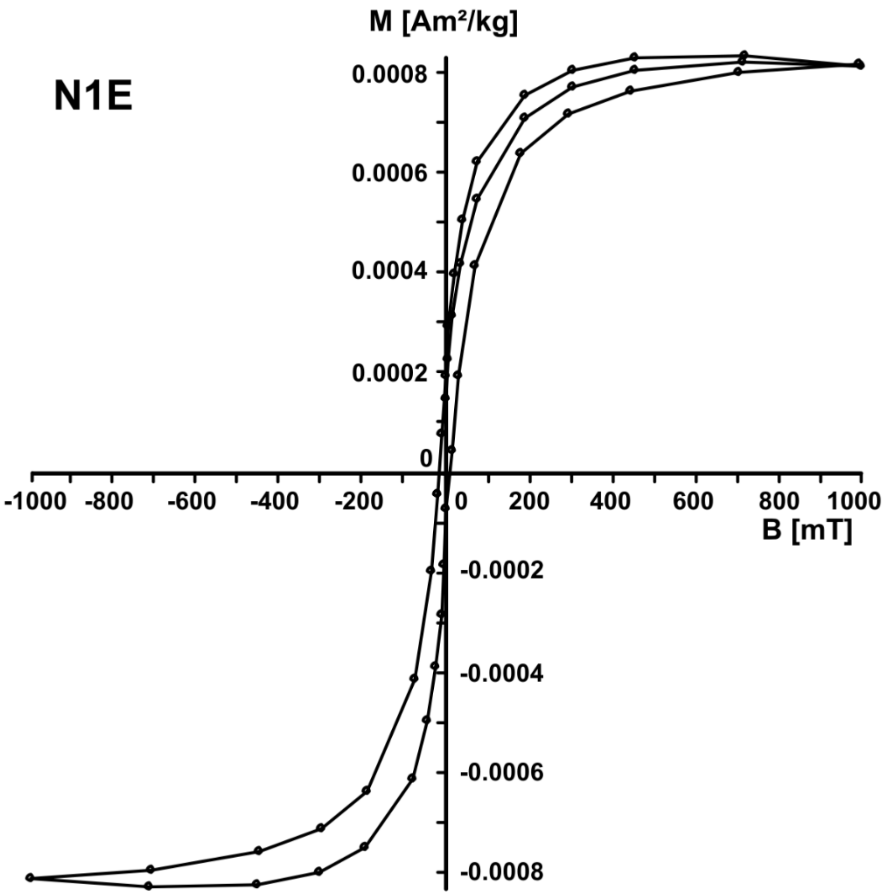
AF demagnetisation up to 100 mT and subsequent thermal demagnetisation of a SIRM imparted at 2T to (a) zoomorphic vessel sample N1E and (b) brick sample LQK2.

Figure 7. *Hysteresis curve results*

(a)

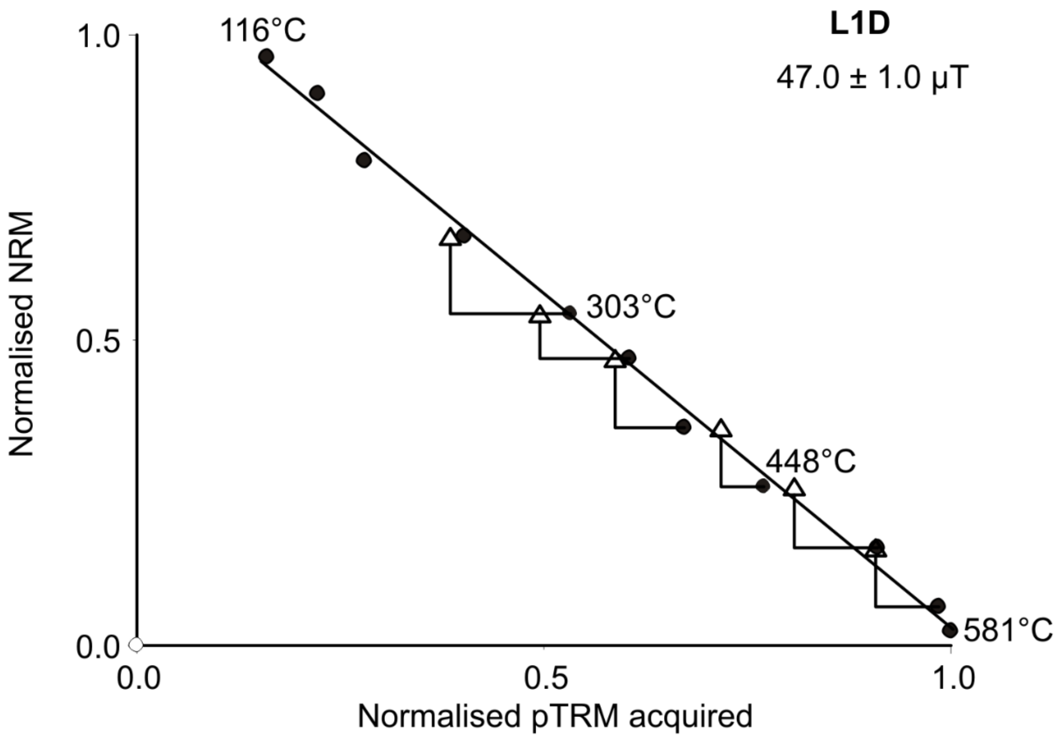


(a) Day-plot; 41, L1, M, N and R4 are ceramic samples; LN and LQ are brick samples. Theoretical curves (Dunlop, 2002) for SD-MD and SP-SD magnetite mixtures are included in the plot.



(b) Hysteresis curve of zoomorphic vessel sample N1E.

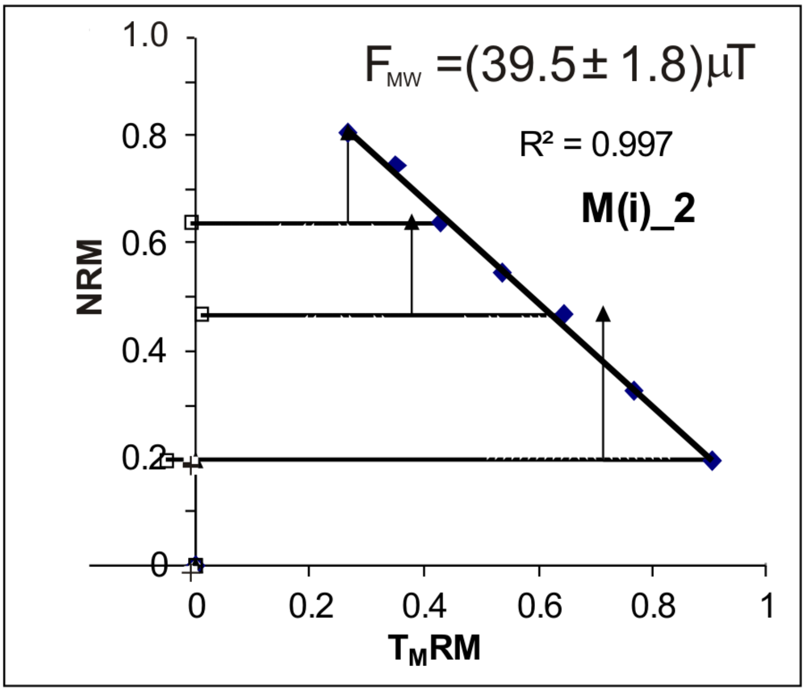
Figure 8. *Thellier-Coe archaeointensity determinations*



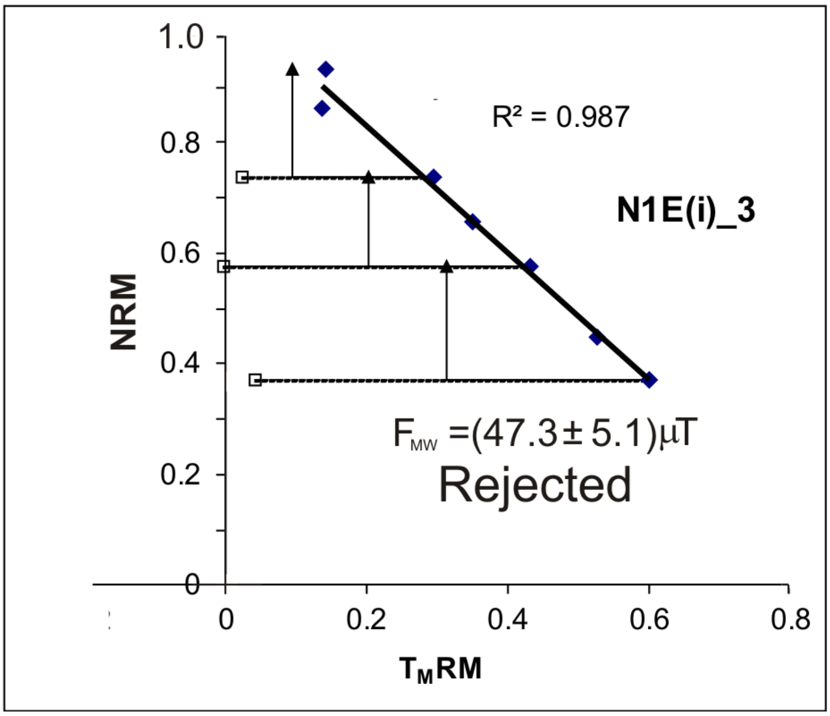
Archaeointensity determination on clay-pot sample L1D.

Figure 9. *Microwave archaeointensity determinations*

(a)

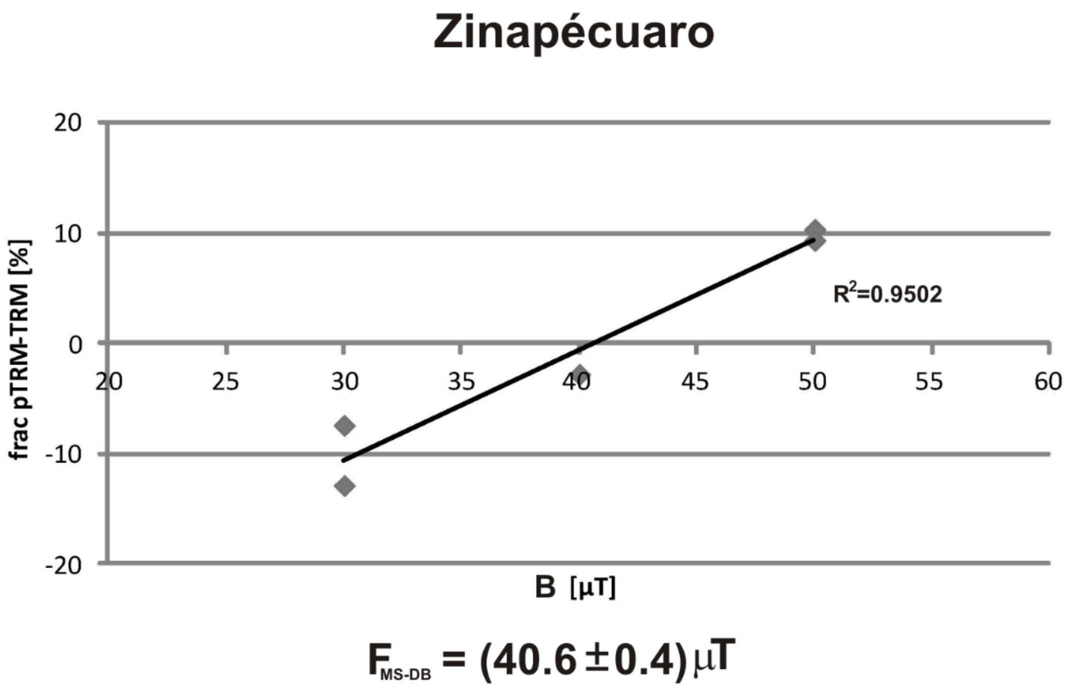


(b)



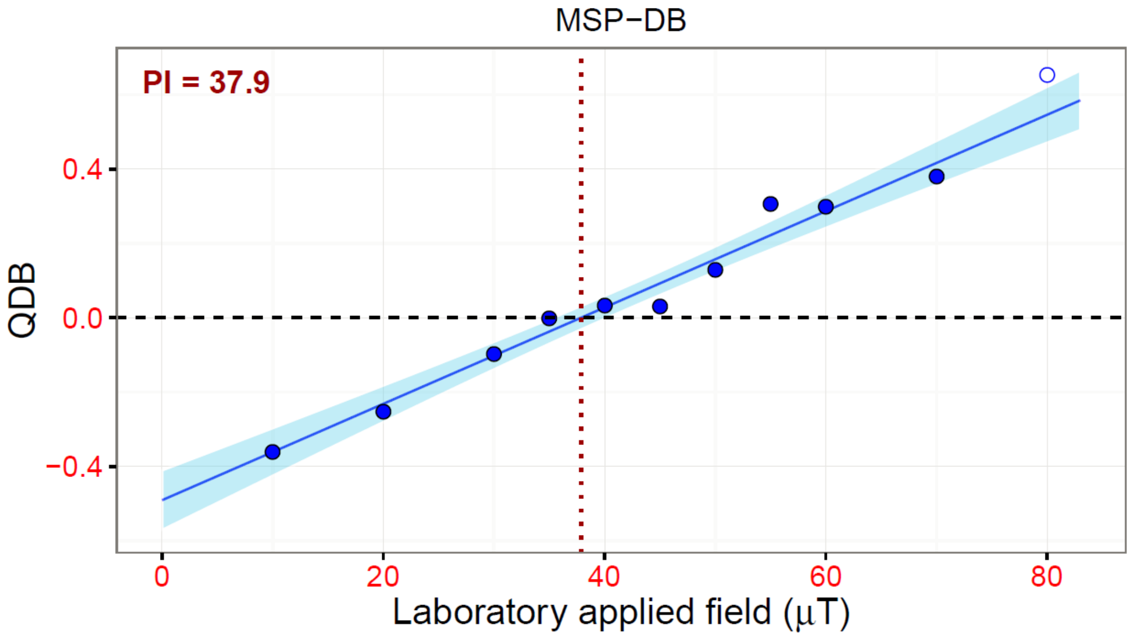
(a) Successful determination on flowerpot sample M; *(b)* Unsuccessful determination (see text) on zoomorphic vessel N1E. Full triangles: pTRM-checks; Open squares; pTRM-tail checks.

Figure 10. *MSP-DB multispecimen archaeointensity determinations*

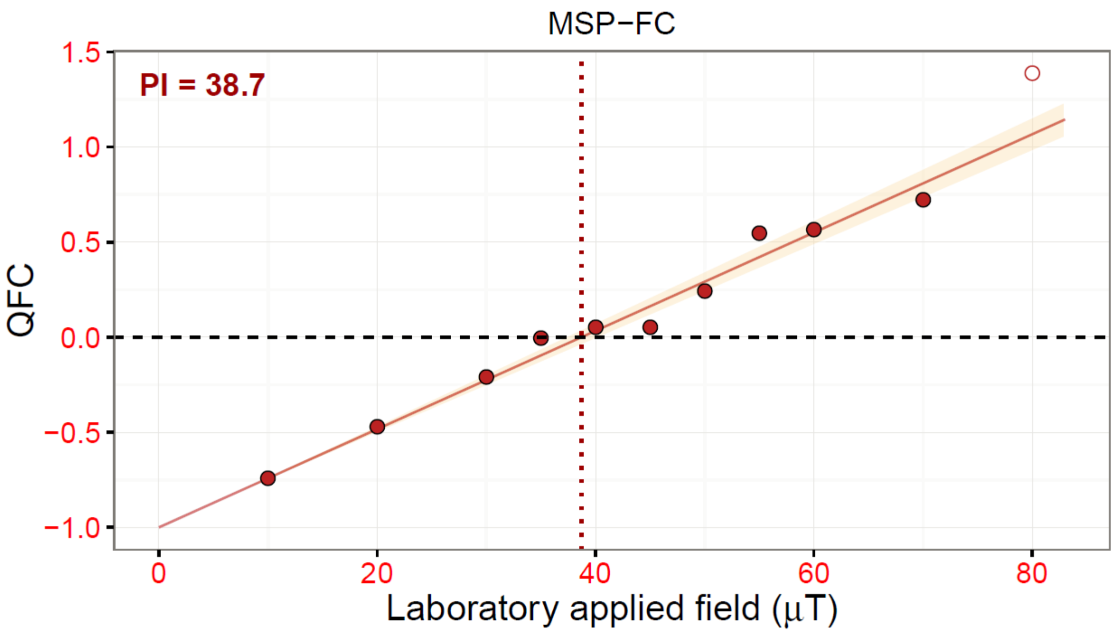


Archaeointensity determination on flowerpot sample M using the original multispecimen method (Dekkers and Böhnel; 2006).

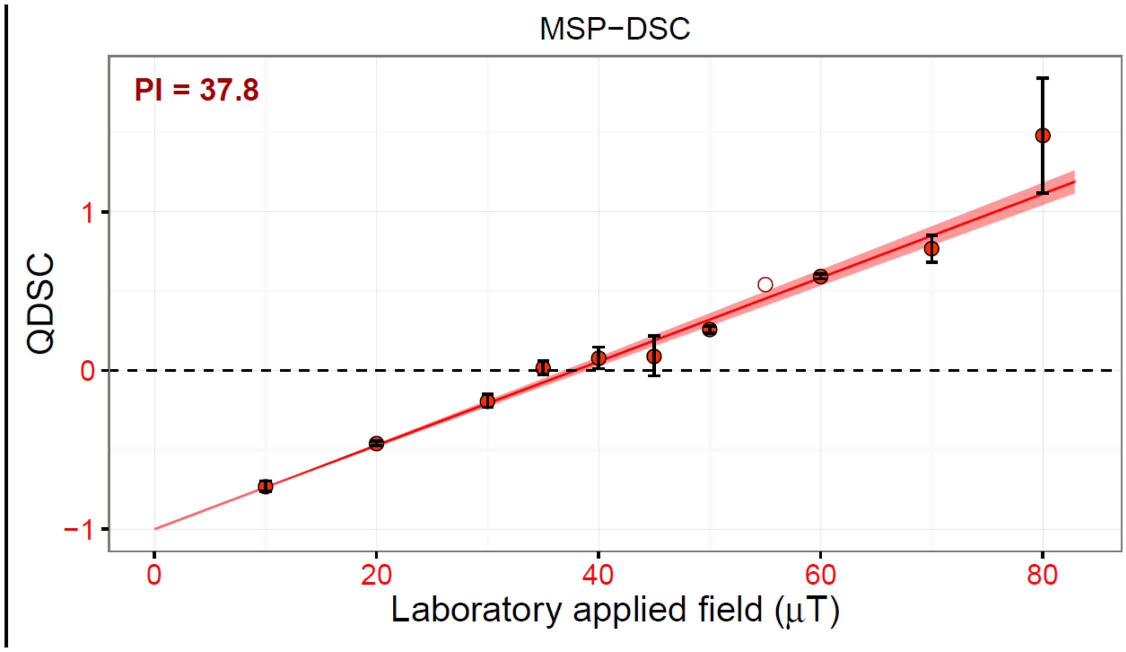
Figure 11. *Extended protocol multispecimen archaeointensity determinations* Archaeointensity determination on flowerpot sample (M1 and M2) using the multispecimen method with corrections steps (Fabian and Leonhardt, 2010); Closed (open) dots represent used (rejected) data from the analysis of the Cook's distance (see text for detail).



(a) uncorrected archaeointensity determination MSP-DB

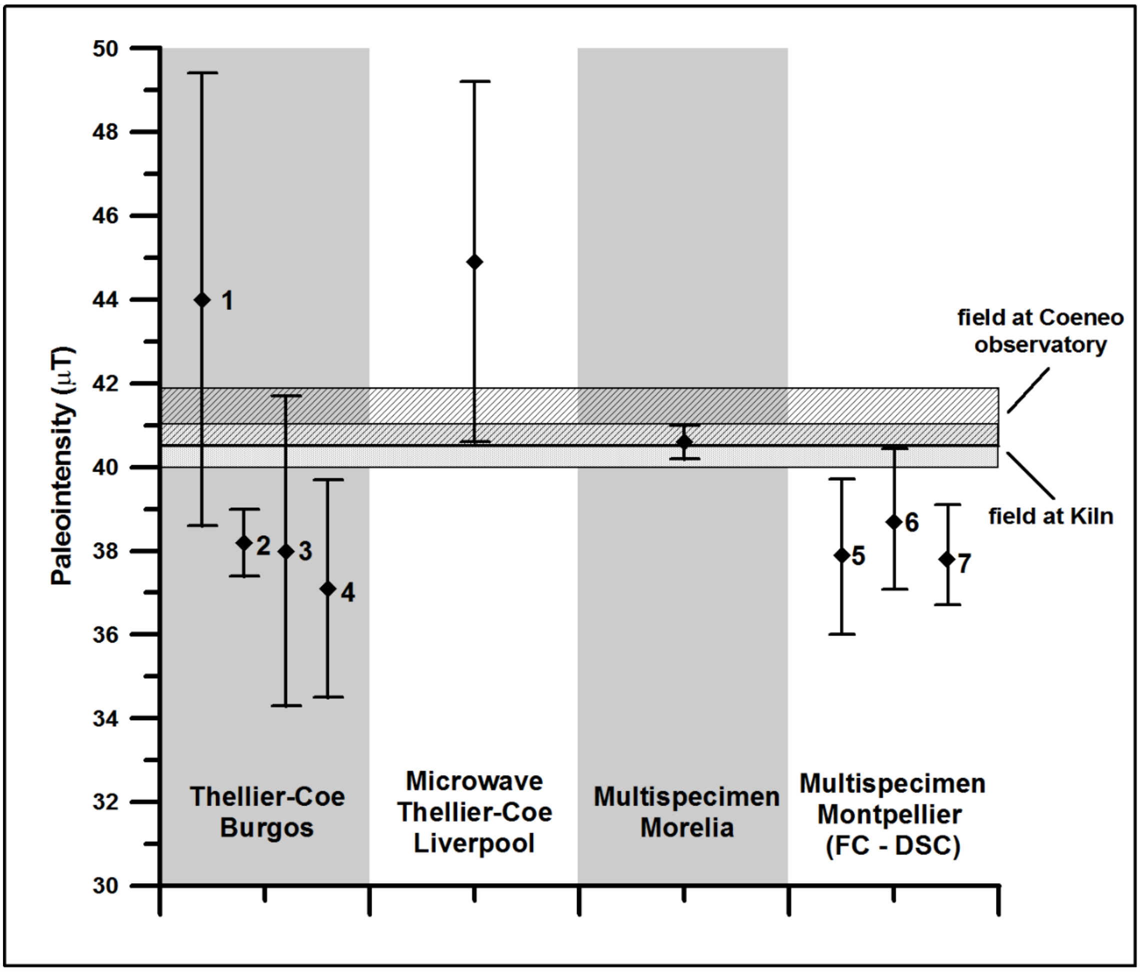


(b) fraction corrected archaeointensity determination MSP-FC



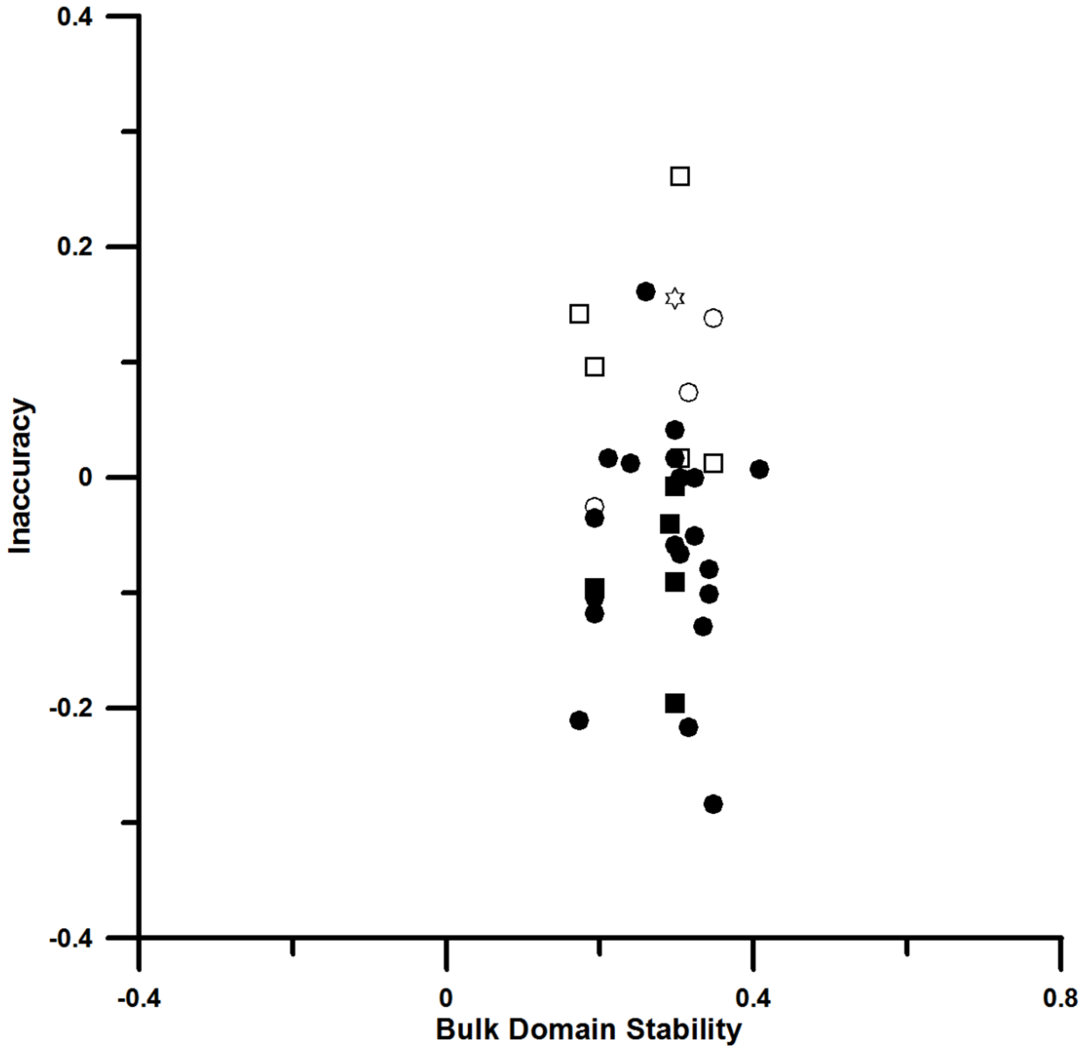
(c) domain-state corrected archaeointensity determination MSP-DSC. The shaded area represents the 95% confidence interval of the best fit slope.

Figure 12. *Summary of palaeointensity results obtained with four methods*



Thellier-Coe (Burgos laboratory), microwave with Thellier-Coe protocol (Liverpool laboratory), original multispecimen method (Morelia laboratory) and fraction and domain-state corrected multispecimen method. (1) Thellier-Coe palaeointensity averaged over all specimens, without anisotropy correction; (2) Anisotropy corrected Thellier-Coe palaeointensity averaged over all specimens; (3) Thellier-Coe palaeointensity averaged over the six analysed archaeological pieces, without anisotropy correction; (4) Anisotropy corrected Thellier-Coe palaeointensity averaged over the six analysed archaeological pieces; (5) Uncorrected multispecimen results (original method); (6) Fraction-corrected multispecimen results; (7) Domain-state corrected multispecimen results. Field values with experimental uncertainties at the study site (kiln) and at near-lying Coeneo observatory are shown.

Figure 13. *Archaeointensity determination quality in Thellier-Coe and microwave experiments*



Relation between palaeointensity inaccuracy ($\ln(B_{anc}/B_{exp})$, B_{anc}: archaeointensity result; B_{exp}: expected field value), bulk domain state (BDS) (Paterson et al., 2017) and quality class of the palaeointensity determination. *Solid circles*: Class-A Thellier-Coe determinations; *Solid squares*: Class B Thellier-Coe determinations; *Open circles*: Class-A microwave determinations; *Open squares*: Class B microwave determinations; *Open star*: Rejected microwave determinations.

Table 1. *Measured geomagnetic field values*

	H [μT]	V [μT]	F [μT]
KILN			
Upper centre	27.1	30.4	40.7
Lower centre	26.0	30.3	39.9
Lower rim	26.0	30.8	40.3
Yard (near kiln)	27.6	30.5	41.1
Kiln (mean)	**26.7 ± 0.8**	**30.5 ± 0.2**	**40.5 ± 0.5**
COENEO OBSERVATORY (100 KM), 5 MEASUREMENTS			
Observatory (mean)	**27.9 ± 1.6**	**30.3 ± 0.7**	**41.2 ± 0.7**

Geomagnetic field values measured in 2011 at the kiln in which samples were baked in an artisans' workshop in Zinapécuaro (Mexico) and and in Coeneo observatory at approximately 100 km distance. *H:* intensity of the horizontal field component; *V:* intensity of the vertical field component; *F:* Total field intensity.

Table 2. *Selection criteria and quality levels*

Class \ Criterion	Thellier-Coe		Microwave	
	A	B	A	B
N	≥ 5	≥ 5	≥ 5	≥ 5
f	≥ 0.35	≥ 0.35	≥ 0.35	≥ 0.35
σ/slope	≤ 0.1	≤ 0.15	≤ 0.1	≤ 0.15
q	≥ 5	≥ 2	≥ 5	≥ 2
MAD	≤ 6	≤ 15		
α	≤ 15	≤ 15		
δ(CK)	≤ 7	≤ 9	≤ 7	≤ 9
δ(pal)	≤ 10	≤ 18		
δ(TR)	≤ 10	≤ 20	≤ 10	≤ 20
δ(t*)	≤ 9	≤ 99		

Selection criteria and threshold values for class A and class B determinations are shown for Thellier-Coe determinations and microwave determinations with the Coe protocol. *Class:* quality class A or B of each determination (see text); *N:* number of NRM-pTRM points used for archaeointensity determination. *f:* fraction of extrapolated NRM used; f is referred to the so-called "true NRM", which is the intersection between linear fit and y-axis (Leonhardt et al., 2004); *σ/slope:* Ratio of the standard error of the slope and the slope of the NRM-TRM diagram *q:* quality factor (Coe et al., 1978). *MAD:* Mean angular deviation of NRM end-point directions at each step obtained from palaeointensity experiments. *α:* angle between the vector average of the data selected for palaeointensity determination and the principal component of the data. *δ(CK):* Difference between the pTRM check and original TRM value at a given temperature normalized to the TRM (Leonhardt et al., 2000); *δ(pal)*: cumulative check error (Leonhardt et al., 2003); *δ(TR)*: relative intensity difference in pTRM-tail check*; δ(t*)*: normalised tail of pTRM (Leonhardt et al., 2004)*;*

Table 3. *Thellier-Coe palaeointensity results*

Sample	Type	Range	N	f	σ/slope	q	MAD	α	δ(CK)	δpal	δ(TR)	δ(t*)	Class	F	ΔF	f-anis	Fcor
2010 experiment (see text)																	
L1-1	vessel	263-581	8	0.62	0.08	6.84	5.6	0.8	6.62	7.47	2.76	5.32	A	**48.9**	**3.7**	0.8282	**40.5**
L1-2	vessel	220-581	9	0.79	0.04	19.19	3.9	1.7	4.33	7.23	1.97	4.64	A	**43.7**	**1.6**	0.8679	**37.9**
L1-3	vessel	116-581	1	0.86	0.04	18.00	7.4	2.9	4.67	0.48	3.94	3.93	B	**52.2**	**2.2**	0.7885	**41.2**
L1D	vessel	116-581	11	0.83	0.02	34.28	5.4	0.4	4.45	1.05	4.45	6.81	A	**47.0**	**1.0**	0.8614	**40.5**
L1E	vessel	162-581	10	0.80	0.03	21.30	1.7	0.6	6.75	0.17	4.74	5.46	A	**42.0**	**1.4**	0.9320	**39.1**
L1H	vessel	220-581	9	0.71	0.03	19.00	3.1	1.6	3.92	3.09	4.42	6.88	A	**42.9**	**1.4**	0.8967	**38.5**
Unweighted Mean intensity L1																	**39.6**
LQI	brick	220-581	8	0.64	0.07	8.44	4.7	1.0	6.41	2.02	0.88	1.21	A	**48.1**	**2.5**	0.9900	**47.6**
LQJ	brick	162-581	10	0.58	0.02	21.33	3.4	1.3	2.78	4.26	0.56	1.05	A	**42.0**	**0.9**	0.9724	**40.8**
LQK	brick	263-581	8	0.62	0.04	10.78	2.2	0.2	2.06	1.44	1.21	1.56	A	**43.1**	**1.7**	0.9548	**41.2**
Unweighted Mean intensity LQ																	**43.2**
M1	flowerpot	116-581	11	0.82	0.02	39.14	4.1	0.2	3.30	9.70	2.70	3.78	A	**39.2**	**0.7**	0.9320	**36.5**
M2	flowerpot	116-581	11	0.82	0.04	17.11	5.4	3.1	6.24	3.31	6.84	17.76	B	**39.7**	**1.7**	0.9279	**36.8**
M3	flowerpot	116-581	11	0.82	0.04	17.92	5.7	2.2	4.86	3.31	6.29	8.35	A	**38.0**	**1.5**	0.9462	**36.0**
MF	flowerpot	162-581	8	0.60	0.05	11.18	4.6	6.2	4.51	0.35	2.43	3.00	A	**48.3**	**2.2**	0.8498	**41.0**
MH	flowerpot	116-499	9	0.74	0.04	16.22	7.1	6.6	4.58	7.95	1.52	3.31	B	**46.5**	**1.8**	0.8363	**38.9**
Unweighted Mean intensity M																	**37.8**
N1A	zm. pot	116-401	7	0.56	0.04	12.64	3.6	4.6	1.52	6.16	4.23	5.53	A	**35.2**	**1.3**	0.8669	**30.5**
N1D	zm. pot	116-581	11	0.90	0.02	35.72	2.1	0.3	5.10	0.85	3.54	4.89	A	**37.8**	**0.8**	0.8669	**32.8**
N1F	zm. pot	162-581	10	0.71	0.02	34.81	5.0	1.8	3.15	2.00	0.76	1.39	A	**37.6**	**0.7**	0.8669	**32.6**
Unweighted Mean intensity N																	**32.0**
Previously fired brick (in 2010) but not heated in the 2010 experiment (see text)																	
LNE	brick	220-581	9	0.73	0.04	15.70	3.4	1.5	5.96	1.87	2.19	2.64	A	**38.4**	**1.6**	0.9736	**37.4**
LNI	brick	116-581	11	0.90	0.03	24.92	2.3	0.7	5.18	8.50	0.60	1.01	A	**37.6**	**1.2**	0.9736	**36.6**
Unweighted Mean intensity LN																	**37.0**
2009 experiment (see text)																	
R4-1	flowerpot	116-351	6	0.46	0.12	2.98	2.4	4.0	4.74	8.36	2.45	4.75	B	**53.4**	**6.4**	0.6934	**37.0**
R4-2	flowerpot	263-581	8	0.60	0.07	7.04	4.3	3.4	7.90	17.37	2.19	4.73	B	**47.4**	**3.3**	0.7025	**33.3**
R4-3	flowerpot	116-401	7	0.49	0.05	8.47	2.8	3.5	3.88	0.99	2.60	5.10	A	**56.2**	**2.6**	0.7509	**42.2**

R4A	flowerpot	116-401	6	0.55	0.04	10.87	4.9	8.5	3.70	0.53	0.79	0.53	A	**41.7**	**1.6**	0.8540	**35.6**
R4G	flowerpot	116-581	11	0.85	0.03	29.25	4.7	2.8	4.25	5.84	5.23	7.70	A	**44.5**	**1.1**	0.8589	**38.2**
R4P	flowerpot	116-351	6	0.50	0.03	11.36	5.1	10.9	3.96	6.60	6.34	9.07	B	**45.9**	**1.5**	0.8752	**40.2**
R4R	flowerpot	116-581	11	0.849	0.05	14.45	5.0	3.4	6.24	1.31	4.49	6.07	A	**46.9**	**2.4**	0.8779	**41.2**
Unweighted Mean intensity R4																	**38.2**

Sample: Sample name. *Type:* Type of piece; zm.pot (zoomorphic pot); *Range:* Temperature interval in °C used for archaeointensity determination. *N, f,* σ*/slope, q, MAD,* α*,* δ*(CK),* δ*(pal),* δ*(TR),* δ*(t*)* and Class as in table 2. Values of δ(CK) and δ(TR) are maximum values in the accepted data points. $F \pm \Delta F$: uncorrected archaeointensity estimate for a single specimen and its standard error; standard error of the archaeointensity estimate is calculated by the product of the standard error of the best-fit line in the Arai plot and the laboratory field; *Fcor*: anisotropy-corrected archaeointensity estimate for a single specimen; *f-anis.*: anisotropy correction factor; *Fcor* (sample mean): mean archaeointensity for each ceramic or brick sample.

Table 4. *Mean Thellier-Coe palaeointensity results. Group*

GROUP	N	CORRECTION	F (μT)	ΔF (μT)
Mean calculation with specimens				
Type A	20	No correction	43.0	5.2
Type A	20	Anisotropy-corrected	38.3	3.6
Type B	6	No correction	47.5	4.9
Type B	6	Anisotropy-corrected	37.9	2.8
2009	7	No correction	48.0	5.1
2009	7	Anisotropy-corrected	38.2	1.1
2010	19	No correction	42.5	4.8
2010	19	Anisotropy-corrected	38.2	3.8
2010 without LN (see text)	17	No correction	43.1	4.8
2010 without LN (see text)	17	Anisotropy-corrected	38.4	4.0
All	26	No correction	44.0	5.4
All	26	Anisotropy-corrected	38.2	3.6
Mean calculation with samples (ceramic/brick pieces)				
All	6	Anisotropy-corrected	38.0	3.7
All	6	Anisotropy-corrected & weighted	37.1	2.6

Specimen or sample group (type A, type B, 2010, 2016, all) used for average calculation (for explanation of different sample groups, see text); *N:* Number of specimens or samples used for calculation of the mean; several specimens were taken from each of six samples (ceramics or bricks) for the palaeointensity experiments; *Correction:* Non-corrected or anisotropy-corrected results; *F ± ΔF*: mean archaeointensity for each specimen or sample group and its error given by standard deviation.

Table 5. *Microwave palaeointensity results*

Sample	Type	Range	N	f	σ/slope	q	δ(CK)	δ(TR)	Class	F	ΔF
M(i)_2	Flowerpot	200-560	7	0.61	0.02	21.65	6.19	4.48	A	**39.5**	**0.9**
M(i)_3	Flowerpot	75-300	9	0.78	0.02	47.95	9.41	1.06	B (*)	**44.6**	**0.6**
N1A(i)_2	zm. pot	150-960	7	0.34	0.06	4.27	3.26	1.89	B	**41.0**	**2.5**
N1A(i)_3	zm. pot	100-300	9	0.73	0.02	48.05	4.73	1.82	A	**46.5**	**0.6**
N1E(i)_3	zm. pot	300-960	7	0.56	0.06	9.20	12.61	4.50	rejected	**47.3**	**2.4**
N1F(i)_3	zm. pot	250-680	7	0.61	0.05	10.35	6.09	4.67	A	**43.6**	**2.1**
N1D(i)_2	zm. pot	250-960	11	0.84	0.04	22.87	7.87	0.75	B	**46.7**	**1.5**
L1(i)_2	vessel	250-720	7	0.39	0.06	5.04	7.90	1.80	B	**52.6**	**3.3**
Mean paleointensity value (N = 7)										**44.9**	**4.3**

Sample: Sample name. *Type:* Type of piece; zm.pot (zoomorphic pot); *Range:* Microwave energy in W the sample was exposed to for the archaeointensity determination. *N, f, σ/slope, q, δ(CK), δ(TR)* and Class as in table 2. Values of δ(CK) and δ(TR) are maximum values in the accepted data points. *F ± ΔF*: archaeointensity estimate for a single specimen and its standard error, calculated by the product of the standard error of the best-fit line in the Arai plot and the laboratory field. Acceptance of type B samples marked with an asterisk is discussed in the text. The mean palaeointensity value is calculated without the rejected result from specimen N1E(i)_3. Results obtained using the horizontal Betty MWS apart from M(i)_3 and N1A(i)_3 where the vertical Tristan system was used.

Table 6. *Extended protocol multispecimen archaeointensity results*

Experimental Protocol	N	n	Paleointensity (μT)	95% Confidence Interval	R^2
DB	11	10	37.9	[36.0 - 39.7]	0.9723
FC	11	10	38.7	[37.1 - 40.4]	0.9966
DSC	11	10	37.8	[36.7 - 39.1]	0.9975

N: number of specimens used in the experimental procedure; *n:* number of specimens used for archaeointensity determination. *DB*: uncorrected determination; *FC*: fraction corrected determination; *DSC*: domain-state corrected determination.