Refining the archaeomagnetic dating curve for the Near East: new intensity data from Bronze Age ceramics at Tell Mozan, Syria

Michele D. Stillinger a, *, Joshua M. Feinberg a, b, Ellery Frahm a, b

a Institute for Rock Magnetism, Department of Earth Sciences, University of Minnesota, 310 Pillsbury Drive SE, Minneapolis, MN 55455, United States
b Department of Anthropology, University of Minnesota, 310 19th Avenue SE, Minneapolis, MN 55455, United States

1. Introduction

Chronological resolution is a fundamental issue in any archaeological investigation and has often focused on relative dating techniques, such as ceramic seriation or textual evidence. These methods rely heavily on interpretations of cultural change and technological dispersion. The introduction of radiocarbon dating more than six decades ago provided archaeologists with an absolute dating technique removing much of the subjective nature of relative dating and resulting in the aptly named Radiocarbon Revolution, which undermined certain theories of cultural diffusion and civilization (Libby et al., 1949; Renfrew, 1979). While radiocarbon dating is now widely accepted and utilized in archaeology, the natural fluctuations in the production rate of $^{14}C$ in the atmosphere results in “wiggles” and plateaus in the global reference curve, which occur on timescales of one decade up to a few centuries (Damon et al., 1978; De Vries, 1958; Suess, 1965). This can result in complicated, even multi-modal, distributions of ages. These problems are especially apparent during the first millennium BCE and have resulted in continued chronology debates in Near Eastern archaeology, where tightly constrained site chronologies are necessary for interpreting site associations with biblical events and locations (e.g., Finkelstein and Piasetzky, 2011).

Archaeomagnetic dating provides both a complementary absolute dating technique that can refine broad chronologies and an alternative in situations where materials suited to radiocarbon analysis are not abundant. As a subfield of paleomagnetism, archaeomagnetism investigates the record of geomagnetic field direction and/or intensity stored within archaeological materials such as fired mudbrick, ceramics, stone, and metal slags. Early archaeomagnetic research was partially responsible for the development of standard paleomagnetic techniques used today due to the ability of heat-treated archaeological materials containing common iron oxide minerals, such as magnetite and hematite, to carry a strong fixed remanence (Sternberg, 2001, 1990). Archaeomagnetism enables the study of small secular variations in the Earth’s magnetic field on scales from decades to millennia. These changes can be plotted through time to create regional reference curves suitable for dating archaeological materials carrying a stable magnetic remanence (Sternberg and McGuire, 1990).

Here we discuss the fundamental methods, sampling procedures, and applications of archaeomagnetism in the context of
magnetic remanence held by pottery at the site of Tell Mozan in northeastern Syria. Detailed discussions of geomagnetism and paleomagnetic techniques are found in Butler (1992), Hulot et al. (2010), and Tauxe (2014), and their application to archaeomagnetism in Eighmy and Sternberg (1990). Interested readers are forwarded to Courtillot and Le Mouel (2007) regarding the development of these fields. Recent archaeointensity data for the Near East, derived from separate studies, have been sufficiently consistent to yield a relatively robust reference curve of field intensity for the last five millennia. Our principal focus here is to demonstrate that the archaeologically defined chronology at Tell Mozan is consistent with that curve. Following our detailed measurement and analysis protocols, all of our well-constrained archaeological dates corresponded to the regional reference curve within a 2σ confidence interval. We also discuss the importance of analyzing and reporting each sample’s archaeointensity results separately, rather than averaging results across a stratigraphic phase. This more nuanced approach may allow researchers to identify and resolve inconsistencies that are due to poor temporal constraints or artifact–feature association.

2. Archaeological overview of Tell Mozan

Tell Mozan has been identified as the ancient city of Urkesh, a political and religious center of the Hurrian culture during at least the third and second millennia BCE (Buccellati and Kelly-Buccellati, 2005; Buccellati, 2005). The 130-ha site (37°03′24″N, 40°59′45″E, elev. 463 m) is located in the Syrian piedmont region, the interface between the Taurus Mountains in southern Anatolia and the Mesopotamian lowlands, along the northern edge of the Fertile Crescent (Fig. 1).

Historically, Urkesh was considered a small vassal of the Akkadian empire (circa 2300–2200 BCE); however, the city’s strategic geographic location might have enabled its residents to exert trade control over copper and other goods entering from the north as well as the local agricultural economy (Buccellati and Kelly-Buccellati, 2005; Buccellati, 2005). The prominence of Urkesh is evidenced by monumental architecture, specifically an Early Dynastic temple complex (circa 2800–2300 BCE) and a royal palace structure constructed during the Early Bronze Age (EBA) III (2269–2240 BCE), as well as cuneiform tablets, seals, and textual inscriptions belonging to the royal household with uniquely Hurrian nomenclature (Buccellati, 2005). Seals belonging to Tar’am-Agade, daughter of the Akkadian king Naram-Sin, suggest that she may have been a queen or Akkadian administrator at Urkesh (Buccellati, 2005). The origins of obsidian artifacts at this time suggest that the city was especially cosmopolitan, having connections throughout Central and Eastern Anatolia (Frahm and Feinberg, 2013a, 2013b; Frahm, 2014).

The chronology of Tell Mozan represents more than 20 years of excavations, which have identified five millennia of occupation (Frahm, 2010:171, and citations within). The period considered in this study is represented by five site-specific phases beginning with an EBA III pre-palace construction in Phase 1 (2334–2270 BCE) and ending with scattered occupations during Phase 6 (1600–1200 BCE) corresponding to the Middle Bronze Age (MBA) IIIC to Late Bronze Age (LBA)/Mitanni period (Table 1). Phases 3 and 4 at the site are further constrained into two distinct time periods each, resulting in a total of seven sequential strata from which archaeomagnetic samples were taken. Phase 2 (2269–2240 BCE) pottery fragments were unavailable at the time of this study.

3. Materials & sampling procedures

All samples used in this research were pottery fragments. The remanent magnetization held in pottery is typically thermal in origin (thermoremanent magnetization, TRM). The best ceramic materials for archaeomagnetic experiments are those that have been fully fired to high temperatures (over 650 °C), presumably above the unblocking temperatures of their magnetic constituents.
which is the temperature at which magnetization becomes randomized. While it may be possible to extract meaningful archaeomagnetic information from ceramics fired at lower temperatures, in this study only fragments that lacked a gray carbon-rich core and appeared fully oxidized, as evidenced by a clear orange, red, or buff coloration, were accepted. While color alone is not a conclusive test of oxidation, it is a helpful first order selection criterion for ensuring that the original pottery had been heated to a high enough temperature necessary to fully combust all carbonaceous materials in the clay (at least 500 °C). Fragments containing large pore spaces and/or lithic temper inclusions were also rejected to avoid crumbling during heating and to minimize the effects of large magnetic grains, respectively. Samples were further chosen based on their size suitability for four replicate 2 × 2 cm specimens. These strict selection criteria resulted in a total of 14 suitable samples (56 specimens) from an original sample size of 97 pottery fragments. See Fig. 2 for photographs of representative fragments. As the target magnetic material is the ceramic body, specimens were cleaned of all debris and surface treatments (i.e., paint and slip), which might alter magnetic measurements and, thus, the resulting dates.

4. Rock magnetic analysis

All pottery fragments (hereafter referred to as samples) underwent a series of standard rock magnetic measurements to determine their suitability for high-temperature experiments, including magnetic susceptibility, hysteresis, and room temperature remanence measurements. The initial natural remanent magnetization (NRM), the sum of all naturally occurring remanence carried by the pottery, as well as all subsequent demagnetizations and partial thermal remanent magnetizations (pTRMs), were measured using either a 2G Enterprises 755 Long Core Magnetometer (referred to hereafter as 2G LC) or a 2G Enterprises 765 Superconducting Rock Magnetometer (SRM). NRM measurement indicated that all specimens carried a strong initial magnetic remanence and subsequent rock magnetic analyses indicated that all samples were resistant to mineral alteration and suitable for thermal experiments necessary for final ancient intensity determination. Rock magnetic results are presented in Supplementary Table S1 and the experiments are outlined here.

4.1. Susceptibility

Low-field magnetic susceptibility (χ) was measured on all four replicate specimens at room temperature to verify the internal homogeneity of each 2 × 2 cm specimen. One specimen from each sample (Specimen Set #1) was tested for the presence of superparamagnetic (SP) grains, nanometer-scale magnetic grains which do not hold a remanence and can spontaneously change their magnetization with small changes in applied field and/or temperature. Susceptibility experiments on all specimens indicated a high degree of internal homogeneity of the pottery samples and, with the exception of one sample, less than 10% SP contribution to magnetization (See Supplementary Table S1). Susceptibility experiments were conducted on a Kappabridge KL4-2 Magnetic Susceptibility Meter in a field of 300 A/m and a frequency of 920 Hz with a sensitivity of 4 × 10⁻⁸ SL. SP measurements were conducted on a Magnon Variable Frequency Susceptibility Meter in a field of 300 A/m at a low frequency (χ₄₆₅) of 465 Hz and a high frequency (χ₄₆₅₀) of 4650 Hz. The frequency dependency of susceptibility (χfd), which corresponds to the percentage of susceptibility held by SP grains, was calculated as:

\[
χ_{fd} = 100\% \left( \frac{χ_{465} - χ_{4650}}{χ_{465}} \right)
\]

4.2. Magnetic hysteresis

Room temperature hysteresis loops, which depict the induced magnetization of a substance in response to an applied magnetic field of varying strength, were generated for Specimen Set #1. How easily magnetic grains respond to an external field (by changing the direction of their magnetization parallel to that field) indicates the stability of the magnetic grains and their ability to carry a remanence. These loops can also be used to interpret mineralogical characteristics of the material, including the concentration of ferromagnetic material and average magnetic domain state behavior. Magnetic domains are regions of uniform magnetization within a magnetic particle or grain and are related to their size and shape. Ideally, for paleointensity experiments specimens would contain small elongated magnetic grains, approximately 30–100 nm for pure magnetite, for example (Dunlop and Özdemir, 2007), and display non-interacting single domain (SD) behavior where all the magnetization in a grain lies in one direction. SD grains carry a strong stable remanence and are ideally suited to archaeomagnetic experiments; however, perfect SD grains are rare in nature and remanence is ultimately affected by grain size, shape, and magnetic interactions (Butler, 1992; Day et al., 1977; Dunlop and Özdemir, 1997; Muxworthy, 2003). Instead, most materials used for paleointensity experiments display a range of magnetic domain states, from very small SP grains (<30 nm) to large multidomain (MD) grains (>10–20 μm).
Day plots (Day et al., 1977) are often used to display the general domain state in specimens based on hysteresis measurements (Fig. 3). While the Day plot parameters are based on specific grain sizes and compositions, the plot can still be used as a first order determination of specimen suitability for paleointensity measurements. Hysteresis measurements indicated that all Tell Mozan specimens generally fell in the pseudo-single domain (PSD) range (~100 nm to 10–20 µm for pure magnetite), which appears typical of materials made of levigated or refined clays used in ceramics (Fig. 3). PSD grains display SD-like qualities under archaeomagnetic experiments, but may also contain a proportion of larger MD grains, which do not carry a stable remanence. Hysteresis loops were measured on a Princeton Applied Research Vibrating Sample Magnetometer (VSM) from 0 to 1 tesla (T) with a nominal sensitivity of $2 \times 10^{-8}$ Am$^2$. Results are displayed in Supplementary Table S1.

4.3. Alternating field (AF) demagnetization

While archaeointensity methods work best on samples that carry a single magnetic remanence, most geologic and archaeologic materials are capable of recording more than one component of magnetization. This frequently occurs in cooking vessels that are reheated to a lower temperature than the initial kiln firing. The result is multiple components of remanence in different directions and/or intensities that affect the total NRM. These secondary components are usually minor and their effects can often be eliminated from final intensity calculations. To ascertain which

![Sample images](image-url)
samples, if any, contained more than one component of magnetization, the NRMs of specimen Set #1 were demagnetized (prior to hysteresis measurements) in a stepwise alternating field (AF) from 2.5 to 170 mT using the 2G LC. Analysis of vector endpoint diagrams, which display the directional decay of remanence during demagnetization (see insets of Fig. 4 for examples), indicated that all Tell Mozan samples contained only weak secondary components that were removed quickly at low fields. The demagnetization spectra generated for each sample were also used to calculate their median destructive field (MDF), which is the field strength required to demagnetize half of the NRM. An elevated MDF value indicates that the sample’s magnetic mineral assemblage contains a higher fraction of SD or PSD grains. Results from rock magnetic experiments are listed in Supplementary Table S1.

5. Archaeointensity methods

5.1. Thermal analysis

Because all pottery samples were magnetically unoriented (not found in their original firing position), only the strength or intensity of the ancient field could be determined. All archaeointensity experiments were conducted at the Institute for Rock Magnetism at the University of Minnesota in a shielded room with a background field less than 200 nT. Archaeointensity data were obtained using the IZZI protocol of Tauxe and Staudigel (2004), a modification of the Thellier-style double heating method (Thellier and Thellier, 1959). This method involves repeated heating and cooling of a specimen over successive temperature intervals, where each temperature step (Ti) demagnetizes a portion of the NRM held by grains whose blocking temperatures are <Ti; and replaces it with a laboratory or partial thermal remanent magnetization (pTRM) created with a user-specified field strength. Processes that impart these pTRMs are linearly related to the applied magnetic field; therefore, the ancient magnetic field intensity can be determined in materials that contain SD-like magnetic grains using the following relationship (simplified based on Thellier laws of SD behavior):

\[ B_{anc} = \left( \frac{M_{NRM}}{M_{lab}} \right) B_{lab} \]

where \( B_{anc} \) is the unknown ancient field intensity, \( B_{lab} \) is the laboratory-applied field, \( M_{lab} \) is the laboratory-acquired magnetization, and \( M_{NRM} \) is the original NRM of the specimen.

Three unaltered specimens from each sample were heated in air using an ASC Model TD-48SC Thermal Demagnetizer Furnace from 150 °C to a maximum of 650 °C and then fan-cooled to room temperature. In-field steps were performed at 30 μT. Remanent magnetization was measured after each heating step with the 2G LC. Ideally, heating steps should be chosen such that the percentage of pTRM gained is the same as the NRM lost after each temperature interval. Temperature intervals that are too large may remove the majority of remanence so quickly that paleointensity estimates are inaccurate. Conversely, temperature intervals that are too narrow can be too time-consuming. In this study, three specimens from each sample were heated in 25 °C intervals beginning at 150 °C and continuing to 650 °C, or until <5% of the NRM remained. Each heating cycle was applied such that specimens were first warmed to 20 °C below Ti, held at this temperature for 10 min, then slowly heated to Ti where they were held for 15 min before being fan-cooled to room temperature. This cautious approach to heating allowed us to avoid extreme heat differentials with the samples that may have caused them to fracture.

To identify mineralogical changes and the effects of multidomain grains during the procedure, pTRM checks (Coe et al., 1978) and pTRM tail checks (Riisager and Riisager, 2001), respectively, were performed every other step starting from the third temperature interval. These checks involve reheating the specimen to a previous lower temperature interval to see if the pTRM gained at that blocking temperature is repeatable.

5.2. Correction for magnetic anisotropy

The direction of magnetization in archaeological ceramics can be highly anisotropic depending on the method of fabrication. Wheel-thrown pottery is particularly susceptible to such effects, as the preferential alignment of clay platelets and the shape of magnetic grains can alter the acquisition of magnetization during stepped heating procedures (Aitken et al., 1981; Rogers et al., 1979). To quantify and correct for this effect, the anisotropy of anhysteretic remanent magnetization (ARM) was calculated as a second rank tensor for each specimen in Set #1 and then used to correct the paleointensity data accordingly. ARM is qualitatively similar to a TRM in that it is a magnetization acquired in response to a stable field in an environment with progressively decreasing amounts of randomizing energy. However, the energy in an ARM arises from rapidly fluctuating electromagnetic fields rather than thermal energy, thus, ARM do not generate the same thermochemical alteration in a sample that is often associated with a TRM. Prior to ARM experiments, specimens were completely demagnetized at room temperature along their X, Y, and Z-axes using the 1.1 T AF demagnetization function on the VSM. Samples were further AF demagnetized along all axes in a 200 mT field with a 0.01 mT decay rate using a Precision Instruments D-2000 DTECH AF Demagnetizer (DTECH). Each specimen’s demagnetized state was measured using the SRM. The specimens were then given an ARM using a 30 μT DC biasing field and a 150 mT AF demagnetization field with 0.01 mT decay rate on the DTECH. ARM intensities and directions were measured on the SRM and the procedure was repeated along all six coordinate axes.

5.3. Cooling rate correction

The difference between laboratory cooling and natural kiln cooling rates can also alter the magnitude of a TRM, with slower cooling rates resulting in a higher TRM when more SD sized grains
are present (Biggin et al., 2013; Dodson and McClelland-Brown, 1980; Halgedahl et al., 1980; Yu, 2011). To correct for this cooling rate effect, one previously heated specimen from each sample was heated to 600 °C and cooled to room temperature three times: first at a “fast” laboratory cooling rate of ~60 min (T1), then at a slower cooling rate of 24 h (T2) at ~25 °C per hour, to approximate kiln cooling, and then again at the fast rate (T3). The addition of a second fast cooling cycle allows for the identification of any mineralogical changes that might be occurring during successive heating. The ratio of the average of the fast cooling rates and the slow cooling is the cooling rate correction ($F_c$) calculated as:

$$F_c = \frac{(T1+T3)}{2T2}$$

The initial thermal demagnetization of the samples revealed that most of the remanence was carried by magnetite and titanomagnetite, or their partially oxidized equivalents (Curie temperature ≤580 °C). Therefore, by cooling the specimens from 600 °C we can determine the cooling rate effects on the minerals holding the primary remanence. The difference between the two fast cooling rate steps was less than 5% for all samples with the exception of one (sample TM51 at 6.4%). The majority of samples also displayed little difference (less than 9%) between slow and fast cooling rates, with the exception of sample TM40 (12%), an indication that PSD grains carry the majority of remanence in most specimens (Yu, 2011). Final $F_c$ values are given in Supplementary Table S2 and the correction was used in final intensity determinations.

5.4. Final selection criteria

As outlined in Section 5.1, the final ancient intensity recorded by a specimen can be determined from the product of the laboratory field and the ratio between the natural (NRM) and laboratory remanences (pTRM). This ratio is the absolute value of the slope of a normalized Arai plot (Nagata et al., 1963), which shows a specimen's remaining NRM on the y-axis versus the pTRM gained on the x-axis as the sample is heated in successive steps in a known

![Arai plots and vector endpoint diagrams](image-url)
laboratory field (Fig. 4). The choice of which temperature steps to include in the ancient intensity calculation can be subjective, and requires the consideration of secondary remanence and multidomain contributions. To make such analyses more objective, the best-fit line of the data (bounded by a series of temperature intervals) is typically subjected to a number of selection criteria to ensure that only the highest quality intensity data are accepted (see Paterson et al. (2014) for a summary of current paleointensity statistics). Intensity results generated for each specimen in this study were subjected to the following five acceptance criteria:

1. A Deviation Angle (DANG) of <5° (the angle between the best-fit line of the directional data used in the slope calculation and the best-fit line from the origin through the center of mass of the directional data as represented in the vector endpoint diagram) (Tauxe and Staudigel, 2004).

2. A maximum angular deviation (MAD) of <10° (representing the variance of the points within a particular temperature interval used to define the direction) (Kirschvink, 1980).

3. An f_{VDS} value of >0.8 (fraction of the total NRM used to calculate the slope of the line normalized by the vector difference sum (VDS) of the zero field demagnetization data) (Gee et al., 1993). The VDS of the magnitudes between successive demagnetization steps more closely estimates the actual total NRM by aligning all remanence components in the same direction. A smaller f_{VDS} value indicates multiple components of remanence.

4. A mean difference ratio (Mean DRAT) value of <10% (the average difference between the pTRM and pTRM) checks normalized by the best-fit line (Selkin and Tauxe, 2000), which is an indication of alteration during heating.

5. A mean multidomain (Mean MD) value of <5% (here calculated as the average of the absolute difference between pTRM and pTRM tail checks normalized by the best-fit line of the zero-step NRM data) indicating the percent contribution from multidomain grains.

If a specimen failed two or more criteria, it was immediately rejected. If a specimen failed only one of the selection criteria, the result of that criterion was scrutinized and compared to the results for the remaining specimens. If a specimen failed two or more criteria, it was immediately rejected. If a specimen failed only one of the selection criteria, the result of that criterion was scrutinized and compared to the results for the remaining specimens. If a specimen failed one of the selection criteria, it was immediately rejected. Specimen TM09C failed the mean DRAT criterion by only 4%; however, results for the remaining criteria were strong enough that we felt this specimen should be accepted in final intensity calculations. Therefore, the final intensities for samples TM09, TM64, TMm01, and TMm04 were calculated on the average of only two specimens each.

Average intensity results for each sample, corrected for anisotropy and cooling rate, are presented in Table 2 along with their corresponding archaeological phase dates and the virtual axial dipole moment (VADM), which is the estimate of the strength of the magnetic dipole aligned with the rotational axis that would generate the ancient intensity observed at specific location. Final intensity results and error rates are based on a weighted average of the specimens for each sample. Individual results from each sample were kept separate to recognize variability or concurrence of intensity within each phase and to identify unusual intensity values that might be the result of artifact contamination between phases.

Results from this study are juxtaposed with existing regional archaeointensity data for Syria, Israel, Jordan, Egypt, and Turkey in Fig. 5 (Aitken et al., 1984; Ben-Yosef et al., 2009, 2008; Ertepinar et al., 2012; Gallet and Al-Maqdissi, 2010; Gallet and Le Goff, 2006; Gallet et al., 2014, 2008, 2006; Genevey and Gallet, 2003; Hussain, 1987; Odah, 1999; Odah et al., 1995; Shaar et al., 2011; Walton, 1990, 1986). Regional data were compiled from the GEOMAGIA50 online database (Donadini et al., 2006; Korhonen et al., 2008) or the individual study if not available online. Regional data chosen for the base curve were constrained to those archaeomagnetic studies that employed the IZTI protocol or other double heating methods which

### Table 2
Average magnetic intensity and archaeological correlation.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>n</th>
<th>Site phase</th>
<th>Near East chronology</th>
<th>Phase attributed age (BCE)</th>
<th>Corrected intensity (μT)</th>
<th>VADM (ZAm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM09</td>
<td>2</td>
<td>6</td>
<td>MBA IIC/LBA</td>
<td>1595–1200</td>
<td>38.9 ± 0.8</td>
<td>69.8 ± 1.4</td>
</tr>
<tr>
<td>TM23</td>
<td>3</td>
<td>6</td>
<td>MBA IIC/LBA</td>
<td>1595–1200</td>
<td>49.7 ± 2.2</td>
<td>89.2 ± 4.0</td>
</tr>
<tr>
<td>TM13</td>
<td>3</td>
<td>5</td>
<td>MBA II</td>
<td>1900–1595</td>
<td>42.1 ± 0.7</td>
<td>75.6 ± 1.1</td>
</tr>
<tr>
<td>TM14</td>
<td>3</td>
<td>5</td>
<td>MBA II</td>
<td>1900–1595</td>
<td>38.9 ± 1.4</td>
<td>69.8 ± 2.5</td>
</tr>
<tr>
<td>TM51</td>
<td>3</td>
<td>4b</td>
<td>MBA IIA</td>
<td>2000–1990</td>
<td>52.6 ± 2.2</td>
<td>94.4 ± 4.0</td>
</tr>
<tr>
<td>TM32</td>
<td>3</td>
<td>4a/4b</td>
<td>MBAI/MABA IIA</td>
<td>2112–1900</td>
<td>52.2 ± 2.8</td>
<td>93.6 ± 5.0</td>
</tr>
<tr>
<td>TM72</td>
<td>3</td>
<td>4a/4b</td>
<td>MBAI/MABA IIA</td>
<td>2112–1900</td>
<td>42.7 ± 2.2</td>
<td>76.6 ± 4.0</td>
</tr>
<tr>
<td>TMm01</td>
<td>2</td>
<td>3b/4a</td>
<td>EBA IV/MABA I</td>
<td>2192–2004</td>
<td>48.9 ± 2.0</td>
<td>87.8 ± 3.5</td>
</tr>
<tr>
<td>TMm04</td>
<td>2</td>
<td>3b/4a</td>
<td>EBA IV/MABA I</td>
<td>2192–2004</td>
<td>40.0 ± 1.7</td>
<td>71.9 ± 3.0</td>
</tr>
<tr>
<td>TM35</td>
<td>3</td>
<td>3a/3b</td>
<td>EBA III/EBA IV</td>
<td>2240–2112</td>
<td>48.6 ± 1.7</td>
<td>87.3 ± 3.0</td>
</tr>
<tr>
<td>TM64</td>
<td>2</td>
<td>3a</td>
<td>EBA III</td>
<td>2240–2193</td>
<td>60.4 ± 3.3</td>
<td>108.4 ± 5.9</td>
</tr>
<tr>
<td>TM65</td>
<td>3</td>
<td>3a</td>
<td>EBA III</td>
<td>2240–2193</td>
<td>42.4 ± 1.5</td>
<td>76.0 ± 2.7</td>
</tr>
<tr>
<td>TM40</td>
<td>3</td>
<td>1</td>
<td>EBA III</td>
<td>2334–2270</td>
<td>42.7 ± 1.8</td>
<td>76.7 ± 3.3</td>
</tr>
<tr>
<td>TM56</td>
<td>3</td>
<td>1</td>
<td>EBA III</td>
<td>2334–2270</td>
<td>44.9 ± 1.1</td>
<td>80.7 ± 1.9</td>
</tr>
</tbody>
</table>

**VADM (ZAm²)** — Virtual axial dipole moment × 10²¹ Am², **n** — number of specimens per sample.
applied pTRM checks. To include only the highest quality data for the base curve, the regional data were further reduced to include only those studies applying some form of anisotropy correction, with two exceptions: 1) intensity data from metal slags, where no correction was necessary (Ben-Yosef et al., 2009, 2008; Shaar et al., 2011), and 2) recent high-quality results from studies where anisotropy was considered negligible due to selective alignment of the samples during measurements (Ertepinar et al., 2012; Gallet and Al-Maqdissi, 2010; Gallet et al., 2014, 2008).

In Fig. 5, regional data points (gray open circles) represent average archaeointensity vs. archaeological date. These dates are derived from relative dating techniques such as pottery seriation, stratigraphy, and/or textual evidence. Black open squares represent average archaeointensity results where archaeomagnetic samples were calibrated using radiocarbon-dated strata (Ben-Yosef et al., 2009, 2008; Shaar et al., 2011; Walton, 1986). The purple dashed line (in web version) is the CALS7K.2 global paleomagnetic field model generated using archaeomagnetic full vector and lake sediment data for the past seven millennia (Korte et al., 2005). This model was chosen for comparison as it most closely captures the low-frequency paleointensity trends for the time period in question. The CALS10k.1b model (Korte et al., 2011) covering the past 10 ka has superseded this model; however, it is controlled by sediment data and tends to strongly smooth the curve when compared to such a narrow time interval, therefore it was not included here. It should be noted that the CALS7K.2 model also underestimates the dipole moment, particularly for the period around 1000 BCE.

Fig. 5 also illustrates a new regional curve model based on a 25nd order polynomial fit of the high-quality regional data. A 25th order polynomial was selected because it produced the lowest residual standard deviation without introducing spurious features in the reference curve that were unsupported by data (See Supplementary Figure S1). In essence, it captures the major intensity trends while eliminating the contributions from outlying data points. The strength of the ceramic seriation ages for Tell Mozan was confirmed if the intensity results correlated with the model within one standard deviation (σ) (represented by the shaded region about the line). Fig. 5b focuses on the region of the curve representing the Tell Mozan results. From this figure, it can be seen that all samples fell within 2σ of the model (dashed line), with the majority of intensities from Phases 1 through 5 (2334–1600 BCE) falling within 1σ.

Sample TM09, from Phase 6 (1600–1200 BCE), yielded a paleointensity estimate with very low uncertainty but fell well outside the 1σ confidence interval. Three interpretations of this result may apply. First, the actual age of this sample may be closer to its oldest possible archaeological age (1600 BCE as defined by the phase) in order to correspond with a period of intensity minimum that occurred around 1800 BCE. Second, this sample may record an extension of the regional intensity minimum into the first part of the 15th century BCE for the northeastern portion of Syria. Third, this sample may be simple contamination from Phase 5. We favor this final scenario as the most likely explanation as there was no documented feature association for the samples from Phase 6, the style of pottery is similar to those found in Phase 5 deposits, and the paleointensity estimate for TM09 is identical to that of TM14 and nearly identical to TM13, both from Phase 5. Sample TMm04 also falls just outside the 1σ error envelope. Upon reexamination of the feature association, it was determined that this sample was from a pit-like feature that cut down through 4b to 3a strata but contained primarily materials from Phase 3b/4a. Therefore, this sample could quite easily be a slightly younger contamination from Phase 4a/4b or incorrect phase association in the database. Based on these assessments, both of these samples were ultimately rejected as representative intensities for their time period.

7. Discussion

From Fig. 5, it is clear that our results display somewhat lower archaeointensity measurement uncertainties than previous studies, due in part to detailed methodology and strict selection criteria. Additionally, we have plotted each sample (from independent vessels) separately. Some of the regional results are from studies that have averaged multiple samples within a phase, which can result in high intensity uncertainties. Because the magnetic field is dynamic and fluctuates yearly, variability in intensity across broadly defined phases would be expected. Reporting only the mean of archaeointensities for a phase may not capture the mode in the data and result in a curve with larger standard deviations. Individual sample results from Mozan were plotted separately on the model to recognize this variability and highlight the differences that can occur in even relatively short phases. This also allows for the scrutiny of possible discrepancies in artifact association at the site level. For example, if only the average of the Phase 3b/4a (2192–2004 BCE) samples were considered, the resulting intensity

Fig. 5. Regional curve (a) of field intensity for the Near East for the Bronze and Iron Ages as modeled in this study. Regional data1 are those intensities calibrated with traditional archaeological dating techniques, regional data2 are those intensities calibrated with radiocarbon dating (see article text). (b) Close-up of Mozan intensity results.
would more closely match the curve. By keeping the results separate in our initial analysis of the results, we can see that sample TMm04 may be slightly younger than TMm01, and is most likely the result of an artifact from Phase 4b being displaced from its original depositional context.

The results of samples TMm04 and TM09 highlight the importance of assessing the strength of the original archaeological interpretation of the artifacts or features themselves. For example, by plotting intensity results for each sample separately, we were able to identify samples that appear to be in disagreement with their defined chronology. In one instance, this likely reflects contamination of strata from the re-use of previous occupational materials in later construction, a typical practice at ancient tells in the Near East. However, this object may also represent an item that was used long after the original firing, such as an heirloom object or long-lived storage vessel passed down through the generations. The careful analysis of individual sample results highlights the importance of increasing sample size per phase. Ultimately, the combination of relative and absolute dating techniques, detailed site records, and strict sample selection is necessary for the construction of an accurate site chronology.

The regional trend calculated in Fig. 5 also assumes that the previously published data are relatively accurate; however, the dates used to fix the regional intensity values are themselves based primarily on relative archaeological chronologies. Only four of the archaeomagnetic studies represented in Fig. 5 utilized radiocarbon dating along with stratigraphic seriation to date their samples (Ben-Yosef et al., 2009, 2008; Shaar et al., 2011; Walton, 1986), which in some cases results in narrower chronological association when compared to the other regional data. If only radiocarbon dated studies were considered, they would not supply enough data to create a meaningful intensity curve. If radiocarbon or other absolute dating is not available or utilized in cross calibration, then the archaeologically defined dates will be based solely on relative...
Dating techniques, such as pottery seriation, which is common in Near East archaeology. In many cases, phases or site dates are broadly defined, resulting in equally broad archaeomagnetic results. For example, Phase 6 at Tell Mozan contained 400 years of deposits representing occupations that have not yet been as thoroughly evaluated as the first five phases. As many of the regional intensity studies used in the base model are a decade old or more, a reassessment of their associated chronologies and an update of the online paleointensity databases with more recent results may be in order.

Finally, it should be noted that the standard deviation around the curve in our model is constant. As can be seen, certain periods are better represented by archaeointensity data than others. An alternative approach would be to apply a moving window average to capture the variability in the data set (Kovacheva and Toshkov, 1994; Sternberg and McGuire, 1990). Here we apply a 300-year moving window average in 20-year increments to the regional data (Fig. 6b) to compare against our polynomial model (Fig. 6a) and expand the models out to 800 CE. Both intensity models (bold lines) are similar and capture the intensity maxima at ~900 and ~450 BCE and the relative plateau between ~2600 and 2200 BCE. The moving window model does appear to minimize the standard deviation in periods with more data; however, it also displays slightly more noise, particularly around 2800 BCE and 50 BCE where there is a lack of data. Fig. 6c displays the calculated difference between the two curves. Here we see very minor variance between the models’ VADM values, with only $\pm 1 \times 10^{21}$ Am² for dates between ~3000 and 50 BCE, indicating both models generate a similar average intensity curve. A comparison of the difference between the standard deviations in the models (Fig. 6d) highlights where there is a great deal of archaeomagnetic data available (area above the horizontal line) and where more data are needed (area below the line). Important to note is the area between ~1000 and 800 BCE in Fig. 6d (shaded) where there appears to be plenty of data, but the broad variability in that data results in a larger standard deviation for the moving window model. These data are represented by extremely narrow chronological associations and two short-lived spikes in intensity. A 300-year moving window average is not appropriate to capture such narrow geomagnetic behavior. From these two model comparisons, we can see that more data are needed for the first millennium BCE but that both model curves are relatively similar.

An alternative approach is to apply a cubic spline model to the regional data; however, this type of model suffers from the same problems as the moving window model in that they both have time-dependent errors, which are more difficult to broadly disseminate. The polynomial model provides a single continuous function that sufficiently captures the same geomagnetic behavior as the standard moving window model and eliminates the effects of uncertain data points, providing a smoother curve more suitable for archaeomagnetic dating. We argue that the polynomial-based reference curve is more useful and accessible to the archaeological community because the model and its associated error can be readily reproduced by other researchers for comparison (see Supplementary Figure S1 for polynomial coefficients).

8. Conclusion

Archaeomorphic measurements were conducted on fourteen pottery samples from seven well-defined and constrained strata dated between 2334 and 1200 BCE at Tell Mozan, Syria. Our results are plotted against prior archaeointensity data for the Near East obtained from recently published studies using similar paleointensity measurement and correction techniques. Initial sample selection and intensity measurements underwent strict acceptance criteria before estimating final archaeointensity, which resulted in a 90% success rate using the detailed paleointensity techniques outlined. We also introduce a new archaeointensity model for the region derived from a 25th order polynomial fit of the regional data and spanning the first three millennia BCE. This model reduces the standard deviation of the residuals about the best-fit line of the data, minimizes the effects of outlying data points, and provides a smooth reference curve suitable for dating artifacts of unknown age. We obtain an 86% concurrence between the archaeologically defined ages and the new model within 1σ.

The results of our study highlight three important considerations for future archaeomagnetic research. First, consistent measurement techniques and strict acceptance criteria result in lower error rates for archaeomagnetic measurements. Second, the application of rigorous sampling criteria can significantly reduce a large sample population, requiring that more samples from each site phase be taken; however, the relatively high NRM held by most of the pottery samples in this study indicates that specimens smaller than 8 cm² could be utilized, which would allow researchers to increase their overall sample populations. Finally, we propose that future archaeointensity results are always uploaded to the regional GEOMAGIA50 database by individual sample (or by individual heated object) for each time period they represent, instead of the reporting only an average of multiple samples from a specific phase. This would allow for a more detailed interpretation of the data, such as identifying modes in intensity distributions. This study adds to the literature confirming the applicability of archaeomagnetism as an independent and complementary tool in constructing archaeological chronologies, provides a new regional reference curve model for archaeomagnetic dating, and contributes twelve new results to the increasingly dense paleointensity data for Syria and the greater Near Eastern region during the Bronze Age.

Acknowledgments

The authors wish to thank Giorgio Buccellati and Marilyn Kelly-Buccellati, directors of the Tell Mozan excavations under the auspices of the International Institute for Mesopotamian Area Studies; the Directorate General of Antiquities and Museums, Syrian Arab Republic for approving the export of Tell Mozan artifacts; Julie Bowles and Mike Jackson; and the Institute for Rock Magnetism (IRM) at the University of Minnesota. This manuscript was improved by the comments of Robert Sternberg and one anonymous reviewer. This is IRM publication number 1405.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jas.2014.10.025.

References


Egypt and western Asia during the second millennium BC. Nature 310, 305–306.


