AGU PUBLICATIONS

Geochemistry, Geophysics, Geosystems

RESEARCH ARTICLE
10.1002/2014GC005311

Key Points:
- The European archeointensity data ranked in four quality categories
- Geomagnetic field reconstructions depend on the quality of the selected data
- Archeointensity regional models for Europe using high-quality data

Supporting Information:
- Readme
- Figures S1-S3
- Tables S1-S2
- Supplementary Material

Correspondence to:
F. J. Pavón-Carrasco,
javier.pavon@ingv.it

Citation:
Pavón-Carrasco, F. J., M. Gómez-Paccard, G. Hervé, M. L. Osete, and A. Chauvin (2014), Intensity of the geomagnetic field in Europe for the last 3 ka: Influence of data quality on geomagnetic field modeling, *Geochem. Geophys. Geosyst.*, *15*, doi:10.1002/2014GC005311.

Received 20 FEB 2014
Accepted 28 MAY 2014
Accepted article online 2 JUN 2014

Intensity of the geomagnetic field in Europe for the last 3 ka: Influence of data quality on geomagnetic field modeling

Francisco Javier Pavón-Carrasco[1], Miriam Gómez-Paccard[2], Gwenaël Hervé[3], María Luisa Osete[4,5], and Annick Chauvin[2]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Geomagnetismo, Aeronomia e Geofisica Ambientale, Rome, Italy, [2]Géosciences-Rennes, CNRS, UMR 6118, Université de Rennes 1, Campus de Beaulieu, Rennes, Cedex, France, [3]Department of Earth and Environmental Sciences, Ludwig-Maximilians Universität, Munich, Germany, [4]Departamento de Física de la Tierra, Universidad Complutense de Madrid, Madrid, Spain, [5]Instituto de Geociencias, IGEO CSIC, UCM, Madrid, Spain

Abstract One of the main challenges of paleomagnetic research is to obtain high-resolution geomagnetic field intensity reconstructions. For the last millennia, these reconstructions are mostly based on archeomagnetic data. However, the quality of the intensity data available in the databases is very variable, and the high scatter observed in the records clearly suggests that some of them might not be reliable. In this work we investigate how the geomagnetic field intensity reconstructions and, hence, our present knowledge of the geomagnetic field in the past, are affected by the quality of the data selected for modeling the Earth's magnetic field. For this purpose we rank the European archeointensity data in four quality categories following widely accepted paleomagnetic criteria based on the methodology used during the laboratory treatment of the samples and on the number of specimens retained to calculate the mean intensities. Four geomagnetic field regional models have been implemented by applying the revised spherical cap harmonic analysis to these four groups of input data. Geomagnetic field models strongly depend on the used data set. The model built using all the available data (without any preselection) appears to be the less accurate, indicating some internal inconsistencies of the data set. In addition, some features of this model are clearly dominated by the less reliable archeointensity data, suggesting that such features might not reflect real variations of the past geomagnetic field. On the contrary, the regional model built on selected high-quality intensity data shows a very consistent intensity pattern at the European scale, confirming that the main intensity changes observed in Europe in the recent history of the geomagnetic field occurred at the continental scale.

1. Introduction

The paleomagnetic studies provide information about the evolution of the ancient geomagnetic field intensity. This is a crucial topic to understand the geometry of the Earth's magnetic field and, consequently, to improve our knowledge of the geodynamo process. Moreover, robust reconstructions of geomagnetic dipole moment at decadal and centennial times scales are needed to address strongly debated questions as the possible link between geomagnetic field secular variation and Earth's climate [e.g., *Bard and Delaygue*, 2008; *Courtillot et al*., 2007; *Gallet et al*., 2005; *Genevey et al*., 2013; *Wollin et al*., 1978]. In addition, detailed dipole moment estimations are necessary for reconstructing the past solar activity based on radionuclide production rates [e.g., *Muscheler et al*., 2007; *Usoskin*, 2013]. Finally, variations in geomagnetic field intensity can potentially be used to provide chronological constraints of baked archeological material, volcanic rocks, and Quaternary sediments [e.g., *Gallet et al*., 2009; *Pavón-Carrasco et al*., 2011; *Roberts et al*., 2013].

Direct instrumental measurements of the direction of the ancient field are available for the last four centuries from shipboard and navigational records [e.g., *Jackson et al*., 2000; *Jonkers et al*., 2003]. Prior to the first direct measurements made by Gauss in 1832 AD, the estimation of past geomagnetic field intensity requires the paleomagnetic analysis of rocks and baked archeological material [*Thellier and Thellier*, 1959]. During the last decades, numerous spot records of the ancient geomagnetic field intensity have been obtained [*Genevey et al*., 2008; *Donadini et al*., 2009, and references therein]. As argued in different studies [*e.g., Chauvin et al*., 2000; *Bowles et al*., 2002; *Genevey et al*., 2008; *Gómez-Paccard et al*., 2008], the scatter of the

paleointensity data is very puzzling and clearly indicates that some of the data available in the databases might not be reliable. During the last years, in an effort to obtain a high-resolution description of past geomagnetic field intensity changes at regional scale, different authors applied different criteria to estimate the reliability of available archeointensity data [*Chauvin et al*., 2000; *De Marco et al*., 2008; *Genevey et al*., 2009; *Gómez-Paccard et al*., 2008; *Hervé et al*., 2013; *Tema et al*., 2012]. In some of these studies the authors even considered some of the data as unreliable intensity estimations and rejected them for interpretation of past geomagnetic field variations [*Gómez-Paccard et al*., 2008, 2012a; *Genevey et al*., 2009, 2013; *Hervé et al*., 2013]. This is, perhaps, an extreme point of view, but it underlines the need to develop robust paleomagnetic criteria for evaluating the reliability of the intensity estimations contained in the databases before using them for modeling the past variation of the geomagnetic field.

During the last years, extensive efforts led to the development of different regional and global geomagnetic field models [e.g., *Korte and Constable*, 2011; *Korte et al*., 2009, 2011; *Pavón-Carrasco et al*., 2009, 2010, 2014; *Licht et al*., 2013]. Obviously, their ability to recover past geomagnetic field intensity changes at decadal and millennial scales strongly depends not only on the abundance of paleomagnetic data in a particular region and time interval or on the type of data used for modeling (only archeomagnetic data or both archeo and sedimentary records), but also on the fidelity of the paleointensity data used for their construction. The need for a greater number of high-quality archeointensity data and for the establishment of selection criteria to assess the reliability of archeointensities is also highlighted by the inconsistencies observed between different global geomagnetic field models [*Jackson et al*., 2000; *Korte and Constable*, 2005; *Gubbins et al*., 2006; *Finlay*, 2008; *Korte et al*., 2011; *Suttie et al*., 2011]. Up to now, and certainly due to the low number of available intensity data, the geomagnetic field models have been constructed using all the available paleointensity data without any preselection of them. Generally, different filters based on intensity measurement and age uncertainties are applied to detect and reject outliers by comparison of the individual mean intensity values and/or their uncertainties with the model results [see *Donadini et al*., 2009]. However, these filters do not depend on the quality of the laboratory protocol, and, hence, geomagnetic field model predictions might be affected by some unreliable data.

Given the increasing interest of geomagnetic field reconstructions, we present an attempt to investigate how the rejection or acceptance of the less reliable intensity data affect geomagnetic field models, and, therefore, our present knowledge of past geomagnetic field intensity changes. For this purpose, we use the most reliable records of geomagnetic field strength (archeomagnetic data) from the best covered region (Europe, North Africa, and the Near East) and the best covered time interval (the last three millennia). We provide four different geomagnetic field intensity reconstructions obtained from different selections of the archeointensity data ranked in four quality categories by means of widely accepted paleomagnetic criteria.

It is worth noting that an additional source of errors is the age uncertainties. Dates are assigned to archeological materials through a variety of chronological methods (archeological information, radiocarbon dating, thermoluminescence, or even archeomagnetic dating). Here we revised the entire databases GEOMAGIA50v2 [*Donadini et al*., 2006; *Korhonen et al*., 2008] and ArcheoInt [*Genevey et al*., 2008] and included recent data carefully checked. We rely on the revision given by *Genevey et al*. [2008], who calibrated all ArcheoInt data, but we excluded the sites dated only by archeomagnetism.

To model the data, we use the revised spherical cap harmonic analysis technique in two dimensions (R-SCHA2D) [*Thébault*, 2008] in space and the classical penalized cubic b-splines in time. We present the different geomagnetic features depicted by the different models, followed by an attempt to decipher if such features are robust or are the result of some inconsistent data contained in the databases. Our overall aim is to underline the crucial role of obtaining and selecting high-quality paleointensities for improving our present understanding of geomagnetic field intensity changes at decadal and centennial time scales.

2. The European Archeointensity Database for the Last Three Millennia

2.1. Absolute Field Intensity Measurements

Since the pioneering work conducted by Emile Thellier [*Thellier*, 1938; *Thellier and Thellier*, 1959], numerous paleomagnetic studies performed on material carrying a thermoremanent magnetization (TRM) provided local snapshots of the ancient geomagnetic field. Paleointensity estimations are based on the assumption that the magnetization of a volcanic rock or baked sample is linear to the geomagnetic field strength

present at the moment of their last cooling from high temperatures [*Thellier*, 1938; *Néel*, 1955; *Thellier and Thellier*, 1959]. This linearity assumption appears to be reasonably well founded for ideal assemblages of single domain grains. However, many problems can occur when experiments are performed on natural assemblages [*Dunlop*, 2011]. The most common problems, which led to unsuccessful experiments, are thermally unstable ferromagnetic phases, the presence of coarse (multidomain) grains, and secondary remanent magnetizations. In order to detect these quite common problems and to check the ability of the specimens to reliable record the ancient geomagnetic field, several different approaches have been used (see *Genevey et al*. [2008] for a detailed review). Up to now, it is generally accepted that the most reliable methods for intensity determination are those based in the original or derived Thellier and Thellier method [*Thellier and Thellier*, 1959] including pTRM (partial thermoremanent magnetization) checks. Those methods are based on the comparison between natural remanent magnetization (NRM) lost and partial thermoremanent magnetization gained in a known laboratory field. The pTRM checks are crucial to detect possible alteration of the magnetic mineralogy of the samples during heating and, thus, are a very valuable way to detect failure of the intensity experiments.

Moreover, it is now accepted that the anisotropy of the TRM is a very important effect that need to be corrected at specimen level (as it can be highly variable from specimen to specimen) for obtaining accurate estimations of paleointensity [*Chauvin et al*., 2000; *Genevey et al*., 2008]. Indeed, many archeomagnetic materials can exhibit an anisotropy effect that, if not corrected, can result in significant errors upon paleointensity estimates. This effect can be particularly important for ceramic fragments, tiles, and for some other archeological baked artifacts (e.g., pipes) and reach, in some cases, very high correction factors of more than 70% of the determined field strength [*Chauvin et al*., 2000; *Genevey et al*., 2008; *Gómez-Paccard et al*., 2012a].

However, for baked clays from archeological kilns, ovens, hearths, bricks, and burnt soils, this effect is generally lower than 5% [*Gómez-Paccard et al*., 2006, 2008; *Kovacheva et al*., 2009b; *Hervé et al*., 2013]. Different approaches to correct intensity estimations for the TRM anisotropy have been proposed in the literature (see *Genevey et al.* [2008] for a review of the different protocols used). One of the most successful approaches is to calculate the TRM anisotropy tensor from the acquisition of a TRM in six different directions, in sample coordinates, X, −X, Y, −Y, Z, −Z; and later correct all the NRM and TRM measurements using the obtained tensor [*Veitch et al*., 1984; *Chauvin et al*., 2000]. In some studies, the approach followed to avoid the TRM anisotropy effect is to apply the TRM in the same direction than the NRM during the laboratory treatment [*Genevey et al*., 2008]. In some cases the correction of the TRM anisotropy effect is performed using the Anhysteretic Remanent Magnetization (ARM), the Isothermal Remanent Magnetization (IRM), or the tensors of Anisotropy of Magnetic Susceptibility (AMS). However, it has been demonstrated that these tensors are often not similar to the TRM anisotropy tensor [e.g., *Chauvin et al*., 2000]. The anisotropy effect can be underestimated if the ARM, IRM, and more especially AMS tensors are used to correct archeointensities instead of the TRM tensor [*Chauvin et al*., 2000; *Kovacheva et al*., 2009b].

Another important consideration in archeointensity studies is the cooling rate dependence of TRM acquisition. This effect is predicted by Néel's theory [*Néel*, 1955; *McClelland-Brown*, 1984] for single domain grain assemblages and has been experimentally observed for both volcanic and archeological material [e.g., *Chauvin et al*., 2000; *Leonhardt et al*., 2006]. This effect is due to the fact that the cooling time of the samples in the laboratory is typically lower (1–1.5 h) than the original cooling of the samples: from several hours to days (months) depending on the size of the archeological (volcanic) structures. When not corrected, it can result in an overestimation of the paleointensity. Experimental studies indicate, however, that this correction is usually lower than 10% [*Genevey et al*., 2008].

Taking into account the diversity of the analyzed materials and laboratory protocols applied, it is irrefutable that all the intensity data cannot be considered as equally reliable. The problem now is how to estimate (and ideally to quantify) the reliability of absolute intensities in a proper way using well-defined and objective paleomagnetic criteria.

In this context, the new archeointensity databases recently developed [*Donadini et al*., 2006; *Genevey et al*., 2008; *Korhonen et al*., 2008] constitute a very powerful tool to make a forward step through the integration of the diversity and complexity of paleointensity data in geomagnetic field studies via the selection and/or ranking of the archeointensity data. These databases include not only mean intensities but also crucial metadata for evaluating their quality.

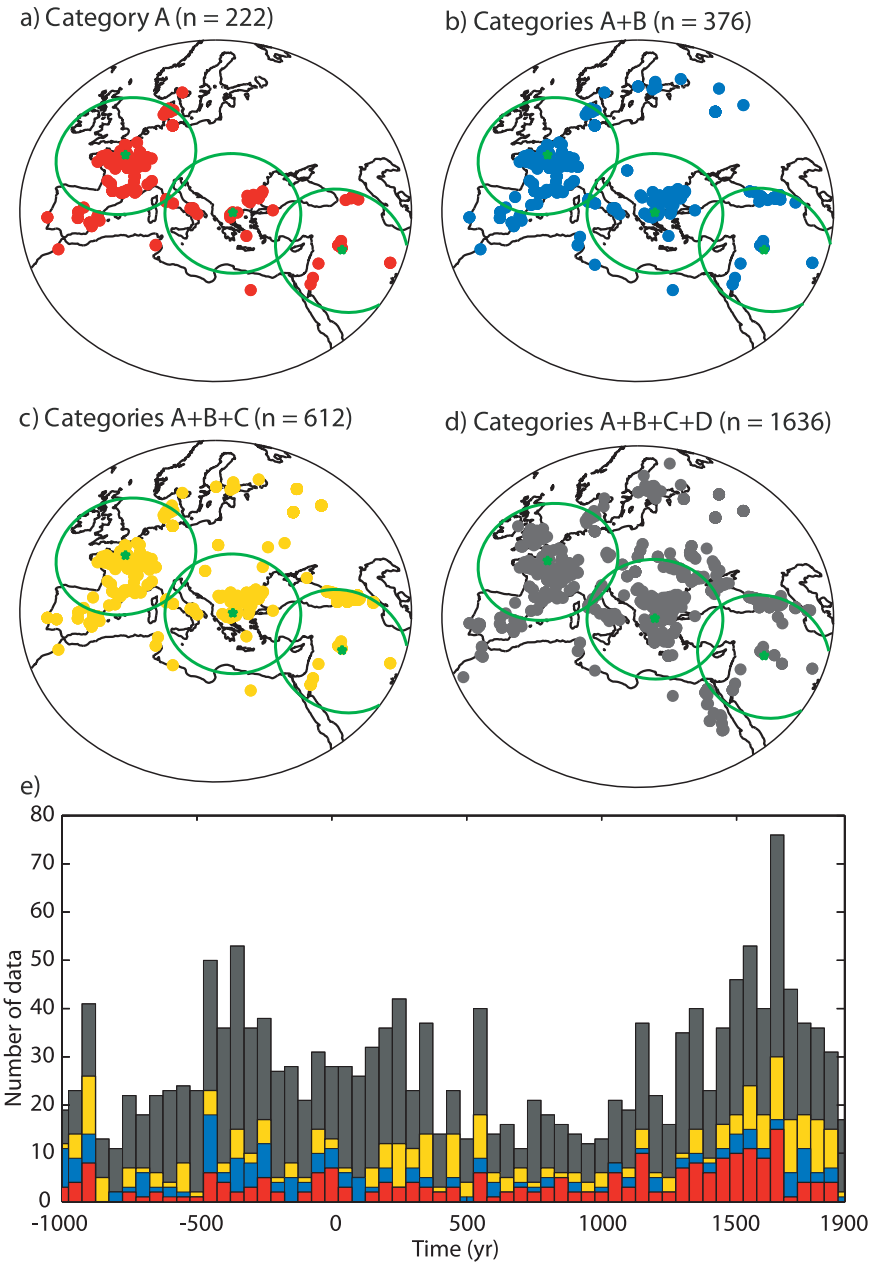

Figure 1. (a–d) Location of the archeointensity data available for Europe, the Near East and North Africa corresponding to the models A, B, C, and D, respectively. (*n*) indicates the number of archeointensity data corresponding to each model. The green circles show the three selected areas analyzed in this work and centered at the green stars. (e) Histogram showing the number of data available for each model at different time intervals (bins of 50 yr).

2.2. Ranking the Reliability of the European Archeointensity Data for the Last Three Millennia

In order to investigate the influence of the data quality on geomagnetic field reconstructions and, hence, on our present knowledge of geomagnetic field intensity changes, we used the archeointensity database from Europe and neighboring areas for the last three millennia (the period and region best covered in terms of archeomagnetic data). We selected archeomagnetic data located between 15°W and 50°E and 30°N and 60°N (Figure 1). This includes data from the ArcheoInt compilation [*Genevey et al.*, 2008] and some new archeointensity data recently published [*Ben-Yosef et al.*, 2009; *Catanzariti et al.*, 2012; *Donadini et al.*, 2010; *Ertepinar et al.*, 2012; *Gallet et al.*, 2009; *Genevey et al.*, 2013; *Gómez-Paccard et al.*, 2012a, 2012b, 2013; *Hervé et al.*, 2011, 2013; *Kovacheva et al.*, 2009a; *Nachasova and Burakov*, 2009; *Prevosti et al.*, 2013; *Schnepp et al.*, 2009; *Shaar et al.*, 2011; *Spatharas et al.*, 2011; *Tema et al.*, 2012, 2013a, 2013b; see special issue on Absolute

geomagnetic field intensity in Georgia during the past 6 millennia (*Latinmag Letter*, *3*, 1–4 pp., 2013)]. The overall used European and neighboring areas data set includes 1636 archeointensities (see Figures 1d and 1e).

Although the definition of criteria for classifying archeointensities in different quality categories is not a simple task, we believe that it is already possible to define objective paleomagnetic criteria generally accepted by the overall paleomagnetic community. As noted by *Tauxe* [2009], it is safe to say that the more stability checks performed (and passed), the greater confidence in the results, and the more specimens and samples measured and retained to calculate mean intensities, the more confidence we can have in the final result. This later point is crucial to derive the reliability of the different archeointensity data as the standard deviation of the mean is often about 10% [*Genevey et al*., 2008]. Therefore, two main criteria have been used to rank the European archeointensities: the intensity protocol used in the laboratory to infer the archeointensities at specimen level and the number of archeointensities used to estimate the mean. As the number of independent samples retained to calculate mean intensities is often missing in the databases due to the different definitions of a site in archeomagnetic studies [*Genevey et al*., 2008], we have decided to use the number of specimens retained to compute the mean in order to rank the archeointensity data selected.

Taking into account the previous discussion, we classify the data in four quality categories (A, B, C, and D) according to these two basic criteria:

Category A (the most reliable data). Category A includes mean archeointensities derived from at least five specimens and obtained from the laboratory protocols considered as the most reliable by the paleomagnetic community. This includes the original or derived Thellier and Thellier method including pTRM checks and anisotropy correction via determination of the TRM anisotropy tensor, whatever the material analyzed is. The archeointensities determined by the Triaxe method also belong to this category, as the laboratory TRM is parallel to the NRM [*Le Goff and Gallet*, 2004]. Slag-derived data for which nonlinear TRM acquisition have been monitored and corrected when necessary are included in category A if five or more specimens were used to derive the final means [*Shaar et al*., 2010, 2011]. Only 222 archeointensity data have been obtained using these high-standard protocols. Most of them are located in Western Europe and correspond to the last millennium (see Figures 1a and 1e).

Category B. Category B includes mean archeointensities derived from at least four specimens and obtained by the original or derived Thellier and Thellier method with pTRM checks without TRM anisotropy correction but obtained from few anisotropic objects (as kilns, hearths, burnt soils, and bricks). The data contained in this category are also highly reliable as the TRM anisotropy effect in such materials is generally lower than 5% [*Kovacheva et al*., 2009b]. Data corresponding to the intensity protocol described in category A but derived from four specimens are included in category B. This category includes 154 archeointensities (Figures 1b and 1e).

Category C. This category includes data obtained by the original or derived Thellier and Thellier method with pTRM checks but without TRM anisotropy corrections and obtained from generally highly anisotropic objects (pottery fragments, tiles, or small baked archaeological artifacts). Data obtained with protocols described in category A and B but derived from three specimens are included in category C. *Shaar et al*. [2010] indicate that the nonlinear TRM acquisition must be monitored and controlled during paleointensity experiments of slag materials. We therefore consider that slag-derived data without this control can be considered as less reliable than data included in A and B categories and we include them in category C [*Gram-Jensen et al*., 2000; *Ben-Yosef et al*., 2008, 2009]. This category includes 236 archeointensities (Figures 1c and 1e).

Category D (the less reliable data). Category D includes all the other intensity data obtained from other intensity protocols (e.g., Games method for adobe bricks [*Games*, 1977], microwave-based method, Shaw-based methods [*Shaw*, 1974], see *Genevey et al*. [2008] for a complete overview), or data derived from (only) one or two specimens. The paleomagnetic community generally accepts that these paleointensity methods are less well constrained than the Thellier-derived methods for which the physical theory of TRM acquisition is well established [*Néel*, 1955; *Thellier and Thellier*, 1959]. Data obtained from Thellier-derived methods not including pTRM checks or based in less than three specimens are also included in this category. Moreover, the dispersion of the intensity data obtained from these methods is clearly higher than for data corresponding to A, B, and C categories [*Chauvin et al*., 2000; *Gómez-Paccard et al*., 2008] (see also Figures 2–5 from this

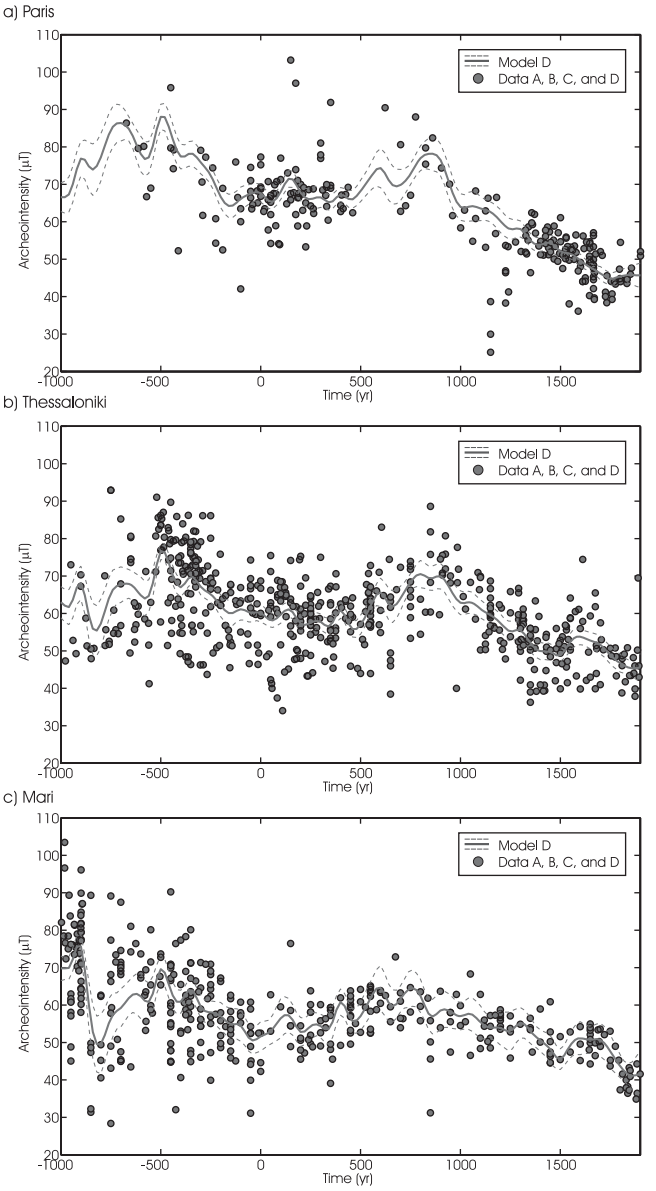

Figure 2. Intensity paleosecular variation curves (and error band at 95% of confidence) at (a) Paris, (b) Thessaloniki, and (c) Mari coordinates from the model D. Note that only archeointensity data from sites within 1000 km from the selected locations are plotted in the figure (see text for details).

study). Category D can therefore be considered as the less reliable category from the methodological point of view. It is worth noting that 1024 data (about 60% of the database) correspond to category D (see Figures 1d and 1e). This very high proportion enhances the need to investigate the impact of low-quality archeointensity estimations on geomagnetic field intensity reconstructions.

It should be noted that cooling rate corrections have not been considered in order to classify the data in different categories. This effect is generally lower than 10% [*Genevey et al.*, 2008] and it is very difficult to properly estimate the original real cooling time of the samples from an archeological point of view and, therefore, the final value of the cooling rate effect applied to the samples can be biased. Finally, we notice that in this study we do not include any criteria related to the age estimations of the sites. We certainly agree that the precision and reliability of the dates ascribed for each site are fundamental, but the evaluation of quality of the datings remains a very difficult task that is out of the scope of this work. Instead of a ranking criterion related to dating methodologies and precision, age uncertainties are considered during the modeling process.

The previously described categories (from A to D) constitute a powerful tool to rank the available data in different quality groups. Four geomagnetic field models have been calculated accordingly with these groups. The model A includes data from category A (222 data). The model B includes categories A and B (376 data). The model C includes categories A, B, and C (612 data), and the model D includes all available data (1636 data). In next sections we present the modeling approach used and the results obtained for the different models.

3. Modeling Approach

In order to model the archeointensity data at continental scale, we have applied the revised spherical cap harmonic analysis technique (R-SCHA2D) [*Thébault*, 2008] in space and the penalized cubic b-splines in time. The revised version of the classical SCHA [*Haines*, 1985] technique is a mathematically complete

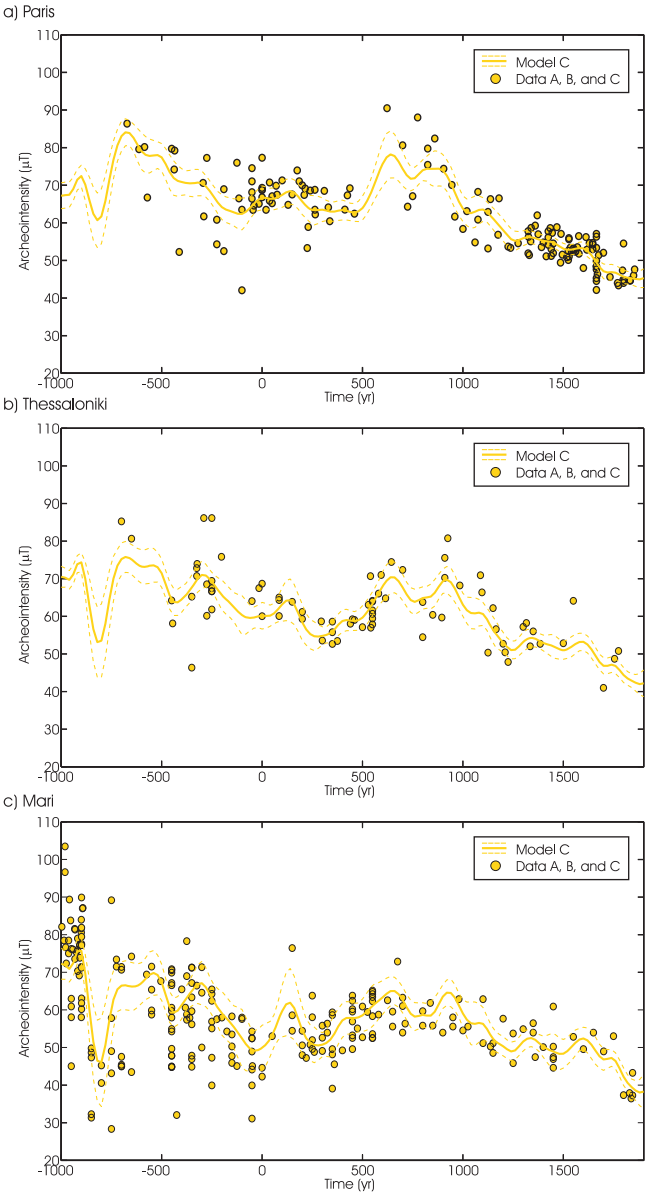

Figure 3. Intensity paleosecular variation curves (and error band at 95% of confidence) at (a) Paris, (b) Thessaloniki, and (c) Mari coordinates from the model C. Note that only archeointensity data from sites within 1000 km from the selected locations are plotted in the figure (see text for details).

solution that turns to be equivalent to global spherical harmonics when the series expansion is infinite [*Thébault et al.*, 2004, 2006; *Thébault*, 2008]. This method provides not only the temporal evolution of the paleofield, but also its spatial behavior which cannot be analyzed when using the classical paleosecular variation curves (PSVC).

The intensity of the geomagnetic field cannot be expressed as a linear combination of the coefficients of the potential geomagnetic field. To solve this nonlinear problem, we have used the Taylor series applied to the intensity data as follows:

$$F(t,\vec{r})=F(t,\vec{r})\Big|_{m=m_0}+\frac{\partial F(t,\vec{r})}{\partial m}\Big|_{m=m_0}\cdot(\vec{m}-\vec{m}_0) \quad (1)$$

where F is the intensity data for a location $\vec{r}$ and epoch t, $\vec{m}$ is the vector containing the Spherical Cap Harmonic (SCH) coefficients, and $\frac{\partial F(t,\vec{r})}{\partial m}$ is the Frechet derivative vector for the intensity component. The subindex 0 corresponds to an initial geomagnetic field model, denoted by the vector $\vec{m}_0$, and fixed as a constant axial dipole field.

To obtain the best estimation of the SCH coefficients, we have applied the regularized least square inversion to equation (1)

$$\delta\vec{m}=\left(\hat{A}'\cdot\hat{C}_e^{-1}\cdot\hat{A}+\alpha\cdot\hat{\Psi}+\tau\cdot\hat{\Phi}\right)^{-1}\hat{A}'\cdot\hat{C}_e^{-1}\cdot\vec{\gamma} \quad (2)$$

where $\delta\vec{m}=\vec{m}-\vec{m}_0$ is the vector of the SCH coefficients, $\hat{A}$ is the matrix of parameters which depends on the spatial and temporal functions in terms of SCH and cubic b-splines, and $\hat{A}'$ is the transpose of $\hat{A}$. The data error covariance matrix, $\hat{C}_e$, is the inverse matrix of weights and $\vec{\gamma}$ is the vector of input data. The $\hat{\Psi}$ and $\hat{\Phi}$ matrices are the spatial and temporal regularization matrices, respectively, with damping parameters α and τ. These two matrices and the corresponding damping parameters control the trade-off between the data misfit and the model roughness. Higher values of α and τ increase the smoothness of the geomagnetic field components in terms of spatial and temporal wavelengths, also in areas well covered by data. Smaller values provide a better fit to the data, but the model complexity increases. In the same way, the maximum spatial expansion of the R-SCHA2D potential field (see supporting information, section S1), the size of the spherical cap, and the temporal position of the cubic splines effect

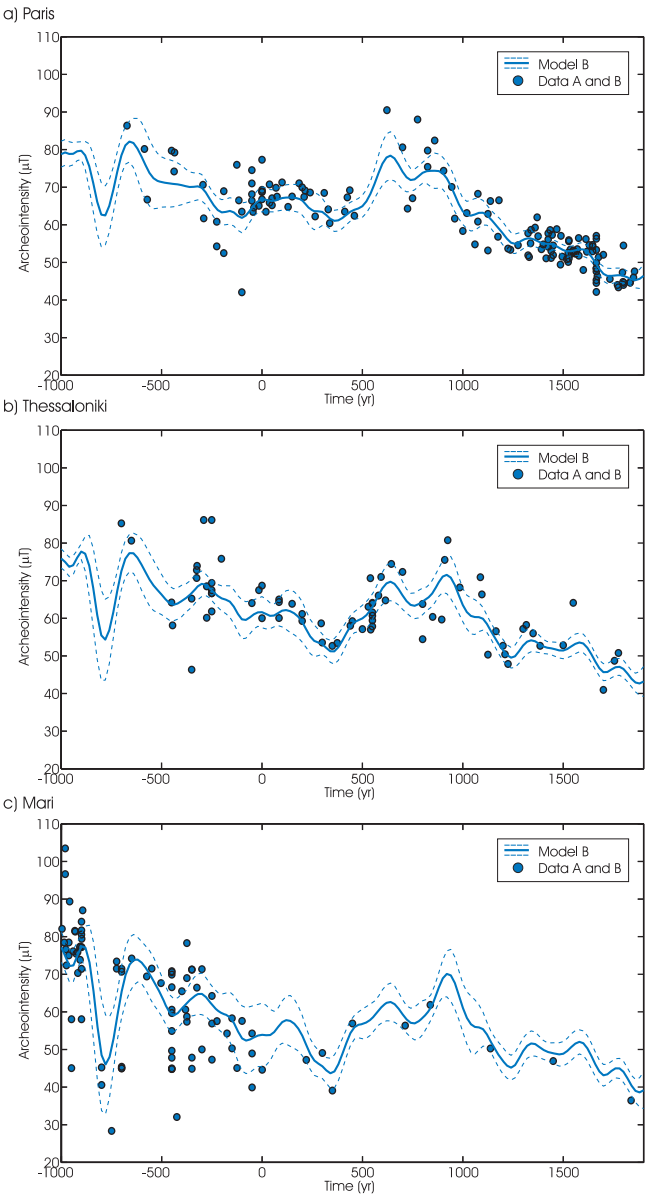

Figure 4. Intensity paleosecular variation curves (and error band at 95% of confidence) at (a) Paris, (b) Thessaloniki, and (c) Mari coordinates from the model B. Note that only archeointensity data from sites within 1000 km from the selected locations are plotted in the figure (see text for details).

to the spatial and temporal resolution of the model. For these reasons, different tests must be carried out to set the most appropriate values of all these parameters.

Supporting information, section S1, of this paper includes additional information about the R-SCHA2D approach and the different tests performed to define the most appropriate values of the above mentioned parameters. The maximum spatial expansions in the Haines potentials of R-SCHA2D (see equation (1) in the section S1 of the supporting information), K_{int} and K_{ext}, was set to 2, and the maximum spatial expansion for the Mehler potential, M, was 1. In a spherical cap of 26°, this corresponds to an approximately 8° in the global spherical harmonic analysis. The spatial parameter α was $5.0 \cdot 10^{-1}$ μT^{-2}. In time, the knot points were fixed every 50 years, from 1000 BC to 1900 AD and the temporal parameter τ was 1 yr^4 μT^{-2}. The archeointensity data were weighted in equation (2) according to the reciprocal of the variance of the intensity uncertainty. The error in modeling has been estimated by applying a bootstrap method taking into account both measurement and age uncertainties [e.g., *Korte et al.*, 2009]. The uncertainties of the model coefficients are given by the standard deviation of 5000 sets of individual model coefficients. A detailed description about the bootstrap method applied to the input data is also given in section S1 of the supporting information.

4. Results

In order to analyze the consistency and coherence of the four archeointensity data sets described in section 2 (see also Figure 1), we have applied the approach described above to generate four regional models of the intensity element of the geomagnetic field. The obtained models, named A, B, C, and D according to the data classification, have been compared with the input data at three representative locations: Paris (France, 48.9°N, 2.3°E) for Western Europe, Thessaloniki (Greece, 40.6°N, 23.0°E) for Eastern Europe, and Mari (modern Tell Hariri, Syria, 34.5°N 40.8°E) for the Near East. For each location and to illustrate the comparison, we have relocated the intensity data within a circular area of 1000 km of radius using the virtual

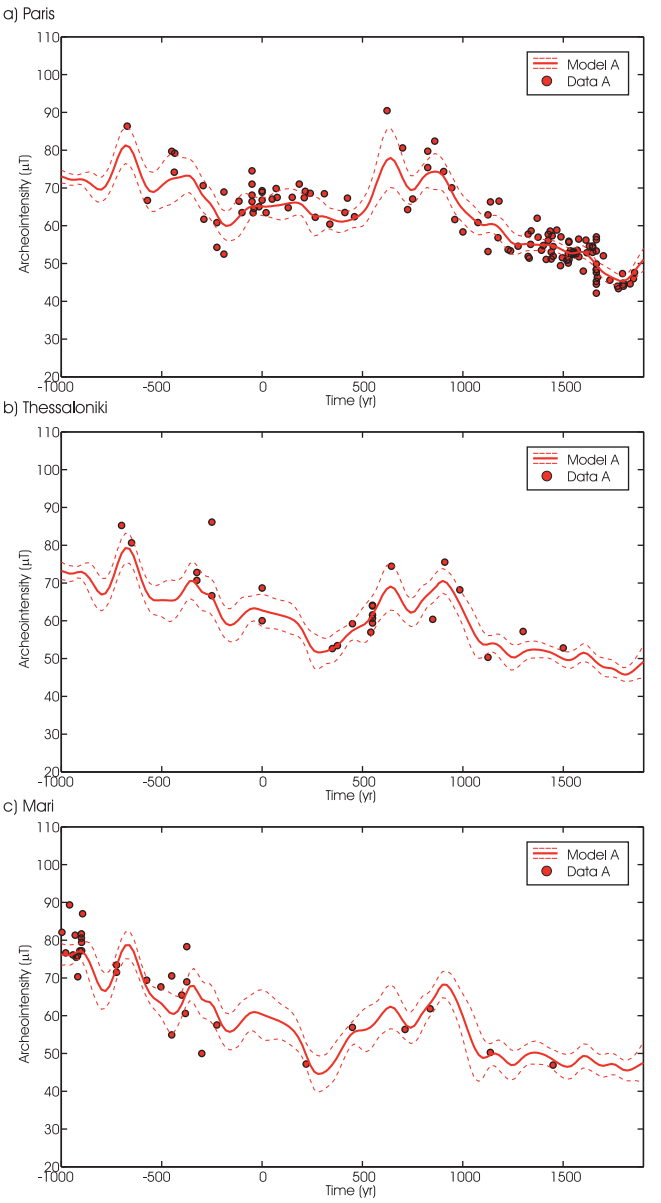

Figure 5. Intensity paleosecular variation curves (and error band at 95% of confidence) at (a) Paris, (b) Thessaloniki, and (c) Mari coordinates from the model A. Note that only archeointensity data from sites within 1000 km from the selected locations are plotted in the figure (see text for details).

axial dipole moment (VADM). The three locations and 1000 km radius regions are plotted in Figure 1 (green stars and circles, respectively). The total number of archeointensity data used to implement each model and the number of data located within the 1000 km radius regions are given in Table 1. To quantify the fit of different developed models to the input data, we have used the normalized root mean square error (n-rms) given by [see *Korte and Constable*, 2003]

$$n-rms=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-\hat{x}_i}{\varepsilon_i}\right)^2} \quad (3)$$

where N is the total number of input data, x_i and $\hat{x}_i$ are the input and output data, respectively, and ε_i is the intensity uncertainty (the standard error of the intensity data).

The first model, model D (Figure 2), was implemented considering all the available data (the 1636 data plotted in Figure 1d). The obtained n-rms has a value of 8.4. This value is higher than the expected n-rms for a well-constrained model for which the n-rms should be close to 1. The combination of two effects can explain the high n-rms obtained: (a) the high dispersion of the intensity data contained in category D, and (b) the underestimation of the uncertainties of the

Table 1. Number of Data and Statistical Parameters[a]

Model	n	n (Paris)	n (Thessaloniki)	n (Mari)	n-rms	Mean Intensity Uncertainty (μT)	Mean Residuals (μT)	Std Residuals (μT)	ε_T (μT)
D	1636	248	570	403	8.4	3.7	0.0	8.6	2.7
C	612	157	79	236	2.7	4.1	−0.1	7.5	1.9
B	376	148	73	91	2.5	3.9	−0.2	6.5	1.8
A	222	127	24	33	1.8	3.6	−0.3	4.6	1.3

[a]Name of model; number of data used to obtain each model; number of data within the 1000 km radius regions (green circles in Figure 1) centered at Paris, Thessaloniki, and Mari; normalized root mean square errors (n-rms); mean value of the intensity uncertainty for the data used for each model; mean residual data; standard deviation of the residual data; time-average accumulative error of the SCH coefficients.

intensity data as pointed out by *Donadini et al*. [2009]. The degree of dispersion can be quantified by the mean and the standard deviation of the residual data [*Korte and Constable*, 2003] because they take into account how the input data differ from the model predictions. The n-rms and mean/standard deviation of the residuals for the model D are given in Table 1. For this model the residuals present a mean value of 0.0 μT with a standard deviation of 8.6 μT. This standard deviation is higher than the average archeointensity uncertainty (3.7 μT, given also in Table 1) characterizing the high dispersion presented in the data from category D (Figure 2). A direct consequence of this high dispersion is the appearance of short nonreal temporal variations in the intensity synthetic PSVCs given by the model D. The dispersion is higher in Eastern Europe and in the Near East than in Western Europe, certainly due to the lower number of category D data in this last region (see the number of data for each region in Table 1).

The second model, model C (Figure 3), was constructed using data from categories A, B, and C. From the model D to the model C, the number of archeointensity data sharply decreases (62.6% from 1636 to 612). The high reduction of data is more significant in Eastern Europe (region of Thessaloniki) where only 79 data are available for constructing the model (see Table 1). This model presents a final n-rms of 2.7 with mean residuals of −0.1 μT and a standard deviation of 7.5 μT. The reduction of the n-rms indicates that the model C is better constrained than the model D. This is also corroborated by the lower standard deviation of residuals obtained (see Table 1). The better agreement between the selected data (as indicated by the standard deviation) suggests a more accurate model (in comparison with the model D) and, therefore, better constrained synthetic PSVCs (Figure 3). However, the standard deviation is still higher than the average archeointensity uncertainty (Table 1, column 7) and the n-rms remains high.

The third model, model B, was developed using a total of 376 input data from categories A and B. This represents ∼23% of the initial number of data. The model does not present important differences with the model C in Western and Eastern Europe (Figures 4a and 4b). However, the Near East suffers the biggest reduction of data, from 236 in the model C to 91 in this case (see Figures 3c and 4c). The n-rms is 2.5 and the residuals have values of −0.2 μT and 6.5 μT of mean and standard deviation, respectively.

The last model, model A, was developed using only data from category A (222 data, ∼13.6% of the total data set). The statistical parameters present an n-rms of 1.8 and a mean residual of −0.3 μT, with a standard deviation of 4.6 μT. From the model B to the model A, the number of data from Eastern Europe and the Near East decreases dramatically (Figure 5 and Table 1). However, for Western Europe (Figure 5a) the difference is only 21 data. For the model A, the n-rms tends to 1 and the standard deviation is closer to the average archeointensity uncertainty (3.6 μT).

5. Discussion

5.1. Robustness of the Models

To identify the most accurate model, the study of the robustness of the different models given by the n-rms and the residuals has been completed by the analysis of the SCH coefficient errors. As the basis functions involved in the R-SCHA2D technique are not orthogonal [see *Thébault*, 2008], the error of the SCH coefficients cannot be estimated from the diagonal of the covariance matrix $\hat{C}_e$. However, the bootstrap method provides the standard deviations of the SCH coefficients that quantify the quality of the model fitting. Figure 6a shows the accumulative error (sum of standard deviations) for all the coefficients as a function of time, and Figure 6b shows the time-average error for each SCH coefficient. The time-average values of the accumulative error clearly decrease from the model D to the model A, with values of 2.7, 1.9, 1.8, and 1.3 μT, respectively (see Table 1). This points out the higher internal consistency of the database when only data from category A are considered. In contrast, the higher value obtained when the whole database is used suggests internal inconsistencies of the overall European paleointensity database.

The accumulative error of the SCH coefficients along with the n-rms and the standard deviation of the residuals indicate a better agreement between the data from category A and, consequently, the model fitting (the model A). On the contrary, the same statistical parameters indicate that model D, developed with all the available paleointensity data, presents the worst fitting (see also section 4). Therefore, our results suggest that mean paleointensities computed with one or two specimens per site or obtained with Thellier derived without no pTRM checks or with non Thellier-derived protocols (data from category D) should not be used for modeling the past geomagnetic field intensity variations.

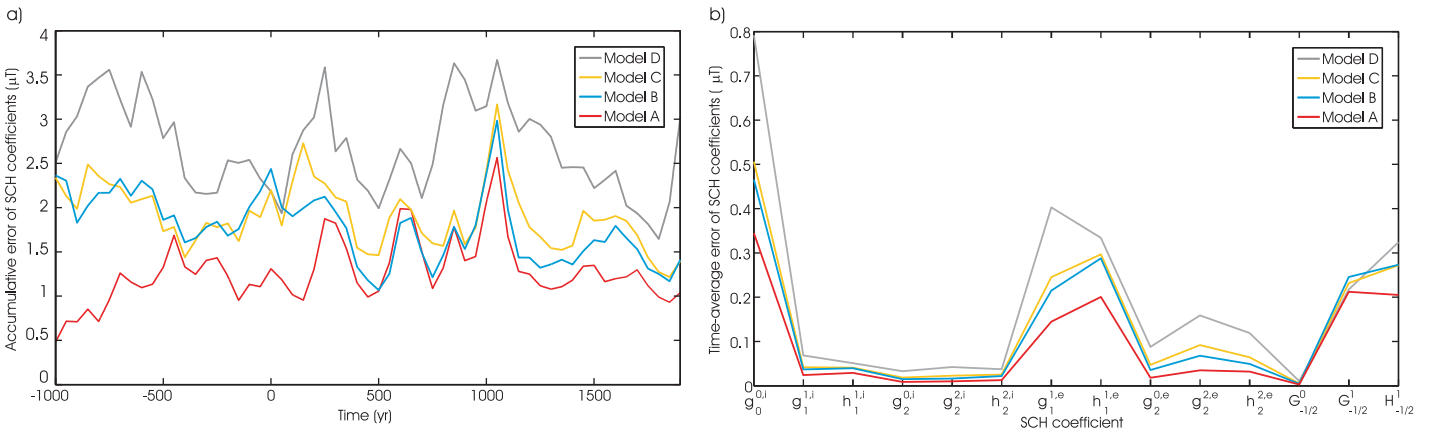

Figure 6. (a) Accumulative error of the SCH coefficients for each model and (b) time-average error for each SCH coefficient.

5.2. Discrepancies Between Model Predictions

According to the results given in the previous section, the model A seems to show the best fitting. However, this model is based on a small data set that is hardly dominated by the data from the western area of the spherical cap (∼75% of the A data are located in western longitudes of 20°E (Figure 1a)). The model B presents lower levels of confidence according to the statistical parameters (see Table 1), but it is developed using a data set that shows a better spatial and temporal distribution into the spherical cap (Figure 1b). To investigate the discrepancies produced by the use of different data sets, we have compared the synthetic curves generated by the four models at three different locations: Paris, Thessaloniki, and Mari. The results are plotted in Figure 7 along with the difference between the model B, C, and D predictions and the synthetic PSVC of the model A. As the differences between models are higher for the first millennium BC, we discuss separately the two millennia AD and the first millennium BC.

5.2.1. The Last Two Millennia

In Western Europe, the models A, B, and C give similar PSVCs (Figure 7a). This similarity can be explained by the high percentage of data from category A in comparison with data from categories B and C. These three models predict two clear intensity maxima at 600 AD and 850 AD, after a wide minimum around 400–500 AD, and since 850 AD a decreasing trend of the intensity (with plateau values or small maxima at 1100, 1350, and 1600 AD) up to 1900 AD. The model D shows a slightly higher temporal variability, with a higher number of relative maxima and minima, and presents differences with the model A (and B) around ±5 to 10 μT (that is 5–10% of the present day intensity of the field at Paris).

In Eastern Europe (Figure 7b), the predictions given by the models A, B, and C are consistent in the general trend, although the time locations of the maxima and minima differ slightly as well as their values. The two characteristic maxima around 650 AD and 900 AD, similar to those observed in Western Europe, are only well recorded after rejecting the data of category D. In this region a pronounced minimum is observed at 300 AD in models A and C which is slightly shifted in the model B. Around 1300 and 1600 AD the models B and C suggest maxima of higher amplitude than expected from the model A. Predictions of model D present differences with the model A (and B) reaching values of ±10 μT.

For the Near East (Figure 7c) there is an important decay of input data from model C to model A (see Table 1, column 5). Models A, B, and C also show the two maxima around 650 and 900 AD and a sharp minimum around 400 AD (less pronounced in the model C). A maxima around 1350 AD and 1600 AD of high amplitude is suggested by the models B and C (attenuated in the model A), and contradictory predictions are obtained for the first years AD between the model A and the models B/C. The model D presents a higher variability in comparison with the Western and Eastern Europe results.

5.2.2. The First Millennium BC

All the models predict at Paris, Thessaloniki, and Mari a common minimum of intensity around 800 BC (more accentuated in the models B and C; see Figure 7). This feature is controlled by data coming from the

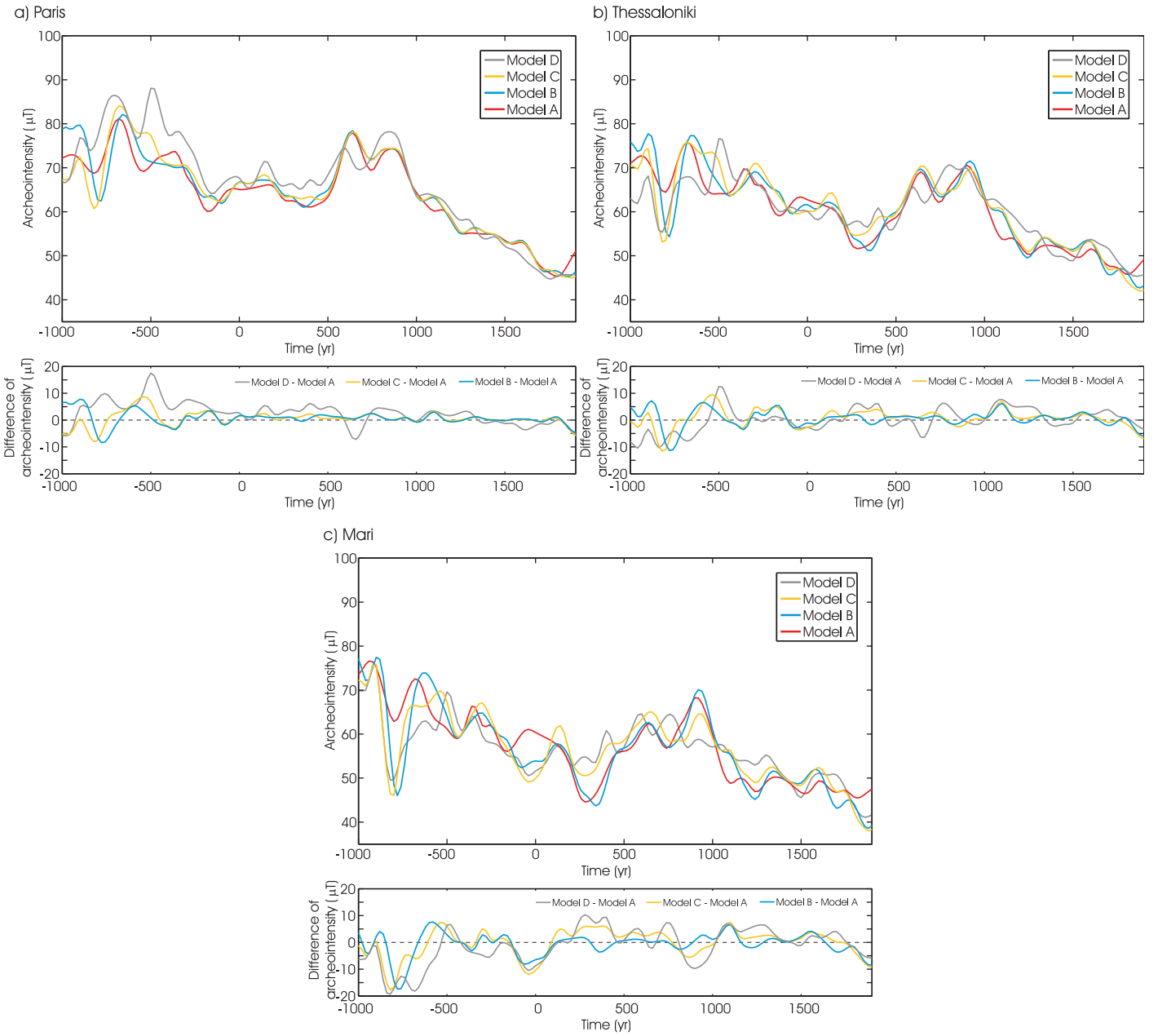

Figure 7. (top) Intensity paleosecular variation curves of the different models and (bottom) the difference between the B/C/D models curves and the curve given by the model A for the three locations: (a) Paris, (b) Thessaloniki, and (c) Mari.

Near East region between 1000 and 750 BC (see Figures 3c and 4c). New high-quality data are clearly needed to confirm this minimum which is essentially constrained by a Georgian data set [*Burlatskaya and Chelidze*, 1987]. It is important to notice that this feature is associated to very high secular variation rates (∼20 μT/century).

The discrepancies between predictions given by the models C/D and those given by the models A/B increase significantly and reach values up to ±15–20 μT at the three selected locations. This is certainly due to the small amount of data included in category A (see Figures 1 and 5). In Western Europe, an intensity maximum around 700 BC is observed in the four synthetic PSVCs, followed by a decaying trend (with a probably second-order maximum around 400 BC and a minimum around 200 BC) observed in models A, B, and C. The model D is the only one that predicts intensities higher than 85 μT at 750 and 500 BC. Important

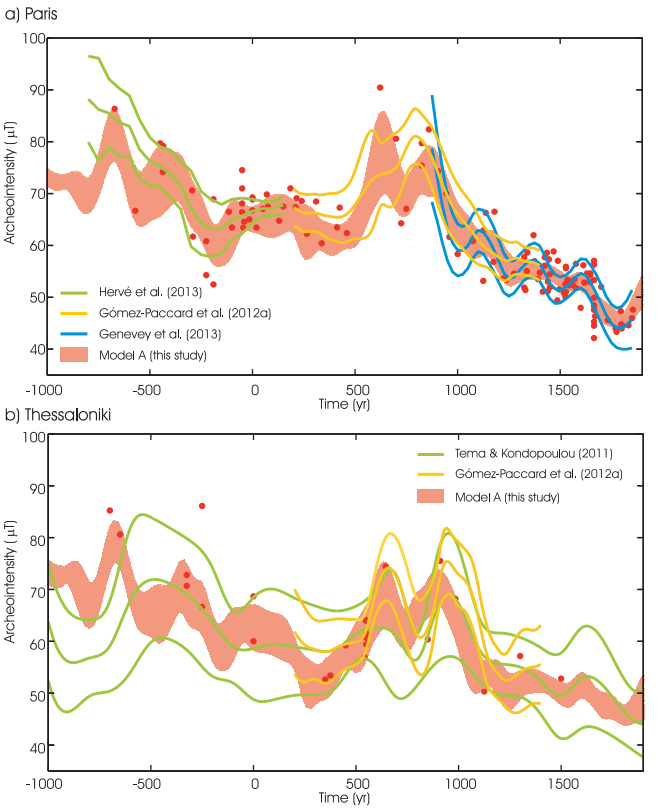

Figure 8. Comparison between the intensity curve (given by the error band at 95% of confidence) of the model A and already published regional intensity PSVCs in (a) Paris and (b) Thessaloniki. The red dots correspond to the intensity data of category A from sites within 1000 km from the selected locations.

differences are also observed in Eastern Europe between the different models predictions (Figure 7b). The maximum around 500 BC suggested by the model D disappears in the other PSVCs that show a clear maximum around 650 BC which agrees pretty well with the maximum at 700 BC observed in Western Europe (Figure 7a). The models A, B, and C suggest also a second maximum around 250 BC, followed by a relative minimum around 100 BC.

The complexity, in terms of temporal variations of the different intensity PSVCs for the Near East, decreases from the model D to the model A. The PSVC of the model A presents the smoothest behavior showing maxima around 900 and 700 BC, a second-order maximum around 300 BC, and a minimum around 150 BC.

For all regions and for the whole time interval, the model D predictions present a higher temporal variability than the models A and B. This seems to be an artifact related to the high dispersion and inconsistencies on the global database. Finally, we observed that when the whole data are used (model D), different patterns of the intensity geomagnetic variations are observed in Western Europe, Eastern Europe, and the Near East. However, the regional model obtained from a selection of high-quality intensity data (models A and B) predicts a more consistent pattern at the continental scale. For the last century, the IGRF11 model [*Finlay et al*., 2010] also shows a consistent pattern of intensity variations at continental scale modulated by the westward drift (see Figure S3). Consequently, anomalous secular variations of the geomagnetic field intensity can be obtained by using a paleointensity database without any quality filter.

5.3. Comparison With Other European Intensity PSVCs

We have compared the predictions of the model A with local PSVCs recently published. Figure 8a shows the PSVC at Paris generated by the model A along with other local PSVCs calculated for different time intervals: (a) from 800 BC to 200 AD [*Hervé et al*., 2013]; (b) from 200 AD to 1400 AD [*Gómez-Paccard et al*., 2012a]; and from 900 AD to 1850 AD [*Genevey et al*., 2013]. All these PSVCs were calculated from selected high-quality intensity data. It is interesting to see how all the curves present, in general, a good agreement with the curves of the model A, but some differences are also observed. These differences can be explained by the different approaches followed to obtain the PSVCs. The curve provided by *Hervé et al*. [2013] used sliding overlapping windows of 160 years shifted every 50 year. For this reason, the curve is more smoothed than the model A prediction (Figure 8a). The curve of *Gómez-Paccard et al*. [2012a] was calculated by the Bayesian approach [*Lanos*, 2004]. As noted by *Gómez-Paccard et al*. [2012a] the Bayesian curve is clearly very smoothed, probably due to the low number of data available between 500 and 1000 AD. Finally, the *Genevey et al*. [2013] curve for the last 1000 year (that gathers 50% of data in Western Europe) was developed applying an iteratively reweighted least squares algorithm combined with a bootstrap approach [*Thébault and Gallet*, 2010]. For its construction, the authors used temporal knot points every 70 year. This

curve fits pretty well with the model A predictions showing clearly the three proposed relative maxima for the last millennium at 1100 AD, 1350 AD, and around 1600 AD.

For Eastern Europe (Figure 8b), the PSVC at Thessaloniki of the model A has been compared with the curve of *Gómez-Paccard et al*. [2012a] covering the temporal interval from 200 AD to 1400 AD, and the curve for the Balkan region of *Tema and Kondopoulou* [2011] which cover the last eight millennia. The curve proposed by *Gómez-Paccard et al*. [2012a] was obtained from 42 selected data (Thellier and Thellier-derived data including pTRM checks) from Bulgaria, Greece, and Italy and by Bayesian modeling. The Bayesian and the model A curves present a well agreement and show the occurrence of two intensity maxima around 650 and 950 AD. However, the Balkan curve [*Tema and Kondopoulou*, 2011] presents smoother intensity changes and higher confidence limits. This can be explained by the approach used for constructing this curve. The authors used a 200 year overlapping sliding windows method shifted every 100 year applied to all intensity data located in a circular area of 700 km around Thessaloniki without any preselection of them [*Tema and Kondopoulou*, 2011]. Instead of a selection criterion based on the quality of the data, the authors applied the typical filters used commonly in geomagnetic field modeling, i.e., all the data with standard deviation σ_F (intensity uncertainty) 3 times the mean σ_F [see *Donadini et al*., 2009], or data with age uncertainties greater than 300 year were rejected. As a consequence of the kind of data used, the Balkan curve is more similar to the predictions of the model D (see Figure 2b) than that of the model A.

6. Conclusions

In this work, we rank the European archeointensity data in four quality categories based on the paleointensity technic used and on the number of samples retained to compute the mean values per site. We demonstrate that the regional geomagnetic field model obtained from all the available data (without any selection) is the less accurate. Moreover it shows significant discrepancies with the predictions given by the model performed with high-quality data. These differences highlight the necessity of establishing a preselection of high-quality data in order to obtain a robust description of the geomagnetic field in the past.

On the other hand, new high-quality archeointensity data are crucial to obtain a detailed reconstruction of the evolution of the geomagnetic field strength during the last millennia. Data are especially welcome for the first millennium BC to investigate the suspected intensity minimum at 800 BC that is only supported by data coming from Caucasus [*Burlatskaya and Chelidze*, 1987]. If this intensity low is confirmed, it would imply a very high secular variation rate ($\sim$20 μT/century). In the same way, intensity maximum around 700 BC should be also confirmed, since it is based on few data coming from the Near Eastern region only. Finally, we would like also point out the need of high-quality data from the so-called European "Dark-ages" (between 500 AD and 1000 AD) to better constrain the double oscillation of the intensity field observed at the European scale [*Gómez-Paccard et al*., 2012a].

Acknowledgments

We would like to express our thanks to the paleomagnetic community for the effort of providing the archeo/paleomagnetic data and the construction and updating of useful databases. We are very grateful for the review and helpful comments of Ron Shaar and a second anonymous referee. This work has been carried out within the Spanish Research Project CGL2011–24790 financed by the Spanish Ministry of Economy and Competitiveness. Financial support to this research was also given by the post doc positions ME-Fulbright CT-2010-0663 and "Assegno di Ricerca" INGV - Roma2 (FJPC) and Marie Curie Intra European Fellowship IEF-2012 (MGP). E. Thébault is acknowledged for providing the software of the Mehler functions of the R-SCHA2D.

References

Bard, E., and G. Delaygue (2008), Comment on "Are there connections between the Earth's magnetic field and climate?" by V. Courtillot et al. (Earth and Planetary Science Letters, 253, 328, 2007), *Earth Planet. Sci. Lett.*, *265*, 302–307.

Ben-Yosef, E., H. Ron, L. Tauxe, A. Agnon, A. Genevey, T. E. Levy, and U. Avner (2008), Application of copper slag in archeointensity research, *J. Geophys. Res.*, *113*, B08101, doi:10.1029/2007JB005235.

Ben-Yosef, E., L. Tauxe, T. E. Levy, R. Shaar, H. Ron, and M. Najjar (2009), Geomagnetic intensity spike recorded in high resolution slag deposit in southern Jordan, *Earth Planet. Sci. Lett.*, *287*, 529–539.

Bowles, J., J. Gee, J. Hildebrand, and L. Tauxe (2002), Archaeomagnetic intensity results from California and Ecuador: Evaluation of regional data, *Earth Planet. Sci. Lett.*, *203*, 967–981.

Burlatskaya S.P., and Z. A. Chelidze (1987), Geomagnetic field variations in Georgia from the third millennium B.C. to the first millennium A.D., Izv. *Earth Phys., Engl. Transl.*, *23*, 774–778.

Catanzariti, G., M. Gómez-Paccard, G. McIntosh, F. J. Pavón-Carrasco, A. Chauvin, and M. L. Osete (2012), New archaeomagnetic data recovered from the study of Roman and Visigothic remains from central Spain (3rd-7th centuries), *Geophys. J. Int.*, *188*, 979–993.

Chauvin, A., Y. Garcia, P. Lanos, and F. Laubenheimer (2000), Paleointensity of the geomagnetic field recovered on archaeomagnetic sites from France, *Phys. Earth Planet. Inter.*, *120*, 111–136.

Courtillot, V., Y. Gallet, J.-L. Le Mouel, F. Fluteau, and A. Genevey (2007), Are there connections between the Earth's magnetic field and climate?, *Earth Planet. Sci. Lett.*, *253*, 328–339.

De Marco, E., V. Spatharas, M. Gómez-Paccard, A. Chauvin, and D. Kondopoulou (2008), New archaeointensity results from archaeological sites and variation of the geomagnetic field intensity for the last 7 millennia in Greece, *Phys. Chem. Earth*, *33*, 578–595.

Donadini, F., K. Korhonen, P. Riisager, and L. Pesonen (2006), Database for Holocene geomagnetic intensity information, *Eos Trans. AGU*, *87*(14), 137.

Donadini, F., M. Korte, and C. G. Constable (2009), Geomagnetic field for 0–3 ka: 1. New data sets for global modeling, *Geochem. Geophys. Geosyst.*, *10*, Q06007, doi:10.1029/2008GC002295.

Donadini, F., M. Kovacheva, and M. Kostadinova (2010), Archaeomagnetic study of ancient Roman lime kilns (1c. AD) and one pottery kiln (1c. BC–1c. AD) at Krivina, Bulgaria, as a contribution to archeomagnetic dating, *Archeol. Bulgarica*, *XIV*(2), 213–225.
Dunlop, D.J. (2011), Physical basis of the Thellier-Thellier and related paleointensity methods, *Phys. Earth Planet. Inter.*, *187*(3–4), 118–138.
Ertepinar, P., C. G. Langereis, A. J. Biggin, M. Frangipane, T. Matney, T. Ökse, and A. Engin (2012), Archaeomagnetic study of five mounds from Upper Mesopotamia between 2500 and 700 BCE: Further evidence for an extremely strong geomagnetic field ca. 3000 years ago, *Earth Planet. Sci. Lett.*, *357–358*, 84–98.
Finlay, C. C. (2008), Historical variation of the geomagnetic axial dipole, *Phys. Earth Planet. Inter.*, *170*, 1–14.
Finlay, C. C., et al. (2010), International Geomagnetic Reference Field: The eleventh generation, *Geophys. J. Int.*, *183*, 1216–1230.
Gallet, Y., A. Genevey, and F. Fluteau (2005), Does Earth's magnetic field secular variation control centennial climate change?, *Earth Planet. Sci. Lett.*, *236*, 339–347.
Gallet, Y., A. Genevey, M. Le Goff, N. Warmé, J. Gran-Aymerich, and A. Lefèvre (2009), On the use of archeology in geomagnetism, and vice-versa: Recent developments in archeomagnetism, *C. R. Phys.*, *10*, 630–648.
Games, K. P. (1977), The magnitude of the paleomagnetic field: A new non-thermal, non-detrital method using sun-dried bricks, *Geophys. J. R. Astron. Soc.*, *48*, 315.
Genevey, A., Y. Gallet, C. G. Constable, M. Korte, and G. Hulot (2008), ArcheoInt: An upgraded compilation of geomagnetic field intensity data for the past ten millennia and its application to the recovery of the past dipole moment, *Geochem. Geophys. Geosyst.*, *9*, Q04038, doi:10.1029/2007GC001881.
Genevey, A., Y. Gallet, J. Rosen, and M. Le Goff (2009), Evidence for rapid geomagnetic field intensity variations in Western Europe over the past 800 years from new French archeointensity data, *Earth Planet. Sci. Lett.*, *284*, 132–143.
Genevey, A., Y. Gallet, E. Thébault, S. Jesset, and M. Le Goff (2013), Geomagnetic field intensity variations in Western Europe over the past 1100 years, *Geochem. Geophys. Geosyst.*, *14*, 2858–2872.
Gómez-Paccard, M., A. Chauvin, P. Lanos, J. Thiriot, and P. Jiménez-Castillo (2006), Archeomagnetic study of seven contemporaneous kilns from Murcia (Spain), *Phys. Earth Planet. Inter.*, *157*, 16–32.
Gómez-Paccard, M., A. Chauvin, P. Lanos, and J. Thiriot (2008), New archeointensity data from Spain and the geomagnetic dipole moment in western Europe over the past 2000 years, *J. Geophys. Res.*, *113*, B09103, doi:10.1029/2008JB005582.
Gómez-Paccard, M., et al. (2012a), Improving our knowledge of rapid geomagnetic field intensity changes observed in Europe between 200 and 1400 AD, *Earth Planet. Sci. Lett.*, *355–356*, 131–143.
Gómez-Paccard, M., G. McIntosh, A. Chauvin, E. Beamud, F. J. Pavón-Carrasco, and J. Thiriot (2012b), Archaeomagnetic and rock magnetic study of six kilns from North Africa (Tunisia and Morocco), *Geophys. J. Int.*, *189*(1), 169–186.
Gómez-Paccard, M., E. Beamud, G. McIntosh, and J. C. Larrasoaña (2013), New archaeomagnetic data recovered from the study of three Roman kilns from North-East Spain: A contribution to the Iberian palaeosecular variation curve, *Archaeometry*, *55*(1), 159–177.
Gram-Jensen, M., N. Abrahamsen, and A. Chauvin (2000), Archeomagnetic intensity in Denmark, *Phys. Chem. Earth*, *25*, 525–531.
Gubbins, D., A. L. Jones, and C. C. Finlay (2006), Fall in Earth's magnetic field is erratic, *Science*, *312*(5775), 900–902.
Haines, G. V. (1985), Spherical cap harmonic analysis, *J. Geophys. Res.*, *90*(B3), 2583–2591.
Hervé, G., E. Schnepp, A. Chauvin, P. Lanos, and N. Nowaczyk (2011), Archaeomagnetic results on three Iron Age salt-kilns from Moyenvic (France), *Geophys. J. Int.*, *185*, 144–156.
Hervé, G., A. Chauvin, and P. Lanos (2013), Geomagnetic field variations in Western Europe from 1500 BC to 200 AD. Part II: New intensity secular variation curve, *Phys. Earth Planet. Inter.*, *218*, 51–65.
Jackson, A., A. R. T. Jonkers, and M. R. Walker (2000), Four centuries of geomagnetic secular variation from historical records, *Philos. Trans. R. Soc. London A*, *358*, 957–990.
Jonkers, A. R. T., A. Jackson, and A. Murray (2003), Four centuries of geomagnetic data from historical records, *Rev. Geophys.*, *41*(2), 1006, doi:10.1029/2002RG000115.
Korhonen, K., F. Donadini, P. Riisager, and L. Pesonen (2008), GEOMAGIA50: An archeointensity database with PHP and MySQL, *Geochem. Geophys. Geosyst.*, *9*, Q04029, doi:10.1029/2007GC001893.
Korte, M., and C. G. Constable (2003), Continuous global geomagnetic field models for the past 3000 years, *Phys. Earth Planet. Inter.*, *140*(1), 73–89.
Korte, M., and C. G. Constable (2005), Continuous geomagnetic field models for the past 7 millenia: 2. CALS7K, *Geochem. Geophys. Geosyst.*, *6*, Q02H16, doi:10.1029/2004GC000801.
Korte, M., F. Donadini, and C. G. Constable (2009), Geomagnetic field for 0–3 ka: 2. A new series of time-varying global models, *Geochem. Geophys. Geosyst.*, *10*, Q06008, doi:10.1029/2008GC002297.
Korte, M., and C. G. Constable (2011), Improving geomagnetic field reconstructions for 0–3 ka, *Phys. Earth Planet. Inter.*, *188*, 247–259.
Korte, M., C. G. Constable, F. Donadini, and R. Holmes (2011), Reconstructing the Holocene geomagnetic field, *Earth Planet. Sci. Lett.*, *312*, 497–505.
Kovacheva, M., Y. Boyadziev, M. Kostadinova-Avramova, N. Jordanova, and F. Donadini (2009a), Updated archeomagnetic data set of the past 8 millennia from the Sofia laboratory, Bulgaria, *Geochem. Geophys. Geosyst.*, *10*, Q05002, doi:10.1029/2008GC002347.
Kovacheva, M., Chauvin, A., Jordanova, N., Lanos, P. and Karloukovski, V. (2009b), Remanence anisotropy effect on the palaeointensity results obtained from various archaeological materials, excluding pottery, *Earth Planets Space*, *61*, 711–732.
Lanos, Ph. (2004), Bayesian inference of calibration curves: Application to archaeomagnetism, in *Tools for Constructing Chronologies: Crossing Disciplinary Boundaries*, edited by C. Buck and A. Millard, vol. 177, pp. 43–82, Springer, London, U. K.
Le Goff, M., and Y. Gallet (2004), A new three-axis vibrating sample magnetometer for continuous high-temperature magnetization measurements: Applications to paleo- and archeointensity determinations, *Earth Planet. Sci. Lett.*, *229*, 31–43.
Leonhardt, R., J. Matzka, A. R. L. Nichols, and D. B. Dingwell (2006), Cooling rate correction for paleointensity determination for volcanic glasses by relaxation geospeedometry, *Earth Planet. Sci. Lett.*, *243*, 282–292.
Licht, A., G. Hulot, Y. Gallet, and E. Thébault (2013), Ensembles of low degree archeomagnetic field models for the past three millennia, *Phys. Earth Planet. Inter.*, *224*, 38–67.
McClelland-Brown, E. (1984), Experiments on TRM intensity dependence on cooling rate, *J. Geophys. Res.*, *11*, 205–208.
Muscheler, R., F. Joos, J. Beer, S. A. Muller, M. VonMoos, and I. Snowball (2007), Solar activity during the last 1000 yr inferred from radionuclide records, *Q. Sci. Rev.*, *26*, 82–97.
Nachasova, I. E., and K. S. Burakov (2009), Variation of the intensity of the Earth s magnetic field in Portugal in the 1st millenium BC, *Phys. Solid Earth*, *45*, 54–62.
Néel, L. (1955), Some theoretical aspects of rock-magnetism, *Adv. Phys.*, *4*, 191–243.

Pavón-Carrasco, F. J., M. L. Osete, J. M. Torta, and L. R. Gaya-Pique (2009), A regional archeomagnetic model for Europe for the last 3000 years, SCHA.DIF.3K: Applications to archeomagnetic dating, *Geochem. Geophys. Geosyst.*, *10*, Q03013, doi:10.1029/2008GC002244.

Pavón-Carrasco, F. J., M. L. Osete, and J. M. Torta (2010), Regional modeling of the geomagnetic field in Europe from 6000 to 1000 BC, *Geochem. Geophys. Geosyst.*, *11*, Q11008, doi:10.1029/2010GC003197.

Pavón-Carrasco, F. J., J. Rodríguez-González, M. L. Osete, and J. M. Torta (2011), A Matlab tool for archaeomagnetic dating, *J. Archaeol. Sci.* *38*, 408–419.

Pavón-Carrasco, F. J., M. L. Osete, J. M. Torta, and A. De Santis (2014), A geomagnetic field model for the Holocene based on archaeomagnetic and lava flow data, *Earth Planet. Sci. Lett.*, *388*, 98–109.

Prevosti, M., L. Casas, J. F. Roig Pérez, B. Fouzai, A. Álvarez, and A. Pitarch (2013), Archaeological and archaeomagnetic dating at a site from the ager Tarraconensis (Tarragona, Spain): El Vila-sec Roman pottery, *J. Archaeol. Sci.*, *40*(6), 2686–2701.

Roberts, A. P., L. Tauxe, and D. Heslop (2013), Magnetic paleointensity stratigraphy and high-resolution Quaternary geochronology: Successes and future challenges, *Quat. Sci. Rev.*, *61*, 1–16.

Schnepp, E., P. Lanos, and A. Chauvin (2009), Geomagnetic Paleointensity between 1300 and 1750 AD derived from a bread oven floor sequence in Lubeck, Germany, *Geochem. Geophys. Geosyst.*, *10*, Q08003, doi:10.1029/2009GC002470.

Shaar, R., H. Ron, L. Tauxe, R. Kessel, A. Agnon, E. Ben-Yosef, and J. M. Feinberg (2010), Testing the accuracy of absolute intensity estimates of the ancient geomagnetic field using copper slag material, *Earth Planet. Sci. Lett.*, *290*, 201–213.

Shaar, R., E. Ben-Yosef, H. Ron, L. Tauxe, A. Agnon, and R. Kessel (2011), Geomagnetic field intensity: How high can it get? How fast can it change? Constraints from Iron copper slag, *Earth Planet. Sci. Lett.*, *301*, 297–306.

Shaw, J. (1974), A new method of determining the magnitude of the palaeomagnetic field: Application to five historic lavas and five archaeological samples, *Geophys. J. R. Astron. Soc.*, *39*, 133–141.

Spatharas, V., D. Kondopoulou, E. Aidona, and K. G. Efthimiadis (2011), New magnetic mineralogy and archaeointensity results from Greek kilns and baked clays, *Stud. Geophys. Geod.*, *55*, 132–157.

Suttie, N., R. Holme, M. J. Hill, and J. Shaw (2011), Consistent treatment of errors in archaeointensity implies rapid decay of the dipole prior to 1840, *Earth Planet. Sci. Lett.*, *304*, 13–21.

Tauxe, L. (2009), *Essentials of Paleomagnetism: Web Edition*, Scripps Inst. of Oceanogr., La Jolla, Calif.

Tema, E., and D. Kondopoulou (2011), Secular variation of the Earth's magnetic field in the Balkan region during the last eight millennia based on archaeomagnetic data, *Geophys. J. Int.*, *186*, 603–614.

Tema, E., M. Gómez-Paccard, D. Kondopoulou, and Y. Almar (2012), Intensity of the Earth's magnetic field in Greece during the last five millennia: New data from Greek pottery, *Phys. Earth Planet. Inter.*, *202–203*, 14–26.

Tema, E., C. Fantino, E. Ferrara, A. Lo Giudice, J. Morales, A. Goguitchaichvili, P. Camps, F. Barello, and M. Gulmini (2013a), Combined archaeomagnetic and thermoluminescence study of a brick kiln excavated at Fontanetto Po (Vercelli, Northern Italy), *J. Archaeol. Sci.*, *40*, 2025–2035.

Tema, E., J. Morales, A. Goguitchaichvili, and P. Camps (2013b), New archaeointensity data from Italy and geomagnetic field intensity variation in the Italian Peninsula, *Geophys. J. Int.*, *193*, 603–614.

Thébault, E. (2008), A proposal for regional modelling at the Earth's surface, R-SCHA2D, *Geophys. J. Int.*, *174*, 118–134.

Thébault, E., and Y. Gallet (2010), A bootstrap algorithm for deriving the archeomagnetic field intensity variation curve in the Middle East over the past 4 millennia BC, *Geophys. Res. Lett.*, *37*, L22303, doi:10.1029/2010GL044788.

Thébault, E., J. J. Schott, M. Mandea, and J. P. Hoffbeck (2004), A new proposal for Spherical Cap Harmonic Analysis, *Geophys. J. Int.*, *159*, 83–105.

Thébault, E., J. J. Schott, and M. Mandea (2006), Revised spherical cap harmonic analysis (R-SCHA): Validation and properties, *J. Geophys. Res.*, *111*, B01102, doi:10.1029/2005JB003836.

Thellier, E. (1938), Sur l'aimantation des terres cuites et ses applications géophysiques, *Ann. Inst. Phys. Globe Paris*, *16*, 157–302.

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

Usoskin, I. G. (2013), A history of solar activity over millennia, *Living Rev. Solar Phys.*, *10*, 1.

Veitch, R. J., G. Hedley, and J. J. Wagner (1984), An investigation of the intensity of the geomagnetic field during Roman times using magnetically anisotropic bricks and tiles, *Arch. Sc. Geneve*, *37*, 359–373.

Wollin, G., D. B. Ericson, and W. B. F. Ryan (1978), Climatic changes, magnetic intensity variations and fluctuations of eccentricity of Earth's orbit during past 2,000,000 years and mechanism which may be responsible for relationship, *Earth Planet. Sci. Lett.*, *41*, 395–397.