An end-member modelling approach to pseudo-Thellier palaeointensity data

Liz van Grinsven [ORCID],[1] Tristan van Leeuwen [ORCID][2,3] and Lennart V. de Groot [ORCID][1]

[1]*Paleomagnetic Laboratory Fort Hoofddijk, Utrecht University, Budapestlaan 17, 3584 CD Utrecht, the Netherlands, E-mail: l.b.vangrinsven@uu.nl*
[2]*Centrum Wiskunde and Informatica, Science Park 123, 1098 XG Amsterdam, the Netherlands*
[3]*Mathematical Institute, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, the Netherlands*

SUMMARY

Absolute palaeointensities are notoriously hard to obtain, because conventional thermal Thellier palaeointensity experiments often have low success rates for volcanic samples. The thermal treatments necessary for these experiments potentially induce (magnetic) alteration in the samples, preventing a reliable palaeointensity estimate. These heating steps can be avoided by pseudo-Thellier measurements, where samples are demagnetized and remagnetized with alternating fields. However, pseudo-Thellier experiments intrinsically produce relative palaeointensities. Over the past years, attempts were made to calibrate pseudo-Thellier results into absolute palaeointensities for lavas by mapping laboratory induced anhysteretic remanent magnetizations (ARMs) to the thermally acquired natural remanent magnetizations (NRMs). Naturally occurring volcanic rocks, however, are assemblages of minerals differing in grain size, shape and chemistry. These different minerals all have their own characteristic mapping between ARMs and thermal NRMs. Here, we show that it is possible to find these characteristic mappings by unmixing the NRM demagnetization and the ARM acquisition curves into end-members, with an iterative method of non-negative matrix factorization. In turn, this end-member modelling approach (EMMA) allows for the calculation of absolute palaeointensities from pseudo-Thellier measurements. We tested our EMMA using a noise-free numerical data set, yielding a perfect reconstruction of the palaeointensities. When adding noise up to levels beyond what is expected in natural samples, the end-member model still produces the known palaeointensities well. In addition, we made a synthetic data set with natural volcanic samples from different volcanic edifices that were given a magnetization by heating and cooling them in a controlled magnetic field in the lab. The applied fields ranged between 10 and 70 μT. The average absolute difference between the calculated palaeointensity and the known lab field is around 10 μT for the models with 2–4 end-members, while the palaeointensity of almost all flows can be retrieved within a deviation of ±20 μT. The deviations between the palaeointensities and the known lab fields are almost Gaussian distributed around the expected values. Although the two data sets in our study show that there is potential for using this end-member modelling technique for finding absolute palaeointensities from pseudo-Thellier data, these synthetic data sets cannot be directly related to natural samples. Therefore, it is necessary to compile a data set of known palaeointensities from different volcanic sites that recently cooled in a known magnetic field to find the universal end-members in future studies.

Key words: Palaeointensity; Palaeomagnetism; Rock and mineral magnetism; Numerical modelling.

1. INTRODUCTION

The Earth's magnetic field is generated by convection in the liquid outer core of our planet. Since the 1970s, this magnetic field has been measured continuously around the world by satellites. To understand how the geomagnetic field behaves over time and possibly

predict its future, it is important to have an accurate understanding of its more distant past. This would, for example, provide boundary conditions for models describing the behaviour of the geodynamo (e.g. Aubert *et al.* 2013; Sprain *et al.* 2019; Meduri *et al.* 2021), and will benefit models of (regional) palaeosecular variation that can be used for dating of archaeological artefacts and volcanic products

(e.g. Korte *et al.* 2011; Nilsson *et al.* 2014; Pavón-Carrasco *et al.* 2014).

The Earth's magnetic field is recorded by iron-bearing minerals in volcanic products, often basaltic lavas, when they cool below their Curie temperature. These volcanic products, therefore, may provide spot readings of the past state of the Earth's magnetic field for their moment of cooling and location on the planet. Full-vector information of the Earth's magnetic field from these lavas consist of a direction of the magnetic field, a palaeodirection and the strength of the field, a palaeointensity. While palaeodirections are usually relatively easy to obtain if the material has not undergone movement since its cooling, palaeointensity measurements are much more complicated (Tauxe & Yamazaki 2015). In principle, it is possible to estimate the intensity of an ancient magnetic field because the mechanism by which volcanic rocks acquire their primary magnetization is in theory linearly related to the ambient field for low fields such as the Earth's (Folgeraiter 1899; Koningsberger 1938). An important assumption of all palaeointensity techniques is therefore that the natural remanent magnetization (NRM) of the samples is a pristine thermoremanent magnetization (TRM), acquired at the time of cooling.

The most common method for obtaining absolute palaeointensities from samples with a TRM, is the Thellier–Thellier method (Thellier & Thellier 1959) and variations of their initial protocol (e.g. Coe 1967; Aitken *et al.* 1988; Tauxe & Staudigel 2004; Yu *et al.* 2004). The concept of a Thellier-style experiment is to stepwise replace the NRM by a laboratory induced TRM. Then, the assumed linear relationship between the NRM and the laboratory induced TRM, since they are both thermally induced, can be used to obtain the palaeointensity with the help of an Arai plot (Arai 1963; Nagata *et al.* 1963). In practice, however, Thellier-style measurements are hampered in many ways, often arising from the complex mineralogy of the iron-bearing minerals. For instance, chemical alteration during the experiment, complex magnetic behaviour of large iron-oxide minerals (i.e. 'multidomain behaviour), and magnetic anisotropy of the sample may all prevent a successful determination of the palaeointensity (Tauxe & Yamazaki 2015). Some of these problems arise due to the subsequent heating steps that are necessary for a Thellier-style experiment. Avoiding heating the samples may prevent chemical alteration and to some extend multidomain behaviour to occur (e.g. Yu *et al.* 2002a). A Thellier-based experimental technique that does not use heating is the 'pseudo-Thellier' technique (Tauxe *et al.* 1995), which instead uses alternating fields (AF) to remove the NRM and impart laboratory magnetizations, in this case anhysteretic remanent magnetizations (ARMs), in the samples.

The pseudo-Thellier Technique was originally developed to help determining relative palaeointensity records from sediment cores (Tauxe *et al.* 1995), but the rationale of using AF steps instead of heating was later also applied to volcanic rocks. The main challenge with substituting the heating steps for AF steps in a Thellier-style experiment is that TRM demagnetization and remagnetization behaviour is not necessarily (linearly) proportional to AF demagnetization and remagnetization behaviour of the same minerals. Therefore, the slope of the linear fit in a pseudo-Thellier Arai diagram cannot directly be used to calculate the palaeointensity, and a conversion or 'mapping' of the AF demagnetization and remagnetization behaviour to the thermal demagnetization and remagnetization behaviour is necessary. Yu *et al.* (2002b) studied the analogy between ARM and TRM behaviour in volcanic rocks. They found

that although the behaviour of ARMs and TRMs in volcanic rocks show some resemblance, it is difficult to map the behaviour one-to-one. Moreover, the mapping between ARMs and TRMs differs for iron oxides that differ in chemistry, size and shape. Since the magnetic carriers in volcanic rocks consists of assemblages of iron oxides with differing properties, the mapping between the TRM and ARM behaviour of the entire sample may very well be complex and nonlinear.

de Groot *et al.* (2013a) applied a strict rock-magnetic selection criterion to choose a group of samples with relatively similar ARM and TRM behaviour for a pseudo-Thellier study subjecting Hawaiian lavas. This allowed establishing a calibration equation by using a group of recent (1840–2010 AD) lavas that cooled in a known magnetic field for pseudo-Thellier experiments. This calibration relation was defined as a linear relation between the slopes of pseudo-Thellier measurements and their known palaeointensities. Surprisingly, this linear calibration relation has a non-zero *y*-axis intercept. This is problematic because this implies that if there is no magnetization present when cooling, the sample still has a palaeointensity of around 15 μT. Furthermore, this calibration relation is only applicable to samples that pass a strict selection criterion, which means that it is not applicable to all samples. Paterson *et al.* (2016) addressed the issue of the non-zero intercept of the *y*-axis, by using samples that were given a known TRM in the laboratory before the pseudo-Thellier experiments were done and anchoring the calibration relation to the origin. Theoretically, this calibration relation is more correct, but unfortunately it does not reproduce known palaeointensities from recent lavas better than the experimentally estimated calibration relation of de Groot *et al.* (2013a, 2015, 2016). Previous efforts to define a calibration relation for pseudo-Thellier experiments, that is, a mapping between ARM and TRM demagnetization and remagnetization behaviour, are hampered by the fact that they attempt to provide a single mapping between the TRM and ARM behaviour for bulk samples that may consist of assemblages of iron oxides that differ in magnetic properties. Because of these differences in magnetic properties, different minerals in a bulk sample may exhibit different relations between ARM and TRM behaviour, while they are measured simultaneously. This adds complexity to finding a proper mapping between ARMs and TRMs in a sample.

Here, we apply an end-member modelling approach (EMMA) to describe the magnetization of bulk samples as a combination of several empirical end-members with their own characteristic mapping between their ARM and TRM behaviour. After these end-members are defined, it is possible to unmix the measurements of bulk samples and obtain absolute palaeointensities from their pseudo-Thellier results. Such an EMMA has been successfully used in palaeomagnetic studies before; it was used to unravel and characterize curves of Isothermal remanent magnetization (IRMs, Kruiver *et al.* 2001) and ARMs (Robertson & France 1994; Egli 2003; Heslop & Dillon 2007). Here, we expand the unmixing of Heslop & Dillon (2007) to optimize for the two data sets in a pseudo-Thellier study (NRM demagnetization and ARM acquisition) simultaneously with the ultimate aim of retrieving the absolute palaeointensities from lavas. To illustrate the potential of the EMMA technique for pseudo-Thellier experiments on lavas we first apply it to a numerically created synthetic data set. Then, we apply it to a data set of basalts that were given a full TRM in the laboratory prior to measuring the pseudo-Thellier experiments.

2. METHODS

A palaeomagnetic sample used for palaeointensity studies typically consists of a mixture of different magnetic minerals that differ in size, shape and chemistry. This implies that these magnetic minerals also differ in magnetic behaviour. The rationale of the EMMA pseudo-Thellier technique is to identify common magnetic components, that is, end-members, in AF demagnetization and ARM acquisition measurements of volcanic samples. These end-members can be obtained by unmixing a large data set of AF measurements from volcanic samples with a known palaeointensity. Once these common end-members have been obtained they can be used to unmix measurements from volcanic samples, with an unknown palaeointensity, as a combination of these end-members with defined behaviour to obtain palaeointensity estimates.

The unmixing modelling technique used in this paper largely follows Heslop & Dillon (2007). They presented a modelling procedure that uses linear combinations of end-members to represent coercivity spectra. Importantly, the procedure of Heslop & Dillon (2007) unmixes and finds end-members for only one data set; we will expand on this and make the routine suitable to unmix two data sets simultaneously.

In general, EMMA starts with the measured (magnetization) data, for which the stepwise measurements of all the samples are contained in data matrix X. This matrix is formed with all measurements of each sample in a row and the measurements of the same demagnetization steps in each column. The rationale is to unmix data matrix X into the product of two matrices, S and A. Therein, matrix S contains the end-member curves, the common magnetic components, and matrix A the (non-negative) abundances, which contains the number of end-members present in each sample and their relative abundance (eq. 4). In practice, the measurements contain errors, and so error ϵ must be added, where epsilon denotes the additive measurement noise.

$$X = AS + \epsilon. \tag{4}$$

To find the combination of end-members and abundances that represent the data best, a squared Euclidean distance cost function needs to be minimized to estimate the end-members and abundances, matrix S and A. The cost function gives a measure of the difference between data X and the modelled data AS (eq. 5):

$$||X = AS||^2. \tag{5}$$

This cost function can be minimized with respect to A and S (eqs 6 and 7), by using an iterative method of non-negative matrix factorization (Lee & Seung 2000). This is an optimization problem, with the aim to find the best-fitting solution, that is, a combination of matrices A and S that explain the data X best.

$$S \leftarrow S \frac{A^T X}{A^T AS}, \tag{6}$$

$$A \leftarrow A \frac{XS^T}{ASS^T}. \tag{7}$$

Here, we aim to find the absolute palaeointensity from pseudo-Thellier measurements, which consists of two data sets; NRM demagnetization and ARM acquisition data. Therefore, we will unmix two data sets simultaneously by building on concept of unmixing only one data set introduced by Heslop & Dillon (2007) (eq. 4), namely:

$$X = B_{ancient} AS_X, \tag{8}$$

$$Y = B_{lab} AS_Y, \tag{9}$$

where X and Y represent the NRM demagnetization and the ARM acquisition data sets, respectively. The absolute palaeointensity and laboratory field strength are indicated by $B_{ancient}$ and B_{lab}. Both data sets have their own set of end-members, S_X and S_Y, which are different but correlated to each other, since it is the expression of the same magnetic component but in a different measurement. Lastly, both data sets have the same matrix of abundances A, since the abundance of the magnetic components of a sample is the same in the AF demagnetization as in the ARM acquisition measurement.

Expanding EMMA from one to two data-series requires changes to eqs (5)–(7). This is done in two steps: step 1 (calibration) aims to estimate the common end-members, S_X and S_Y, from a data set that acquired its magnetizations in a known magnetic field (in other words $B_{ancient}$ is known). In step 2 (classification), we use the end-members obtained in step 1, to check how well they are able to produce absolute palaeointensities ($B_{ancient}$) from a data set that acquired its magnetizations in an unknown magnetic field.

2.1 Step 1: calibration

We measure X and Y, and $B_{ancient}$ and B_{lab} are known → we estimate S_X and S_Y (to find them we also need to find A).

In a very similar way to eq. (5), we obtain two Euclidean distance cost functions from eqs (8) and (9), which we must minimize simultaneously with respect to A, S_X and S_Y. Note that in this equation X and Y have been divided by the known palaeointensities $B_{ancient}$ and B_{lab}.

$$||X - AS_X||^2 + ||Y - AS_Y||^2. \tag{10}$$

By using the same iterative method of non-negative matrix factorization as Lee and Seung (2001), the following equations for A, S_X and S_Y can be obtained:

$$S_X \leftarrow S_X \frac{A^T X}{A^T AS_X}, \tag{11}$$

$$S_Y \leftarrow S_Y \frac{A^T Y}{A^T AS_Y}, \tag{12}$$

$$A \leftarrow \alpha_1 A \frac{XS_X^T}{AS_X S_X^T} + \alpha_2 A \frac{YS_Y^T}{AS_Y S_Y^T}, \tag{13}$$

where α_1 and α_2 are weighting factors which together sum to 1. These weighting factors determine the dependency on the AF demagnetization data set with respect to the ARM acquisition data set for the calculation of matrix A. These weighting factors can be empirically tested, to see which values give the most accurate calculation of the palaeointensity. The unmixing scheme starts with an initial guess for each end-member and the abundances, which should be slightly different for each end-member, otherwise the unmixing does not start properly. Any adequate initial guess gives similar results, observed from testing different initial guesses for the end-members.

2.2 Step 2: classification

We measure X and Y, and S_X, S_Y and B_{lab} are known → we estimate $B_{ancient}$ (to find $B_{ancient}$ we also need to find A)

The aim of step 2 is to find the absolute palaeointensity ($B_{ancient}$) for each pseudo-Thellier measurement from the common end-members that were defined in step 1. This can easily be obtained from eqs (8) and (9), since we have two equations and two unknowns, namely; A and $B_{ancient}$.

2.3 Data sets

To illustrate the potential of EMMA for pseudo-Thellier experiments we use two different data sets in our study: a numerical data set and a synthetic data set based on basaltic samples that were given a full TRM in the laboratory. First, for 90 per cent of the data set (the training data set), step 1 is performed to find the end-members for that specific training data set. Second, step 2 is performed on the remaining 10 per cent of the data set (the test data set) to see how well absolute palaeointensities can be found from the end-members obtained in step 1. These two steps are repeated 100 times to get an average for the absolute calculated palaeointensities and an uncertainty estimate. For the EMMA technique, the data sets need to be normalized: all measurements, both the NRM demagnetization and ARM acquisition measurements, are normalized by the maximum value of the corresponding ARM measurement. In addition, during step 1, the data set is divided by the field strength of the ancient field ($B_{ancient}$) in case of the NRM measurements and by the laboratory field strength (B_{lab}) for the ARM acquisition measurement.

The first data set to be tested is numerical, to check whether EMMA works for optimizing two data-series at once. This numerical data set is constructed from a set of end-members (S_X and S_Y), chosen from the results of an arbitrary run of the model with the synthetic data set described below, which is multiplied by a randomly chosen A and $B_{ancient}$ to obtain the numerical data-series X and Y. In addition, we observe how well the model is able to handle this numerical data set with increasing amounts of Gaussian noise added to the data-series X and Y. This is done by defining the standard deviation of the Gaussian noise that is added to the data set by a percentage, σ error per cent, (2, 4, 6, 8 and 10 per cent) of the norm of the corresponding data set.

The second data set is synthetic and contains real pseudo-Thellier measurements of laboratory magnetized samples. The data set contains a total of 247 cores that were taken from five different volcanic edifices, namely the island of Réunion (France), Hawaii (USA), and Pico (Azores, Portugal); Mt. Etna on Sicily (Italy) and Iceland. The cores were remagnetized in an oven going up to $700\,^\circ$C and were given a palaeointensity ($B_{ancient}$) between 10 and $70\,\mu$T. The samples from each volcanic location were divided into six sets of about eight samples, while a set from each volcanic location was heated and cooled in one of the following field strengths ($B_{ancient}$): 10, 22, 34, 46, 58 or $70\,\mu$T. Lastly, to check if the end-members of the synthetic data set have a physical meaning, that is, correspond to certain magnetic minerals and/or magnetic behaviour, a Curie temperature measurement was performed on a specimen from each set of samples.

3. RESULTS

To determine how well the EMMA technique is able to find the common end-members and calculate absolute palaeointensities for pseudo-Thellier measurements, first the numerical data set is used. The accuracy of the different models is described by two parameters. First, we use the difference in palaeointensity, Δ_{int}, which is defined as the average difference between the calculated palaeointensity by EMMA and the reference palaeointensity, for all samples; that is, Δ_{int} shows a potential systematic bias of the calculated palaeointensities from the model. Second, we define the absolute difference in palaeointensity $|\Delta|_{int}$ which is calculated in the same way as Δ_{int}, but takes the absolute values of the difference between calculated palaeointensity and reference palaeointensity, before calculating the average of all samples. This parameter shows on average how well
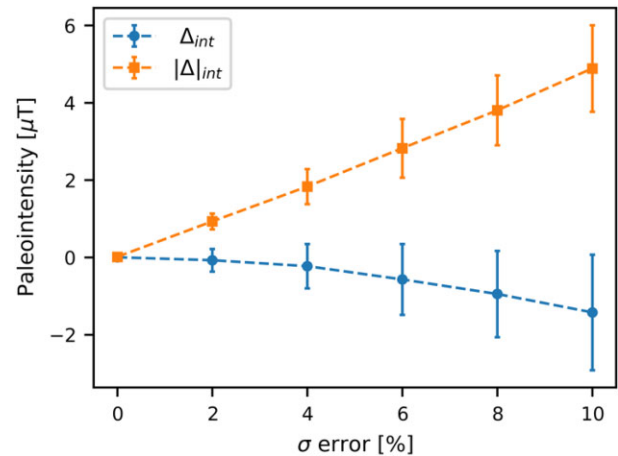
Figure 1. The performance of the 3 end-member model for a numerical data set with 255 measurements and $\alpha_1 = \alpha_2 = 0.5$: illustrated by the average difference between calculated palaeointensity and reference palaeointensity for all samples ($\Delta_{\mathbf{int}}$) and the average difference in absolute calculated palaeointensity ($|\Delta|_{\mathbf{int}}$) plotted against an increasing amount of added error, averaged over 100 iterations. Where the standard deviation of the Gaussian noise added to the data set is determined by a percentage (2, 4, 6, 8 and 10 per cent) of the norm of the data matrix (*X* and *Y,* respectively).

the model is able to calculate the palaeointensities for different samples. When EMMA performs well, both these values should be close to zero. If the performance of EMMA breaks down, it is important to know whether the errors in predicted palaeointensities are random, as would be indicated by a near zero Δ_{int}, but a high $|\Delta|_{int}$; or whether there is a bias in the results, in which case Δ_{int} would be negative for a systematic underestimate of the palaeointensity calculated by EMMA, or a positive Δ_{int} would show a systematic overestimate of the palaeointensity.

3.1 Numerical data set

For the numerical data set without added noise, EMMA is able to retrieve the palaeointensities perfectly: both Δ_{int} and $|\Delta|_{int}$ are close to zero (Fig 1). When adding Gaussian error to the numerical data set the calculation of the palaeointensity gradually becomes less accurate; $|\Delta|_{int}$ becomes linearly higher with increasing noise; Δ_{int}, which indicates systemic bias in the results is still very close to zero for an σ error *per cent* of 2 and 4, but then slowly decreases to negative values. It is important to note, however, that when adding error with a standard deviation of 10 per cent of the original data set, the model is still able to calculate the palaeointensities with an average absolute difference ($|\Delta|_{int}$) of $4.9\,\mu$T, which is better than the uncertainties in many palaeointensity studies on natural samples.

3.2 Synthetic data set

The synthetic data set uses pseudo-Thellier results from 247 samples. The entire data set, consisting of an NRM demagnetization measurement and an ARM acquisition measurement for each sample, is normalized by the corresponding last ARM measurement. This normalization means, for example, that a maximum NRM value for a sample of 5, implies a five time as high NRM demagnetization measurement compared to its corresponding ARM acquisition measurement. Fig. 2 displays the distribution of maximum NRM values of the synthetic data set, normalized by the
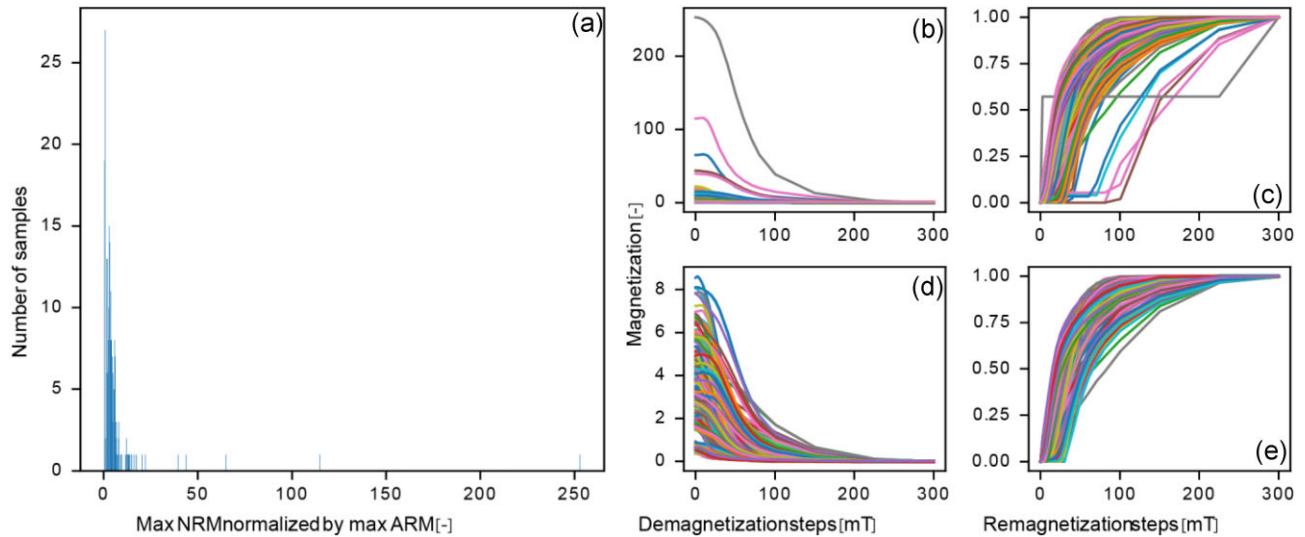
Figure 2. (a)The distribution of the maximum NRM values, obtained from the first measurement of each NRM demagnetization measurement normalized by the last measurement of the corresponding ARM acquisition measurement. (b)a nd (c) The synthetic data set (247 samples), NRM demagnetization (X) and ARM acquisition (Y) data. (d) and (e) The data set (225 samples) selected from the distribution of the normalized maximum NRM values, the data with and normalized maximum NRM value between 0 and 9.

corresponding maximum ARM acquisition value. This shows that the bulk of the data set (89 per cent) has a maximum normalized NRM value between 0 and 9. The calculation of the palaeointensity becomes twice as accurate by removing the outliers which have a maximum normalized NRM value higher than 9. This choice is validated by the observation that measurements which have a maximum NRM value that is much higher than the maximum ARM values exhibit anomalous shapes. In addition, these extreme outliers are very dominant in an unmixing scheme, because EMMA is trying to find common end-members into which all the different NRM demagnetization measurements can be unmixed. For example, a very strong NRM demagnetization plot will require a strong end-member, which can be very disturbing for finding the common end-members which make up the bulk of the measurements. The difference in the NRM demagnetization data set (X) and ARM acquisition data set (Y) with and without the outliers is shown in Fig. 2.

Both the numerical and synthetic data set give the most accurate palaeointensity calculation for an equal contribution of both the NRM demagnetization and ARM acquisition data sets for the calculation of A, in other words $\alpha_1 = \alpha_2 = 0.5$. For each iteration of the 3 end-member model, the average difference in palaeointensity (Δ_{int}) and the absolute difference in palaeointensity ($|\Delta|_{int}$) is calculated (Fig. 3a). In addition, the variance is calculated for each iteration to observe how well the calculated end-members (S_X and S_Y) and abundances (A) are able to explain the data matrices (X and Y), which EMMA is unmixing. The NRM demagnetization variance (NRM_v) and the ARM acquisition variance (ARM_v) are calculated by the difference between the original X and Y matrices and the X and Y matrices calculated from the unmixed end-members and abundances (Fig. 3b). It becomes clear from the values per iteration in Fig. 3 that the model quickly converges to a (local) minimum, with Δ_{int}, $|\Delta|_{int}$, NRM_v and ARM_v all moving towards fairly low values.

The synthetic data set can be unmixed into different numbers of end-members. The unmixed end-members for the 2, 3 and 4 end-member model of the synthetic data set are in Fig. 4. The (absolute) average error for the subsequent calculation of the palaeointensity

(Δ_{int} and $|\Delta|_{int}$) for the test data set after 100 iterations, is in Table 1. It is evident that for every end-member model, Δ_{int} shows a slight bias towards negative values and $|\Delta|_{int}$ is around 10 µT. Fig. 5, which displays the error distribution of $|\Delta|_{int}$ for each sample of the 3 end-member model, illustrates that the palaeointensity of almost all flows can be retrieved within a deviation of ± 20 µT, where most flows can be found with a much smaller error.

4. DISCUSSION

4.1 Performance of EMMA in pseudo-Thellier studies

The results using the numerical data set illustrate that the method used is mathematically sound. Without any noise, EMMA produces the palaeointensities of all samples without error. The model is also robust against Gaussian error, up to percentages of noise which are higher than expected in natural samples. However, it must be noted that the Gaussian error added to the numerical model may not be entirely representative of natural noise.

To test how well EMMA performs on samples that resemble natural samples, we made a synthetic data set using natural samples. This synthetic data set contains measurements of lab-magnetized samples from different volcanic edifices. The EMMA method is able to retrieve most palaeointensities (94 per cent) within a deviation of ±20 µT for a 3 end-member model of the synthetic data set. Most calculated palaeointensities are more precise, with 84 per cent being within a deviation of ±15 µT, 64 per cent within ±10 µT and 32 per cent within ±5 µT. In other words, the EMMA method is at least able to give an indication of most palaeointensities. Besides looking at the precision of the calculated palaeointensities, it is also important that the method is capable of calculating palaeointensities for a wide range of field strengths. Fig. 6 shows the distribution of the measured/calculated palaeointensity against the laboratory induced palaeointensity of the synthetic data set, for each sample. It is clear that, although far from perfect, EMMA is able to roughly estimate the palaeointensity for a wide range of field strengths. The average palaeointensities of the groups that were given palaeointensities 10, 22, 46 and 58 µT are very accurate, with 10.7, 21.3,
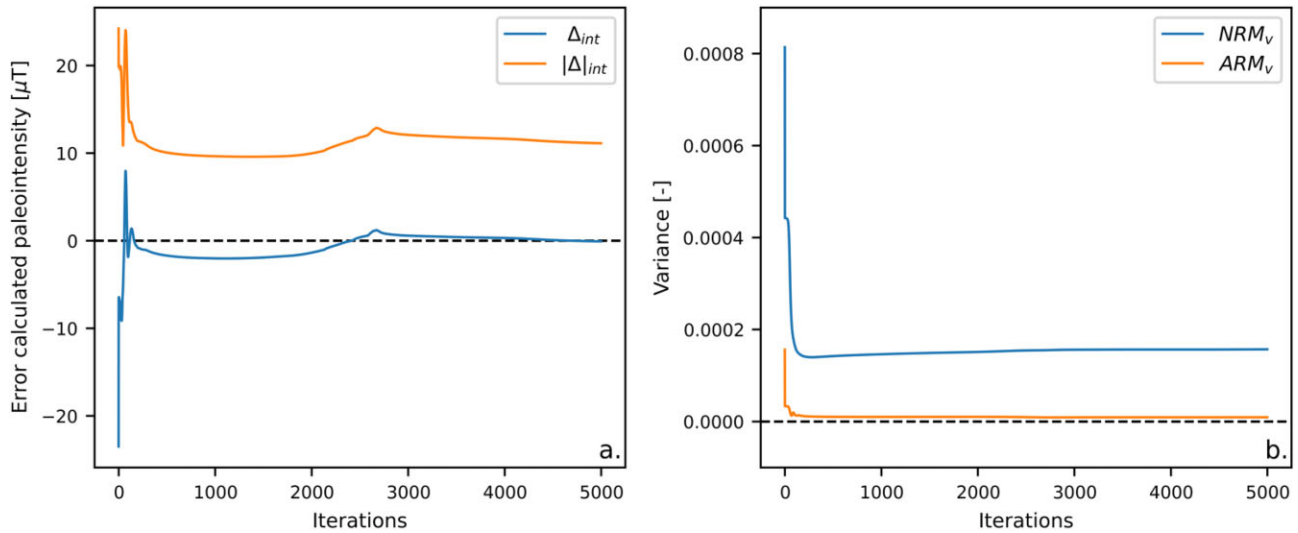
Figure 3. (a) Difference in palaeointensity (Δ_{int}) and absolute difference in palaeointensity ($|\Delta|_{int}$) for the 3 end-member model of the synthetic data set for the first 5000 iterations of step 1, the unmixing. (b) The NRM demagnetization (**NRM$_v$**) and ARM acquisition error (**ARM$_v$**) calculated, for the first 5000 iterations of the 3 end-member model, by the variance between the original **X** and **Y** matrices and the **X** and **Y** matrices calculated from the unmixed end-members (**S$_X$** and **S$_Y$**) and abundances (**A**).
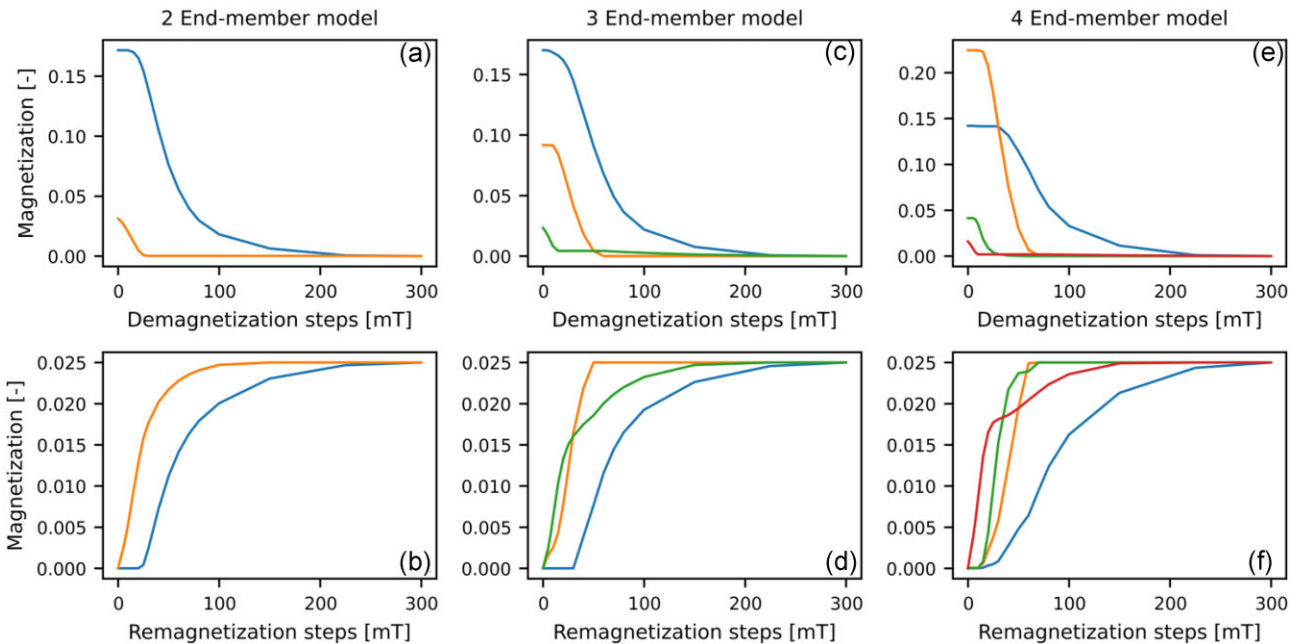

Figure 4. The end-members, for the NRM demagnetization and ARM acquisition data set, calculated in step 1, unmixing, for the synthetic data set for the(a) and (b) 2, (c) and (d) 3, and (e) and (f) 4 end-member model.

Table 1: Difference in calculated palaeointensity (Δ_{int}) and absolute difference in calculated palaeointensity ($|\Delta|_{int}$) calculated for the test data set averaged over 100 iterations with standard deviation (σ) for the 2, 3 and 4 end-member model of the synthetic data set in μ**T**.

	Δ_{int} (μT)	σ	$|\Delta|_{int}$ (μT)	σ
2 end-member model	−1.65	3.11	9.44	1.74
3 end-member model	−2.38	3.15	9.41	2.08
4 end-member model	−3.95	2.74	10.00	1.73

44.0 and 57.5 μT, respectively. The average of the 34 and 70 μT palaeointensity sets, however, deviate more with 27.1 and 61.3 μT,

respectively. The deviation of the 34 and 70 μT palaeointensity sets is most likely because of the distribution of data in the synthetic data set. Due to removing the outliers in the beginning, more samples have been removed from these sets than from the other groups of samples. What is furthermore important to note is that the trendline in Fig. 6 goes almost through the origin, the *y*-intercept is only 0.89 μT. This implies that EMMA overcomes the problem of the non-zero *y*-axis intercept of the calibration relations in de Groot *et al.* (2013a, 2015, 2016), and does not need to be forced through the origin as in Paterson *et al.* (2016). The slope of the fit through all points, however, is slightly low (0.91), where we would expect 1.
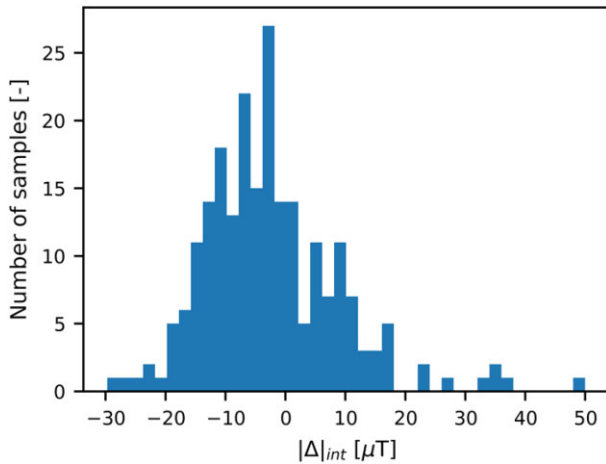
Figure 5. The distribution of the difference in calculated palaeointensity ($|\Delta|_{int}$) of each sample in the test data set, averaged over 100 iterations, for the 3 end-member model of the synthetic data set.
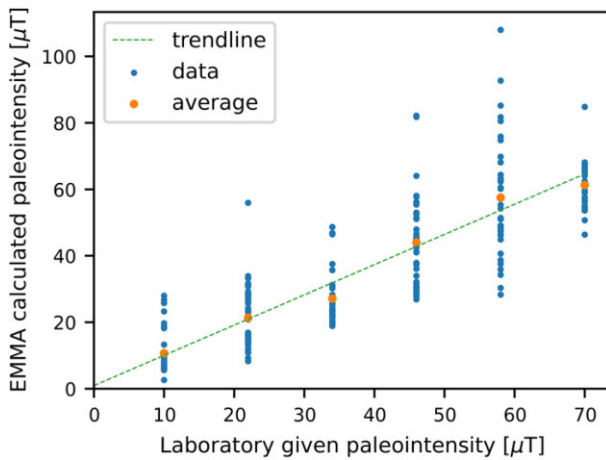
Figure 6. The blue dots give the distribution of the laboratory given palaeointensity [μT] against the measured/calculated palaeointensity [μT] of the test data set over 100 iterations by EMMA. The orange dots are the average measured/calculated palaeointensity [μT] for the entire palaeointensity set of 10, 22, 34, 46, 58 and 70 μT and the green line is the trendline through all the data points.

4.2 Physical representation of the end-members

To assess whether the end-members as produced by EMMA have any physical meaning, that is, represent groups of grains with similar magnetic behaviour, we measured the Curie temperature of each set of samples. Fig. 7 shows the correlation between the average abundances of each sample group (A) and the dominant Curie temperature (T_C) of the 3 end-member model in two steps over 100 iterations. In step 1, where the training data set is unmixed into end-members, and in step 2, when the palaeointensity of the test data set is calculated from the end-members. The relationship between the Curie temperatures and abundances are notably similar for steps 1 and 2, for the 3 end-member model (Fig. 7). This means that the distribution between end-members calculated in step 1, is the same distribution of end-members found when determining the palaeointensity in step 2. A second observation, which was not seen for the 2 and 4 end-member models, is a difference in the relationship between the Curie temperature and the abundance of each end-member. End-member 3 has higher abundances for

low Curie temperatures ($\pm$110–270 °C), whilst end-member 2 has higher abundances for higher Curie temperatures ($\pm$480–580 °C). End-member 1 shows no clear relationship between the Curie temperature and abundances for both steps 1 and 2. It must, however, be noted that the relationship between Curie temperatures and abundances is a trend and certainly not a rule which is illustrated by flows that clearly deviate from this trend in both plots.

4.3 Limitations of our model

In an ideal scenario, the end-members that are found by EMMA would have a physical meaning, that is, represent a specific type of mineral with well-defined magnetic behaviour. Based on the ternary diagram of common iron oxides in lavas (Fig. 8), a 4 end-member model should select ilmenite, ulvöspinel, magnetite and haematite. In practice, however, the Curie temperatures of end-members ilmenite and ulvöspinel are below room temperature and are therefore not distinguishable when measuring palaeomagnetic samples. Moreover, the variation in Curie temperatures, and therefore the magnetic behaviour of iron oxides in the samples, is a continuous spectrum governed by Ti-content and oxidation state. Therefore, unmixing the data set into a certain number of end-members still is a simplification of the system of iron oxides, and EMMA has to define end-members that represent somewhat random points in the ternary diagram. Lastly, the magnetic behaviour of iron oxides is not only determined by magnetic chemistry but also by grain size and shape, which makes a physical representation of end-members even more complex. Nevertheless, the end-members that were defined are somewhat related to magnetic minerology, as illustrated by the trend between abundance of end-members and the Curie temperatures of the sample (Fig. 7).

Another limitation of the unmixing technique is the large range in values between the NRM curves, which makes it difficult to unmix the entire data set into a few end-members. For this reason, we choose to remove strong outliers, measurements of samples with a very high NRM curve, before running the EMMA routine. Despite knowing that the ratio between the maximum NRM and ARM values is determined by the chosen laboratory field strength during the ARM measurements, which in our case is always 40 μT. This inherently favours a certain range of palaeointensities for the data sets presented in this paper. When further developing EMMA, it would be interesting to use measurements with a larger range of laboratory field strengths used during the ARM measurement.

Lastly, the method currently has no parameter or check that validates the accuracy of the palaeointensity calculation. The method is able to give an indication of the palaeointensity for almost all samples in the synthetic data set. However, the range of error for these calculations in palaeointensity is large. For the future development of this method, it would be very useful if such a validity parameter can be added to the EMMA method.

5. CONCLUSION AND OUTLOOK

We have shown, despite some limitations, that by unmixing a training data set into end-members and using these end-members to calculate the palaeointensity for a test data set we are able to at least provide a reasonable, first-order, approximation of the reference palaeointensity (Fig. 5), at least for our numerical and synthetic data sets. Our approximation of palaeointensities using EMMA is also better than previous attempts to calibrate pseudo-Thellier results from lavas into absolute estimates of the palaeointensity (de
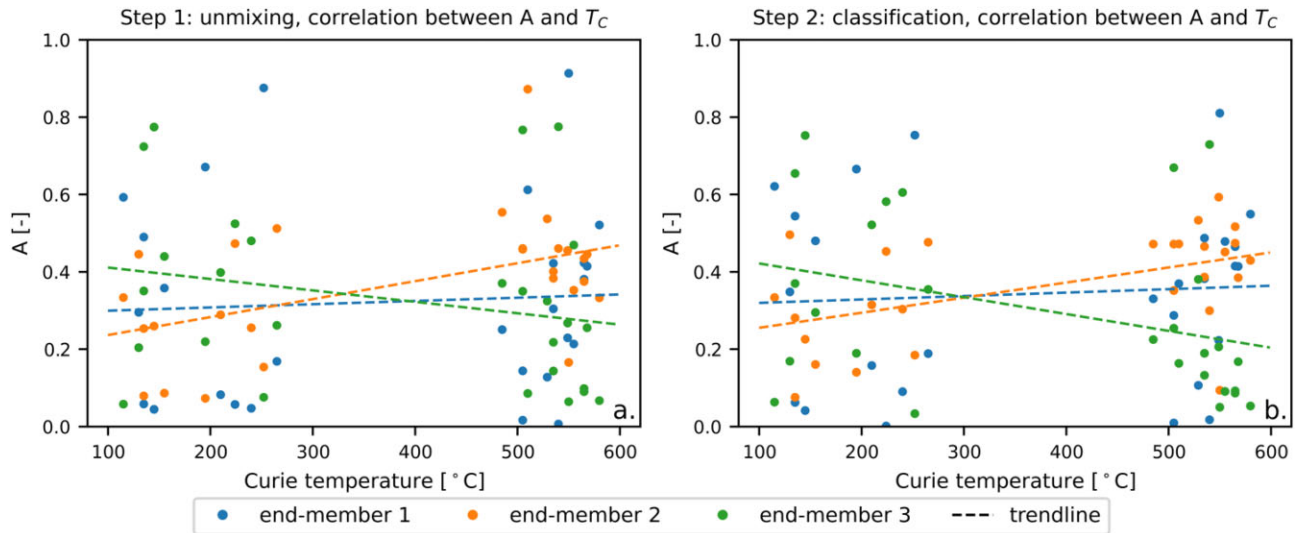
Figure 7. The correlation between the most dominant Curie temperatures (T_C) and the average abundances per sample group of each end-member (**A**) for step 1, (a) calibration and step 2, (b) classification. A linear trendline has been fitted for each end-member.
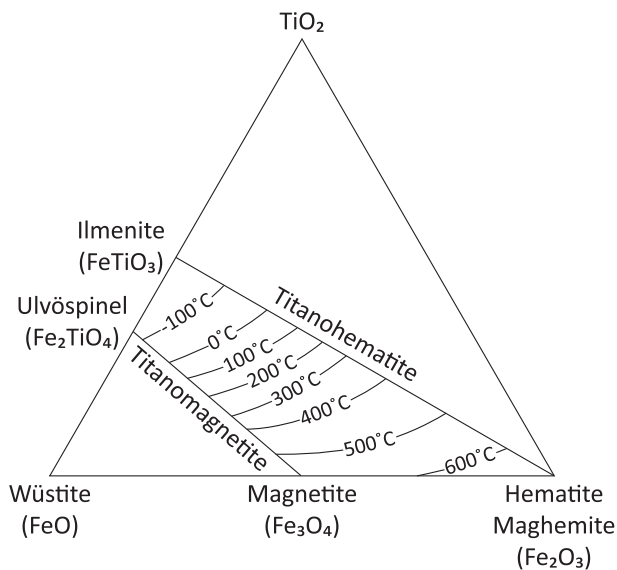
Figure 8. Ternary diagram showing the ulvôspinel–magnetite solid solution series with increasing temperature and the ilmenite–haematite solid solution (after Readman *et al*. 1972).

Groot *et al.* 2013a, 2015, 2016; Paterson *et al.* 2016). It is also important to emphasize that the EMMA method does not exclude samples based on rock-magnetic properties before the analysis and is applicable to all volcanic samples measured.

Interpreting pseudo-Thellier data from lavas using EMMA improves both the accuracy and the amount of data that can be interpreted. The numerical and synthetic data set are the first steps into defining a set of end-members that can be generally applied to volcanic or basaltic rocks from different volcanic edifices. Nevertheless, we know that palaeointensity experiments on samples that were just given a full TRM in the laboratory prior to the palaeointensity experiments always perform better than natural samples (e.g. de Groot *et al.* 2013b). Therefore, the data sets presented here cannot be directly related to natural samples, although they show that there is potential for using this end-member modelling technique for finding an approximation of the absolute palaeointensities from

pseudo-Thellier data. To define generally applicable end-members, it is necessary to compile a sample set of volcanics that recently cooled in known palaeointensities from different volcanic sites, that is, that cooled in the Earth's magnetic field with different field strengths.

CODE AND DATA AVAILABILITY

The EMMA.py code and the laboratory created test data set can be found on https://doi.org/10.5281/zenodo.8367125.

References

Aitken, M.J., Allsop, A.L., Bussell, G.D. & Winter, M.B., 1988. Determination of the intensity of the Earth's magnetic field during archaeological times: reliability of the Thellier Technique, *Reviews of Geophysics* **26**(1), 3–12.

Arai, Y., 1963. Secular variation in the intensity of the past geomagnetic field. *MSc thesis,* Univ. Tokyo, **84**.

Aubert, J., Finlay, C.C. & Fournier, A., 2013. Bottom-up control of geomagnetic secular variation by the Earth's inner core. *Nature,* **502**(7470), 219–223.

Coe, R.S., 1967. The determination of paleo-intensities of the Earth's magnetic field with emphasis on mechanisms which could cause non-ideal behavior in Thellier's method. *J. Geomag. Geoelec.,* **19**(3), 157–179.

de Groot, L. V., Béguin, A. Kosters, M. E. van Rijsingen, E. M. Struijk, E. L. Biggin, A. J. & Dekkers, M. J. 2015. High paleointensities for the Canary Islands constrain the Levant geomagnetic high., *Earth and Planetary Science Letters,* **419,** 154–167.

de Groot, L. V., Pimentel, A. & Di Chiara, A. 2016. The multimethod palaeointensity approach applied to volcanics from Terceira: full-vector geomagnetic data for the past 50 kyr., *Geophysical Journal International,* **206**(1), 590–604.

de Groot, L.V., Mullender, T.A. & Dekkers, M.J., 2013a. An evaluation of the influence of the experimental cooling rate along with other thermomagnetic effects to explain anomalously low palaeointensities obtained for historic lavas of Mt Etna (Italy). *Geophys. J. Int.,* **193**(3), 1198–1215.

de Groot, V., Biggin, A.J., Dekkers, M.J., Langereis, C.G. & Herrero-Bervera, E., 2013b. Rapid regional perturbations to the recent global geomagnetic decay revealed by a new Hawaiian record. *Nat. Commun.,* **4**(1).2727. 10.1038/ncomms3727.

Egli, R., 2003. Analysis of the field dependence of remanent magnetization curves. *J. geophys. Res.: Solid Earth,* **108**(B2). doi: 10.1029/2002JB002023.

Folgerhaiter, G., 1899. Sur le variations seculaires de l'inclinaison magnetique dans antiquite. *J. Phys.,* **8**, 5–16.

Heslop, D. & Dillon, M., 2007. Unmixin g magnetic remanence curves without a priori knowledge. *Geophys. J. Int.,* **193**(2), 1198–1215.

Koenigsberger, J.G., 1938. Natural residual magnetism of eruptive rocks. *Terr. Magn. Atmos. Electr.,* **43**(3), 299–320.

Korte, M., Constable, C., Donadini, F. & Holme, R., 2011. Reconstructing the Holocene geomagnetic field. *Earth planet. Sci. Lett.,* **312**(3-4), 497–505.

Kruiver, P.P., Dekkers, M.J & Heslop, D., 2001. Quantification of magnetic coercivity components by the analysis of acquisition curves of isothermal remanent magnetisation. *Earth planet. Sci. Lett.,* [online], **189**(3), 269–276.

Lee, D. & Seung, H. 2000. Algorithms for Non-negative Matrix Factorization., *Advances in neural information processing systems,* **13**, 535–541.

Meduri, D.G., Biggin, A.J., Davies, C.J., Bono, R.K., Sprain, C.J. & Wicht, J., 2021. Numerical dynamo simulations reproduce paleomagnetic field behavior. *Geophys. Res. Lett.,* **48**(5). doi: 10.1029/2020GL090544.

Nagata, T., Arai, Y. & Momose, K., 1963. Secular variation of the geomagnetic total force during the last 5000 years. *J. geophys. Res.,* **68**(18), 5277–5281.

Nilsson, A., Holme, R., Korte, M., Suttie, N. & Hill, M., 2014. Reconstructing Holocene geomagnetic field variation: new methods, models and implications. *Geophys. J. Int.,* **198**(1), 229–248.

Paterson, G.A., Heslop, D. & Pan, Y., 2016. The pseudo-Thellier palaeointensity method: new calibration and uncertainty estimates. *Geophys. Suppl. Mon. Notices R. Astron. Soc.,* **207**(3), 1596–1608.

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

Readman, P. W. & O'reilly, W. 1972. Magnetic properties of oxidized (cation-deficient) titanomagnetites (Fe, Ti, ￼??￼)43O4., *Journal of geomagnetism and geoelectricity,* **24**(1), 69–90. doi: 10.5636/jgg.24.69.

Robertson, D.J. & France, D.E., 1994. Discrimination of remanence-carrying minerals in mixtures, using isothermal remanent magnetisation acquisition curves. *Phys. Earth planet. Inter.,* **388**(3-4), 98–109.

Sprain, C.J., Biggin, A.J., Davies, C.J., Bono, R.K. & Meduri, D.G., 2019. An assessment of long duration geodynamo simulations using new paleomagnetic modeling criteria (QPM). *Earth planet. Sci. Lett.,* [online], **526**, 115758. doi: 10.1016/j.epsl.2019.115758.

Tauxe, L., Pick, T. & Kok, Y.S., 1995. Relative paleointensity in sediments: a Pseudo-Thellier Approach. *Geophys. Res. Lett.,* **22**(21), 2885–2888.

Tauxe, L. & Staudigel, H., 2004. Strength of the geomagnetic field in the Cretaceous Normal Superchron: new data from submarine basaltic glass of the Troodos Ophiolite. *Geochem. Geophys. Geosyst.,* **5**(2), doi: 10.1029/2003GC000635.

Tauxe, L. & Yamazaki, T., 2015. Paleointensities. In *Treatise on geophysics,* pp. 461–509. Elsevier.

Thellier, E., 1959. Sur l'intensite du shamp magnetique terrestre dans le passe historique et geologique. *Ann. Geophys.,* **15**, 285–376.

Yu, Y., Dunlop, D.J. & Özdemir, Ö., 2002a. Partial anhysteretic remanent magnetization in magnetite 1. Additivity. *J. geophys. Res.: Solid Earth,* **107**(B10), EPM 7–1-EPM 7-9.

Yu, Y., Dunlop, D.J. & Özdemir, Ö., 2002b. Partial anhysteretic remanent magnetization in magnetite 2. Reciprocity. *J. geophys. Res.: Solid Earth,* **107**(B10), EPM 8–1-EPM 8-9.

Yu, Y., Tauxe, L. & Genevey, A., 2004. Toward an optimal geomagnetic field intensity determination technique. *Geochem. Geophys. Geosyst.,* **5**(2), doi: 10.1029/2003GC000630.