An end-member modeling approach (EMMA) to pseudo-Thellier paleointensity data

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

[1]Paleomagnetic laboratory Fort Hoofddijk, Utrecht University, Budapestlaan 17, 3584 CD, Utrecht, the Netherlands

[2]Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, The Netherlands

[3]Mathematical Institute, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands

Correspondence to: Liz van Grinsven (l.b.vangrinsven@uu.nl)

Short title: End-member modeling of pseudo-Thellier data

Summary

Absolute paleointensities are notoriously hard to obtain, because conventional thermal Thellier paleointensity 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 paleointensity 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 paleointensities. Over the past years attempts were made to calibrate pseudo-Thellier results into absolute paleointensities 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 modeling approach (EMMA) allows for the calculation of absolute paleointensities from pseudo-Thellier measurements. We tested our end-member modeling approach using a noise-free numerical data set, yielding a perfect reconstruction of the paleointensities. When adding noise up to levels beyond what is expected in natural samples, the end-member model still produces the known paleointensities well. In addition, we made a synthetic dataset 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 paleointensity and the known lab-field is around 10 μT for the models with 2 to 4 end-members, while the paleointensity of almost all flows can be retrieved within a deviation of $\pm$ 20 μT. The average difference between calculated paleointensities for the 3 end-member model is -1.7 μT. The deviations between the paleointensities and the known lab-fields are almost Gaussian distributed around the expected values. To assess whether the end-members produced by our analysis have a physical meaning, we measured the Curie temperatures of our samples. These Curie measurements show that there is a relationship between the abundances of the end members of the 3 end-member model in the samples and their dominant Curie temperatures. This indicates that even whilst the spectrum of Curie temperatures and hence composition of iron-oxides in the sample set is continuous, and the magnetization is also related to mineral size and shape, the calculated end-members of the 3 end-member model are somewhat related to magnetic mineral composition present in the samples. Although the two datasets in our study show that there is potential for using this end-member modeling technique for finding absolute paleointensities from pseudo-Thellier data, these synthetic datasets cannot be directly related to natural samples. Therefore, it is necessary to compile a dataset of known

paleointensities 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 1970's, 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 behavior of the geodynamo (e.g., Aubert et al. 2013; Sprain et al. 2019; Meduri et al. 2021), and will benefit models of (regional) paleosecular variation that can be used for dating of archaeological artefacts and volcanic products (e.g., Korte et al. 2011; Pavón-Carrasco et al. 2014; Nilsson 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 paleodirection, and the strength of the field, a paleointensity. While paleodirections are usually relatively easy to obtain if the material has not undergone movement since its cooling, paleointensity measurements are much more complicated (Tauxe and 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 paleointensity 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 paleointensities from samples with a TRM, is the Thellier-Thellier method (Thellier and 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 paleointensity 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 behavior of large iron-oxide minerals (i.e., 'multidomain behavior), and magnetic anisotropy of the sample may all prevent a successful determination of the paleointensity (Tauxe and 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 behavior to occur (e.g., Yu et al. 2002). 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 paleointensity 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 behavior is not necessarily (linearly) proportional to AF

demagnetization and remagnetization behavior of the same minerals. Therefore, the slope of the linear fit in a pseudo-Thellier Arai diagram cannot directly be used to calculate the paleointensity, and a conversion or 'mapping' of the AF demagnetization and remagnetization behavior to the thermal demagnetization and remagnetization behavior is necessary. Yu and Dunlop (2002b) studied the analogy between ARM and TRM behavior in volcanic rocks. They found that although the behavior of ARMs and TRMs in volcanic rocks show some resemblance, it is difficult to map the behavior 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 behavior of the entire sample may very well be complex and non-linear.

De Groot et al. (2013a) applied a strict rock-magnetic selection criterion to choose a group of samples with relatively similar ARM and TRM behavior 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 paleointensities. 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 paleointensity 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 paleointensities 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, i.e., a mapping between ARM and TRM demagnetization and remagnetization behavior, are hampered by the fact that they attempt to provide a single mapping between the TRM and ARM behavior 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 behavior, 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 Modeling 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 behavior. After these end-members are defined, it is possible to unmix the measurements of bulk samples and obtain absolute paleointensities from their pseudo-Thellier results. Such an end-member modeling approach has been successfully used in paleomagnetic studies before; it was used to unravel and characterize curves of Isothermal Remanent Magnetization (IRMs, Kruiver et al. 2001) and ARMs (Robertson and France 1994; Egli 2003; Heslop and Dillon, 2007). Here we expand the unmixing of Heslop and Dillon (2007) to optimize for the two datasets in a pseudo-Thellier study (NRM demagnetization and ARM acquisition) simultaneously with the ultimate aim of retrieving the absolute paleointensities 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 dataset. Then we apply it to a dataset of basalts that were given a full TRM in the laboratory prior to measuring the pseudo-Thellier experiments.

2. Methods

A paleomagnetic sample used for paleointensity 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 behavior. The rationale of the EMMA pseudo-Thellier technique is to identify common magnetic components, i.e., end-members, in AF demagnetization and ARM acquisition measurements of volcanic samples. These end-members can be obtained by unmixing a large dataset of AF measurements from volcanic samples with a known paleointensity. Once these common end-members have been obtained they can be used to unmix measurements from volcanic samples, with an unknown paleointensity, as a combination of these end-members with defined behavior to obtain paleointensity estimates.

The unmixing modeling technique used in this paper largely follows Heslop and Dillon (2007). They presented a modeling procedure that uses linear combinations of end-members to represent coercivity spectra. Importantly, the procedure of Heslop and Dillon (2007) unmixes and finds end-members for only one dataset; we will expand on this and make the routine suitable to unmix two datasets 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 (equation 4). In practice, the measurements contain errors, and so error ϵ must be added, where epsilon denotes the additive measurement noise.

$$X = AS + \epsilon \quad (4)$$

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 modeled data AS (equation 5).

$$\|X - AS\|^2 \quad (5)$$

This cost function can be minimized with respect to A and S (equation 6 and 7), by using an iterative method of non-negative matrix factorization (Lee and Seung, 2001). This is an optimalization problem, with the aim to find the best fit solution, i.e., a combination of matrices A and S that explain the data X best.

$$S \leftarrow S\frac{A^T X}{A^T AS} \quad (6)$$

$$A \leftarrow A\frac{XS^T}{ASS^T} \quad (7)$$

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

$$X = B_{ancient} A S_X \quad (8)$$

$$Y = B_{lab} A S_Y \quad (9)$$

Where X and Y represent the NRM demagnetization and the ARM acquisition datasets, respectively. The absolute paleointensity and laboratory field strength are indicated by $B_{ancient}$ and B_{lab}. Both datasets 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 datasets 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 equations 5 to 7. This is done in two steps: step 1 (calibration) aims to estimate the common end-members, S_X and S_Y, from a dataset 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 paleointensities ($B_{ancient}$) from a dataset 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 equation 5, we obtain two Euclidean distance cost functions from equation 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 paleointensities $B_{ancient}$ and B_{lab}.

$$\|X - AS_X\|^2 + \|Y - AS_Y\|^2 \quad (10)$$

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:

$$S_X \leftarrow S_X \frac{A^T X}{A^T A S_X} \quad (11)$$

$$S_Y \leftarrow S_Y \frac{A^T Y}{A^T A S_Y} \quad (12)$$

$$A \leftarrow \alpha_1 A \frac{X S_X{}^T}{A S_X S_X{}^T} + \alpha_2 A \frac{Y S_Y{}^T}{A S_Y S_Y{}^T} \quad (13)$$

Where α_1 and α_2 are weighting factors which together sum to 1. These weighting factors determine the dependency on the AF demagnetization dataset with respect to the ARM acquisition dataset for the calculation of matrix A. These weighting factors can be empirically tested, to see which values give the most accurate calculation of the paleointensity. 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 paleointensity ($B_{ancient}$) for each pseudo-Thellier measurement from the common end-members that were defined in step 1. This can easily be obtained from equation 8 and 9, since we have two equations and two unknowns, namely; A and $B_{ancient}$.

2.3 Datasets

To illustrate the potential of EMMA for pseudo-Thellier experiments we use two different datasets in our study: a numerical dataset and a synthetic dataset based on basaltic samples that were given a full TRM in the laboratory. First, for 90 % of the dataset (the training dataset), step 1 is performed to find the end-members for that specific training dataset. Second, step 2 is performed on the remaining 10% of the dataset (the test dataset) to see how well absolute paleointensities 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 paleointensities and an uncertainty estimate. For the EMMA technique the datasets 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 dataset 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 dataset to be tested is numerical, to check whether EMMA works for optimizing two data-series at once. This numerical dataset 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 dataset 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 dataset 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 dataset by a percentage, σ error [%], (2, 4, 6, 8 and 10 %) of the norm of the corresponding dataset.

The second dataset is synthetic and contains real pseudo-Thellier measurements of laboratory magnetized samples. The dataset contains a total of 247 cores that were taken from 5 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 °C and were given a paleointensity ($B_{ancient}$) between 10 and 70 μT. 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 μT. Lastly, to check if the end-members of the synthetic dataset have a physical meaning, i.e., correspond to certain magnetic minerals and/or magnetic behavior, 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 paleointensities for pseudo-Thellier measurements, first the numerical dataset is used. The accuracy of the different models is described by two parameters. First, we use the difference in paleointensity, Δ_{int}, which is defined as the average difference between the calculated paleointensity by EMMA and the reference paleointensity, for all samples; i.e., Δ_{int} shows a potential systematic bias of the calculated paleointensities from the model. Second, we define the absolute difference in paleointensity $|\Delta|_{int}$ which is calculated in the same way as Δ_{int}, but takes the absolute values of the difference between calculated paleointensity and reference

paleointensity, before calculating the average of all samples. This parameter shows on average how well the model is able to calculate the paleointensities 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 paleointensities are random, as would be indicated by a near zero Δ_{int}, but a high $|\Delta|_{\mathrm{int}}$; or whether there is a bias in the results, in which case Δ_{int} would be negative for a systematic underestimate of the paleointensity calculated by EMMA, or a positive Δ_{int} would show a systematic overestimate of the paleointensity.

3.1 Numerical dataset

For the numerical dataset without added noise, EMMA is able to retrieve the paleointensities perfectly: both Δ_{int} and $|\Delta|_{\mathrm{int}}$ are close to zero (Fig 1). When adding gaussian error to the numerical dataset the calculation of the paleointensity gradually becomes less accurate; $|\Delta|_{\mathrm{int}}$ becomes linearly higher with increasing noise; Δ_{int}, which indicates systemic bias in the results is still very close to zero for an $\sigma\,\mathrm{error}\,[\%]$ 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% of the original dataset, the model is still able to calculate the paleointensities with an average absolute difference ($|\Delta|_{\mathrm{int}}$) of 4.9 μT, which is better than the uncertainties in many paleointensity studies on natural samples.

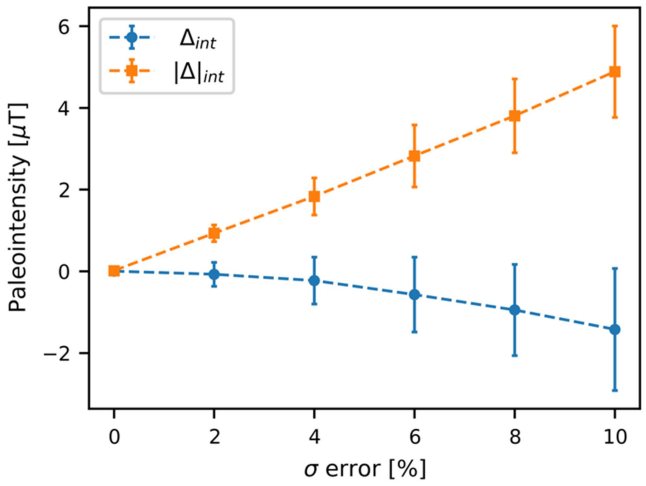

Figure 1: The performance of the 3 end-member model for a numerical dataset with 255 measurements and $\alpha_1 = \alpha_2 = 0.5$: Illustrated by the average difference between calculated paleointensity and reference paleointensity for all samples (Δ_{int}) and the average difference in absolute calculated paleointensity ($|\Delta|_{int}$) plotted against an increasing amount of added error, averaged over 100 iterations. Where the standard deviation of the Gaussian noise added to the dataset is determined by a percentage (2, 4, 6, 8 and 10 %) of the norm of the data matrix (X and Y respectively).

3.2 Synthetic dataset

The synthetic dataset uses pseudo-Thellier results from 247 samples. The entire dataset, 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. Figure 2 displays the distribution of maximum NRM values of the synthetic dataset, normalized by the corresponding maximum ARM acquisition value. This shows that the bulk of the dataset (89%) has a maximum normalized NRM value between 0 and 9. The calculation of the paleointensity

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 dataset (X) and ARM acquisition dataset (Y) with and without the outliers is shown in Fig. 2.

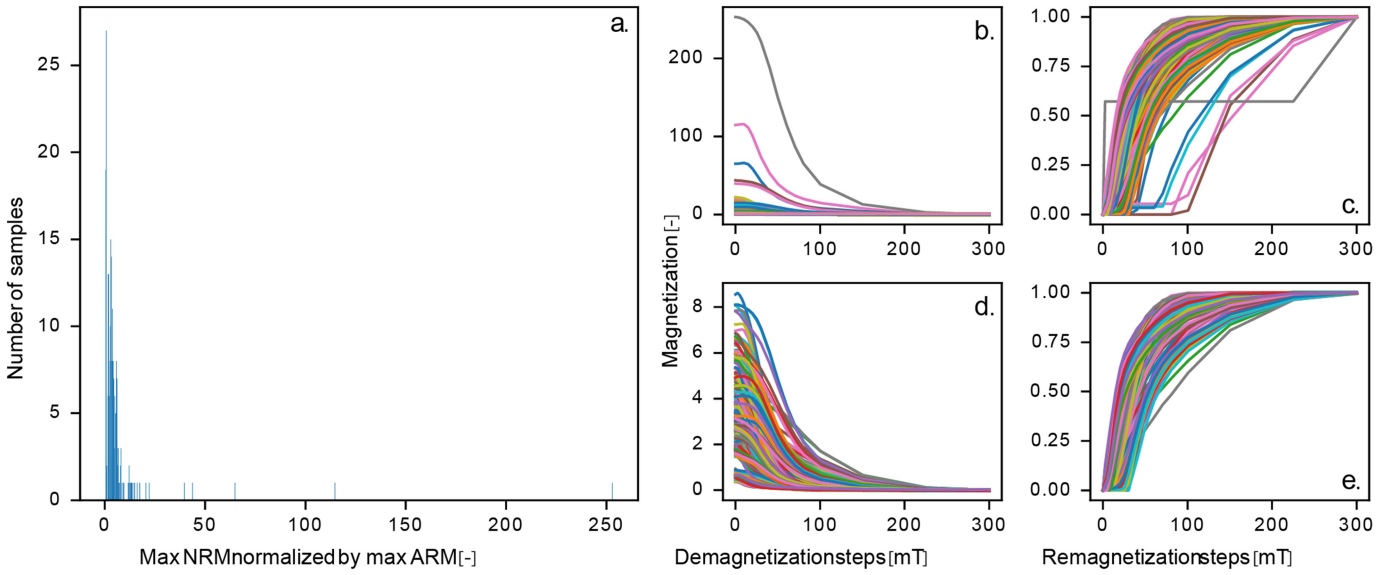

Figure 2: a. The distribution of the maximum NRM values, obtained from the first measurement of each NRM demagnetization measurement normalized by the last measurement of the corresponding ARM acquisition measurement. b./c. The synthetic dataset (247 samples), NRM demagnetization (X) and ARM acquisition (Y) data. d./e. The dataset (225 samples) selected from the distribution of the normalized maximum NRM values, the data with and normalized maximum NRM value between 0 and 9.

Both the numerical and synthetic dataset give the most accurate paleointensity calculation for an equal contribution of both the NRM demagnetization and ARM acquisition datasets 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 paleointensity (Δ_{int}) and the absolute difference in paleointensity ($|\Delta|_{int}$) 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.

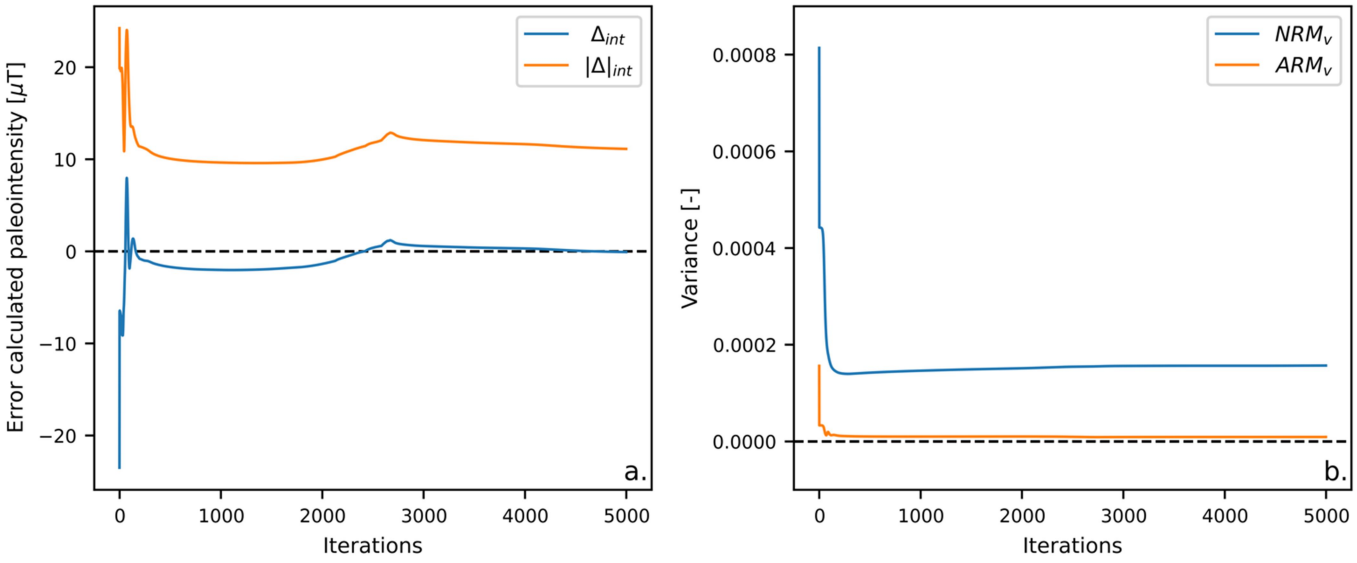

Figure 3: a. Difference in paleointensity (Δ_{int}) and absolute difference in paleointensity ($|\Delta|_{int}$) for the 3 end-member model of the synthetic dataset for the first 5000 iterations of step 1, the unmixing. b. The NRM demagnetization (NRM_v) and ARM acquisition error (ARM_v) calculated, 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).

The synthetic dataset 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 dataset are in Fig. 4. The (absolute) average error for the subsequent calculation of the paleointensity (Δ_{int} and $|\Delta|_{int}$) for the test dataset after a 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. Figure 5, which displays the error distribution of $|\Delta|_{int}$ for each sample of the 3 end-member model, illustrates that the paleointensity of almost all flows can be retrieved within a deviation of ±20 µT, where most flows can be found with a much smaller error.

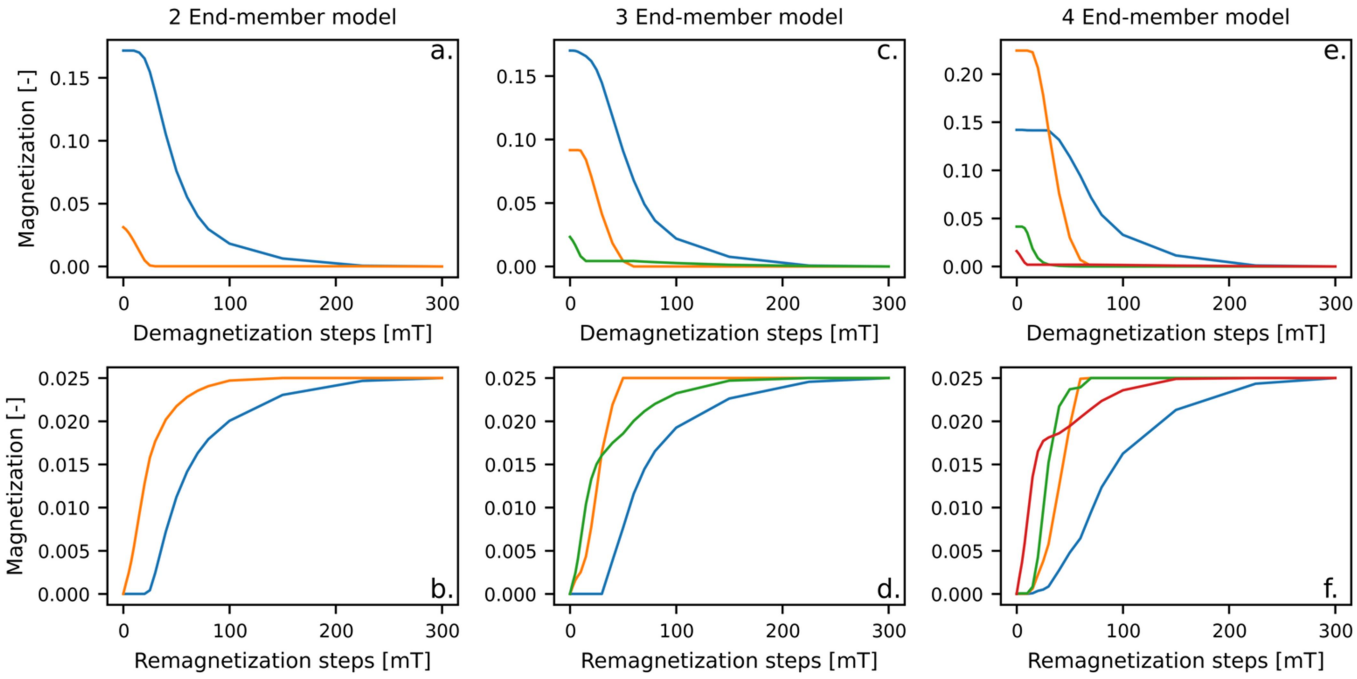

Figure 4: The end-members, for the NRM demagnetization and ARM acquisition dataset, calculated in step 1, unmixing, for the synthetic dataset for the 2 (a. and b.), 3 (c. and d.) and 4 (e. and f.) end-member model.

	Δ_{int} [μT]	σ	$\lvert\Delta\rvert_{int}$ [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

Table 1: Difference in calculated paleointensity (Δ_{int}) and absolute difference in calculated paleointensity ($\lvert\Delta\rvert_{int}$) calculated for the test dataset averaged over 100 iterations with standard deviation (σ) for the 2, 3, and 4 end-member model of the synthetic dataset in μT.

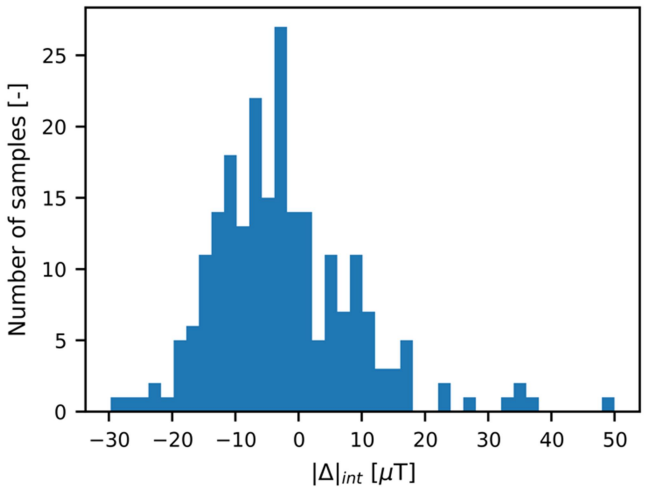

Figure 5: The distribution of the difference in calculated paleointensity ($\lvert\Delta\rvert_{int}$) of each sample in the test dataset, averaged over 100 iterations, for the 3 end-member model of the synthetic dataset.

4. Discussion

4.1 Performance of EMMA in pseudo-Thellier studies

The results using the numerical dataset illustrate that the method used is mathematically sound. Without any noise, EMMA produces the paleointensities 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 dataset using natural samples. This synthetic dataset contains measurements of lab-magnetized samples from different volcanic edifices. The EMMA method is able to retrieve most paleointensities (94%) within a deviation of ±20 μT for a 3 end-member model of the synthetic dataset. Most calculated paleointensities are more precise, with 84% being within a deviation of ±15 μT, 64% within ±10 μT, and 32% within ±5 μT. In other words, the EMMA method is at least able to give an indication of most paleointensities. Besides looking at the precision of the calculated paleointensities, it is also important that the method is capable of calculating paleointensities for a wide range of field strengths. Figure 6 shows the distribution of the measured/calculated paleointensity against the laboratory induced paleointensity of the synthetic dataset, for each sample. It is clear that, although far from perfect, EMMA is able to roughly estimate the paleointensity for a wide range of field strengths. The average paleointensities of the groups that were given paleointensities 10, 22, 46 and 58 μT are very accurate, with 10.7, 21.3, 44.0, 57.5 μT respectively. The average of the 34 and 70 μT paleointensity sets, however, deviate more with 27.1 and 61.3 μT respectively. The deviation of the 34 and 70 μT paleointensity sets is most likely because of the distribution of data in the synthetic dataset. 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.

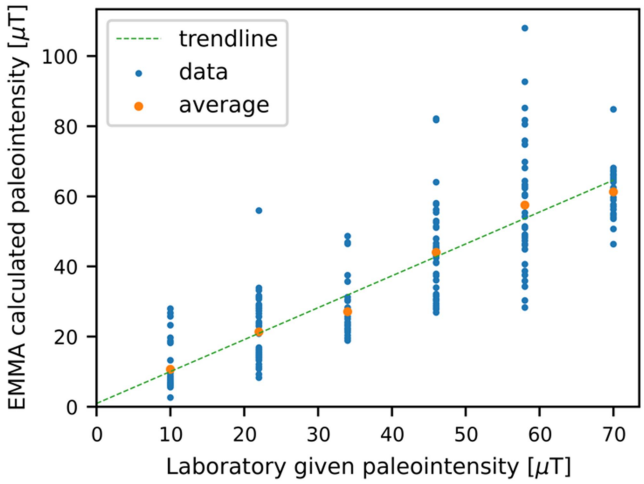

Figure 6: The blue dots give the distribution of the laboratory given paleointensity [µT] against the measured/calculated paleointensity [µT] of the test dataset over 100 iterations by EMMA. The orange dots are the average measured/calculated paleointensity [µT] for the entire paleointensity set of 10, 22, 34, 46, 58 and 70 [µT], 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, i.e. represent groups of grains with similar magnetic behavior, we measured the Curie temperature of each set of samples. Figure 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 dataset is unmixed into end-members, and in step 2, when the paleointensity of the test dataset is calculated from the end-members.

The relationship between the Curie temperatures and abundances are notably similar for step 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 paleointensity 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 (±110-270°C), whilst end-member 2 has higher abundances for higher Curie temperatures (±480-580°C). End-member 1 shows no clear relationship between the Curie temperature and abundances for both step 1 and step 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.

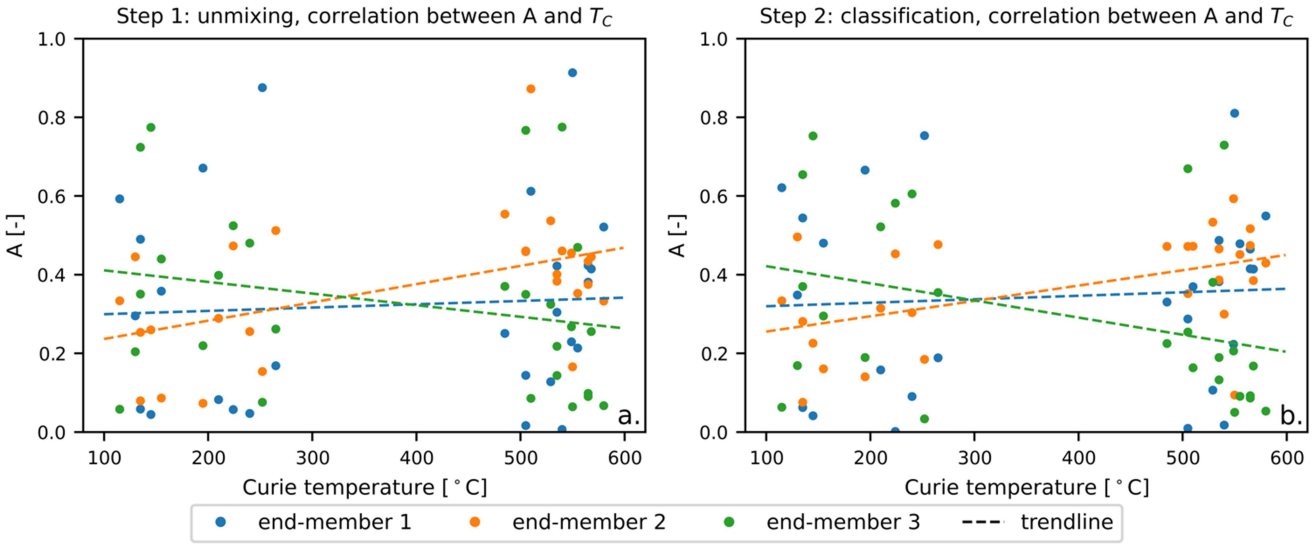

Figure 7: The correlation between the most dominant Curie temperatures (T_C) and the average abundances per sample group of each end-member (A) for step 1, calibration (Figure 7a.) and step 2, classification (Figure 7b.). A linear trendline has been fitted for each end-member.

4.3 Limitations of our model

In an ideal scenario, the end-members that are found by EMMA would have a physical meaning, i.e., represent a specific type of mineral with well-defined magnetic behavior. Based on the ternary diagram of common iron-oxides in lavas (Fig. 8), a four end-member model should select Ilmenite, Ulvöspinel, Magnetite and Hematite. In practice, however, the Curie temperatures of end-members Ilmenite and Ulvöspinel are below room temperature and are therefore not distinguishable when measuring paleomagnetic samples. Moreover, the variation in Curie temperatures, and therefore the magnetic behavior of iron-oxides in the samples, is a continuous spectrum governed by Ti-content and oxidation state. Therefore, unmixing the dataset 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 behavior 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 dataset 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 favors a certain range of paleointensities for the datasets 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.

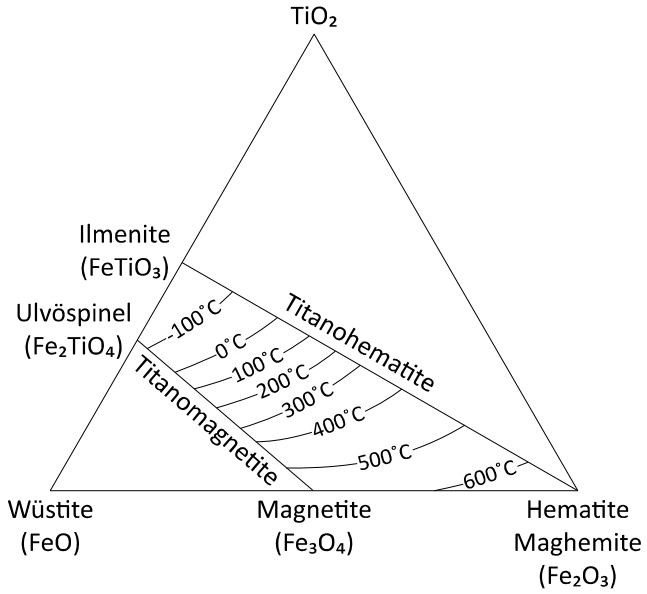

Figure 8: Ternary diagram showing the ulvôspinel-magnetite solid solution series with increasing temperature and the ilmenite-hematite solid solution (after Readman et al. 1972).

Lastly, the method currently has no parameter or check that validates the accuracy of the paleointensity calculation. The method is able to give an indication of the paleointensity for almost all samples in the synthetic dataset. However, the range of error for these calculations in paleointensity 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 dataset into end-members and using these end-members to calculate the paleointensity for a test dataset we are able to at least provide a reasonable, first-order, approximation of the reference paleointensity (Fig. 5), at least for our numerical and synthetic datasets. Our approximation of paleointensities using EMMA is also better than previous attempts to calibrate pseudo-Thellier results from lavas into absolute estimates of the paleointensity (de 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 dataset 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 paleointensity experiments on samples that were just given a full TRM in the laboratory prior to the paleointensity experiments always perform better than natural samples (e.g., de Groot et al. 2013b). Therefore, the datasets presented here cannot be directly related to natural samples, although they show that there is potential for using this end-member modeling technique for finding an approximation of the absolute paleointensities 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 paleointensities from different volcanic sites, i.e., 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 dataset can be found on https://doi.org/10.5281/zenodo.8367125.

Author contribution

Lennart V. de Groot conceived of the presented idea. Liz van Grinsven developed the theory, carried out the experiments, performed the computations and took the lead in writing the manuscript. Lennart V. de Groot and

Tristan van Leeuwen provided critical feedback and helped shape the research, analysis, and manuscript. Lennart V. de Groot supervised the project.

Acknowledgements

We thank our colleagues at paleomagnetic laboratory Fort Hoofddijk for their help, lively discussions, and comradeship. This project is funded by the Dutch Science Foundation (NWO) VIDI grant VI.Vidi.192.047 to LVdG.

References

Aitken, M.J., Allsop, A.L., Bussell, G.D. and Winter, M.B. (1988). Determination of the intensity of the Earth's magnetic field during archaeological times: Reliability of the Thellier Technique. 26(1), pp.3–3. doi:https://doi.org/10.1029/rg026i001p00003.

Arai, Y. (1963). Secular variation in the intensity of the past geomagnetic field. *M. Sc. Thesis, Univ. Tokyo*, *84*.

Aubert, J., Finlay, C.C. and Fournier, A. (2013). Bottom-up control of geomagnetic secular variation by the Earth's inner core. *Nature*, 502(7470), pp.219–223. doi:https://doi.org/10.1038/nature12574.

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. *Journal of geomagnetism and geoelectricity*, 19(3), pp.157–179. doi:https://doi.org/10.5636/jgg.19.157.

Egli, R. (2003). Analysis of the field dependence of remanent magnetization curves. *Journal of Geophysical Research: Solid Earth*, 108(B2). doi:https://doi.org/10.1029/2002jb002023.

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

Koenigsberger, J. G. (1938). Natural residual magnetism of eruptive rocks. *Terrestrial magnetism and atmospheric electricity*, *43*(3), 299-320.

De Groot, V., Biggin, A.J., Dekkers, M.J., Langereis, C.G. and Herrero-Bervera, E. (2013). Rapid regional perturbations to the recent global geomagnetic decay revealed by a new Hawaiian record. *Nature Communications*, 4(1). doi:https://doi.org/10.1038/ncomms3727.

de Groot, L. V., Mullender, T. A. and Dekkers, M. J. (2013). 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). *Geophysical Journal International*, *193*(3), 1198-1215.Heslop, David, and Melanie Dillon. "Unmixing magnetic remanence curves without a priori knowledge." *Geophysical Journal International* 170.2 (2007): 556-566. https://doi.org/10.1093/gji/ggt065.

Korte, M., Constable, C., Donadini, F. and Holme, R. (2011). Reconstructing the Holocene geomagnetic field. *Earth and Planetary Science Letters*, 312(3-4), pp.497–505. doi:https://doi.org/10.1016/j.epsl.2011.10.031.

Kruiver, P.P., Dekkers, M.J. and Heslop, D. (2001). Quantification of magnetic coercivity components by the analysis of acquisition curves of isothermal remanent magnetisation. *Earth and Planetary Science Letters*, [online] 189(3), pp.269–276. doi:https://doi.org/10.1016/S0012-821X(01)00367-3.

Meduri, D.G., Biggin, A.J., Davies, C.J., Bono, R.K., Sprain, C.J. and Wicht, J. (2021). Numerical Dynamo Simulations Reproduce Paleomagnetic Field Behavior. *Geophysical Research Letters*, 48(5). doi:https://doi.org/10.1029/2020gl090544.

Nagata, T., Arai, Y. and Momose, K. (1963). Secular Variation of the Geomagnetic Total Force during the Last 5000 Years. *Journal of Geophysical Research*, 68(18), pp.5277–5281. doi:https://doi.org/10.1029/j.2156-2202.1963.tb00005.x.

Nilsson, A., Holme, R., Korte, M., Suttie, N. and Hill, M. (2014). Reconstructing Holocene geomagnetic field variation: new methods, models and implications. *Geophysical Journal International*, 198(1), pp.229–248. doi:https://doi.org/10.1093/gji/ggu120.

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

Paterson, G. A., Heslop, D., and Pan, Y. (2016). The pseudo-Thellier palaeointensity method: New calibration and uncertainty estimates. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *207*(3), 1596-1608. doi:https://doi.org/10.1093/gji/ggw349.

Pavón-Carrasco, F. J., Osete, M. L., Torta, J. M., and De Santis, A. (2014). A geomagnetic field model for the Holocene based on archaeomagnetic and lava flow data. *Earth and Planetary Science Letters*, *388*, 98-109.Robertson, D. J., and D. E. France. "Discrimination of remanence-carrying minerals in mixtures, using isothermal remanent magnetisation acquisition curves." *Physics of the Earth and Planetary interiors* 82.3-4 (1994): 223-234. doi:https://doi.org/10.1016/j.epsl.2013.11.046.

Tauxe, L., Pick, T. and Kok, Y.S. (1995). Relative paleointensity in sediments: A Pseudo-Thellier Approach. *Geophysical Research Letters*, 22(21), pp.2885–2888. doi:https://doi.org/10.1029/95gl03166.

Tauxe, L. and Staudigel, H. (2004). Strength of the geomagnetic field in the Cretaceous Normal Superchron: New data from submarine basaltic glass of the Troodos Ophiolite. *Geochemistry, Geophysics, Geosystems*, 5(2), p.n/a-n/a. doi:https://doi.org/10.1029/2003gc000635.

Tauxe, L., and Yamazaki, T. (2015). Paleointensities. In Treatise on geophysics (pp. 461–509). Elsevier. https://doi.org/10.1016/B978-0-444-53802-4.00107-X.

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. and Özdemir, Ö. (2002). Partial anhysteretic remanent magnetization in magnetite 1. Additivity. *Journal of Geophysical Research: Solid Earth*, 107(B10), p.EPM 7-1-EPM 7-9. doi:https://doi.org/10.1029/2001jb001249.

Yu, Y., Dunlop, D.J. and Özdemir, Ö. (2002). Partial anhysteretic remanent magnetization in magnetite 2. Reciprocity. *Journal of Geophysical Research: Solid Earth*, 107(B10), p.EPM 8-1-EPM 8-9. doi:https://doi.org/10.1029/2001jb001269.

Yu, Y., Tauxe, L. and Agnès Genevey (2004). Toward an optimal geomagnetic field intensity determination technique. *Geochemistry Geophysics Geosystems*, 5(2), p.n/a-n/a. doi:https://doi.org/10.1029/2003gc000630.