Local magnetic anomalies explain bias in paleomagnetic data: consequences for sampling

Romy Meyer[1], Lennart V. de Groot[1]

[1]Paleomagnetic laboratory Fort Hoofddijk, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands

Key Points:

- Paleomagnetic data from Mt. Etna does often not reproduce the known geomagnetic field well
- Local magnetic anomalies explain bias in paleomagnetic data as function of topography
- Optimizing the paleomagnetic sampling strategy may suppress this bias in paleomagnetic data

Corresponding author: Romy Meyer, `r.meyer@uu.nl`

Abstract

Volcanic rocks are considered reliable recorders of past changes in the Earth's magnetic field. Recent flows, however, sometimes fail to produce the known magnetic field at the time of cooling. Here, we tested the accuracy of paleomagnetic data recorded by Mt. Etna lavas by comparing paleomagnetic data from historical flows to direct measurements of the magnetic field above the current topography. The inclinations and intensities in both data sets are biased towards lower values. They vary as a function of topography; both are higher above ridges and lower in gullies. To suppress this paleomagnetic data bias it is important to take samples several meters apart and from different parts of the flow whenever possible. While this leads to a higher degree of scatter in paleodirections, the results will better represent the Earth's magnetic field at the time of cooling. This emphasises the importance of reporting paleomagnetic sampling strategies in detail.

Plain Language Summary

Paleomagnetic data from lavas is routinely used in the Earth Sciences to e.g. reconstruct the past behavior of the Earth's magnetic field, or make models of past plate motions. Very young flows for which the ambient magnetic field at the time of cooling is known, however, sometimes fail to produce the known reference values. Here we show that the topography of volcanic terrain may influence the magnetic signal of new, overlying, flows, and we make recommendations for sampling strategies that suppress these terrain effects as much as possible.

1 Introduction

For decades magnetic signals from volcanic rocks have been used as a source to study the ancient behavior of the Earth's magnetic field. Upon cooling, volcanic rocks obtain a natural remanent magnetization which reflects the direction and intensity of the ambient geomagnetic field at that specific moment in time. Paleomagnetic data from well-dated flows (e.g. historical observations, radiocarbon dating) are used to create regional paleosecular variation (PSV) curves, and models that describe the global behavior of the Earth's magnetic field through time. With PSV curves, lava flows from unknown ages may be dated, which is vital for volcanic hazard assessment. An important prerequisite of the reliability of these models is the accuracy of the input data; volcanic rocks are often considered to be excellent recorders of the Earth's magnetic field. Paleomagnetic data obtained from recent volcanic rocks, however, regularly fail to produce their known field values (e.g. Cromwell et al., 2015) or their reference value from the International Geomagnetic Reference Field (IGRF, (Alken et al., 2021)).

Recent lavas from Mt Etna, Italy, have been extensively studied in terms of paleodirections and paleointensities. As a result there is a large paleomagnetic dataset, which is regularly inconsistent with the reference values. Moreover, the scatter in paleodirections from a single lava flow is often inexplicably large (Speranza et al., 2006), with inclinations around 2° to shallow (Tanguy et al., 1985; Rolph & Shaw, 1986; Rolph, 1997; Tanguy et al., 1999; Calvo et al., 2002; Incoronato et al., 2002; Tanguy et al., 2003; Lanza et al., 2005). Likewise, paleointensities are found to be generally to low (Rolph & Shaw, 1986; Sherwood, 1991; Biggin et al., 2007; de Groot et al., 2012, 2013). These deviations were attributed to 'multi-domain behavior' (Hill & Shaw, 1999; Biggin et al., 2007), differences between natural and laboratory cooling rates (Hill & Shaw, 1999; Biggin et al., 2007), 'magnetic refraction' (Rolph & Shaw, 1986; Rolph et al., 1987) or 'transdomain processes' occurring in paleointensity experiments (de Groot et al., 2013). Alternatively, the bias in paleomagnetic data might be explained by the presence of local magnetic anomalies, i.e. a local disturbance of the magnetic field induced by the magnetic field from underlying lava flows.

Mt. Etna is characterized by irregular topography; virtually all lava flows are classified as aa' type and the terrain is rough with rubble up to boulder size on the surface (Calvari & Pinkerton, 1998; Kilburn & Lopes, 1988). Mt. Etna lavas are also strongly magnetized. The remanent magnetization of specimens at Mt. Etna sometimes exceeds 20A/m and there is a large deviation between sun and magnetic compass readings (Speranza et al., 2006). The earliest volcanic products of Mt. Etna are dated around 500 ka ago (Branca et al., 2011), therefore all lava flows must be of normal polarity. Previously, measurements of the ambient geomagnetic field above the surface of lava flows were performed on La Palma and Tenerife (Valet & Soler, 1999), Hawaii (Baag et al., 1995) and on Mt. Etna (Tanguy & Le Goff, 2004). Valet and Soler (1999) and Baag et al. (1995) found significant deviations from the IGRF value and attributed these to local magnetic anomalies arising from the underlying terrain. In contrast, Tanguy and Le Goff (2004) concluded from averaging over 124 measurements above 12 sites on Mt. Etna that their results are close to the actual geomagnetic field and there is no global effect on either direction or intensity. The averages per site, however, show small deviations from the main field ($\pm 3\%$ in intensity and $\pm 1.5°$ in direction (Tanguy & Le Goff, 2004)). Furthermore, they only took 10 measurements per site and avoided obvious terrain features during measuring which may have smoothed their results.

Here we test whether the strongly magnetized terrain of Mt. Etna influences the ambient magnetic field directly above it. First, we compile an overview of paleomagnetic literature data to characterize a potential bias in the data, while also paying attention to which sampling strategy is used. Second, we add new paleomagnetic directional data from 12 sites sampled from 7 different historical flows. Third, we measure the magnetic field above 4 recent lava flows of Mt. Etna, 3 of which were also sampled for paleomagnetic measurements. Combining these datasets allows us to characterize the expression of local magnetic anomalies in paleomagnetic measurements, quantify the impact on paleomagnetic statistics, and provide recommendations for paleomagnetic sampling strategies in volcanic terrain.

2 Paleomagnetic data

2.1 Data from previous studies

To characterize a possible bias in paleomagnetic data from Mt. Etna, we compiled an overview of all paleomagnetic results reported by previous studies of lava flows younger than 1850CE. The directional dataset (Supp. Table 1) consists of the declination, inclination and corresponding precision parameter (k) and α_{95} of 14 flows, which were deposited between 1853 and 1983. The other dataset (Supp. Table 2) consists of the paleointensities of 20 flows between 1853 and 2002, including their standard deviation and paleointensity method used.

How samples are generally obtained in the field, i.e. the sampling strategy, differs between studies. Studies aiming to produce paleodirections often take samples spread out over a flow, and measurements are deemed reliable when there is a low scatter, a small α_{95}, and/or a high k (Fisher, 1953). For paleointensity studies samples are sometimes taken closer together to ensure homogeneity between the samples, and results are found reliable when the standard deviation of the paleointensity results is low. These sampling strategies are, however, not universally defined and not all studies report their sampling strategy in detail. Previous studies on Mt. Etna that do report their sampling strategies are: Tanguy et al. (1985, 1999, 2003), who use the 'big sample method', taking samples spread out over a larger area. In contrast, Rolph (1997), Calvo et al. (2002) and Biggin et al. (2007) take their samples from top to bottom at one location of one single flow.

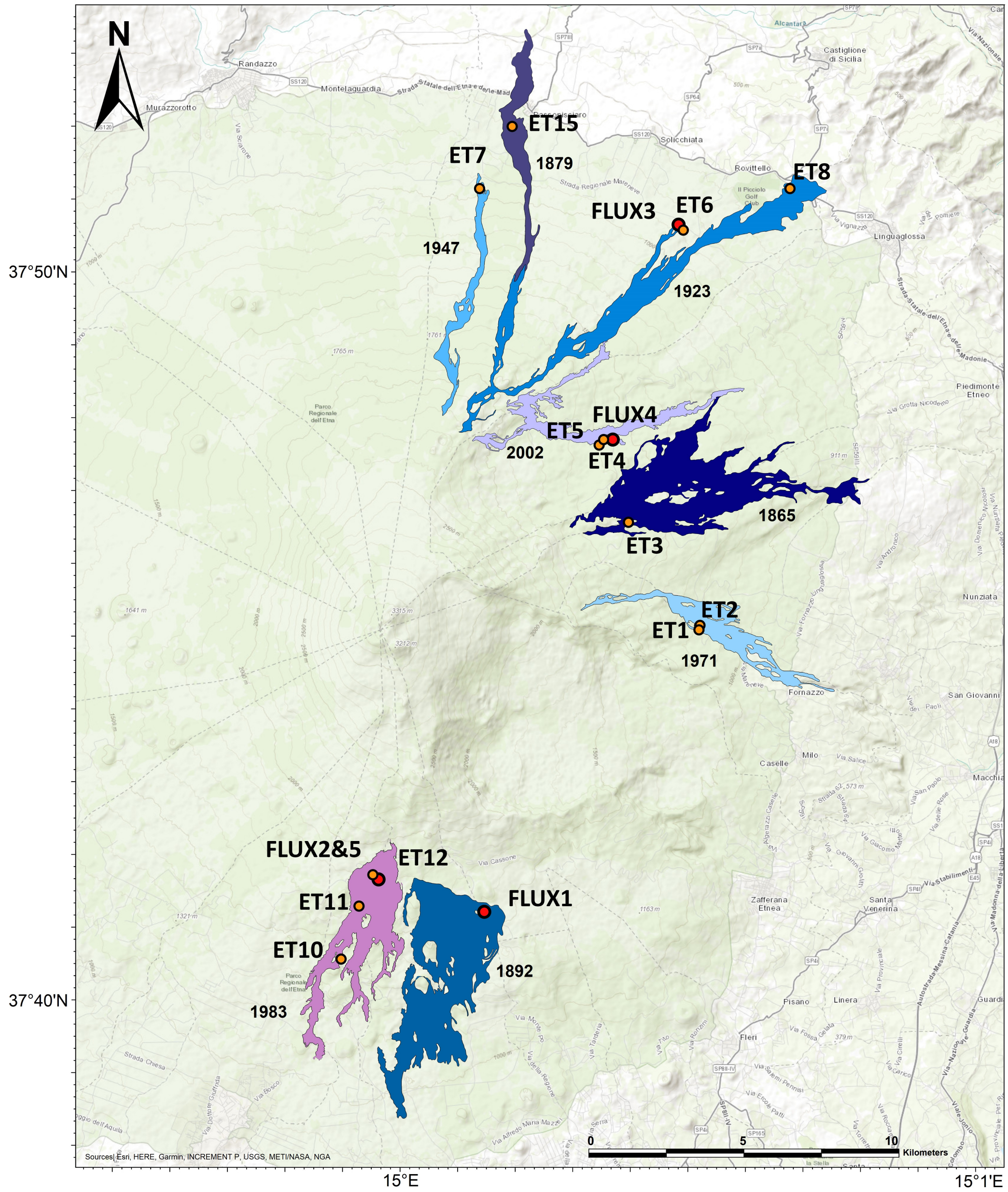

Figure 1: Sampling locations on Mt. Etna, Sicily, Italy. ET sites are where paleomagnetic samples were taken and FLUX are AnomalyMapper measurement sites. Outlines of lava flows from Branca et al. (2011).

Intensity results from Calvo et al. (2002) come from three different sites of the 1928 flow. Lastly, de Groot et al. (2013) used closely spaced drill cores, 8-12 samples taken less than 1m of each other to ensure sampling homogeneity.

2.2 Directional data from recent flows

To complement the existing paleomagnetic data set, we sampled twelve new sites (Fig. 1, named ET) from seven historical flows with ages between 1865 and 2002 during a fieldwork in April 2016. Flow 1923, 1971 and 2002 were sampled twice at different locations and flow 1983 was sampled at three different locations. Some sites were sampled at the same location as in de Groot et al. (2013) and most samples were taken along road cuts. For each site, standard paleomagnetic cores (2.5cm in diameter, up to 10cm in length) were taken using a petrol powered drill. Cores were drilled several meters apart, at different heights in the flow, and differed in borehole orientations. To orientate the cores the use of a sun compass is preferred to avoid the influence from the surrounding magnetized rock. Unfortunately, the weather did not permit the use of a sun compass during the fieldwork. Instead the samples were oriented using a magnetic compass and readings were corrected for the current declination of the IGRF.

Between four to ten cores per site, depending on the amount of cores available, were selected for paleodirection experiments. Four samples per site were thermally demagnetized in 11 temperature steps: 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600°C and measured on a 2G cryogenic magnetometer. Some samples were magnetically so strong that they exceeded the measurement range of the magnetometer, and could not be interpreted. A further four to nine samples were subjected to alternating field demagnetization experiments. Because the samples were strongly magnetized they were sliced in half (A and B specimens). The A and B specimens should have the exact same result, differences between them can be attributed to measurement or sample orientation errors in the machine. The samples were demagnetized in a robotized 2G DC-SQUID magnetometer (Mullender et al., 2016) with stepwise increasing alternating fields of 2.5, 5, 7.5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 100, 150, 225 and 270mT. All demagnetization results were analyzed in paleomagnetism.org (Koymans et al., 2016). Afterwards, site mean directions were calculated using Fisher statistics (Fisher, 1953) and the outliers are identified in the VGP distribution with the fixed 45° cut-off (Koymans et al., 2016). All other samples were retained for calculating site means (Fig. 2; Table 1). The precision parameter k ranges from 23.3 to 207.8, resulting in α_{95} values between 3.2° and 7.3°. Our k-values are on average lower than those from previous studies, in existing data k-values as high as 1070 have been reported (e.g. Tanguy et al., 2003).

Some flows (1923, 1971, 1983 and 2002) were sampled at multiple sites. The directions of these sites were grouped together to calculate 'age means' (Table 1). The k-values for these age means are lower than the k-values for individual sites. As the number of samples increases for the age means, the α95-values are also lower than the α_{95}-values of the individual sites. The age means averages out the effect of sites with large deviations from the expected reference values. Therefore these age means might be considered better estimates of the paleomagnetic vector, although the data from some individual sites are closer to the expected field value.

2.3 Bias in paleomagnetic data

All above results are compared with their expected values according to the IGRF-13 model (Alken et al., 2021), or for flows prior to 1900CE with the gufm1 model (Jackson et al., 2000). The reference geomagnetic field is obtained for every lava flow at the corresponding sampling location and elevation. In older papers the GPS coordinates are not always given. In this case, the reference value was determined using a location from the same flow from another research paper, or the geological map of Branca et al. (2011).

Table 1: Sampling sites and directional results this study

Site	Year(CE)	Lat(N)	Long(E)	Elv(m)	c/n/N	Dec(°)	Inc(°)	k	α_{95}(°)
ET1	1971	37.752	15.087	1185	8/14/14	6.46	49.96	156.37	3.19
ET2	1971	37.753	15.087	1200	5/8/11	-4.16	48.26	212.62	3.81
ET3	1865	37.777	15.066	1606	10/18/20	-9.04	53.46	94.51	3.57
ET4	2002	37.796	15.062	1544	6/11/11	0.59	49.63	83.82	5.02
ET5	2002	37.795	15.057	1606	8/20/20	-2.54	48.07	23.69	6.85
ET6	1923	37.845	15.081	866	4/9/9	-17.52	51.54	55.58	6.97
ET7	1947	37.854	15.023	928	7/10/14	-3.45	48.11	101.59	4.82
ET8	1923	37.854	15.113	641	4/9/9	-8.79	43.27	51.22	7.26
ET10	1983	37.676	14.982	1423	8/20/20	0.64	53.49	86.84	3.52
ET11	1983	37.688	14.987	1671	7/12/12	-1.3	44.3	130.48	3.81
ET12	1983	37.695	14.991	1833	6/14/14	-7.15	47.11	118.88	3.66
ET15	1879	37.868	15.032	778	7/12/12	-10.28	47.59	115.65	4.05
1923_{mean}	1923				8/18/18	-12.81	47.49	46.01	5.15
1971_{mean}	1971				13/22/25	2.51	49.45	135.03	2.68
1983_{mean}	1983				21/46/46	-2.39	49.19	81.27	2.35
2002_{mean}	2002				14/31/31	-1.43	48.65	32.31	4.62
FLUX1	1892	37.687	15.019	1620					
FLUX2	1983	37.695	14.992	1830					
FLUX3	1923	37.845	15.081	880					
FLUX4	2002	37.796	15.062	1530					
FLUX5	1983	37.694	14.993	1825					

For each site the age of the flow, location and elevation (Elv) of sampling is given. The obtained directions per site are given by the parameters: (c/n/N) number of different cores / number of samples accepted / total amount of samples per site, the declination (Dec), inclination (Inc), precision parameter (k), 95 percent confidence interval α_{95}. Furthermore, the age means of four flows (1923,1971,1983 and 2002) are given. For the fluxgate measurement sites only the age of the flow above which was measured, the coordinates and elevation are given here.

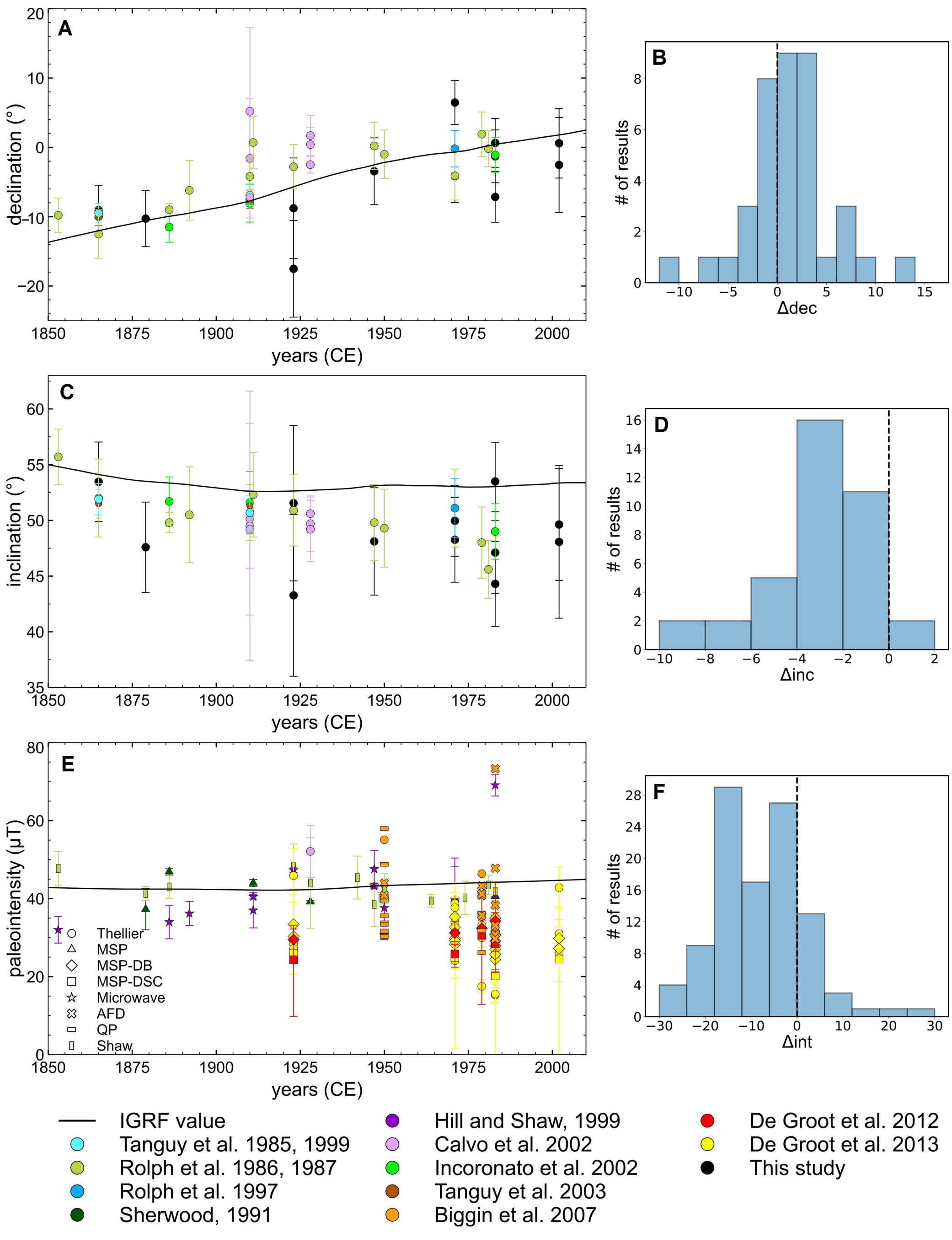

Figure 2: The (a) declination, (c) inclination and (e) intensity measurements of recent (>1850 CE) lava flows of Mt. Etna. In (a) and (b) the error bars are the corresponding α_{95} values and in (c) the error bars are the standard deviations. The histograms on the right-hand-side (b,d,f) show the difference (Δ) of the data points with respect to their expected field value.

Rolph and Shaw (1986) do not provide the exact GPS coordinates but a map with sampling locations, from this map the approximate GPS coordinates and elevations were utilized. In Fig. 2 the reference values are compared with the paleomagnetic data set of Mt. Etna, there is a systematic bias in the paleomagnetic data obtained. The declinations are generally in good agreement with the expected values: the median difference between the declination of a site and the expected value ($\widetilde{\Delta}$dec) is just 0.8° too high (Fig. 2b), and the Δdec is approximately Gaussian distributed around this value. In contrast to the declination, the inclination values are skewed towards lower than expected values. Only two data points yield (slightly) higher than expected values, while the median difference ($\widetilde{\Delta}$inc) is -2.9° (Fig. 2d). The majority of the intensity data is also lower than the reference value: the median difference ($\widetilde{\Delta}$int) is -8.8μT (Fig. 2f). There is no general correlation between the difference with respect to the reference value and the paleointensity method used.

3 Mapping magnetic anomalies

The ambient geomagnetic field, i.e. the magnetic field that would be recorded by a new lava flow, was measured using the AnomalyMapper - a three-axial fluxgate magnetometer (De Groot & De Groot, 2019) - at five sites above four lava flows in April 2018 (Fig. 1, Supp. Table 3). At each site, three 'paths' were measured perpendicular to ridges and gullies to obtain the largest topographic differences possible, with measurement locations being ∼1m apart; the three paths were 20 to 80m apart up/down the slope of the lava flow (Supp. Fig. 1). At FLUX1 to FLUX4 the paths were measured twice, with the magnetometer positioned at 100 and 180cm above the ground. The paths of FLUX5 were measured four times at 25, 75, 125 and 175cm above the ground (Supp. Fig. 3–16). In total, we measured the ambient geomagnetic field above the lava flows of Mt. Etna 1,334 times. The exact topography was obtained from the GPS sensor mounted on the magnetometer.

The AnomalyMapper uses a scope to point the magnetometer towards a reference point with a known (GPS) location (De Groot & De Groot, 2019). Due to the irregular terrain it was not always possible to see the reference point, most often in topographic lows, therefore the declination record is discontinuous for some paths. This did not affect the inclination data, as this is only dependent on the leveling of the magnetometer which is done using a tilt sensor, or the intensity data, that is the length of the total vector measured irrespective of its orientation.

3.1 Local magnetic anomalies

For all paths we observe major variations in declination, inclination and intensity above the lava flows. The reference field according to the IGRF-model in April 2018 was calculated for each site at the corresponding GPS coordinates and average elevation (Table 1). Here we use the results of path 2 of site FLUX3 at 100cm height and path 1 of site FLUX5 at 125cm height as examples (Fig. 3). Of FLUX3, the variation in declination is -6.5 to 5.4°; with a median difference of -3.2° with respect to the expected IGRF-value for measurements done at 100cm above ground. The inclination is on average closer to the IGRF-value, with a median difference of -1.5°, and varies between 46.8 and 58.7°. The intensity varies between 42.1 and 47.9μT, with a median offset of -0.9μT. FLUX5 was measured with most detail, and has for the measurements done at 125cm above the ground similar large fluctuations as FLUX3 has at 100cm. Declination varies between -4.2 and 8.7°, with a median difference of -0.9°. Inclination measurements range from 49.7 to 54.2° and the median difference is -1.3°. Finally, the intensity varies between 40.9 and 47.2μT with a median offset from the IGRF-value of -1.5μT. The data for all paths and sites generally show similar behavior (Supp. Fig. 3–16; Supp. Table 4–7). The median deviations with respect to the expected IGRF values for all paths at 100cm (or in

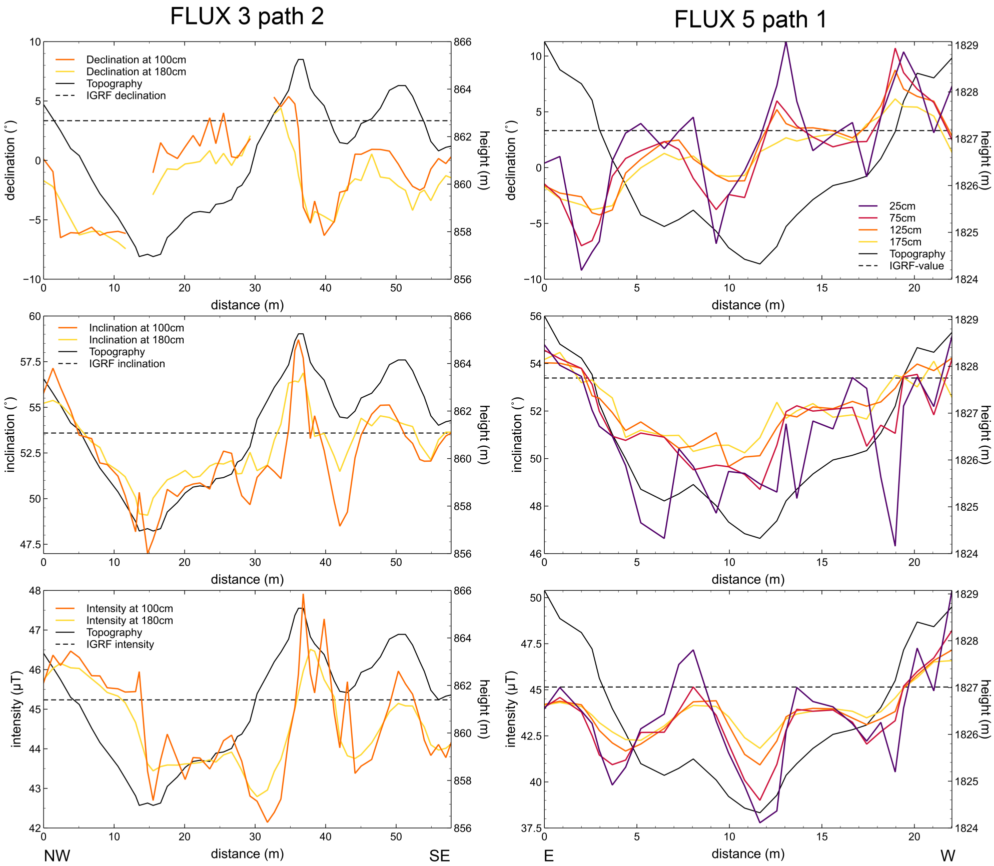

Figure 3: Fluxgate measurements of site FLUX3, path 2 (left) and FLUX5, path 1 (right). The variation in declination (top) does not show a clear correlation with topography (black line). The variations in inclination (middle) and intensity (bottom) correlate with the topography variations. Measurements closest to the ground surface (100cm for FLUX3 and 25cm for FLUX5) have the largest variation.

the case of FLUX5 at 125cm) above ground in the dataset ranges from -5.9 to -0.9° for $\widetilde{\Delta}$dec; -2.2 to 1.1° for $\widetilde{\Delta}$inc; and -2.2 to 0.1μT for $\widetilde{\Delta}$int.

3.2 Variations with height above surface

The deviations from the expected IGRF-values are largest close to the surface and become less pronounced higher above the flow (Supp. Fig. 2). This is most prominent in the inclination and intensity data, and less in the declination data. For all three the standard deviation decreases when measurement height above the flow increases (Supp. Fig. 2). This is also reflected in the $\widetilde{\Delta}$ range of values. For path 2 of site FLUX3 the range of declination values is -9.8 to 2° at 100cm above ground and -10.7 to 1° at 180cm, inclination values are -6.6 to 5.1° at 100cm above ground and -4.5 to 3.3° at 180cm, and for the intensity the variation is -3.1 to 2.7 μT at 100cm and only -2.4 to 1.3 at 180cm (Fig. 3). Site FLUX5 was measured at four different heights above the surface, with the lowest being at 25cm above ground and the highest at 175cm. The largest spikes in the measurement data are at 25cm height, the level closest to the lava flow (Fig. 3). For path

1 of FLUX5, the $\widetilde{\Delta}$dec range decreases from -12.5 to 8.0° at 25cm to -7.1 to 2.8° at 175cm. For the inclination the range at 25cm above the flow is -7.1 to 1.8° and only -3.1 to 1.1° at 175cm. The intensities vary from -7.4 to 5.2 at 25cm, and from 3.3 to 1.4μT at 175cm above the flow. As the intensity of the magnetic field decays with the power of three as function of distance to its source, the observed gradients as function of height above the flow imply that the source of the local magnetic anomalies must be close to the surface. This means that the magnetic signal of the flow(s) closest to the surface have the most impact on the ambient magnetic field above the flow.

3.3 Correlation with topography

Beyond the influence of the height above the flow, both the inclination and intensity variations seem to correlate with changes in topography. All paths are characterized by an irregular topography with at least one distinct gully (Supp. Fig. 3–16). Path 2 of site FLUX3 is a good example of such a distinct gully which is approximately 35m wide and 8m deep (Fig. 3). The gully in FLUX5 path 1 is around 20m wide and 4m deep. Both the inclination and intensity are higher above ridges and lower in gullies. For path 2 of FLUX3 the differences compared to the IGRF are +4.1° in inclination and +2.7μT in intensity with respect to the IGRF-value at 100cm above the highest peak in the profile. At 100cm above the lowest point, i.e. in the gully, the inclination is -5.6° and the intensity -2.4μT with respect to the IGRF-value. For FLUX5 path 2 most measurements are done in the gully, there is not a clear ridge in the profile but the peaks are located at the edges. At 125cm height above the highest peak the difference with the IGRF-value is +0.6° in inclination and -0.9μT for intensity. Above the lowest peak they are -3.3° and -4.2μT, respectively. To statistically assess the correlation between the fluxgate measurements and the topography, the Pearson's correlation coefficient and its corresponding p-value were calculated for each path and at each height. A Pearson's coefficient of +1 is a positive correlation, 0 is no correlation and with -1 there is a negative correlation. In terms of our fluxgate measurements, for a positive correlation the measurement value increases with increasing topography. Supp. Table 8 includes all Pearson correlation coefficients for each site and path. In Fig. 4 the correlation coefficients are grouped for FLUX1-4 (Fig. 4A) based on measurement height above the surface of the lava flows (100 and 180cm). Because FLUX5 was measured at four different levels we consider that site independently in Fig. 4B. The median correlation coefficient of declination is around 0 for FLUX1-4, which statistically suggests there is no trend between the declination and the topography. Inclination and intensity have a medium to strong positive correlation, these appear to have a positive trend with topography. Finally, for some paths the intensity signal seems to be slightly offset with respect to the topography, as illustrated at distance 30 to 35m in path 2 of site FLUX3 (Fig. 3). This offset, however, is small, not always present and does not correlate with the orientation of the gully or with the orientation with respect to the summit of Mt. Etna.

4 Discussion

4.1 Systematic bias due to local magnetic anomalies

Paleomagnetic data produced by this and previous studies were compared with the reference value predicted by the IGRF-13 or gufm1 model, both are estimations of the Earth's magnetic field at that time and might not be fully accurate. We, however, expect minor errors in the prediction of these models. Measured values at three different Italian magnetic observatories show a good correlation with the IGRF-model during the period of 1960-2020 (Di Mauro et al., 2021). This confirms that we can reliably compare our paleomagnetic measurements from the historical lava flows with the predicted reference value, at least the flows for after 1960. Prior to 1960 errors might slightly increase but we assume those to be negligible. The paleomagnetic data set of Mt. Etna shows

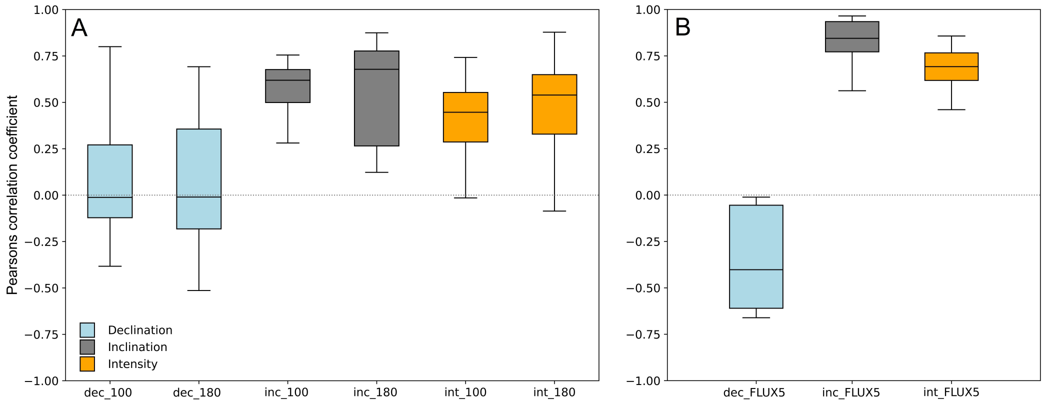

Figure 4: Box-plots for Pearson's correlation coefficient of A) FLUX1 to 4 and B) FLUX5. For each path a Pearson's correlation coefficient was calculated between the topography and the declination, inclination or intensity. For FLUX1-4 the coefficients are grouped for inclination, declination and intensity at 100 or 180cm height. For FLUX5 all four different measurement levels (25, 75, 125 and 175cm) are together. See Supp. Table 8 for the individual correlation coefficients.

a systematic bias in both the inclination and intensity. This bias is also present in our direct measurements of the magnetic field, and the median difference in inclination (-2.9°) is very close to the inclination shallowing that Pavón-Carrasco et al. (2014) reported for paleomagnetic data from volcanic products on the Northern Hemisphere for the past 400 years. Both the inclinations and intensities vary as function of topography: they are even lower in the gullies, where we expect the largest volume of a new flow to be deposited. This may explain the overall bias in paleomagnetic data from Mt. Etna.

The declinations of the paleomagnetic data show variation around the expected IGRF-values (Fig. 2a,b), but there is no systematic offset. The median declinations in our direct measurements, however, are up to 6.5° lower than the expected IGRF-values. Due to the design of the AnomalyMapper, the declination is prone to errors and potentially a bias (De Groot & De Groot, 2019). It relies on aiming the AnomalyMapper to a fixed reference point using a scope, while the inclination is determined using a tilt-sensor, and the intensity is independent of the orientation of the device. If the scope is slightly offset in its mount this would lead to a systematic bias in the declinations and limit their interpretation to describing relative variations. The requirement of having a line of sight to a reference point also sometimes prevents determining a declination. Especially in deeper gullies the reference point is sometimes not visible. If the bias in declinations would be strongly positive deep in the gullies, a lack of declination measurements there may also explain the bias towards negative values for the median declinations. For the sites that do have continuous declination data in the gullies, such a trend may be suggested (e.g. site FLUX5 paths 2 and 3 which have Pearson correlation coefficients at 125cm of -0.4 and -0.6, respectively), but it is not present for all sites, and it is certainly not strong enough to explain the deviations in median declinations fully.

4.2 The impact on paleomagnetic statistics

If a hypothetical new flow on Mt. Etna would record the ambient magnetic field that we measured directly, we can simulate what the effect of local magnetic anomalies would be on a paleomagnetic study. More than 20% of the data points in sites FLUX1 and FLUX4 lack declinations, we therefore exclude these sites from this simulation. For

Table 2: Random sampling of fluxgate measurements

	All measurements						Gully +3m					
Site	N	$\widetilde{\Delta}$dec	$\widetilde{\Delta}$inc	$\widetilde{\Delta}$int	k med	σ	N	$\widetilde{\Delta}$dec	$\widetilde{\Delta}$inc	$\widetilde{\Delta}$int	k med	σ
FLUX2	187	-2.76	-0.81	-0.82	1226	1.26	134	-2.67	-0.96	-1.20	1522	1.12
FLUX3	257	-4.87	-0.38	-0.79	953	1.12	96	-5.12	-1.78	-1.25	1372	1.05
FLUX5	344	-2.27	-0.94	-1.49	461	2.24	188	-0.84	-2.23	-2.69	667	1.89
Average	788	-3.30	-0.71	-1.03	880	1.54	418	-2.88	-1.66	-1.72	1187	1.35

	Gully +2m						Gully +1m					
Site	N	$\widetilde{\Delta}$dec	$\widetilde{\Delta}$inc	$\widetilde{\Delta}$int	k med	σ	N	$\widetilde{\Delta}$dec	$\widetilde{\Delta}$inc	$\widetilde{\Delta}$int	k med	σ
FLUX2	71	-3.03	-1.16	-1.33	1504	1.19	53	-2.82	-1.41	-1.61	1457	1.21
FLUX3	60	-4.89	-2.18	-1.16	1857	1.02	28	-5.15	-2.50	-1.25	2026	1.16
FLUX5	148	-0.55	-2.64	-2.94	908	1.85	88	0.04	-2.94	-3.41	980	1.87
Average	279	-2.82	-1.99	-1.81	1423	1.35	169	-2.64	-2.28	-2.09	1488	1.41

Simulated paleomagnetic data based on AnomalyMapper measurements. N is the amount of measurements available to take random samples from, $\widetilde{\Delta}$dec,inc,int is the difference of the median with the IGRF-value, k med is the median of the precision parameter and σ is the standard deviation of the intensity measurements.

the other sites we randomly drew 10 AnomalyMapper measurements for each site and calculate what the resulting declination, inclination, intensity, k and intensity error (σ) would be. This was repeated a 1000 times and we report the median values for each site (Table 2). Furthermore, we expect the largest volume of a new flow to be deposited in the gullies of the underlying flow, we therefore repeated this analysis by selecting only AnomalyMapper measurements from the gullies. We defined a gully as the lowest point in the topography, the local minimum, and selected the measurements around it up to +1, +2m or +3m height.

The k-value is an expression of how well measurements from individual samples agree. The median k-values are 1226 for FLUX2, 953 for FLUX3, and 461 for FLUX5, when all measurements per site are considered. If a new flow would be deposited deep in the gullies (+1m from the lowest point), the k-values increase to 1457, 2026, and 980, respectively. This illustrates that high k-values in rough volcanic terrain may indicate that a local magnetic anomaly is not averaged out sufficiently. Moreover, it should be emphasized that the AnomalyMapper measurements do not suffer from orientation errors that occur during paleomagnetic sampling and measurements that would certainly lower our simulated k's. k's associated with real paleomagnetic data are therefore expected to be (much) lower than the theoretical upper limits from our simulation. Hence, paleomagnetic studies in rough volcanic terrain should be treated with caution when their results have high k-values, e.g. >1000.

The standard deviation of paleointensity measurements, σ, is a measure of how well paleointensity results from different samples agree. For this parameter we see the same trend as for the k-value, but the σ's reported here are negligible compared to the uncertainties arising from paleointensity experiments (e.g. Biggin et al., 2007; de Groot et al., 2012, 2013).

4.3 Optimal sampling strategies

Our observations have consequences for paleomagnetic sampling strategies. To suppress the influence of local magnetic anomalies arising from the underlying terrain, it is important to take samples for both paleodirectional and paleointensity studies far apart on the outcrop. If possible, take samples at different distances from top and/or bottom of a flow. A sun compass is preferable for sample orientation to avoid the influence of local magnetic anomalies on drill core orientations. Other techniques to suppress this influence are backsighting using distinct landmarks (Tauxe, 2010) or a differential GPS technique (Lawrence et al., 2009), originally developed for high-latitude sampling sites but also useful when due to weather conditions a sun compass cannot be used.

Sometimes, however, non of these orientation methods are available, and one has to revert to using a magnetic compass for orienting the samples. This was also the case for the paleomagnetic data in this study, as weather conditions only allowed using a magnetic compass. This is not the ideal scenario because Speranza et al. (2006) already demonstrated that there might be significant differences between Sun and magnetic compass declination readings on Mt Etna. The use of a magnetic compass would only influence the declination of the sample orientation. When determining a magnetic direction for a site/flow the results of several samples are averaged. We do not find a systematic trend between the magnetic declination and topography (Fig. 4); and the declination of paleomagnetic data from Mt. Etna does not show a systematic deviation from their expected values (Fig. 2B). This implies that the error made by using a magnetic compass can be reduced when samples are taken well spread out over the flow, with different bore hole orientations, and on different sides of an outcrop.

If a paleomagnetic protocol prescribes the use of sister specimens it is necessary to take multiple groups of samples to average out local magnetic anomalies. Then, it is important to avoid using samples from the same group to determine the paleodirection or paleointensity of the entire cooling unit. Finally, if a certain cooling unit is accessible at different locations, e.g. on both sides of a lava flow and/or higher or lower on a mountain, taking multiple sites from a cooling unit and calculating paleomagnetic age means greatly increases the chance of being closer to the 'true' paleomagnetic vector at the time of cooling. It is worth noting that taking samples well spread out over the flow and from different parts also averages out possible variations in the properties of the magnetic minerals present in the sample (e.g. Thellier, 1977; de Groot et al., 2014). In practice it is of course often difficult to use an optimal sampling strategy because of limitations in availability and/or accessibility of outcrops. This emphasizes the need to report the sampling strategy in high detail in forthcoming publications and as metadata in data repositories.

5 Conclusion

Paleomagnetic data from recent flows of Mt. Etna often yield lower inclinations and intensities than expected from the IGRF. This bias in paleomagnetic data can be attributed to local magnetic anomalies due to the underlying irregular terrain of Mt. Etna. Direct measurements above the strongly magnetized flows show that inclination and intensity vary as function of topography. The values are higher above ridges and lower above gullies. The largest deviations are found closest to the surface, which emphasizes the influence the underlying terrain has on the ambient magnetic field that would be recorded by a new flow. Although sampling at a single location will result in a low scatter in paleomagnetic studies, there is a high chance that a local magnetic anomaly was sampled. A high k-value therefore not necessarily reflects accurate paleomagnetic data. This emphasizes the need to take samples spread out over a larger area that will often lead to lower k-values, and always report the sampling strategies used in detail.

Acknowledgments
Esmee Waardenburg, Meeke van Ede, Maartje van de Biggelaar, Lynn Vogel, and Wout Krijgsman are kindly acknowledged for their help in the field and with (the processing of) the measurements. Bertwin de Groot developed the AnomalyMapper. This project has received funding from the Dutch Research Council (NWO) under its Talent program (VENI grant 863.15.003 and VIDI grant VI.Vidi.192.047 to LVdG).

Open Research
All paleomagnetic data measured in this study can be found in MagIC (DOI: 10.7288/V4/MAGIC/19784, private contribution link: https://earthref.org/MagIC/19784/26089666-41d9-4832-98b4-7cba623a4e48). The AnomalyMapper data is in Yoda (DOI: 10.24416/UU01-B6JJC0).

References

Alken, P., Thébault, E., Beggan, C. D., Amit, H., Aubert, J., Baerenzung, J., ... Zhou, B. (2021). International geomagnetic reference field: the thirteenth generation. *Earth, Planets and Space*, *73*(1), 49. Retrieved from `https://doi.org/10.1186/s40623-020-01288-x` doi: 10.1186/s40623-020-01288-x

Baag, C., Helsley, C., Xu, S.-z., & Lienert, B. (1995). Deflection of paleomagnetic directions due to magnetization of the underlying terrain. *Journal of Geophysical Research: Solid Earth*, *100*(B6), 10013–10027.

Biggin, A. J., Perrin, M., & Dekkers, M. J. (2007). A reliable absolute palaeointensity determination obtained from a non-ideal recorder. *Earth and Planetary Science Letters*, *257*(3-4), 545–563.

Branca, S., Coltelli, M., Groppelli, G., & Lentini, F. (2011). Geological map of etna volcano, 1: 50,000 scale. *Italian Journal of Geosciences*, *130*(3), 265–291.

Calvari, S., & Pinkerton, H. (1998). Formation of lava tubes and extensive flow field during the 1991–1993 eruption of mount etna. *Journal of Geophysical Research: Solid Earth*, *103*(B11), 27291–27301.

Calvo, M., Prévot, M., Perrin, M., & Riisager, J. (2002). Investigating the reasons for the failure of palaeointensity experiments: a study on historical lava flows from mt. etna (italy). *Geophysical journal international*, *149*(1), 44–63.

Cromwell, G., Tauxe, L., Staudigel, H., & Ron, H. (2015). Paleointensity estimates from historic and modern hawaiian lava flows using glassy basalt as a primary source material. *Physics of the Earth and Planetary Interiors*, *241*, 44-56. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0031920114002556` doi: https://doi.org/10.1016/j.pepi.2014.12.007

De Groot, B. M., & De Groot, L. V. (2019). A low-cost device for measuring local magnetic anomalies in volcanic terrain. *Geoscientific Instrumentation, Methods and Data Systems*, *8*(2), 217–225.

de Groot, L. V., Dekkers, M. J., & Mullender, T. A. (2012). Exploring the potential of acquisition curves of the anhysteretic remanent magnetization as a tool to detect subtle magnetic alteration induced by heating. *Physics of the Earth and Planetary Interiors*, *194*, 71–84.

de Groot, L. V., Dekkers, M. J., Visscher, M., & ter Maat, G. W. (2014). Magnetic properties and paleointensities as function of depth in a hawaiian lava flow. *Geochemistry, Geophysics, Geosystems*, *15*(4), 1096–1112.

de Groot, L. V., Mullender, T. A., & 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.

Di Mauro, D., Regi, M., Lepidi, S., Del Corpo, A., Dominici, G., Bagiacchi, P., ... Cafarella, L. (2021). Geomagnetic activity at lampedusa island: characterization and comparison with the other italian observatories, also in response to space weather events. *Remote Sensing*, *13*(16), 3111.

Fisher, R. A. (1953). Dispersion on a sphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *217*(1130), 295–305.

Hill, M. J., & Shaw, J. (1999). Palaeointensity results for historic lavas from mt etna using microwave demagnetization/remagnetization in a modified thellier-type experiment. *Geophysical Journal International*, *139*(2), 583–590.

Incoronato, A., Angelino, A., Romano, R., Romano, A., Sauna, R., Vanacore, G., & Vecchione, C. (2002). Retrieving geomagnetic secular variations from lava flows: Evidence from mounts arso, etna and vesuvius (southern italy). *Geophysical Journal International*, *149*(3), 724–730.

Jackson, A., Jonkers, A. R., & Walker, M. R. (2000). Four centuries of geomagnetic secular variation from historical records. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *358*(1768), 957–990.

Kilburn, C. R., & Lopes, R. M. (1988). The growth of aa lava flow fields on mount etna, sicily. *Journal of Geophysical Research: Solid Earth*, *93*(B12), 14759–14772.

Koymans, M. R., Langereis, C. G., Pastor-Galán, D., & van Hinsbergen, D. J. (2016). *Paleomagnetism. org: An online multi-platform open source environment for paleomagnetic data analysis.* Elsevier.

Lanza, R., Meloni, A., & Tema, E. (2005). Historical measurements of the earth's magnetic field compared with remanence directions from lava flows in italy over the last four centuries. *Physics of the Earth and Planetary Interiors*, *148*(1), 97–107.

Lawrence, K., Tauxe, L., Staudigel, H., Constable, C., Koppers, A., McIntosh, W., & Johnson, C. (2009). Paleomagnetic field properties at high southern latitude. *Geochemistry, Geophysics, Geosystems*, *10*(1).

Mullender, T. A., Frederichs, T., Hilgenfeldt, C., de Groot, L. V., Fabian, K., & Dekkers, M. J. (2016). Automated paleomagnetic and rock magnetic data acquisition with an in-line horizontal "2 g" system. *Geochemistry, Geophysics, Geosystems*, *17*(9), 3546–3559.

Pavón-Carrasco, F. J., Tema, E., Osete, M. L., & Lanza, R. (2014). Statistical analysis of palaeomagnetic data from the last four centuries: Evidence of systematic inclination shallowing in lava flow records. *Pure and Applied Geophysics*, *173*(3), 839–848.

Rolph, T. (1997). An investigation of the magnetic variation within two recent lava flows. *Geophysical Journal International*, *130*(1), 125–136.

Rolph, T., & Shaw, J. (1986). Variations of the geomagnetic field in sicily. *Journal of geomagnetism and geoelectricity*, *38*(12), 1269–1277.

Rolph, T., Shaw, J., & Guest, J. (1987). Geomagnetic field variations as a dating tool: application to sicilian lavas. *Journal of archaeological science*, *14*(2), 215–225.

Sherwood, G. J. (1991). Evaluation of a multi-specimen approach to palaeointensity determination. *Journal of geomagnetism and geoelectricity*, *43*(5), 341–349.

Speranza, F., Branca, S., Coltelli, M., D'Ajello Caracciolo, F., & Vigliotti, L. (2006). How accurate is "paleomagnetic dating"? new evidence from historical lavas from mount etna. *Journal of Geophysical Research: Solid Earth*, *111*(B12).

Tanguy, J., Bucur, I., & Thompson, J. (1985). Geomagnetic secular variation in sicily and revised ages of historic lavas from mount etna. *Nature*, *318*(6045), 453–455.

Tanguy, J., & Le Goff, M. (2004). Distortion of the geomagnetic field in volcanic terrains: an experimental study of the mount etna stratovolcano. *Physics of the Earth and Planetary Interiors*, *141*(1), 59–70.

Tanguy, J., Le Goff, M., Chillemi, V., Paiotti, A., Principe, C., La Delfa, S., & Patane, G. (1999). Secular variation of the geomagnetic field direction recorded in lavas from etna and vesuvius during the last two millennia. *COMPTES RENDUS DE L ACADEMIE DES SCIENCES SERIE II FASCICULE A-SCIENCES DE LA TERRE ET DES PLANETES*, *329*(8), 557–564.

Tanguy, J., Le Goff, M., Principe, C., Arrighi, S., Chillemi, V., Paiotti, A., ... Patanè, G. (2003). Archeomagnetic dating of mediterranean volcanics of the last 2100 years: validity and limits. *Earth and Planetary Science Letters*, *211*(1-2), 111–124.

Tauxe, L. (2010). *Essentials of paleomagnetism.* Univ of California Press.

Thellier, E. (1977). Early research on the intensity of the ancient geomagnetic field. *Physics of the Earth and Planetary Interiors*, *13*(4), 241–244.

Valet, J.-P., & Soler, V. (1999). Magnetic anomalies of lava fields in the canary islands. possible consequences for paleomagnetic records. *Physics of the earth and planetary interiors*, *115*(2), 109–118.