Paleointensity Estimates from the Pleistocene of Northern Israel: Implications for hemispheric asymmetry in the time-averaged field

L. Tauxe[1], H. Asefaw[1], N. Behar[2], A.A.P. Koppers[3], R. Shaar[2]

[1]Geosciences Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA
[2]The Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel
[3]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA

Key Points:

- We present 26 ^{40}Ar/^{39}Ar ages from volcanic rocks from Northern Israel (90 ka to 3.3 Ma)
- Twenty-two Pleistocene intensity estimates have a mean paleomagnetic dipole moment of 62.24± 30.6 ZAm^2
- The northern hemisphere had persistently higher fields than the southern during the Pleistocene

Corresponding author: Lisa Tauxe, `ltauxe@ucsd.edu`

Abstract

Twenty-two sites, subjected to an IZZI-modified Thellier-Thellier experiment and strict selection criteria, recover a paleomagnetic axial dipole moment (PADM) of 62.24± 30.6 ZAm^2 in Northern Israel over the Pleistocene (0.012 - 2.58 Ma). Pleistocene data from comparable studies from Antarctica, Iceland, and Hawaii, re-analyzed using the same criteria and age range, show that the Northern Israeli data are on average slightly higher than those from Iceland (PADM = 53.8 ± 23 ZAm^2, n = 51 sites) and even higher than the Antarctica average (PADM = 40.3 ± 17.3 ZAm^2, n = 42 sites). Also, the data from the Hawaiian drill core, HSDP2, spanning the last half million years (PADM = 76.7 ± 21.3 ZAm^2, n = 59 sites) are higher than those from Northern Israel. These results, when compared to Pleistocene results filtered from the PINT database (www.pintdb.org) suggest that data from the Northern hemisphere mid-latitudes are on average higher than those from the southern hemisphere and than those from latitudes higher than 60°N. The weaker intensities found at high (northern and southern) latitudes therefore, cannot be attributed to inadequate spatio-temporal sampling of a time-varying dipole moment or low quality data. The high fields in mid-latitude Northern hemisphere could result from long-lived non-axial dipole terms in the geomagnetic field with episodes of high field intensities occurring at different times in different longitudes. This hypothesis is supported by an asymmetry predicted from the Holocene, 100 kyr, and five million year time-averaged geomagnetic field models.

Plain Language Summary

According to the Geocentric Axial Dipole hypothesis, the geomagnetic field may be approximated by a dipole that is aligned with the spin axis and positioned in the center of Earth. Such a field would produce field strengths that vary with respect to latitude with high latitudes associated with high intensities, or, converted to equivalent 'virtual' dipole moments, would be essentially independent of latitude. It has long been suggested that high latitudes have had lower field strengths than predicted by such a model, when compared to data from mid-latitudes, but these claims have always been accompanied by caveats regarding differences in temporal coverage or methodological approaches. Here we present new data from Pleistocene aged rapidly cooled cinder cones and lava flow tops from Israel. We compare these data to other recent data sets obtained from rapidly cooled materials collected in Hawaii, Iceland and Antarctica. These confirm that virtual dipole moments from mid northern hemisphere latitudes are higher than those from high latitudes and from the southern hemisphere. Global compilations spanning the Pleistocene, when filtered for quality also shows this behavior as do time averaged field models. Therefore, field strengths over even millions of years can have persistent non-dipole field contributions.

1 Introduction

The geomagnetic field changes through time, a phenomenon known as secular variation, or paleosecular variation (PSV) when extended to the more ancient past. The spatial variability is evident in the present field (2022) as represented by predictions of field strength over the globe from the International Geomagnetic Reference Field (IGRF, Alken et al., 2021, see Figure 1a). While the present field is quite variable along lines of latitude, models of the time-averaged field are much smoother and when averaged over sufficient time, the geometry of the field can be represented by that generated by a magnetic dipole centered in the Earth and aligned along the spin axis (Hospers, 1955). This is basis of the 'geocentric axial dipole' (GAD) hypothesis that is fundamental to plate tectonic reconstructions that extend back to the Archean. Yet significant non-dipole contributions to the global field have long been known from directional data (e.g., Wilson, 1970) and more recently suspected from intensity data (e.g., Cromwell et al., 2013).

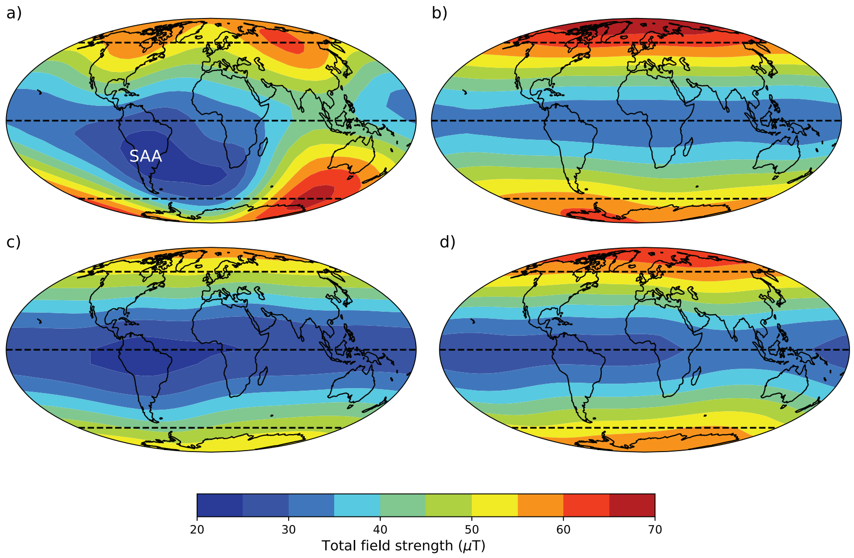

Figure 1. Intensity (in μT) of the geomagnetic field from global field models. a) International Geomagnetic Reference Field (IGRF) for the year 2022 (Alken et al., 2021). b) Average of the Holocene field from CALS10k.2 (Constable et al., 2016). c) Average field for the last 100 ka (Panovska et al., 2018). d) LN3 time averaged field model for the last 5 Ma (Cromwell et al., 2018).

Time-varying field models extend the IGRF like models back to 10 kyr ago (e.g., CALS10k.2 Constable et al., 2016) or even 100 kyr (GGF100k, Panovska et al., 2018). These, when averaged over their entire time span, produce ever smoother models (see Figure 1b-c). Numerous studies over the past decades recovered directions from lava flows over the past 10 million years. Cromwell et al. (2018) compiled these data sets and produced a five million year time averaged geomagnetic field model, LN3. The LN3 field model, although based on directional data alone, can also be used to predict field intensity variations over the Earth (Figure 1d). While the prominent low intensity bulge labeled 'SAA' for South Atlantic Anomaly in Figure 1a may not have persisted over long periods of time, it is interesting that the time averaged models all have an asymmetry between field strengths in the northern and southern hemispheres as suspected by Cromwell et al. (2013). Compare for example the 60°N latitude band with an average of some 65 μT with its southern hemisphere sister, whose average field is $\sim$55 μT. There are, therefore, hemispheric differences in predicted field strength that apparently persisted over millions of years.

To test the idea of persistent hemispheric asymmetry, we need high quality paleointensity data from around the globe. Although there are databases that compile published data (e.g., the PINT, and MagIC databases; Bono et al., 2022 and Tauxe et al., 2016 respectively), these contain data derived from very different sampling, laboratory and data analysis approaches and may not reflect the magnetic field strength in an unbiased way. In this study, we present new paleointensity data from the Pleistocene volcanic units in Northern Israel (32.9°-33.2°N, 35.5°-35.8°E) from rapidly cooled cinder cones and lava flow tops. We compare these new results with those re-interpreted from studies conducted in a similar fashion in Antarctica (Asefaw et al., 2021), Hawaii (Cai et al., 2017; Tauxe & Love, 2003) and Iceland (Cromwell et al., 2015b), and then to those filtered from the the PINT database of Bono et al. (2022), attempting to choose the most reliable results in a consistent fashion. In Section 2 we describe the geological setting for the present study. In Section 3 we lay our out sampling, and laboratory procedures. Results are presented in Section 4 and the implications are discussed in Section 5. Finally, we summarize our conclusions in Section 6.

2 Geological Setting

Our study area is a volcanic province in Northern Israel (Figure 2) located at the western edge of the extensive NW-SE trending Harrat ash Shaam volcanic field which developed during the late Cenozoic. The volcanic activity in the study area occurred in several phases beginning in the Miocene and continuing through the late Pleistocene. The most recent volcanic phase began about 5.3 Ma (Heimann et al., 1996) and continued until 0.1 Ma (Behar et al., 2019; Weinstein et al., 2020). The Plio-Pleistocene volcanism includes basaltic flows and cinder cones, with compositions ranging between alkali basalt, hawaiite, and basanite (Weinstein et al., 2006a; Weinstein, 2006b). The geological and geomorphological processes that shaped the existing landscape includes a progressive migration of the volcanic activity to the northeast and tectonic activity along the Dead Sea Transform (DST) plate boundary. The Golan Heights plateau, east of the DST, is a largely un-faulted area where we collected many samples. The topographic relief led to the development of canyons toward the valley that cut through the geological units and revealed excellent exposures of the entire Plio-Pleistocene volcanic sequences.

3 Methods

3.1 Sample Collection

Samples were collected from cinder cones and lava flows (Figure 2 and Table S1) during two field expeditions. On our first trip in 2015, we drilled oriented cores from 52 lava flows (the GH series of Behar et al., 2019) and took unoriented hand samples from ten cinder cones (GHI sites 01-10 in Figure 2). Behar et al. (2019) demagnetized specimens from the drill cores using alternating field and thermal demagnetization techniques and obtained paleodirections for characterizing the behavior of PSV over the Plio-Pleistocene from Israel. We performed paleointensity experiments on these drilled specimens, but, as is common with lava flows, the data failed our selection criteria. However, six of the ten cinder cones performed well and we therefore returned for a second field trip and targeted cinder cones and quickly cooled lava flow tops, as these seem to perform better in our paleointensity experiments (Cromwell et al., 2015a). In total, we collected 52 sites from the quickly cooled contexts in Northern Israel, spanning the Plio-Pleistocene. Forty-three were from the Golan Heights Plateau itself and nine were from the Eastern Galilee, west of the Dead Sea Transform.

3.2 ^{40}Ar-^{39}Ar Geochronology

Sites that were deemed promising for paleointensity results were selected for dating using the $^{40}Ar/^{39}Ar$ dating method. We sent a total of 29 samples to the Argon Geochronology lab at Oregon State University (OSU). There we conducted $^{40}Ar/^{39}Ar$ incremental heating experiments on groundmass samples. Samples ranging from 200 - 300 μm were prepared, and leached in acid with 1N and 6 N HCl and 1N and 3N HNO_3 in an ultrasonic bath (Koppers et al., 2000). The samples were then irradiated for six hours in a TRIGA CLICIT nuclear reactor at OSU. After irradiation, samples were scanned with a defocused, continuous CO_2 laser beam to incrementally heat the samples. The released argon gas fractions were then purified using ST101 and AP10 SAES getters for 3 - 6 minutes. All gas fractions released were analyzed on an ARGUS-VI multi-collector mass spectrometer.

The ages are interpreted as eruption ages including a consecutive set of incremental heating steps with ages falling within $1.96\sqrt{\sigma_1^2+\sigma_2^2}$. σ_1 (σ_2) is the standard deviation of the lowest (highest) age in the plateau. Plateaus were subjected to the follow-

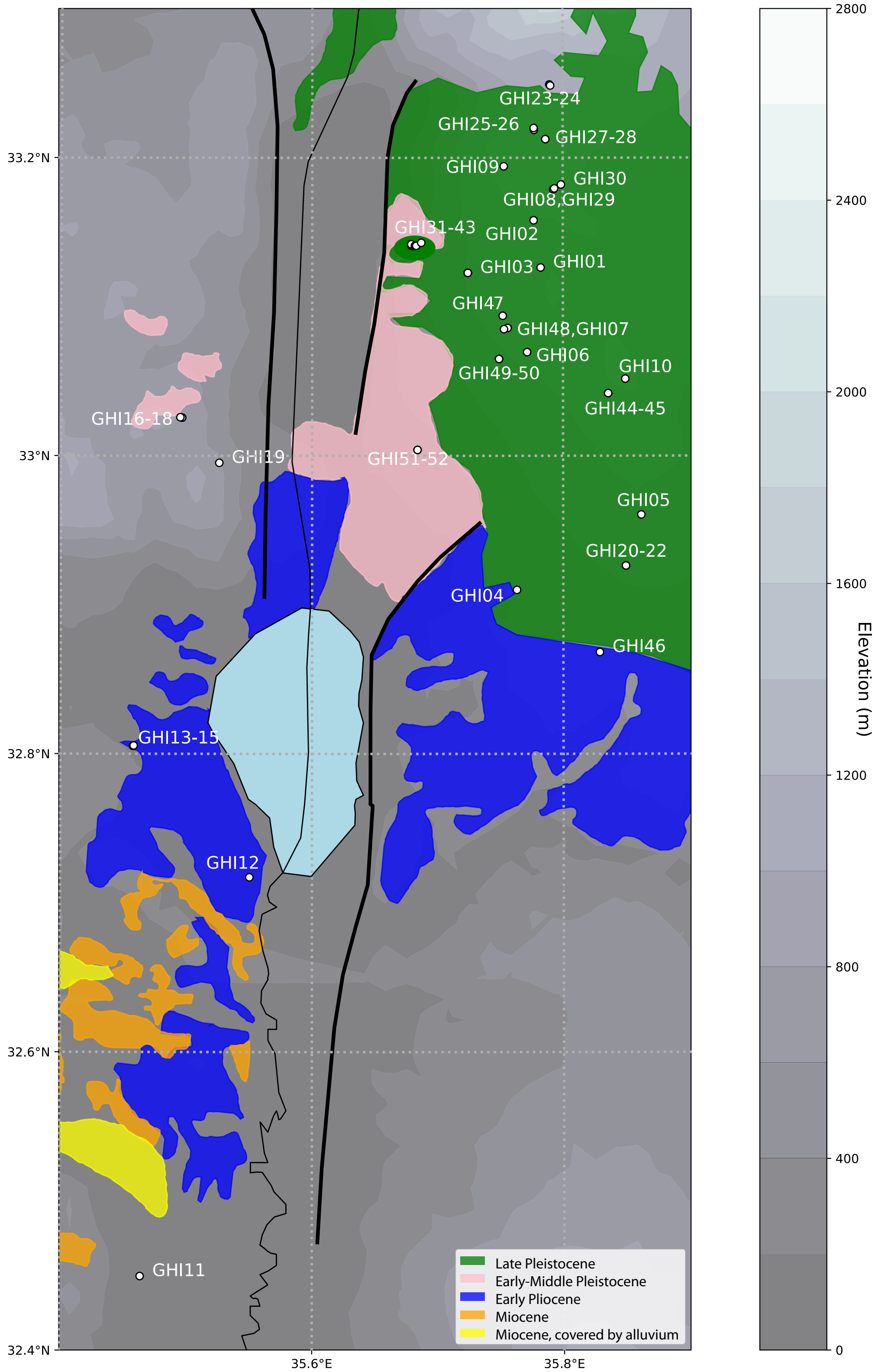

Figure 2. A map of the study region. White circles mark sites that were hand sampled for paleointensity. Volcanic units after Weinstein et al. (2006a) and Heimann et al. (1996) and this paper. Faults are shown as heavy black lines.

ing quality criteria: they must include at least three heating steps and at least 50% of the total ^{39}Ar released and they must be concordant at the 95% confidence level with the isochron and total fusion ages (Koppers et al., 2008).

In addition to the new ages presented here, we use additional age information from Weinstein et al. (2020) for Mt. Bar-On and Tel Sheivan (sites GHI02 and GHI03 respectively in this paper). For the former, we take the arithmetic mean of the two plateau ages and their uncertainties for an age estimate of 0.130 ± 0.012 Ma (2σ). Of the 29 samples analyzed, 26 resulted in robust plateau ages (Figure S1, Table 1).

3.3 Paleointensity experiment

Samples were gently crushed with a ceramic mortar. The fragments were then examined under a binocular microscope to select the finest grained and freshest material. We chose the finest grained material as it likely retains a primary thermal remanent magnetization (TRM) carried by mostly single-domain grains of magnetite as these conform to the assumptions of the Thellier-Thellier experiment (Thellier & Thellier, 1959). Individual specimens up to 0.5 gm were encased in glass microfiber filter paper and affixed inside a borosilicate glass vial with K_2SiO_3. Specimens were kept in the shielded room in the Paleomagnetic Laboratory at Scripps Institution of Oceanography while the experiments were underway.

The specimens were then subjected to the IZZI paleointensity experiment of Yu et al. (2004). A total of 498 specimens from the cinder cones or lava flow tops (GHI series) were subjected to the IZZI experiment in the Scripps Paleomagnetic Laboratory. In this experiment, specimens were heated in a step-wise fashion, cooling either in an applied laboratory field (I steps) or in zero field (Z steps) at each temperature. Temperature steps were at 100°C intervals between 0 and 300°C, 50°C intervals between 300 and 400°C, 25°C intervals between 400 and 575°C and then at 10°C intervals until at least 90% of the natural remanent magnetization (NRM) of each specimen was removed in the zero field steps. Zero-field cooling followed by in-field (ZI) or in-field cooling followed by zero field (IZ) alternate at every subsequant temperature step. In addition, we repeated an in-field step at a lower temperature after every IZ step to monitor for changes in the capacity of the specimens to acquire a partial thermal remanence (pTRM checks of Coe, 1967a).

The ratio of the natural remanence remaining compared to the pTRM gained over the experiment can be assumed to be quasi-linearly related to the strength of the field in which the specimen acquired its NRM (Néel, 1949). This ratio, when multiplied by the laboratory field B_{lab} is taken as an estimate of the ancient field strength, B_{anc}.

4 Results

There are many causes of failure of paleointensity experiments. Here we adopt the approach of Cromwell et al. (2015a) who chose selection criteria (Table 2), called CCRIT by Tauxe et al. (2016). These criteria are designed to test the assumptions of the IZZI experiment. Cromwell et al. (2015a) applied the criteria to specimens taken from historical lava flow tops that cooled quickly in fields known from historical measurements and tabulated in the International Geomagnetic Reference Field models (e.g., Alken et al., 2021). The Cromwell et al. (2015a) study recovered the field strength to within a few μT of the known field. CCRIT has specified threshold values for parameters at the specimen and at the site levels. At the former, CCRIT criteria are meant to test whether the demagnetization direction decays toward the origin using the deviation angle (DANG) and maximum angel of deviation (MAD) parameters (see definitions and original references in Paterson et al., 2014). DANG estimates the angle between the best fit line and the origin for the demagnetization direction. MAD measures the scatter in the NRM di-

Site	Location	Latitude ($^{\circ}N$)	Longitude ($^{\circ}E$)	Age (Ma)	$\pm 2\sigma$ (Ma)	^{39}Ar%	K/Ca	$\pm 2\sigma$	MSWD	n
GHI01	Mt. Bental	33.12635	35.78227	0.1177	0.0358	89	0.175	0.069	0.62	23
GHI05	Nahal Yehudiya, Rd 87	32.96051	35.86224	0.1679	0.0255	100	0.022	0.012	0.63	21
GHI06	Mt. Shifon	33.06958	35.77143	0.1145	0.0085	100	0.063	0.026	0.63	21
GHI07	Ortal	33.08581	35.75589	0.6805	0.0183	46	0.182	0.022	0.37	5
GHI08	Mt. Hermonit	33.17882	35.79236	0.7676	0.0179	56	0.116	0.032	1.11	14
GHI09	Mt. Odem	33.19430	35.75293	0.0894	0.0251	75	0.006	0.006	0.47	8
GHI10	Bashanit	33.05168	35.84968	0.6149	0.0349	100	0.029	0.012	0.97	26
GHI18	Dalton	33.02583	35.49491	1.6700	0.0400	100	0.320	0.070	1.12	25
GHI19	Amuka	32.99528	35.52599	2.4500	0.0226	65	0.656	0.036	0.43	20
GHI20	Givat Orcha	32.92629	35.84994	1.6500	0.0200	66	0.339	0.020	1.50	12
GHI21	Givat Orcha	32.92629	35.84994	1.6765	0.0302	92	0.054	0.015	0.59	22
GHI24	Mt. Ram	33.24848	35.79011	3.3300	0.0200	76	0.145	0.049	0.65	17
GHI25	Mt. Kramin	33.21873	35.77706	0.8723	0.0053	84	0.530	0.058	0.64	7
GHI26	Mt. Kramin	33.22000	35.77683	0.8704	0.0169	97	0.121	0.071	0.75	13
GHI27	Mt. Varda	33.21250	35.78616	1.1498	0.0348	81	0.511	0.036	0.39	18
GHI28	Mt. Varda	33.21250	35.78616	1.1912	0.0152	91	0.130	0.028	1.37	19
GHI29	Mt. Hermonit	33.17944	35.79322	0.7496	0.0945	87	0.272	0.050	0.89	18
GHI30*	Mt. Hermonit	33.18206	35.79858	1.2317	0.0757	80	0.054	0.022	2.93	20
GHI39	Nahal Orvim	33.14100	35.68200	0.8476	0.1165	100	0.320	0.076	0.04	24
GHI40	Nahal Orvim	33.14100	35.68200	0.7736	0.1949	100	0.290	0.053	0.22	23
GHI41	Nahal Orvim	33.14100	35.68300	0.7902	0.0058	70	0.212	0.014	1.01	12
GHI44*	Alonei Habashan	33.04200	35.83600	1.4369	0.0195	85	0.354	0.048	0.67	18
GHI46	Tel Saki	32.86829	35.82905	2.7442	0.0475	100	0.010	0.005	0.86	31
GHI47	Dalawe	33.09400	35.75200	0.9699	0.0636	100	0.038	0.015	0.45	21
GHI48	Dalawe	33.08500	35.75300	0.7231	0.0324	62	0.064	0.019	0.58	5
GHI49	Hashirion Junction	33.06500	35.74900	0.1162	0.0088	97	0.038	0.019	1.34	16

Table 1. Ar-Ar ages from this study. MSWD: mean squared weighted deviation, n: the number of steps in the plateau. * age based on 'mini-plateau' and all others are plateau ages.

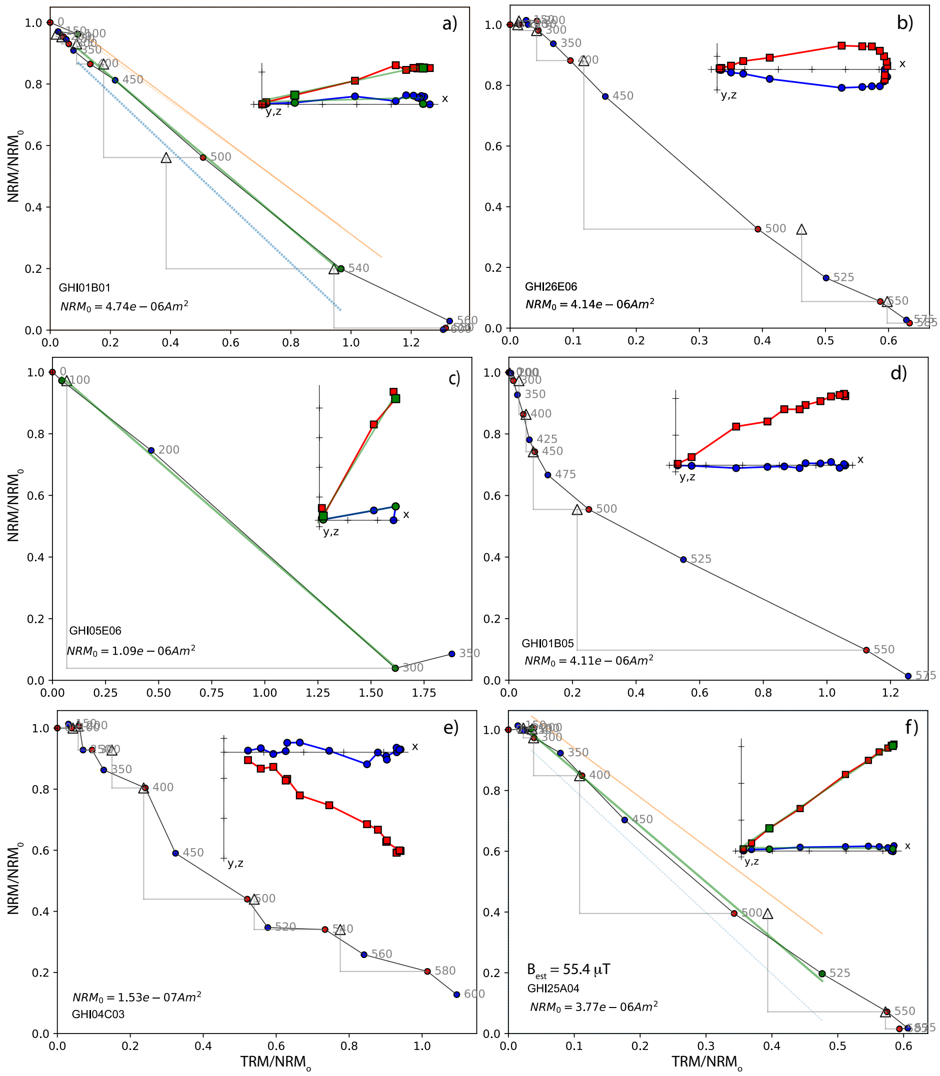

Figure 3. Examples of Arai plots of IZZI experiments and the effect of the CCRIT criteria. Circle color indicates the sequence of treatment steps- ZI (blue) or IZ (red). NRM remaining versus pTRM gained as a function of temperatures (circles). Triangles are pTRM checks. Insets are Zijderveld diagrams for the zero field steps with the magnetization vector projected onto the xz-plane (red) and the xy-plane (blue) for each specimen. The declinations have been rotated to the ‘X’ axis as these are all unoriented specimens. a) Failed the SCAT criterion because the 500° pTRM step falls outside the SCAT box shown as the blue and red lines. b) Failed the MAD criterion with MAD of 12.4. c) Failed the Gap Max criterion with G_{max} of 0.76. d) Failed the curvature criterion with $\vec{k'} = 0.728$. e) Failed the curvature criterion with $\vec{k'} = 0.618$. f) Passed all criteria.

rections during the experiment. The ratio relating the remanence remaining against that acquired is estimated by the best fitting line through a selection of the data. We use the ‘Auto Interpreter’ function of the Thellier GUI program of (Shaar & Tauxe, 2013), part of the PmagPy software package of (Tauxe et al., 2016) to find the portion of the data that passes CCRIT criteria in an objective way. PmagPy is freely available at:

https://github.com/PmagPy/PmagPy.

The fraction of remanence used in the fit (quantified by FRAC) must be large for the intensity estimate to be meaningful and we add an additional constraint, n, the minimum number of measurements used to fit the line. CCRIT also sets G_{max}, the maximum amount of fractional remanence removed between consecutive temperature steps, to 0.6. SCAT is a boolean value that indicates whether the data fall within $2\sigma_{threshold}$ of the best fit slope. Finally, CCRIT screens for non-linearity by applying a parameter that quantifies the curvature of the NRM/pTRM data, $\vec{k}$, as suggested by Paterson (2011); curvature is associated with biased intensity estimates (Krása et al., 2003; Tauxe et al., 2021; Cych et al., 2021). In the CCRIT criteria, we use $|\vec{k}'|$ which calculates curvature over the portion of remanence used in the calculation (hence the requirement of a large FRAC).

n	DANG	MAD	β	SCAT	FRAC	G_{max}	$\left\|\vec{k}\right\|$	N	$B_{\%}$	$B\sigma$
4	$\leq 10°$	$\leq 5°$	0.1	TRUE	0.78	$\geq$0.6	0.164	3	10	4 μT

Table 2. The CCRIT Cromwell et al. (2015a); Tauxe et al. (2016) selection criteria applied to the data from the IZZI experiment. See Paterson et al. (2014) for expanded definitions. n: minimum number of consecutive demagnetization steps, DANG: deviation angle, MAD: maximum angle of deviation, β = the maximum ratio of the standard error to the best fit slope, SCAT: a boolean value that indicates whether the data fall within $2\sigma_{threshold}$ of the best fit slope, FRAC: fractional remanence, G_{max}: maximum fractional remanence removed between consecutive temperature steps, $\vec{k'}$: maximum curvature statistic, N: minimum number of specimens per site, $B_{\%}$: maximum percentage deviation from the site average intensity, B_{σ}: maximum intensity (μT) deviation from the site average intensity.

We observed a wide range of behaviors in our study (Figure 3). A change in the ability to acquire pTRM results in failure to reproduce the original pTRM step and a SCAT value of False (Figure 3a). Some specimens appear to have rotated during cooling resulting in multi-component behavior in the zero field steps. This behavior often results in a failure of the MAD criterion (see inset to Figure 3b). In several specimens, the NRM was entirely unblocked between two consecutive steps (e.g., Figure 3c) violating our G_{max} criterion. In others the Arai plots were excessively curved (Figure 3d, exceeding the $\vec{k}'$ criterion. Others varied as a function of treatment steps (IZ or ZI) (e.g., Figure 3e) resulting in a zig-zagging pattern (Yu et al., 2004). These failed the curvature criterion (and also frequently the MAD threshold). DANG fails when the demagnetization vector bi-passes the origin. In our experiments, no specimens failed DANG that did not also fail MAD. Such behavior suggests the presence of non-ideal magnetic recorders and results from these specimens failed the CCRIT criteria. Of the 498 specimens from the GHI collection that underwent the IZZI experiment, 117 passed our specimen level criteria (see Table S3 and example in Figure 3f).

At the site level, CCRIT tests for consistency between intensity estimates ($B_{\%}$ or B_{σ}). B_{σ} is the standard deviation of the intensity estimates from a given site and B% is the standard deviation of intensity estimates at the site level expressed as a percentage of the mean intensity. A maximum threshold is set for $B_{\%}$ and B_{σ} and sites must meet at least one of the two thresholds to pass the CCRIT criteria. After we applied our site-level criteria, 18 high quality site estimates of paleointensity remained (Table 3).

Sites with specimens showing a range of curvatures such as those shown in Figures 3d and 4d might contain useful information for constraining paleointensity estimates, par-

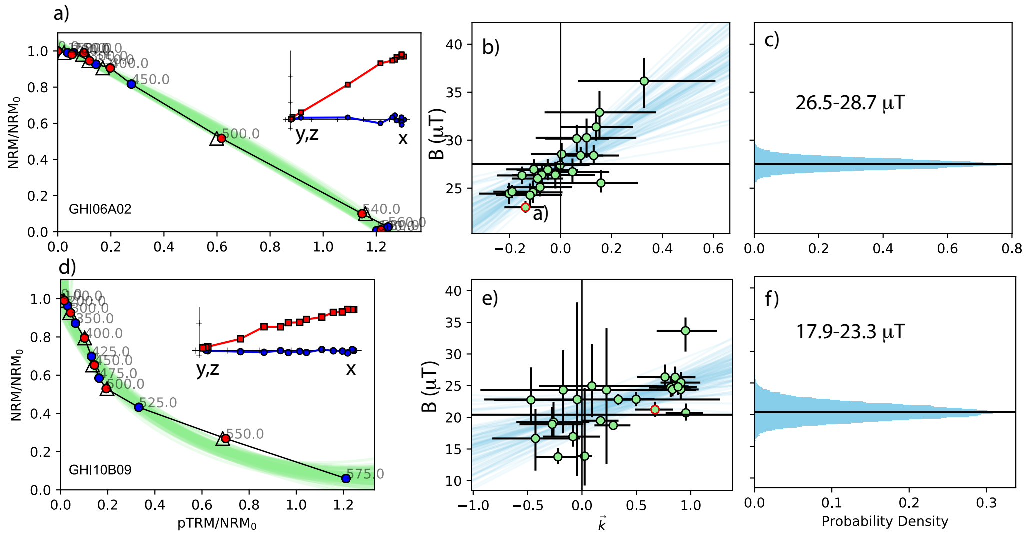

Figure 4. Examples of IZZI experiments and the effect of the BiCEP method. a) Example of Arai plot for specimen GHI06A02. Symbols same as in Figure 3. Green lines are Monte Carlo circle fits from the BiCEP method. b) Plots of intensity estimates from the circle fits against curvature ($\vec{k}$) and Monte Carlo line fits (shown in blue). c) Density plots of estimated intensities from the y-intercepts of the Monte Carlo line fits to the data shown in b). The Bayesian 95% credibility interval on the intensity estimates is 26.5-28.7 μT. d) Same as a) but for specimen GHI10B09. e) same as b) but for site GHI10. f) same as c) but for data shown in e).

ticularly if there are many specimens at the site level. For such sites we used the recently developed Bias-Corrected Estimation of Paleointensity (BiCEP) method of Cych et al. (2021). This method uses a Bayesian statistical approach. It makes the assumption that curved results ($|\vec{k}|$ >0.164) are linearly biased with respect to the true value as suggested by Santos and Tauxe (2019) and Tauxe et al. (2021). As an example of how BiCEP works, we use the data from site GHI06, which passed the CCRIT criteria with 20 specimens, yielding an average intensity value of 27.3 $\pm$ 1.8 μT (see Table 3). When subjected to BiCEP, we get an example of curvature fits to the data from one specimen in Figure 4a as green lines and the collection of estimates at the site level in Figure 4b. The Bayesian probability density plot (Figure 4c) gives a range in estimates of 26.5-28.7 μT, in excellent agreement with the CCRIT results. These bounds are minimum and maximum estimates which are analogous to 95% confidence bounds (so four times the width of our 1σ uncertainties with CCRIT).

The BiCEP method is most appropriate for sites that fail owing to curvature or mulit-component behavior and have at least five specimens. In general, low-temperature components can be removed as well as high temperature steps after the onset of alteration. An example of a site that failed CCRIT (because of a lack of sufficient specimens with low enough curvature) is shown in Figure 4d-f. This site yields a paleointensity estimate ranging from 17.9 to 23.3 μT. All of the BiCEP results are listed in Table 4. Where both CCRIT and BiCEP were successful (GHI06, GHI20 and GHI25), the two methods yielded very similar results and we use the CCRIT results.

Another example of how BiCEP can enhance interpretations at the site level when the CCRIT method fails is shown in Figure 5. This site had seven specimens that passed the CCRIT criteria but the within site scatter exceeded the CCRIT thresholds for both

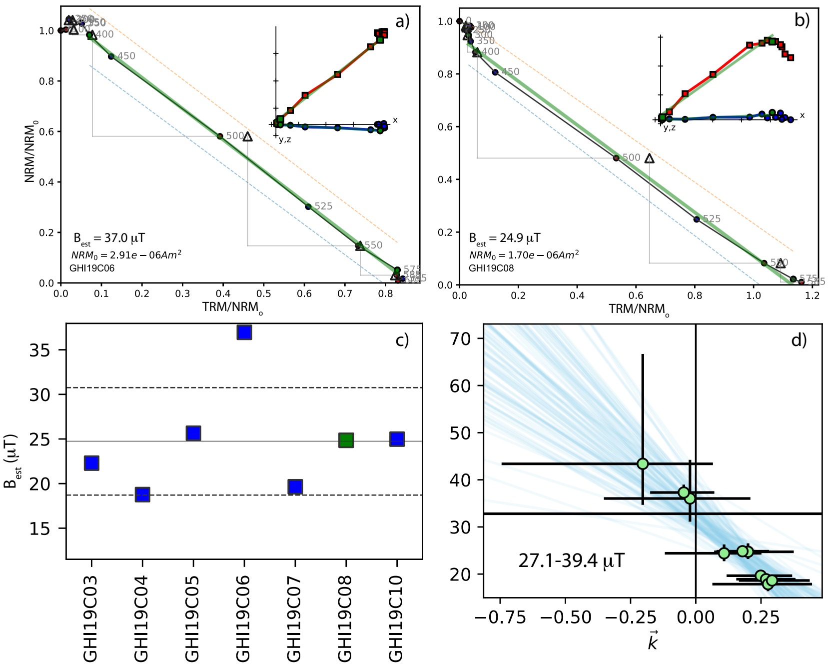

Figure 5. Examples BiCEP on a site with specimens that passed CCRIT but failed at the site level for being too scattered (GHI19). a) Data for specimen GHI19C06. Symbols same as in Figure 3. b) Same as a) but for sister specimen GHI19C08. c) Estmated intensities for all specimens passing CCRIT from site GHI19. d) Data from GHI19C treated using the BiCEP method. Symbols same as in Figure 4.

B_σ and $B_\%$ (see Figure 5c), hence was rejected by CCRIT. Of course we could arbitrarily exclude results deemed to be 'outliers', for example, the specimen shown in Figure 5a, which has the best specimen level statistics of the entire site. Arbitrary exclusion of specimens in this fashion well lead to misleading conclusions as we would be relying on data from specimens like that shown in Figure 5b, which is more curved than the 'outlier' and has a low temperature overprint. Instead of arbitrary data selection, we consider all the experimental data from the site using the BiCEP method (Figure 5d).

By standard paleomagnetic convention, a 'site' is a unit that forms over a short period or time and so records a uniform paleointensity and paleodirection. We would expect, for example, all specimens from a single lava flow to record the same paleomagnetic field. However, a cinder cone may have erupted over a period of time so, while we treat most cinder cones as a 'site', averaging all specimens together, there are two exceptions. GHI03 is composed of separate bombs scattered across the outcrop, so it may have erupted over a period of time. Samples from three of the bombs gave excellent, yet distinct, results so GHI03B, GHI03C and GHI03D could be treated as different sites. We also calculate the average of these three samples for a GHI03 average (star in Figure 6). This average has a standard deviation which fails the site level CCRIT criteria, however. In addition, specimens from GHI07C behaved consistently so we exclude the few specimens from GHI07A and GHI07E (which were distinct), but too few to pass at the site level criteria. All other cinder cones were treated as sites and all specimens were averaged at the cone (site) level.

Figure 6 shows the site mean data in equivalent virtual axial dipole moments (VADMs) in ZAm^2. Most of the data have intensities well below the present axial dipole field value

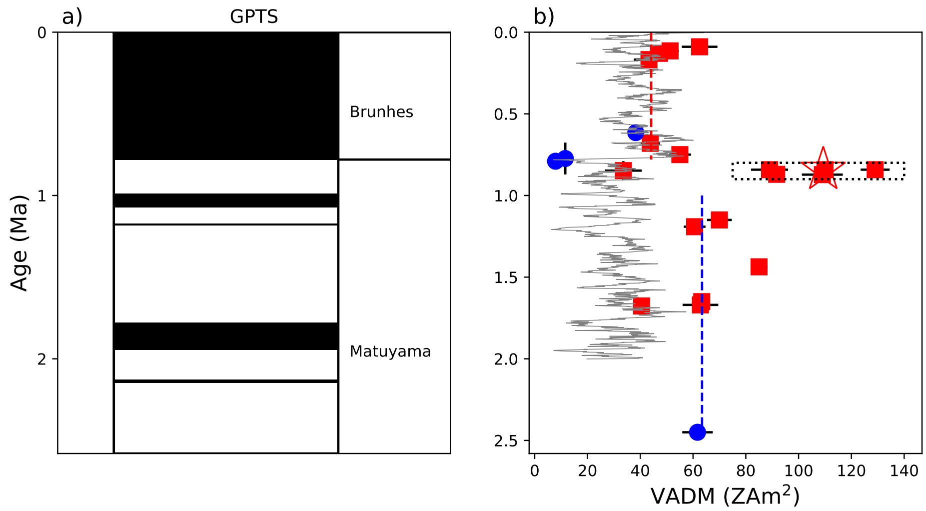

Figure 6. a) The Geomagnetic Polarity Timescale (Gradstein et al., 2012) for the Pleistocene. b) The VADM estimates and uncertainties from successful sites in this study along with their age constraints. Red (blue) squares (dots) are the CCRIT (BiCEP) site means and 1σ uncertainties. The grey line is PADM2M record of Ziegler et al. (2011). The box encloses five sites, three from the GHI03 cinder cone and GHI25, and GHI26. The average of the GHI03 sites is shown as a star. The red dashed line is the average value of the VADMs for the Brunhes Chron (44 ZAm^2) and the blue dashed line is that for the period 1-2.5 Ma (63 ZAm^2).

of ~80 ZAm^2, but there is a cluster of values just before the Brunhes/Matuyama boundary (surrounded by a dotted line box) with values higher than 80 ZAm^2. Three sites are from the GHI03 cinder cone, all assigned to the same age (0.842 Ma) but with distinct paleointensities. We have treated these three data points as separate sites because of their distinct paleointensity values, but they were erupted very close in time and it is likely that we have over-sampled a very brief interval of rapidly changing and high field values, similar to the so-called Levantine 'spikes' that occurred some 3000 years ago (e.g. Ben Yosef et al., 2009; Shaar et al., 2011, 2016) in the same part of the world. For this reason, we also averaged together the three sites from the GHI03 cinder cone (star in Figure 6). The mean paleointensity is 33.1 μT and the mean PADM is 62.2 ZAm^2 using the 22 sites that passed CCRIT or BiCEP criteria. Although one of our sites, GHI24, has an age of 3.3 Ma, all successful sites were from the Pleistocene (maximum age of 2.58 Ma, Gradstein et al., 2012).

Selkin & Tauxe, 2000, suggested that there may be a change in the average PADM sometime in the Brunhes Chron whereby data preceding about 0.3 Ma had an average of some 50 ZAm^2, while younger data had a higher average. This notion of a change in average moment was amplified by the work of Ziegler et al., 2011 who suggested a step-change in PADM at the Brunhes/Matuyama boundary in their PADM2M record (gray line in Figure 6b). Therefore, we calculate a Brunhes age (0-0.78 Ma) average (44 ZAm^2, red dashed line in the figure), which is in close agreement with the PADM2M curve. The average from 1-2.5 Ma (dashed blue line in Figure 6b) is 63 ZAm^2, or higher than the Brunhes average. These averages exclude the extrema just prior to and coincident with the Brunhes/Matuyama boundary (sites GHI03, GHI25, GHI26 and GHI39). The PADM2M curve was based on stacking of many marine sediment cores from around the world, calibrating the relative paleointensity stack with absolute ages from lava flows of known age.

Site	n	Intensity (μT)	B_σ (μT)	$B_\%$ (%)	VADM (ZAm^2)	Age (Ma)	(1σ) (Ma)	Latitude (°N)	Longitude (°E)
GHI02	3	25.2	2.2	8.8	47.3	0.1296	0.0006	33.1580	35.7767
GHI03B	7	68.7	2.9	4.3	129.0	0.842	0.01165	33.1228	35.7242
GHI03C	4	47.4	3.7	7.9	89.0	0.842	0.01165	33.1228	35.7242
GHI03D	3	58.8	0.3	0.4	110.4	0.842	0.01165	33.1228	35.7242
GHI03*	3	58.0	0.1	18.3	109.3	0.842	0.01165	33.1228	35.7242
GHI05	8	23.0	3.0	13.2	43.3	0.1679	0.01275	32.9605	35.8622
GHI06	20	27.3	1.8	6.6	51.3	0.1145	0.00425	33.0696	35.7714
GHI07C	6	23.3	1.9	8.3	43.8	0.6805	0.00915	33.0858	35.7559
GHI09	4	33.3	3.6	10.8	62.5	0.0894	0.00125	33.1943	35.7529
GHI18	3	33.4	3.6	10.8	62.8	1.67	0.02	33.0258	35.4949
GHI20	7	33.6	1.6	4.9	63.3	1.65	0.01	32.9263	35.8499
GHI21	4	21.5	1.4	6.3	40.5	1.6765	0.0151	32.9263	35.8499
GHI25	4	58.2	4.1	7.1	109.2	0.8723	0.00265	33.2187	35.7771
GHI26	6	48.9	1.4	2.9	91.7	0.8704	0.00845	33.2200	35.7768
GHI27	6	37.3	2.5	6.7	70.0	1.1498	0.0174	33.2125	35.7862
GHI28	5	32.3	2.2	6.8	60.6	1.1912	0.0076	33.2125	35.7862
GHI29	6	29.3	2.3	7.7	55.0	0.7496	0.04725	33.1794	35.7932
GHI39	3	17.9	3.7	20.7	33.6	0.8476	0.05825	33.1410	35.6820
GHI44	4	45.2	1.7	3.8	85.0	1.4369	0.00975	33.0420	35.8360

Table 3. Paleointensity results from this study that passed the CCRIT criteria. n: number of specimens per site, Intensity: site average intensity, B_σ: standard deviation , $B_\%$: percent error, VADM: site average VADM. GHI03* is the average of the three individual layers within the GHI03 cinder cone.

The curve is therefore thought to be a reflection of the dipole (global) strength while our data from Northern Israel are limited in geographic extent and represent spot readings of the field in a restricted area.

5 Discussion

5.1 Age of the Brunhes/Matuyama boundary

Two sites (GHI40, GHI41) shown in Figure 6 have very low intensities of 11.6 and 7.9 ZAm^2 with ages of 0.7736 and 0.7902 Ma, respectively. The age for the Brunhes/Matuyama boundary is 0.781 Ma in the Gradstein et al. (2012) time scale used here. Singer et al. (2019) suggested a younger age for the global reversal of 0.773 $\pm$ .002 Ma but with a long low intensity period prior to the actual reversal. Our new data are therefore consistent with revised age estimates of Singer et al. (2019).

5.2 Geologic map of the Golan Heights

With the new ages presented here, we have an opportunity to examine the generalized geological map for the Golan Heights region shown in Figure 2. The current age estimates for the Plio/Pleistocene boundary are 2.54 Ma of Gradstein et al. (2020) or 2.58 Ma from Gradstein et al. (2012). We are using the latter for consistency with our earlier studies as the differences for our purposes are negligible. Two locations in the Eastern Galilee (Dalton, GHI18, 1.67 Ma; Amuka, GHI19, 2.45 Ma), which were previously estimated to be between 2.7-1.7 Ma based on K-Ar dating (Mor, 1993; Heimann, 1990), late Pliocene using the earlier Pliocene/Pleistocene age boundary, yielded a similar age

Site	n	Intensity (μT)	B$_{min}$ (μT)	B$_{max}$ (μT)	VADM (ZAm^2)	Age (Ma)	1σ (Ma)	Latitude (°N)	Longitude (°E)
*GHI06	43	27.5	26.5	28.7	51.7	0.1145	0.00425	33.0696	35.7714
GHI10	28	20.4	17.9	23.3	38.3	0.6149	0.01745	33.0517	35.8497
GHI19	18	32.8	27.1	39.4	61.7	2.45	0.0113	32.9953	35.5260
*GHI20	15	35.7	33.5	39.2	67.2	1.65	0.01	32.9263	35.8499
*GHI25	14	52.3	43.9	60.5	98.1	0.8723	0.00265	33.2187	35.7771
GHI40	16	6.2	3.8	8.6	11.6	0.7736	0.09745	33.1410	35.6820
GHI41	8	4.2	1.1	8.0	7.9	0.7902	0.0029	33.1410	35.6830

Table 4. Paleointensity results from this study subjected to BiCEP intensity estimation of Cych et al. (2021). n: number of specimens per site, Intensity: site average intensity, B$_{min}$, B$_{max}$: minimum and maximum intensity values from BiCEP. VADM: site VADM. Starred sites also passed CCRIT and we use those results in the rest of the paper.

Study	Specimen n	Site n	Intensity (μT)	1σ (μT)	VADM (ZAm^2)	1σ (ZAm^2)	Latitude (°)
This Study	173	22	33.1	16.3	62.2	30.6	33
Asefaw et al. 2021	158	43	30.3	12.8	39.8	16.8	-78
Cromwell et al. 2015b	232	51	38.6	16.4	53.8	22.9	64
HSDP2 combined	199	56	34.1	9.2	76.1	20.4	20

Table 5. Paleointensity results from similar studies that investigate the paleomagnetic field over the Pleistocene. Specimen n: number of specimen that pass our specimen-level selection criteria, Site n: number of sites that pass our specimen and site-level selection criteria, Intensity: average intensity of all the successful sites in the study, σ: standard deviation, VADM: PADM of all the successful sites from the study. HSDP2 combined is the composite record of both the subaerial (Cai et al., 2017) and submarine (Tauxe & Love, 2003) portions of the Hawaii Scientific Drilling Project core HSDP2. See Figure 7 for locations.

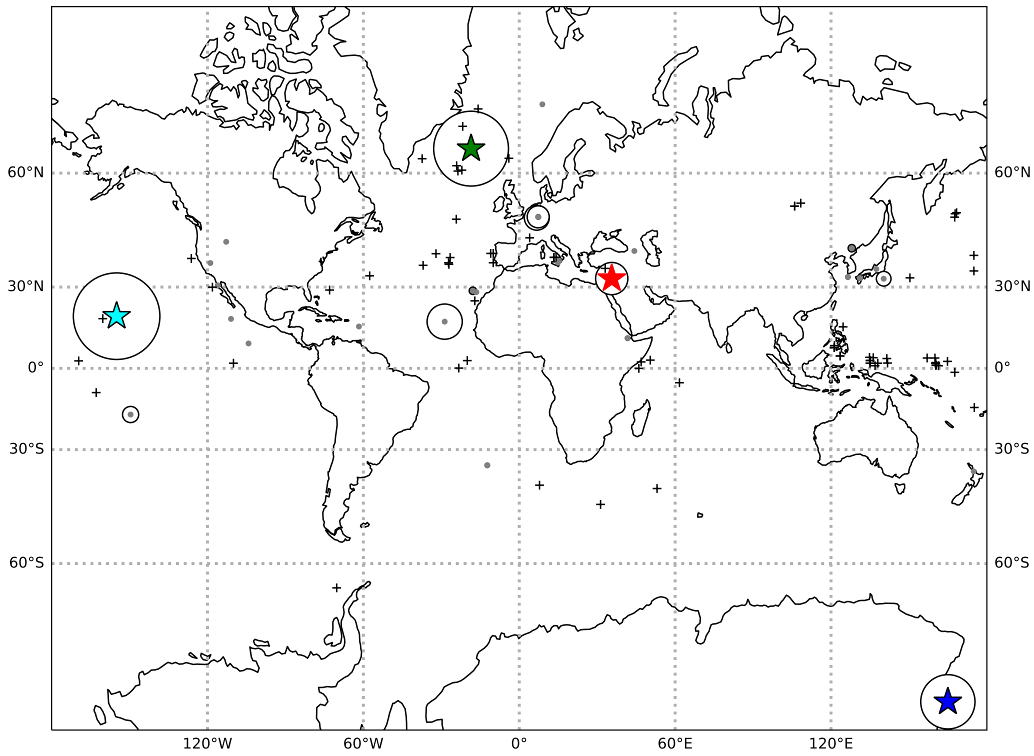

Figure 7. Map of site locations from the studies used here. PINT locations are plotted in grey and the size of the surrounding circles is proportional to the number of sites in each reference. The locations of the studies with measurement level data available that passed the CCRIT/BiCEP criteria are plotted as stars (Northern Israel: red; Antartica: blue; Iceland: green, HSDP2: cyan). Plus signs are locations of cores included in PADM2M (Ziegler et al., 2011).

range to that found here. We therefore mark these basalts in Figure 2 as Early to Middle-Pleistocene. Also, we confirm here that all Nahal Orvim sites (GHI39-41), previously dated with K-Ar (Mor, 1986; Heimann & Ron, 1993) are Late Pleistocene. Site GHI46 (Tel Saki), which appears at the boundary between the Early Pliocene (> 3.5 Ma) and the Pleistocene (< 1.76 Ma) basalts, gave a similar age as in Behar et al. (2019) of 2.74 Ma (see Figure S1) and therefore associated with an unrecognized Late Pliocene volcanic phase. Our youngest age (GHI09, 0.089 Ma from Mount Odem) provides new constraints to the age of the latest volcanic phase in the area (∼0.1 Ma; Weinstein et al., 2013; Shaanan et al., 2011; Behar et al., 2019).

5.3 Comparison of intensities with similar studies from elsewhere

Paleointensity studies conducted at different latitudes and over the same time interval should recover similar average VADMs (here called PADMs), if the field structure is a GAD field. To compare PADM estimates with different latitudes, we identified studies that span the Pleistocene and focussed on 'ordinary' PSV, avoiding targeting abnormal field behavior such as excursions or reversals. We selected studies that applied a Thellier-Thellier variant (Thellier & Thellier, 1959; Coe, 1967b) to measure paleointensity and included a pTRM check to monitor lack of reproducibility. Cromwell et al. (2015a), among others, suggested that different paleointensity methods applied to the same lava flow can produce a large range in paleointensities. And, applying 'looser' or 'stricter' selection criteria to calculate paleointensity can also result in different paleointensity estimates for the same specimen. Therefore, we focused on studies for which the measurement level data were avalable, and applied the same selection criteria. Seven studies met these requirements (Leonhardt et al., 2003; Wang et al., 2015; Cromwell et al., 2015b; Asefaw et al., 2021; Cai et al., 2017; Tauxe & Love, 2003; Biasi et al., 2021). The data were either in the MagIC database (earthref.org/MagIC) already, or the authors agreed to share their measurement level data.

The study of Leonhardt et al. (2003) presented data from volcanic units in Brazil (3.85°S) that span 1.8 – 3.3 Myr. They published a 75 ZAm^2 PADM based on data from nine discrete units. No sites passed the CCRIT criteria. Cromwell et al. (2015b) reported on paleointensity estimates from Iceland (64.4°N). They found a 78.1 $\pm$ 22 ZAm^2 PADM from four sites that formed 0 - 11 ka and a 47 $\pm$ 11.6 ZAm^2 PADM from 37 sites that span 11 ka - 3.35 Ma. Thirty-nine of these sites from Pleistocene units of Iceland met our CCRIT selection criteria (Table S3) and an additional 12 were successfully analyzed with BiCEP (Table S4). The new Pleistocene PADM for Iceland is 53.8 $\pm$ 22.9 ZAm^2. Asefaw et al. (2021) investigated paleointensities in Antarctica that range in age from the Miocene to the Late Pleistocene. The authors applied a modified CCRIT criteria and recovered a 44 ZAm^2 PADM from 26 sites. We re-interpreted their data using the same CCRIT parameters as for this study as well as BiCEP (Tables S5 and S6 respectively). The Pleistocene mean PADM from the 42 sites is 40.3 $\pm$ 17 ZAm^2. There are two studies that analyzed quenched horizons from the Hawaii Scientific Drilling Project HSDP2 core, one targeting the submarine sequence (Tauxe & Love, 2003) and a second study focused on quenched margins of the subaerial sequence (Cai et al., 2017). These were re-analyzed here. The sites from Tauxe and Love (2003) that passed CCRIT are listed in Table S7. No sites had a sufficient number of specimens for the BiCEP method. The sites from Cai et al. (2017) that passed CCRIT are listed in Table S8 and BiCEP are in Table S9. The results from the two studies were combined together and the mean PADM from the resulting 59 sites from the Pleistocene (spanning from 0.03 to 0.553 Ma) is 76.7 $\pm$ 21. Wang et al. (2015) published paleointensities from the Galapagos Islands (1° S) with ages ranging between 0 – 3 Myr. In their study, the authors used a new approach known as the Multi-Domain Correction method (Wang & Kent, 2013) to their data. This was intended to correct for non-ideal magnetic recorders. They produced a PADM of 55.9 $\pm$ 2.9 ZAm^2 based on 27 independent lava flows. We found that only two sites met our CCRIT selection criteria (see Table S10). The two successful sites from the Galapagos are insufficient for a meaningful average. Biasi et al. (2021) sampled 31 sites from the James Ross Island in the Antarctic Peninsula and subjected them to the IZZI protocol (Yu et al., 2004), Tsunakawa-Shaw (Yamamoto & Yamaoka, 2018) and the pseudo-Thellier method (Tauxe et al., 1995) 'calibrated' using the approach of de Groot et al. (2013). None of these data passed the CCRIT or BiCEP criteria used here, so we proceed with the data from Iceland, Hawaii and Antarctica (see Figure 7 for locations).

Figure 8a displays the new and re-analyzed results from the four locations against their mean latitudes (see also Table 5 and Figure 7 for locations). In order to ensure that we are considering only Pleistocene data, all data sets have been filtered to include only

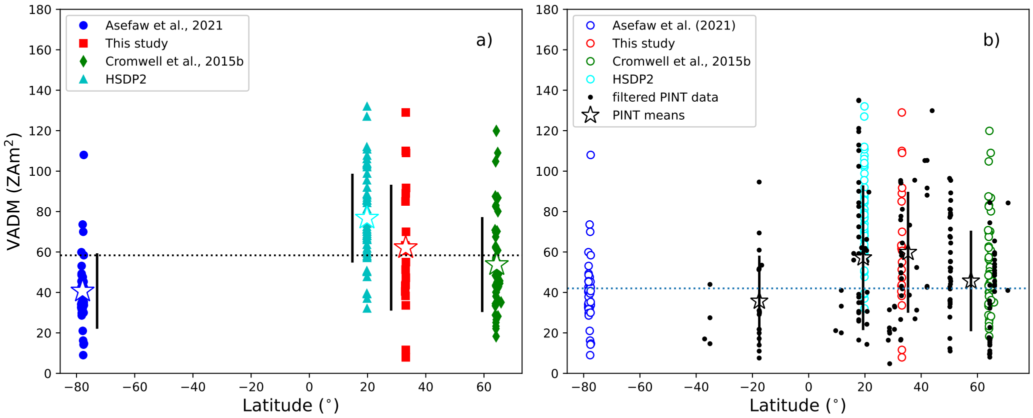

Figure 8. a) VADM estimates from four similar studies: Asefaw et al. (2021) (blue circles), this study (red squares), Cromwell et al. (2015b) (green diamonds) and Cai et al. (2017); Tauxe and Love (2003) (HSDP2 combined: cyan triangles). Only Pleistocene sites that passed our CCRIT set of selection criteria or BiCEP are included. The stars mark the average VADM in each study. Error bars are one standard deviation. Dotted line is the grand mean of the four locations. b) Filtered data from the PINT database of Bono et al. (2022) (black dots). Stars are averages from 10° latitudinal bins along with the standard deviations (black lines). Colored circles are from a). Data from (Lawrence et al., 2009), (Cromwell et al., 2015b) and HSDP2 from Cai et al. (2017); Tauxe and Love (2003) are superseded by the Antarctic, Icelandic and HSDP2 data re-analyzed here. They were replaced in the PINT data plotted here.

those with ages with standard deviations less than 0.2 Ma. All four study means are within one standard deviation of the grand mean of the four. To consider whether or not the data sets were drawn from a single distribution of dipole moments, we plot the cumulative distributions of the VADMs from the four studies in Figure 9a. In this plot, it appears that each of the data sets and latitude bands are distinct from each other.

We need some statistical test for the null hypothesis that the four data sets are the same or different, for example, the Student's t-test. The p-values from a two-sided Student's t-test for the Northern Israel data versus the Icelandic data is 0.2, which does not allow us to reject the null-hypothesis that they were drawn from the same distribution. All other comparisons gave p-values less than 0.05. However, there is an inherent assumption in the t-test that the data are normally distributed, which may not be true. So we examined the four data sets with the non-parametric approach of using Kolmogorov-Smirnov (KS) tests on the cumulative distributions. Here we use a two-sample Kolmogorov-Smirnov (K-S) test. These gave similar results. Therefore, each of these data sets performed the same experiment and were subjected to an identical set of selection criteria (including age) but recover different distributions, with Antarctica being lower and Hawaii being higher.

A key assumption here is that the data sets span the same time interval. We plot the data against age in Figure 10. Of course the exact same ages cannot be identified in separate studies because the field can change very fast within the uncertainty of the dating method, so any two lava flows with identical ages could very well yield very different results (see for example the GHI03 cinder cone considered in Section 4). Despite the fact that these data sets are the largest available in the public record that have the original measurements available, it is still possible that they are under-sampled with respect to the variation in field strength with time and that more data will sharpen regional

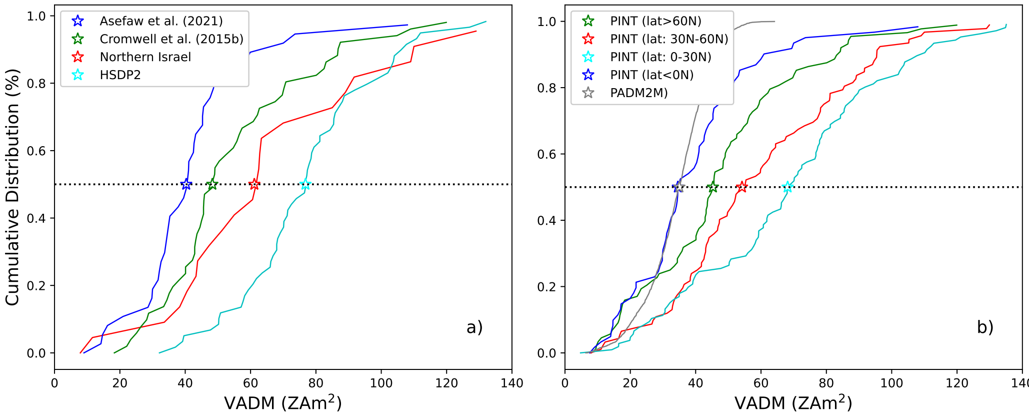

Figure 9. Cumulative distribution functions of VADMs from Pleistocene paleointensity data. a) Data sets (re-)analyzed here. Red line: Northern Israel (this study), Blue line: Antarctica (Asefaw et al., 2021); Green line: Iceland (Cromwell et al., 2015b); Cyan line HSDP2 (Cai et al., 2017; Tauxe & Love, 2003). b) Data from the PINT database and PADM2M (Ziegler et al., 2011) (grey line). Data in PINT from Antarctica (Lawrence et al., 2009), from HSDP2 (Cai et al., 2017; Tauxe & Love, 2003), and from Iceland (Cromwell et al., 2015b) were replaced with the re-analyzed data from this study. Red line: mid-latitude data (30-60°N), green line: high-latitude data ($\geq$ 60°N); cyan line: low latitude northern hemisphere data (0-30°N), blue line: data from southern hemisphere (latitudes $<$ 0°N). Stars are median values for each subset of the data.

differences. Support for this view comes with a comparison of the paleointensity estimates considered here with estimates of the globally averaged data set.

5.4 Comparison with the PINT database

So far we have focused our attention on studies that applied a a similar, proven, experimental technique and subjected the data to the same analysis. However, the resulting dataset is limited to those studies with measurement level data available. To increase the number of sites, we use the paleointensities in the PINT database (Bono et al., 2022) (available at http://www.pintdb.org/ Database). As of January, 2022, the PINT database archived results from 4353 absolute paleointensity sites from 296 unique references. The studies included in the PINT database applied a variety of techniques (e.g., Thellier & Thellier, 1959; Hill & Shaw, 1999; van Zijl, Graham, & Hales, 1962), corrections, and quality criteria to estimate paleointensity and range in age from 4.2 Ga to 50,000 years ago. The database does not, however, include measurement level data, so we cannot subject the data to a uniform set of selection criteria as done in the foregoing. The quality of the paleointensity estimates may therefore vary widely between different studies making a direct comparison between different studies challenging. Some authors (e.g., Biggin & Paterson, 2014; Kulakov et al., 2019) address this challenge by creating a quality scale and assigning each site a quality score while others (Bono et al., 2020) apply additional filters to the dataset. In this study, we first filtered the data for the Thellier-Thellier method (Thellier & Thellier, 1959) (T+), the microwave method (Hill & Shaw, 1999) M+, the low-temperature demagnetization with Thellier (Yamamoto & Tsunakawa, 2005) LTD-T+, and the low-temperature demagnetization variation of the Shaw method (Yamamoto et al., 2003) LTD-DHT-S. The addition of a '+' indicates that p-TRM checks were included in the experiment. We then chose only results based on at least three specimens that had a standard deviation of $\leq$ 4 μT or $\leq$10% at the site level, as in CCRIT.

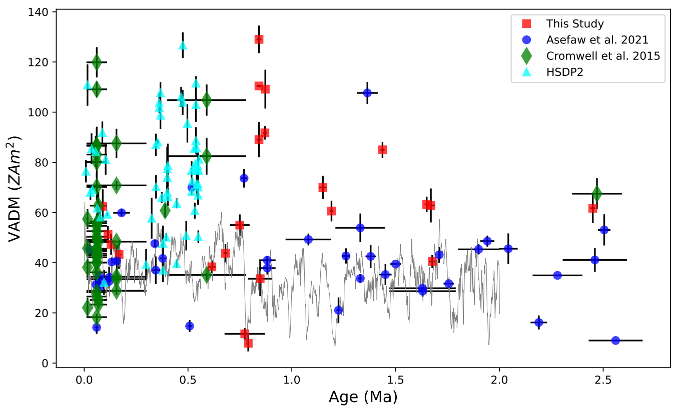

Figure 10. VADM estimates for Pleistocene aged data from the studies re-analyzed here along with the globally averaged estimates from PADM2M of Ziegler et al. (2011) (gray line). All absolute paleointensity studies performed an IZZI-modified Thellier-Thellier experiment and were re-interpreted with a uniform set of selection criteria to estimate paleointensity. Paleointensities were recovered from 42 sites (blue circles) from Antarctica (Asefaw et al., 2021), 22 sites (red squares) from Northern Israel, 51 sites (green diamonds) from Iceland (Cromwell et al., 2015b) and 59 sites from HSDP2 (Cai et al., 2017; Tauxe & Love, 2003). Only data from sites with age uncertainties <0.2 are shown.

Furthermore, we filtered for those studies whose ages had a standard deviation of less than 0.2 Ma and were Pleistocene in age. Finally, we replaced the studies re-analyzed here with the re-interpreted data (REF numbers 639, 663, 707, 210, and 719) to avoid over-weighting those results. The resulting dataset includes 352 results from 35 unique references. The locations of the resulting filtered PINT sites are shown in Figure 7 and the VADMs of the data (recalculated here for consistency) are plotted against latitude in Figure 8b.

Lawrence et al. (2009), in their study of Antarctic paleointensities from the Erebus Volcanic Province in Antarctica (superceded by Asefaw et al., 2021), plotted data from the PINT08 database at the time (Biggin et al., 2009) against latitude. They folded southern latitudes onto the northern equivalent as there were too few southern hemisphere data points for a meaningful comparison. They concluded that the Antarctic data were anomalously low compared to lower (absolute) latitudes for the last five million years. They suggested several possible causes for this departure from a GAD field, including differences in temporal coverage, experimental design and the effect of the 'tangent cylinder' surrounding the inner core on field generation. Asefaw et al. (2021) re-analyzed the data of Lawrence et al. (2009) using stricter criteria which eliminated many sites from consideration, but added many new sites that were sampled targeting rapidly cooled parts of the lava flows, similar to the approach taken here and by Cromwell et al. (2015b) in Iceland. The Asefaw et al. (2021) study supported the contention that Antarctic VADMs were lower on average than lower latitude sites, but they also found that the data were close to those from Iceland published by Cromwell et al. (2015b).

Having discounted experimental design as a probable cause for the 'low' paleointensities in the polar data, one of the motivations for the present study was to assess whether the paleointensity values found in Antarctica and Iceland over the last few million years appeared 'low' because the data from lower latitudes were biased in some way owing to inadequate temporal sampling or experimental design. Here we have found that the data from Northern Israel (mid latitude Northern Hemisphere) appear to be likely higher on average than those from Antarctica. If we include all the data of comparable quality from the PINT database (to the extent that it is possible to assess that), we see from Figure 8b that data from mid-latitudes (northern hemisphere) are in general higher than those from the southern hemisphere or from high northerly latitudes.

Turning again to the plots of cumulative distributions of VADMs (Figure 9b), we see that the data from mid-latitudes (between 30° and 60°N) are higher than those from the low latitude band of 0-30°N. This suspicion is supported by the Student's t-test on subsets of the PINT database (with replacement of re-analyzed studies as described in the foregoing). The p-values for the mid-latitude subset (30-60°N) versus high northerly latitudes (>60°N) is $< 10^{-3}$ allowing us to reject the hypothesis that they are drawn from the same distribution at the 95% level of confidence. Similarly, the p-value for mid-latitudes versus the southern hemisphere data is $<10^{-5}$. Moreover, the p-value for data from mid-latitudes compared to low latitudes (0-30°N) is 0.03, which also allows us to reject the hypothesis that the two data sets are drawn from the same distribution. Therefore, it appears that VADMs from the Pleistocene from the northern hemisphere latitudes less than 30°N are higher than elsewhere. It is also worth pointing out that (Wang et al., 2015) found VADMs from the equatorial sites in Galapagos that were comparable to those from Antarctica. That dataset did not survive our filtering process but methods are being developed which may provide high quality paleointensity estimates from lava flows in the near future (Wang & Kent, 2021). Further support for low intensities from the southern hemisphere came from Engbers et al. (2022), who found low intensities from their Miocene sites from Saint Helena. There also appears to also be a large amount of variability with respect to longitude in the timings of the periods of high intensity (see Figure 8).

Each of the paleointensity data points considered here are 'spot' readings of field strength. The data set we have compiled here is also strongly biased to the northern hemisphere. It is therefore worthwhile considering the so-called paleointensity axial dipole moment (PADM) data set for the last 2 million years (PADM2M of Ziegler et al., 2011; plus signs in Figure 7). Relative paleointensity records from seventy-six cores taken around the globe were placed on a common time scale by Tauxe and Yamazaki (2007). These were combined with absolute paleointensity (API) records from the Geomagia50.v2 database of Donadini et al. (2009) and the PINT08 database of Biggin et al. (2009). The API and RPI data were stacked to create a globally averaged estimate of the PADM. This record is an interesting comparison with the absolute paleointensity data considered here as there is much better representation of the southern hemisphere by using marine sediment cores than available from absolute paleointensity alone.

The generally lower estimates for the dipole moment in PADM2M than those for our low and mid-latitude data from the northern hemisphere, could well be caused by a real difference between northern and southern latitude field strengths. It seems that in the northern hemisphere data sets plotted in Figure 10, there are extended periods of time with high field strengths that persist over periods of time of some 50 kyr, but that these periods of high field strength do not occur at the same time globally. A possible explanation would be to use the so-called South Atlantic Anomaly in the recent geomagnetic field (SAA in Figure 1a) as an example of a strong non-dipolar field structure. While this low intensity dimple does not appear to persist over long periods of time as it is not apparent in a field model calculated by taking the average of the Holocene field models in the CALS10k.2 model of Constable et al. (2016) (Figure 1b), or any others we examined, it is interesting that this model does have an asymmetry between field strengths in the northern and southern hemispheres. It seems likely that a low intensity dimple did exist, perhaps fleetingly, in the southern hemisphere and that would account for the asymmetry observed. Compare for example the 60°N latitude band with an average of some 65 μT with its southern hemisphere sister, whose average field is $\sim$55 μT. This same persistent asymmetry is also seen in the time averaged field model of, for example, Cromwell et al. (2018) who compiled a global database of paleomagnetic directional data and produced a time averaged field model for the past five million years. We show intensities predicted from their LN3 model in Figure 1d. In this model, there are hemispheric differences in predicted field strength that apparently persisted for millions of years.

6 Conclusions

Forty-four sites (out of 52 sampled) from Northern Israel were were subjected to an IZZI Thellier-Thellier experiment. Eighteen sites passed the strict selection criteria (CCRIT) of Tauxe et al. (2016) and a further four gave acceptable results using the BiCEP method of Cych et al. (2021). Taken together, the study yields a 33.1 $\pm$ 16.3 μT mean intensity or 62.2 $\pm$ 30.6 ZAm^2 paleomagnetic axial dipole moment (PADM) for the Pleistocene. We re-analyzed data from four other comparable studies using the same selection criteria and filtering for the same Pleistocene age range. Data from the Hawaii Scientific Drilling Project's HSDP2 of Cai et al. (2017) and Tauxe and Love (2003) yielded 59 sites with a higher PADM of 76.7 $\pm$ 21.3 ZAm^2. In contrast, those from Cromwell et al. (2015b) for Iceland recovered a lower PADM of 53.8 $\pm$ 22.9 ZAm^2 (n=51). That average is higher than results from Antarctica (Asefaw et al., 2021), which when re-analyzed here resulted in 42 sites with a mean of PADM, 40.3 $\pm$ 17.3 ZAm^2.

We compared the results from our new and re-analyzed data sets with those from the paleointensity (PINT) database Bono et al. (2022) and found that in general, low to mid latitude northern hemisphere field strengths are higher than southern hemisphere (mostly Antarctica) and high northerly latitudes (mostly Iceland). The globally averaged PADMs predicted from the PADM2M record of Ziegler et al. (2011) are also much lower than those found here. The PADM2M record, unlike the absolute paleointensity data considered in this paper incorporates a large number of relative paleointensity records, including many from mid-southerly latitudes, suggesting the possibility of a persistent asymmetry in field strengths between the northern and southern hemispheres. This is supported by analysis of field models from the present (2022) field, the Holocene and five million year time averaged fields, which all show an asymmetry between northern and southern hemispheres, with the northern hemisphere predicted to be on average some 10 μT stronger than the same latitude band in the southern hemisphere.

Acknowledgments

This study was funded by the National Science Foundation. H.A. acknowledges support from the National Science Foundation Graduate Research Fellowship Grant No. DGE-1650112 and L.T. acknowledges funding from the National Science Foundation Grants EAR1345003 and EAR1827263. We thank Christeanne Santos and Malana Tabak for their assistance in the lab. We thank Huapei Wang and an anonymous reviewer for their thorough and insightful reviews which greatly improved the manuscript. We also thank Yael Ebert for her hard work in the field and Shuhui Cai for her helpful discussions and guidance. We thank Jeff Gee for his helpful comments on the manuscript. We also thank Lisa Schnepp, Huapei Wang, Geoff Cromwell, Pierre Camps, and Roman Leonhardt for sharing their measurement level data. We thank Brendan Cych for advice on the use of the BiCEP code. All other data referenced in the paper are in the MagIC database (https://earthref.org/MagIC) and PINT database (http://www.pintdb.org). The new and re-analyzed dataset is temporarily available here https://earthref.org/MagIC/19491/b161c048-ff5e-4981-a75b-99ee50a32fa5 for the purposes of review, and will be publicly available upon acceptance of this manuscript at this link: https://earthref.org/MagIC/19491. Code used to perform calculations are in the PmagPy software distribution (Tauxe et al., 2016) (https://github.com/PmagPy/PmagPy) and the BiCEP software package (Cych et al., 2021) (https://github.com/bcych/BiCEP_GUI).

A fully functional Jupyter notebook used to make the calculations and plots is available through https://github.com/ltauxe/Pleistocene-paleointensity-notebook as a supplement to this paper. Thanks to Rupert Minnett for assistance with Binder.

References

Alken, P., Thébault, E., Beggan, C., Amit, H., Aubert, J., Baerenzung, J., ... Zhou, B. (2021). 2021 International Geomagnetic Reference Field: the thirteenth generation. *Earth, Planets and Space*, *73*. doi: 10.1186/s40623-020-01288-x

Asefaw, H., Tauxe, L., Koppers, A., & Staudigel, H. (2021). Four-dimensional paleomagnetic dataset: Plio-Pleistocene paleodirection and paleointensity results from the Erebus Volcanic Province, Antarctica. *Journal of Geophysical Research: Solid Earth*, *126*, e2020JB020834 doi: 10.1029/2020JB020834.

Behar, N., Shaar, R., Tauxe, L., Asefaw, H., Ebert, Y., Heimann, A., ... Hagai, R. (2019). Paleomagnetism and paleosecular variations from the Plio-Pleistocene Golan Heights Volcanic Plateau, Israel. *Geochemistry, Geophysics, Geosystems*, *20*, 4319-4334 doi:10.1029/2019GC008479.

Ben Yosef, E., Tauxe, L., Levy, T., Shaar, R., Ron, H., & Najjar, M. (2009). Archaeomagnetic intensity spike recorded in high resolution slag deposit from historical biblical archaeology site in southern Jordan. *Earth and Planetary Science Letters*, *287*, 529-539.

Biasi, J., Kirschvink, J., & Fu, R. (2021). Characterizing the geomagnetic field at high southern latitudes: Evidence from the Antarctic Peninsula. *Journal of Geophysical Research*, *126*, e2021JB023273. doi: 10.1029/2021JB023273

Biggin, A., & Paterson, G. (2014). A new set of qualitation reliability criteria to aid inferences on paleomagnetic dipole moment variations through geological time. *Frontiers in Earth Science*, *2*, doi: 10.3389/feart.2014.00024.

Biggin, A., Strik, G., & Langereis, C. (2009). The intensity of the geomagnetic field in the late-Archean: new measurements and analysis of the updated IAGA palaeointensity database. *Earth Planets and Space*, *61*, 9-22 doi:10.1186/BF03352881.

Bono, R., Biggin, A., Holme, R., Davies, C., Meduri, D., & Bestard, J. (2020). Covariant giant Gaussian process models with improved reproduction of paleosecular variation. *Geochemistry Geophysics Geosytems.*, *21*, e2020GC008960 doi: 10.1029/2020GC008960.

Bono, R., Paterson, G., van der Boon, A., Engbers, Y., Grappone, J., Handford, B., ... Biggin, A. (2022). The PINT database: a definitive compilation of absolute palaeomagnetic intensity determinations since 4 billion years ago. *Geophysical Journal International*, *229*, 522-545. doi: doi:10.1093/gji/ggab490

Cai, S., Tauxe, L., & Cromwell, G. (2017). Paleointensity from subaerial basltic glasses from the second Hawaii Scientific Drilling Project (HSDP2) core and implications for possible bias in data from lava flow interiors. *J .Geophys. Res.*, *122*, 8664-8674. doi: 10.1002/2017JB014683

Coe, R. (1967a). 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*, 157- 179 doi:10.5636/jgg.19.157.

Coe, R. (1967b). Paleointensities of the Earth's magnetic field determined from Tertiary and Quaternary rocks. *Journal of Geophysical Research*, *72*, 3247-3262 doi:10.1029/JZ072i012p03247.

Constable, C., Korte, M., & Panovska, S. (2016). Persistent high paleosecular variation activity in southern hemisphere for at least 10,000 years. *Earth and Planet. Sci. Lett.*, *453*, 78-86.

Cromwell, G., Tauxe, L., & Halldorsson, S. (2015b). New paleointensity results from rapidly cooled Icelandic lavas: Implications for Arctic geomagnetic field strength. *Journal of Geophysical Research: Solid Earth*, *120*, 2913 - 2934 doi:10.1002/2014JB011828.

Cromwell, G., Tauxe, L., Staudigel, H., Constable, C., Koppers, A., & Pedersen, R. (2013). In search of long-term hemispheric asymmetry in the geomagnetic field: Results from high northern latitudes. *Geochemistry Geophysics Geosystems*, *14*, 3234-3249 doi:10.1002/ggge.20174.

Cromwell, G., Tauxe, L., Staudigel, H., & Ron, H. (2015a). 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 doi:10.1016/j.pepi.2014.12.007.

Cromwell, G., Trusdell, F., Tauxe, L., Staudigel, H., & Ron, H. (2018). Holocene paleointensity of the island of Hawai'i from glassy volcanics. *Geochemistry Geophysics Geosysystems*, *19*, 3224-3245 doi:101002/2017GC006927.

Cych, B., Morzfeld, M., & Tauxe, L. (2021). Bias corrected estimation of paleointensity (BiCEP): An improved methodology for obtaining paleointensity estimates. *Geochemistry Geophysics Geosystems*, *22*, e2021GC009755. doi: 10.1002/2021GC009755

de Groot, L., Biggin, A., Dekkers, M., Langereis, C., & Herrero-Bervera, E. (2013). Rapid regional perturbations to the recent global decay revealed by a new Hawaiian record. *Nature Communications*, *4*, 1-7. doi: 10.1038/ncomms37

Donadini, F., Korte, M., & Constable, C. (2009). Geomagnetic field for 0-3ka: 1. New data sets for global modeling. *Geochemistry Geophysics Geosystems*, *10*(6). doi: 10.1029-2008GC002295

Engbers, Y., Grappone, J., Mark, D., & Biggin, A. (2022). Low paleointensities and Ar/Ar ages from St Helena provide evidence for recurring magnetic field weaknesses in the south atlantic. *Journal of Geophysical Research*, *127*, e2021JB023358. doi: 10.1029/2021JB023358

Gradstein, F., Ogg, J., Schmitz, M., & Ogg, G. (2012). *Geologic Time Scale 2012.* Elsevier.

Gradstein, F., Ogg, J., Schmitz, M., & Ogg, G. (2020). *Geologic Time Scale 2020.* Elsevier.

Heimann, A. (1990). *The development of the Dead Sea Rift and its margins in northern Israel during the Pliocene and the Pleistocene* (PhD).

Heimann, A., & Ron, H. (1993). Geometric changes of plate boundaries along part of the northern dead-sea transform - geochronological and paleomagnetic evidence. *Tectonics*, *12*(2), 477-491. doi: Doi10.1029/92tc01789

Heimann, A., Steinitz, G., Mor, D., & Shaliv, G. (1996). The Cover Basalt Formation, its age and its regional and tectonic setting: Implications from K-Ar and ^{40}Ar/^{39}Ar geochronology. *Israel Journal of Earth Sciences*, *45*, 55 - 71.

Hill, M., & Shaw, J. (1999). Paleointensity results for historic lavas from Mt. Etna using microwave demagnetization/remagnetization in a modified Thellier-type experiment. *Geophysical Journal International*, *139*, 583-590. doi: doi:10.1046/j.1365-246x.1999.00980.x

Hospers, J. (1955). Rock magnetism and polar wandering. *The Journal of Geology*, *63*, 59 - 74 doi: 10.1038/1731183a0.

Koppers, A., Russell, J., Jackson, M., Konter, J., Staudigel, H., & Hart, S. (2008). Samoa reinstated as a primary hotspot trail. *Geology*, *36*, 435 - 438 doi:10.1130/G24630A.1.

Koppers, A., Staudigel, H., & Wijbrans, J. (2000). Dating crystalline groundmass separates of altered Cretaceous seamount basalts by the ^{40}Ar/^{39}Ar incremental heating technique. *Chemical Geology*, *166*, 139-158 doi:10.1016/S0009-2541(99)00188-6.

Krása, D., Heunemann, C., Leonhardt, R., & Petersen, N. (2003). Experimental procedure to detect multidomain remanence during Thellier-Thellier experiments. *Phyics and Chemistry of the Earth*, *28*, 681-687. doi:

10.1016/S1474-7065(03)00122-0

Kulakov, E., Sprain, C., Doubrovine, P., Smirnov, A., Patterson, G., Hawkins, L., ... Biggin, A. (2019). Analysis of an updated paleointensity database (Qpi-PINT) for 65-200 ma: Implications for the long-term history of dipole moment through the Mesozoic. *Journal of Geophysical Research: Solid Earth*, *124*, 9999-10,0222.

Lawrence, K., Tauxe, L., Staudigel, H., Constable, A., C.G. Koppers, McIntosh, W., & Johnson, C. (2009). Paleomagnetic field properties at high southern latitude. *Geochemistry Geophysics Geosystems*, *10*, Q001005 doi: 10.1029/2008GC002072.

Leonhardt, R., Matzka, J., & Menor, E. (2003). Absolute paleointensities and paleodirections of Miocene and Pliocene lavas from Fernando de Noronha, Brazil. *Physics of the Earth and Planetary Interiors*, *139*, 285-303 doi:10.1016/j.pepi.2003.09.008.

Mor, D. (1986). *The volcanism of the Golan Heights* (Tech. Rep. No. GSI/5/86). Isr. Geol. Surv.

Mor, D. (1993). A time-table for the Levant Volcanic Province, according to K-Ar dating in the Golan Heights, Israel. *Journal of African Earth Sciences*, *16*, 223-234 doi:10.1016/0899-5362(93)90044-Q.

Néel, L. (1949). Théorie du trainage magnétique des ferromagneétiques en grains fines avec applications aux terres cuites. *Annals of Geophysics*, *5*, 99-136.

Panovska, S., Constable, C., & Brown, M. (2018). Extending global continuous geomagnetic reild recfonstructions on timescales beyond human civilization. *Geochemistry Geophysics Geosystems*, *19*, 4757-4772. doi: 10.1029/2018GC007966

Paterson, G. (2011). A simple test for the presence of multidomain behavior during paleointensity experiments. *Journal of Geophysical Research*, *116*, doi:10.1029/2011JB008369.

Paterson, G., Tauxe, L., Biggin, A., Shaar, R., & Jonestrask, L. (2014). On improving the selection of Thellier-type paleointensity data. *Geochemistry Geophysics Geosystems*, *15*(4), 10.1002/2013GC005135.

Santos, C., & Tauxe, L. (2019). Investigating the accuracy, precision, and cooling rate dependence of laboratory acquired thermal remanences during paleointensity experiments. *Geochemistry Geophysics Geosystems*, *20*, 383-397. doi: 10.1029/2018GC007946

Selkin, P., & Tauxe, L. (2000). Long-term variations in paleointensity. *Phil. Trans. Roy. Soc. Lond.*, *358*, 1065-1088.

Shaanan, U., Porat, N., Navon, O., Weinberger, R., Calvert, A., & Weinstein, Y. (2011). OSL dating of a Pleistocene maar: Birket Ram, the Golan Heights. *Journal of Volcanology and Geothermal Research*, *201*(1-4), 397-403. doi: Doi10.1016/J.Jvolgeores.2010.06.007

Shaar, R., Ben Yosef, E., Ron, H., Tauxe, L., Agnon, A., & Kessel, R. (2011). Geomagnetic field intensity: How high can it get? how fast can it change? constraints from iron-age copper-slag. *Earth and Planetary Science Letters*, *301*, 297-306. doi: 10.1016/j.epsl.2010.11.013

Shaar, R., & Tauxe, L. (2013). Thellier GUI: an integrated tool for analyzing paleointensity data from Thellier-type experiments. *Geochemistry, Geophysics, Geosystems*, *14*, 677- 692 doi:10.1002/ggge.20062.

Shaar, R., Tauxe, L., Ron, H., Ebert, Y., Zuckerman, S., Finkelstein, I., & Agnon, A. (2016). Large geomagnetic field anomalies revealed in Bronze to Iron Age archeomagnetic data from Tel Megiddo and Tel Hazor, Israel. *Earth and Planetary Science Letters*, *442*, 173-185. doi: 10.1016/j.epsl.2016.02.038

Singer, B., jicha, B., Mochizuki, N., & Coe, R. (2019). Synchronizing volcanic, sedimentary, and ice core records of earth's last magnetic polarity reversal. *Science Advances*, *5*, eaaw4621. doi: 10.1126/sciadv.aaw4621

Tauxe, L., & Love, J. (2003). Paleointensity in Hawaiian Scientific Drilling project Hole (HSDP2): results from submarine basaltic glass. *Geochemistry Geophysics Geosystems*, *4*, doi: 10.1029/2001 GC000276.

Tauxe, L., Pick, T., & Kok, Y. S. (1995). Relative paleointensity in sediments; a pseudo-thellier approach. *Geophysical Research Letters*, *22*, 2885-2888.

Tauxe, L., Santos, C., Cych, B., Zhao, X., Roberts, A., Nagy, L., & Williams, W. (2021). Understanding non-ideal paleointensity recording in igneous rocks: Insights from aging experiments on lava samples and the causes and consequenes of 'fragile' curvature in Arai plots. *Geochemistry Geophysics Geosystems*, *22*, e2020GC009423. doi: 10.1029/2020GC009423

Tauxe, L., Shaar, R., Jonestrask, L., Swanson-Hysell, N., Minnett, R., Koppers, A. A. P., … Fairchild, L. (2016). Pmagpy: Software package for paleomagnetic data analysis and a bridge to the magnetics information consortium (magic) database. *Geochem. Geophys. Geosys.*, *17*. doi: 10.1002/2016GC006307

Tauxe, L., & Yamazaki, T. (2007). Paleointensities. In M. Kono (Ed.), *Geomagnetism* (Vol. 5, pp. 509-563, doi:10.1016/B978-044452748-6/00098-5). Elsevier.

Thellier, E., & Thellier, O. (1959). Sur l'intensité du champ magnétique terrestre dans le passé historique et géologique. *Annals of Geophysics*, *15*, 285-378.

van Zijl, J., Graham, K., & Hales, A. (1962). The paleomagnetism of the Stormberg Lavas, II. The behaviour of the magnetic field during a reversal. *Geophysical Journal International*, *7*, 169-182 doi:10.1111/j.1365-246X.1962.tb00366.x.

Wang, H., & Kent, D. (2013). A paleointensity technique for multidomain igneous rocks. *Geochemistry Geophysics Geosystems*, *14*. doi: 10.1002/ggge.20248

Wang, H., & Kent, D. (2021). RESET: A method to monitor thermoremanent alteration in Thellier-series paleointensity experiments. *Geophysical Research Letters*, *48*, e2020GL091617. doi: 10.1029/2020GL091617

Wang, H., Kent, D., & Rochette, P. (2015). Weaker axially dipolar time-averaged paleomagnetic field based on multidomain-corrected paleointensities from Galapagos Lavas. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, 15,036 – 15,041 doi:10.1073/pnas.1505450112.

Weinstein, Y. (2006b). A transition from strombolian to phreatomagmatic activity induced by a lava flow damming water in a valley. *Journal of Volcanology and Geothermal Research*, *159*, 267-284 doi: 10.1016/j.volgeores.2006.06.015.

Weinstein, Y., Navon, O., Altherr, R., & Stein, M. (2006a). The role of lithospheric mantle heterogeneity in the generation of Plio-Pleistocene alkali basaltic suites from NW Harat Ash Shaam (Israel). *Journal of Petrology*, *47*, 1017-1050 doi:10.1093/petrology/egl003.

Weinstein, Y., Nuriel, P., Inbar, M., Jicha, B., & Weinberger, R. (2020). Impact of the dead sea transform kinematics on adjacent volcanic activity. *Tectonics*, *39*, e2019TC005645. doi: 10.1029/2019TC005645

Weinstein, Y., Weinberger, R., & Calvert, A. (2013). High-resolution ^{40}Ar/^{39}Ar study of Mount Avital, northern Golan: reconstructing the interaction between volcanism and a drainage system and their impact on eruptive styles. *Bulletin of Volcanology*, *75*(5), 1-12.

Wilson, R. L. (1970). Permanent aspects of the Earth's non-dipole magnetic field over upper tertiary times. *Geophysical Journal of the Royal Astronomical Society*, *19*, 417-437.

Yamamoto, Y., & Tsunakawa, H. (2005). Geomagnetic field intensity during the last 5 Myr: LTD-DHT Shaw palaeointensities from volcanic rocks of the Society Islands, French Polynesia. *Geophysical Journal International*, *162*(1), 79-114.

Yamamoto, Y., Tsunakawa, H., & Shibuya, H. (2003). Palaeointensity study of the hawaiian 1960 lava: implications for possible causes of erroneously high intensities. *Geophys J Int*, *153*(1), 263-276.

Yamamoto, Y., & Yamaoka, R. (2018). Paleointensity study on the Holocene surface lavas on the Island of Hawaii using the Tsunakawa-Shaw method. *Frontiers in Earth Science*, *6*. doi: 10.3389/feart.2018.00048

Yu, Y., Tauxe, L., & Genevey, A. (2004). Toward an optimal geomagnetic field intensity determination technique. *Geochemistry Geophysics Geosystems*, *5*, 1-18 doi:10.1029/2003GC000630.

Ziegler, L., Constable, C., Johnson, C., & Tauxe, L. (2011). PADM2M: a penalized maximum likelihood model of the 0-2 ma paleomagnetic axial dipole moment. *Geophysical Journal International*, *184*, 1069-1089 doi:10.111/j.1365-246X.2010.04905.x.