Covariant Giant Gaussian Process Models with Improved Reproduction of Palaeosecular Variation

Richard K. Bono[1], Andrew J. Biggin[1], Richard Holme[1], Christopher J. Davies[2], Domenico G. Meduri[1], and Jack Bestard[1]

[1]Department of Earth, Ocean and Ecological Science, University of Liverpool, Liverpool L69 7ZE, UK
[2]University of Leeds, Leeds, UK

Key Points:

- We present GGP models for the past 10 Myr with improved reproduction of PSV and field strengths
- GGP models with spatial covariance are consistent with geodynamo simulations and observations
- Separating axial dipole variance improves simultaneous PSV and dipole fits

Corresponding author: Richard K. Bono, r.k.bono@liverpool.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2020GC008960

Abstract

A commonly used family of statistical magnetic field models are based on a giant Gaussian process (GGP), which assumes each Gauss coefficient can be realized from an independent normal distribution. GGP models are capable of generating suites of plausible Gauss coefficients, allowing for palaeomagnetic data to be tested against the expected distribution arising from a time-averaged geomagnetic field. But, existing GGP models do not simultaneously reproduce the distribution of field strength and palaeosecular variation estimates reported for the past 10 million years, and tend to under-predict virtual geomagnetic pole (VGP) dispersion at high latitudes unless trade-offs are made to the fit at lower latitudes. Here we introduce a new family of GGP models, *BB18* and *BB18.Z3* (the latter includes non-zero-mean zonal terms for spherical harmonic degrees 2 and 3). Our models are distinct from prior GGP models by simultaneously treating the axial dipole variance separately from higher degree terms, applying an odd-even variance structure, and incorporating a covariance between certain Gauss coefficients. Covariance between Gauss coefficients, a property both expected from dynamo theory and observed in numerical dynamo simulations, has not previously been included in GGP models. Introducing covariance between certain Gauss coefficients inferred from an ensemble of "Earth-like" dynamo simulations and predicted by theory yields a reduced misfit to VGP dispersion, allowing for GGP models which generate improved reproductions of the distribution of field strengths and palaeosecular variation observed for the last 10 million years.

Plain Language Summary

Earth's magnetic field varies on a continuous spectrum of time scales, ranging up to millions of years or longer. Being able to describe and predict these changes helps us understand the processes in Earth's core which give rise to the magnetic field. One way of understanding variations in the magnetic field is to use statistical models which assume that terms used to describe the magnetic field follow independent and identical Gaussian (bell-shaped) distributions, and that Earth's magnetic field averages to a dipole field with poles aligned to the geographic poles (the so-called "Geocentric Axial Dipole" field). However, such models do not simultaneously reproduce the variations in magnetic field direction and strength. We show that these models can be improved by using information on magnetic field behaviour from numerical simulations of the field generation processes. These new models are capable of improving reproduction of the variations of both magnetic field strength and directions, and will improve our ability to characterise the variability of Earth's magnetic field, apply corrections to sedimentary data where magnetic records may have been distorted by post-depositional compaction, and determine whether new data capture a sufficient interval of time to record the average magnetic field.

1 Introduction

Palaeomagnetic statistical field models are descriptions of the time-averaged magnetic field, typically presented as suites of spherical harmonic Gauss coefficients with assumed statistical properties. These models allow for straightforward determinations of the magnetic field and associated metrics, such as dispersion of magnetic directions or field strength distributions anywhere on the globe. The most common field models have previously assumed that the variation in Gauss coefficients can be described by a giant Gaussian process, where Gauss coefficients are normally distributed following a prescribed set of rules (e.g., Constable & Parker, 1988; Quidelleur et al., 1994; Constable & Johnson, 1999). These assume that Gauss coefficients are independently and identically distributed (*i.i.d.*) with (most) non-dipole terms having zero means and standard deviations such that power spectrum at core-mantle boundary is consistent with a white-noise source. A refinement on earlier GGP-style models, *TK03* (Tauxe & Kent, 2004), imposes an additional scaling term for the variance of Gauss coefficients describing the equatorially anti-symmetric field. GGP models have been applied to model secular variation and virtual axial dipole moment (VADM) distributions for palaeomagnetic studies across geologic history. Applications include assessing whether palaeomagnetic data from a given study record the expected amount of dispersion typical for the time-averaged field (as estimated following, e.g., Cox, 1970) and estimating the degree of inclination shallowing recorded in a sedimentary record by examining the observed elongation of directions compared against the directional elongation predicted by *TK03* (Tauxe & Kent, 2004).

Palaeosecular variation (PSV) characterizes how much Earth's field varies around a time-averaged position (often referred to as geocentric axial dipole, GAD) over some interval of time, typically of durations less than $\sim 10^7$ years (Johnson & McFadden, 2015). Assessing PSV requires estimates of the position of geomagnetic poles with respect to the spin axis. Palaeomagnetic observations are measured for individual sites (i.e., instantaneous records of the field at a specific location), which in studies of volcanic units are comprised of individual cooling units which share a similar location. Typically, these observations are reported as directions (declination, inclination) while full vector data are much rarer (due to the increased complexity and challenge in recovering these palaeointensities in the laboratory). To allow for comparison between palaeomagnetic observations from different sites, a geometric transformation of the Fisher (1953) mean palaeomagnetic direction to the geomagnetic pole is often performed (e.g., Butler, 1992). For instantaneous field records (i.e., "spot readings" capturing an instant in time much shorter than needed to average secular variation), this position is referred to as a virtual geomagnetic pole (VGP). The angular dispersion of VGPs (S), which can be used to characterize palaeosecular variation, is defined as:

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\Delta_i - \frac{S_{w_i}^2}{n_i}\right) \tag{1}$$

where Δ_i is the angle between the Fisher mean VGP (Fisher, 1953) and the i^{th} VGP for N sites, and $S_{w_i}^2/n_i$ is the portion of dispersion due to intra-site scatter for n_i samples. Palaeomagnetic analyses of PSV attempt to separate contributions to S from measurement or sample variation ($S_{w_i}^2/n_i$) and temporal variation, since temporal variation is the parameter of interest (see Johnson & McFadden, 2015). In this study, we focus on measures of S which exclude transitional VGPs, identified using the Vandamme (1994) iterative cut-off method, referred to as S_{VD}. A phenomenological model of VGP dispersion, termed Model G, was introduced by McFadden et al. (1988). In their model, the latitude dependence of VGP dispersion is attributed to a combination of equatorially anti-symmetric ("dipole"-family) and symmetric ("quadrupole"-family) terms, yielding a quadratic fit to data. While the dynamical basis relies on idealized dynamo behaviour (Merrill et al., 1996) and the explanatory power has been questioned (Doubrovine et al., 2019), Model G persists as a widely used approach to describe VGP dispersion data.

Hulot and Gallet (1996) show that spatial correlations between Gauss coefficients of the same spherical harmonic order m and shared membership in either symmetric or anti-symmetric families are expected on the basis of field symmetry arguments (Gubbins & Zhang, 1993). Hulot and Gallet (1996) provide the caveat that VGP dispersion alone is not sufficient to distinguish between a variance structure (e.g., anisotropic variance between odd and even terms) and a covariance structure. This was further investigated in Hulot and Bouligand (2005), who defined a covariance structure analytically compatible with the observed breaks in assumed symmetry properties of convective dynamos. Dynamo simulations reveal the predicted correlation pattern (Bouligand et al., 2005; Sanchez et al., 2019), which is expected for dynamos generated in a rotating spherical shell due to the interaction between a dominantly axially dipolar (odd) field and equatorially symmetric (even) core flow. In addition to these theoretical considerations, we show that the application of this correlation matrix, when converted to a covariance matrix with modelled variances assumed using the GGP framework, yields reduced misfit to S_{VD} estimates from the PSV10 dataset (Cromwell et al., 2018). These results suggest that this covariance is a fundamental statistical property of the geodynamo and motivates its inclusion in future GGP models.

Here we first describe GGP models (Section 2) and assess the semblance of selected existing GGP models with palaeomagnetic observations for the last 10 million years, with particular focus on the distribution of field strength estimates, VGP dispersion and magnitude of inclination anomalies (Section 3). Next, the observed covariance between Gauss coefficients from a wide range of numerical dynamo simulations is characterised, from which a mean correlation matrix (Section 4) and GGP model parameters (Section 5) are determined. With this covariance matrix, we introduce two new GGP models, *BB18* and *BB18.Z3*, that yield improved fits to the PSV10 dataset through the application of a prescribed covariance pattern inferred from dynamo simulations and theoretical considerations (Section 6). The first model, *BB18*, assumes that the mean value for all non-GAD terms is zero, while the second model, *BB18.Z3*, allows for non-zero-mean zonal terms to better fit the observed inclination anomaly estimates of PSV10. In the Supplementary Materials, alternative *BB18* models without covariance and variant *TK03* models are considered.

2 Giant Gaussian Process models

Constable and Parker (1988) introduced the first GGP model, *CP88*, which uses a small number of model parameters: mean axial dipole ($\overline{g_1^0}$), mean zonal quadrupole ($\overline{g_2^0}$), and an isotropic scaling term, α, which is used to define the standard deviation for each Gauss coefficient (σ_l^m, where l and m are spherical harmonic degree and order respectively, see equation 2). On its own, isotropic variance of the Gauss coefficients does not yield the observed latitude dependence of VGP dispersion. Quidelleur and Courtillot (1996) and Constable and Johnson (1999) adapted the GGP model by adjusting $\overline{g_2^0}$ and variance for $l = 2$ terms [and in the case of Constable and Johnson (1999), a different α and g_1^0 variance]. Through the introduction of anisotropic variances for degrees $l \leq 2$, a latitude dependence to VGP dispersion is achieved. A fundamental difference between *TK03* (Tauxe & Kent, 2004) and prior GGP models is the usage of a single anisotropic scaling factor, β, for $l - m$ odd terms. The model parameters are:

$$\sigma_l^2 = \frac{(R_c/R_E)^{2l}\alpha^2}{(l+1)(2l+1)} \tag{2}$$

$$\sigma_l^m = \sigma_l \quad \text{for } l - m \text{ even,} \tag{3}$$

$$\sigma_l^m = \beta\sigma_l \quad \text{for } l - m \text{ odd} \tag{4}$$

where R_c/R_E is the ratio between the Earth's core-mantle boundary and surface radii. In effect, GGP models prior to *TK03* have an implicit β of 1.

The models of Constable and Parker (1988) and Constable and Johnson (1999) assign a separate variance of the axial dipole (Table 1), whereas *TK03* uses the scaling terms of equations 2-4 to define σ_1^0. While the reduction of model parameters in *TK03* appeals to parsimony, the resulting simplification to the GGP yields statistical models which do not simultaneously reproduce the observed VGP dispersion and field strength estimates for the past 10 million years (see discussion in Section 3). This suggests that the separate treatment of the axial dipole variance, which is the primary term responsible for the distribution of virtual dipole moments (VDM), may be necessary. It is on this basis that our *BB18* models assign a separate variance for the g_1^0 term (Section 6).

Table 1. Model parameters of selected GGP models

	Model parameters										Misfit statistics					
Model	α	β	$\overline{g_1^0}$	σ_1^0	σ_1^1	$\overline{g_2^0}$	$\sigma_{g_2^1}$	$\sigma_{h_2^1}$	$\overline{g_3^0}$	*cov*	$\chi^2_{S_{VD}}$	$L^2_{S_{VD}}$	$\chi^2_{\Delta I}$	$L^2_{\Delta I}$	p_{KS}	D_{KS}
*Model G**	-	-	-	-	-	-	-	-	-	-	*93*	*2.4*	-	-	-	-
CP88	27.7	1	−30	*3.0*	*3.0*	−1.8	-	-	-	none	390	4.9	72	2.1	0	0.438
CJ98nz	15	1	−30	*11.72*	*1.67*	−1.5	*1.16*	*8.12*	-	none	330	4.5	73	2.1	0	0.306
TK03	7.5	3.8	−18	-	-	-	-	-	-	none	189	3.4	115	2.7	0	0.213
BB18	12.25	2.82	−22.04	*10.80*	-	-	-	-	-	$l \leq 4$	105	2.6	121	2.8	0.764	0.058
BB18.Z3	12.25	2.82	−22.04	*10.74*	-	−0.65	-	-	0.29	$l \leq 4$	103	2.5	70	2.1	0.471	0.073

Model parameters: "-" represent a scaled parameter following equations 2-4, italics denote a fixed parameter which would otherwise be scaled; α, β: scaling parameters following (Constable & Parker, 1988; Tauxe & Kent, 2004); $\overline{g_l^m}$: mean Gauss coefficient of degree l, order m; σ_l^m, $\sigma_{g_l^m}$, $\sigma_{h_l^m}$: standard deviation of specified Gauss coefficient(s); *cov*: covariance applied. All terms except β and *cov* are reported in μT. Misfit statistics: $\chi^2_{S_{VD}}$: misfit compared to PSV10 for S_{VD} data divided into 10° latitude bins; $L^2_{S_{VD}}$: normalised misfit of S_{VD} (Parker, 1994); $\chi^2_{\Delta I}$: misfit compared to PSV10 inclination anomaly estimates divided into 10° latitude bins; $L^2_{\Delta I}$: normalised misfit of ΔI; p_{KS}, D_{KS}: two sample Kolmogorov-Smirnov test p-value and test statistic comparing predicted VDM distribution to PINT10. *Model G only predicts S_{VD} values.

3 Comparing Extant GGP Models with Observations for the Past 10 Myr

Several global databases of palaeomagnetic data for directional analysis have been compiled for the past 5-10 million years (e.g., Lee, 1983; Quidelleur et al., 1994; Johnson & Constable, 1996) which have been used to construct GGP models. Of particular focus was the PSVRL (McElhinny & McFadden, 1997) database, which is an updated compilation of several previous datasets (e.g., Lee, 1983) and is used to constrain *TK03*. Cromwell et al. (2018) revisited this dataset in their compilation of the PSV10 dataset, applying new selection criteria to exclude lower quality data. In their analysis, when requiring that included data apply at least principal component analysis and step-wise experiments to determine remanence directions, only 12% of the PSVRL database meet these criteria. The inclusion of lower quality demagnetization data may bias resulting GGP models which will affect VGP dispersion predictions. We therefore used the PSV10 dataset, which compiled the results of 81 palaeomagnetic studies on volcanic units, representing 2401 total sites. An approximately global distribution of sites is achieved, albeit with a bias towards the Northern hemisphere; temporally, most of the sites were emplaced during the last two chrons, with some data extending to 10 Ma. VGP dispersion estimates, S_{VD}, applying the Vandamme cut-off technique and averaged into 10° latitudinal bins, from PSV10 are systematically higher than the S_{VD} estimates of Tauxe and Kent (2004) and *TK03* model predictions.

To estimate the field strength we use the Palaeointensity Database PINT [Biggin et al. (2009); updated Biggin et al. (2015)]. Broadly, the PINT database for the past 10 Myr mimics the spatio-temporal distribution of the PSV10 dataset and represents the best available database for estimating past field strength. We apply two mild quality filters to the approximately 2000 records for the past 10 million years. First, the experimental protocol should be capable of recognizing non-ideal recording potential (e.g., multidomain contribution or alteration); only studies reporting the following method codes were included: low-temperature Shaw method ("LTD-DHT-S"; Yamamoto & Tsunakawa, 2005), Low-temperature Thellier with partial thermoremanent (pTRM) tail checks ("LTD-T+"; Yamamoto et al., 2003), microwave technique with pTRM checks ("M+"; Shaw, 1974), Multi-Specimen Parallel Differential Technique ("MSPDp"; Dekkers & Böhnel, 2006), Shaw & Thellier ("ST+"), Thellier or variant with pTRM checks ("T+"; Thellier & Thellier, 1959), Thellier with pTRM checks and correction ("T+Tv"; Valet et al., 1996), Wilson (Wilson, 1961) & Thellier with pTRM checks ("WT+"). This reduces the dataset to ∼1350 records. Second, the number of intensity estimates per site mean (N_{int}) must be greater than or equal to 5. Applying both filters reduces the number of observations to 258 sites; however, we note that there are only subtle differences in the distribution of VDMs beyond the number of observations between the dataset filtered by N_{int} and by method alone. While a more thorough examination of the paleointensity record is needed (along with a more considered filtering procedure, such as the Q_{PI} method, Biggin & Paterson, 2014), we view this as a compromise between existing data reliability and availability. From this dataset, a median VDM can be determined for the past 10 Myr of 57 ZAm^2 with 95% confidence intervals (based on a bootstrap resampling) of 54 to 62 ZAm^2, somewhat higher than the 0 to 300 Ma average of Selkin and Tauxe (2000) used to define the mean field strength of *TK03*, but less than the mean field strength of ~ 82 ZAm^2 (Tanaka et al., 1995) used to define *CJ98nz*.

The availability of new data in the PSV10 dataset and substantial new contributions to the palaeointensity database in the last decade affords the opportunity to assess how well extant GGP models predict PSV and field strength behaviour. Three measures are used to compare with palaeomagnetic observations: distribution of VDMs; VGP dispersion grouped into 10° latitude bins; and inclination anomaly estimates (grouped in 10° latitude bins), ΔI (defined as the difference in inclination between an observation and the predicted inclination from a GAD field). In order to establish how well S_{VD} and ΔI are reproduced, a bootstrapping approach (Efron & Tibshirani, 1993) is used, in combination with a χ^2 metric which allows for weighting by observation variance:

$$\chi^2 = \sum_{i=1}^{N_b} \frac{(O_i - E_i)^2}{\sigma_i^2} \tag{5}$$

following the approach of Doubrovine et al. (2019). Here, O_i represents the i^{th} of N_b binned observations from the PSV10 dataset ($N_b = 16$), E_i represents the predicted value from a given field model for the parameter of interest (e.g., S_{VD}), and σ_i^2 is the variance of the i^{th} observed value, which is estimated from the 95% confidence intervals of the PSV10 estimates of S_{VD} and ΔI (assuming normally distributed uncertainties). To assess predicted VDM distributions, a two-sample Kolmogorov-Smirnov (KS) test (Massey, 1951) is applied between the distribution of PINT V(A)DMs and a distribution of VDMs realized from the given GGP model following the same spatial and temporal sampling as PINT. This yields a test statistic (D_{KS}), which measures the maximum absolute difference between each sample's corresponding empirical cumulative distribution functions (cdf), and its p-value (p_{KS}), which is the probability of observing a higher test statistic under the null hypothesis. We have chosen the following three GGP models for comparison whose model parameters are reported in Table 1: *CP88* (Constable & Parker, 1988), *CJ98nz* (Constable & Johnson, 1999), and *TK03* (Tauxe & Kent, 2004).

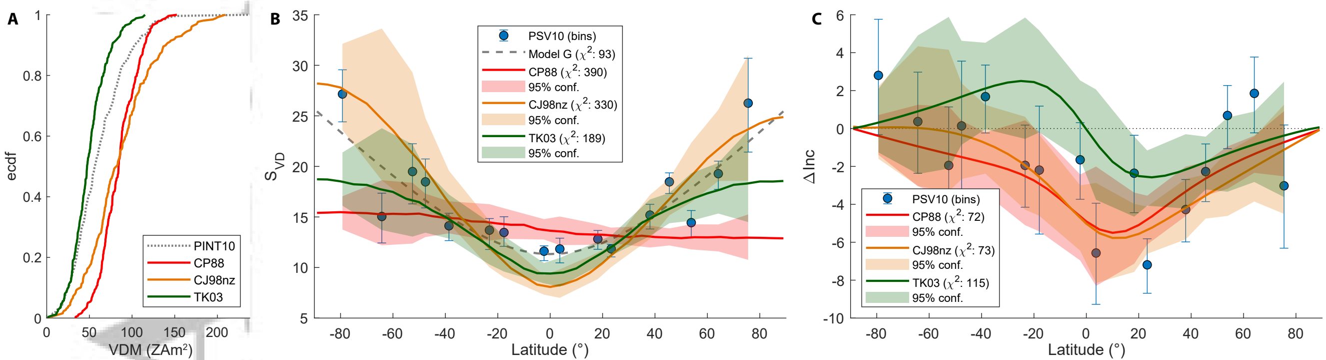


Figure 1. Predictions from three existing GGP models. A) Empirical cumulative density function of VDMs sampled similarly to PINT. B) VGP dispersion using Vandamme (1994) cutoff. Blue circles, PSV10 S_{VD} in 10° bins; dashed line, Model G-style fit of Doubrovine et al. (2019). C) Inclination anomaly predictions. Shaded regions in B and C show 95% confidence intervals from a bootstrap resampling reproducing the number of samples for each latitude bin of PSV10 (Cromwell et al., 2018). χ^2 values reported in the legend.

We show that these models are not able to simultaneously reproduce both PSV and PINT observations for the past 10 million years (Table 1). Distributions of VDMs from all three GGP models yield vanishingly small p_{KS}, suggesting that the PINT10 dataset and these GGP models sample significantly different distributions (Table 1, inferred in Figure 1A). GGP models which yield good fits to VGP dispersion data (low χ^2), such as *TK03*, do not also yield low χ^2 values when considering inclination anomaly (albeit with predominantly overlapping 95% confidence intervals when sampled similarly to the PSV10 data set, Figure 1B-C). Qualitatively it can be seen that even when considering the 95% confidence intervals of VGP dispersion *CP88*, *CJ98nz* and *TK03* over- or under-predict low-latitude to equatorial VGP dispersion, and with the exception of *CJ98nz*, also under-predict high latitude VGP dispersion. Inclination anomaly is less straightforward to assess, due the higher uncertainty in the PSV10 dataset, however, a prominent feature of the PSV10 record is a hemispheric asymmetry between northern and southern hemispheres. The GGP model which best reproduces the VGP dispersion, *TK03*, yields a symmetric inclination anomaly trend because of the GAD assumption used in its formulation, which is inconsistent with the PSV10 inclination anomaly trend (see Section 8 for discussion on estimation of inclination anomalies).

4 Characterising dynamo covariance and the application to GGP models

We consider possible inferences from 21 dynamo simulations which demonstrate "Earth-like" time-averaged behaviour following the Q_{PM} framework of Sprain et al. (2019), here defined as having misfit values of $\Delta Q_{PM} \leq 10$ and a $\tau_t < 0.15$ (where τ_t is the fraction of the total integration time when the absolute dipole latitude is $< 45°$). While we did not explicitly filter simulations by dipolarity, f_{dip}, we wanted to exclude multipolar simulations. Assigning a threshold which delineates stable dipolar from multipolar dynamos is not clear (Christensen & Aubert, 2006; Wicht & Tilgner, 2010; Wicht et al., 2015). Instead, we apply the Q_{PM} framework (Sprain et al., 2019) to define "Earth-likeness" based on palaeomagnetic observations, which uses clearly defined thresholds. We note that the lowest f_{dip} of the simulations selected is 0.28, which is close to, but not clearly within, the multipolar solutions described by Oruba and Dormy (2014). This minimum f_{dip} brackets the multipolar thresholds of Christensen and Aubert (2006) (of 0.35) and Wicht et al. (2015) (of $\sim$0.20), and we therefore consider that Q_{PM} effectively filters out multipolar states without excluding rarely reversing, "Earth-like" dynamos. The Q_{PM} framework compares five measures of PSV between observational data for the past 10 million years and a given numerical dynamo simulation: equatorial VGP dispersion and the latitude dependence of VGP dispersion (through the Model G-style fit of S_{VD}), maximum absolute inclination anomaly, proportion of time spent in transtion (i.e., absolute dipole latitude $< 45°$), including the presence/absence of reversals and VDM variability (VDM inter-quartile range normalized by median VDM). For a complete description of the Q_{PM} framework, see Sprain et al. (2019).

Dynamo simulations used in this analysis (Table 2) have been reported previously (Davies & Constable, 2014; Wicht & Meduri, 2016; Sprain et al., 2019), and are thus only described briefly. These simulations were integrated for at least 4 magnetic diffusion times, representing 300 to 600 kyr dependent on the choice of thermal conductivity (Pozzo et al., 2012; Konôpková et al., 2016), and include both reversing and non-reversing cases.

The simulations consider a convecting, electrically conducting fluid under the Boussinesq approximation, with no-slip boundary flow conditions, and consider an electrically insulating mantle while the inner core is either insulating or conducting. Fixed heat flux or temperature are prescribed at the inner core and core-mantle boundaries. In some simulations a lateral heat flux pattern was imposed at the core-mantle boundary [see Sprain et al. (2019) for additional details]. Simulations previously described by Wicht and Meduri (2016) explored both thermally and purely chemically driven dynamos under different input parameters, while the dynamos of Davies and Constable (2014) and Sprain et al. (2019) were solely thermally driven. Following the definitions of Davies and Gubbins (2011), all dynamo simulations used Ekman numbers spanning 1.2×10^{-4} to 3×10^{-3}, Rayleigh numbers spanning up to 100 times the critical value for non-magnetic convection, and magnetic Prandtl numbers ranging from 2 and 20. The dynamo simulations of Wicht and Meduri (2016) have not been previously assessed under the Q_{PM} framework, thus we have included the relevant Q_{PM} statistics in Supplementary Tables S1 and S2.

We determined the Pearson correlation coefficient (ρ) of all pairs of Gauss coefficients for the dynamo simulations, which are sampled at $\sim$ 10000 year steps, to reduce possible contributions due to auto-correlation (Bouligand et al., 2005). From the 21 dynamo simulation correlation matrices, we determined a mean correlation matrix, $\bar{\rho}$, presented in Figure 2 up to degree 4, with relevant terms reported in Table 3. We find that Gauss coefficients of the same order m and membership to either symmetric ($l-m$ even) or anti-symmetric ($l-m$ odd) families are correlated, consistent with prior descriptions of dynamo covariance (Bouligand et al., 2005; Sanchez et al., 2019); otherwise, correlations cluster close to zero, suggesting that other pairs of Gauss coefficients are independent. The amplitude of correlated terms varies between dynamo simulations; whether any systematic variation in correlation coefficient amplitude can be associated with dynamo control parameters was not explicitly explored in this study (Supplemental Figure S1 shows the variation of selected correlation coefficients for simulations included in this study).

The mean correlation matrix $\bar{\rho}$ is used to define a covariance matrix (Σ) for our new GGP models *BB18* and *BB18.Z3* by scaling correlation with a predefined variance for each Gauss coefficient (equations 2-4, except for the g_1^0 variance discussed below):

$$\Sigma_{ij} = \sigma_i \sigma_j \bar{\rho}_{ij} \tag{6}$$

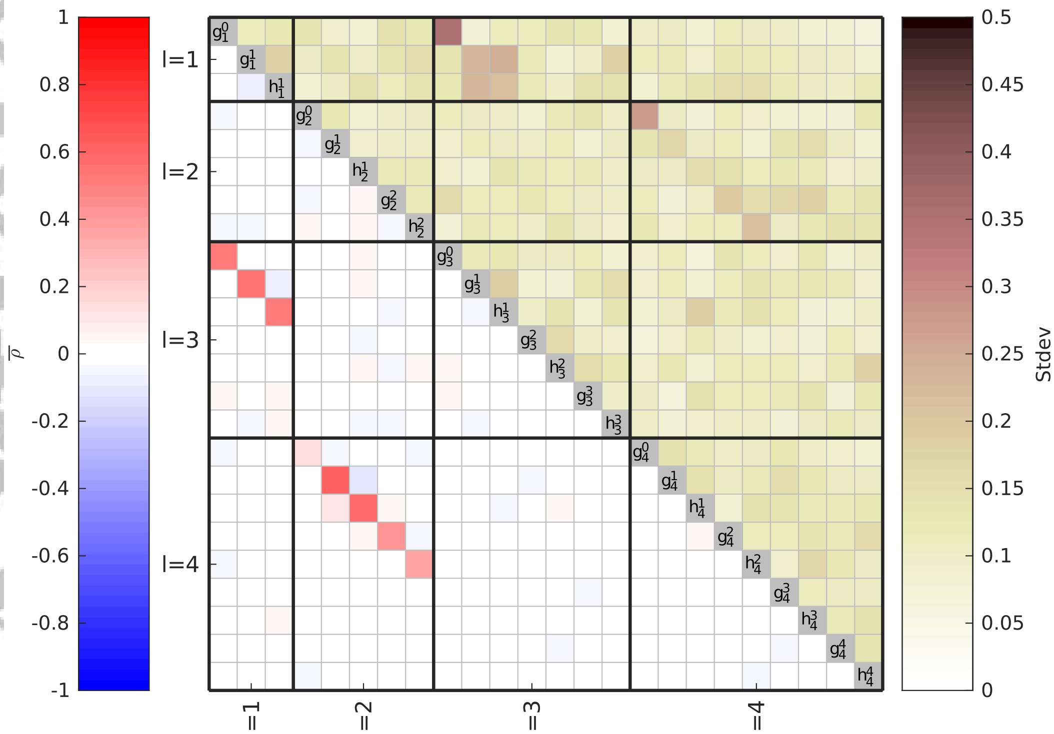


Figure 2. Mean and standard deviation of correlation coefficients determined from dynamo simulations considered in this study (n=21). Gauss coefficients are listed in the following sequence: $g_1^0, g_1^1, h_1^1, ..., g_l^m, h_l^m$, up to spherical harmonic degree $l \leq 4$. The matrix is symmetric; only one triangle is shown, with diagonal terms ($\rho = 1$ by definition) coloured grey. Lower triangle: mean correlation coefficients ($\bar{\rho}$); upper triangle: standard deviation of correlation coefficients for all simulations.

where σ is the standard deviation for each Gauss coefficient, and i and j refer to individual Gauss coefficients. Sensitivity testing suggests that covariances are required for degrees $l \leq 4$, with no substantial change to the latitude dependence of VGP dispersion when covariances with degree $l = 5$ and higher terms are also included. While the covariance matrix applied to the *BB18* models is restricted to spherical harmonic degrees 4 and lower on the basis of parsimony, we note that similar to the studies of Bouligand et al. (2005) and Sanchez et al. (2019), the covariance pattern observed in our simulations extends to all degrees examined.

Table 2. Dynamo simulations selected for defining *BB18* covariance

Name	E($\times 10^{-4}$)	Ra	Pm	BBC	TBC	Conv.	HBC	ϵ	τ	Rm	τ_t	f_{dip}	ΔQ_{PM}	Reference
Model 4	5	350	5	II	FF	T	N	0	13	226.0	0.112	0.28	7.0	Sprain et al. (2019)
Model 5	5	400	5	II	FF	T	N	0	14	226.7	0.114	0.28	8.1	Sprain et al. (2019)
Model 6	5	250	10	II	FF	T	N	0	5	326.8	0.003	0.34	6.2	Sprain et al. (2019)
B2*	5	200	10	II	TF	T	N	0	9	326.1	0.026	0.38	6.6	Sprain et al. (2019)
Model 11	5	400	5	II	TF	T	N	0	6	258.2	0.053	0.31	7.2	Sprain et al. (2019)
Model 19	5	100	10	II	TF	T	R	1.5	4	218.6	0	0.48	6.5	Sprain et al. (2019)
Model 20	5	100	10	II	FF	T	R	1.5	4	210.3	0	0.52	5.6	Sprain et al. (2019)
C1-4	1.2	100	2	CI	TF	T	N	0	4	264.4	0	0.64	8.2	Davies and Constable (2014)
C3-3	1.2	50	2	CI	TF	T	N	0	10	102.7	0	0.71	8.4	Davies and Constable (2014)
Model 30	10	60	10	II	TF	T	N	0	19	118.9	0	0.62	8.3	Sprain et al. (2019)
Model 31	10	70	10	II	TF	T	N	0	14	134.1	0	0.60	7.6	Sprain et al. (2019)
Model 32	10	90	10	II	TF	T	N	0	13	160.6	0	0.57	7.0	Sprain et al. (2019)
Model 51	5	100	20	II	TF	T	N	0	4	332.2	0	0.49	8.4	Sprain et al. (2019)
E4R53C	1.5	1500	3	CI	TF	C	N	0	11	264.0	0	0.66	9.3	Wicht & Meduri (2016)
E4R78C	1.5	2250	3	CI	TF	C	N	0	*37*	340.0	0	0.59	6.2	Wicht & Meduri (2016)
E4R106C	1.5	3000	3	CI	TF	C	N	0	87	408.0	0.064	0.40	3.2	Wicht & Meduri (2016)
E3R23C	5	625	10	CI	TF	C	N	0	431	442.0	0.080	0.30	5.6	Wicht & Meduri (2016)
E3R5	5	125	10	CI	TT	T	N	0	*935*	202.0	0	0.60	6.4	Wicht & Meduri (2016)
E3R7	5	200	10	CI	TT	T	N	0	*58*	350.0	0	0.44	5.8	Wicht & Meduri (2016)
E3R8	5	225	10	CI	TT	T	N	0	*87*	393.0	0.007	0.38	5.9	Wicht & Meduri (2016)
E3R9	5	250	10	CI	TT	T	N	0	*693*	436.0	0.051	0.31	4.8	Wicht & Meduri (2016)

Columns two to four detail the input model parameters which are: the Ekman number $E = \nu/2\Omega d^2$ where ν is the fluid kinematic viscosity, Ω the shell rotation rate and d the shell gap; In thermally driven dynamos, the Rayleigh number is $Ra = \alpha g \Delta T d/2\Omega\kappa$ where α and κ are the fluid thermal expansivity and thermal diffusivity respectively, g is gravity at the outer boundary and ΔT denotes a temperature scale that depends on the specified boundary conditions and heating mode (see Davies and Constable (2014); Sprain et al. (2019)). Chemically driven dynamos employ a standard codensity formulation (Wicht and Meduri (2016)) and the Rayleigh number is $Ra = g\Delta C d/2\Omega\kappa$ where ΔC is the codensity jump across the shell. Note that the two definitions of Ra coincide when considering $\Delta C = \alpha\Delta T$. In all cases the shell aspect ratio is 0.35 and the Prandtl number $Pr = \nu/\kappa = 1$. BBC: magnetic boundary conditions, "I" for insulating, "C" for conducting, first letter for inner core boundary, second letter for core-mantle boundary; TBC: thermal boundary conditions, "F" for fixed heat flux, "T" for fixed temperature, first letter for inner core boundary, second letter for core-mantle boundary; Conv.: convection type, "T" for thermally driven convection, "C" for chemically driven convection; HBC: heterogeneous thermal boundary condition, "N" for none, "T" for tomographic boundary after Masters et al. (1996), "R" for recumbent Y_2^0 following Dziewonski et al. (2010); ϵ: amplitude of heterogeneous thermal boundary condition following Sprain et al. (2019); τ: simulation duration reported in outer core magnetic diffusion times, italicized values have different durations than reported in the original study; Rm: magnetic Reynolds number; τ_t: proportion time in a transition state, following Sprain et al. (2019); f_{dip}: time averaged ratio of the mean dipole field strength to the field strength in degrees $l \leq 12$ evaluated at the core mantle boundary; ΔQ_{PM}: total misfit of simulation to Earth's time-averaged field, following Sprain et al. (2019). *N.B., model B2 was previously reported in Sprain et al. (2019) as Model 7 erroneously.

5 Model construction

Our strategy to determine model parameters considered here ($\overline{g_1^0}$, α, β, σ_1^0, and zonal terms for the non-GAD model) was to apply an iterative approach to find the best-fitting values which minimises $\chi^2_{S_{VD}}$, $\chi^2_{\Delta I}$, and D_{KS}. We estimated the model parameter $\overline{g_1^0}$ directly from PINT (Section 5.1). Next, we determined α and β terms which yield the lowest misfit to PSV10 (Section 5.2). We then determined the variance of g_1^0 which best reproduces the distribution of VDMs (Section 5.3). For models with non-zero-mean zonal terms, $\overline{g_2^0}$ and $\overline{g_3^0}$ were determined using a multi-objective genetic algorithm (Deb & Kalyanmoy, 2001) to minimise total power at the core-mantle boundary and misfit to PSV10 (Section 5.4).

5.1 Estimating g_1^0 Mean

From the PINT dataset, we estimated a $\overline{g_1^0}$ term which would yield the observed median VDM using the following equation:

$$g_1^0 = \frac{\mu_0 \mathrm{VDM}}{4\pi {R_E}^3} \tag{7}$$

where μ_0 is magnetic permeability of free space. Here we assume the median VDM can be used to approximate the mean g_1^0 for the time-averaged field (under the assumptions that the VDM is entirely described by the dipole field and that the time-averaged equatorial terms are zero-mean). The assumption that $\overline{g_1^0}$ can be approximated by the median VDM is not strictly accurate because VDMs include all non-axial-dipole contributions and the distribution is not Gaussian. However, for a reasonably dipolar field (i.e., $\overline{g_1^0} > 10$ μT, $\overline{g_2^0} < 3$ μT and $\alpha < 30$ μT), the amount of over-estimation due to these assumptions is small, and for the chosen $\overline{g_1^0}$ we estimate the possible misfit to be $< 1\mu T$ (Supplementary Materials Section S1, Figure S3). We determined an estimated $\overline{g_1^0}$ of -22.04 μT.

5.2 GGP Model Minimisation

Our minimisation approach applied the following procedure: generate a *TK03*-style model, varying α, β and $\overline{g_1^0}$ and compare the S_{VD} at the equator with the Model G a parameter of the PSV10 data as fit by Doubrovine et al. (2019), which acts as an estimate of equatorial S_{VD} ($S_{VD}^{D19}(\lambda = 0)$). While the Model G fit to PSV10 data does not satisfy the strict statistical threshold defined by Doubrovine et al. (2019) to predict S_{VD}, we feel that the estimation of minimum S provided by Model G is a good proxy for equatorial VGP dispersion. We note that currently no GGP model considered here (or even Model G-style fit) adequately reproduces the PSV10 observations. This can be shown using a normalised χ^2 misfit, L^2 , defined as the χ^2 misfit divided by the number of observation bins (N_b) (Parker, 1994); here, the expected $L^2 \sim 1$ is not achieved by any model (Table 1), which is consistent with the observations of (Doubrovine et al., 2019) of Model G-style fits. Untangling the contributions to misfit from biases in the PSV10 dataset and issues inherent in GGP models is non-trivial; however, it is clear that the PSV10 dataset may contain some biases which affect model construction (cf. ΔI in PSV10 vs. Behar et al. (2019)). The approach we have taken prevents the biasing of S by individual studies which may be affected by unrecognised tectonic effects (Opdyke et al., 2015). For a given set of α, β and $\overline{g_1^0}$, the S_{VD} at the equator ($S^m_{VD}(\lambda = 0)$) was determined and the square of the residual (ERS) between $S^{D19}_{VD}(\lambda = 0)$ and the estimated $S^m_{VD}(\lambda = 0)$ was calculated.

We find there is a clear relationship between $\overline{g_1^0}$ and α (for a given value of β, explored here from 1 to 5) which describes the relative variance of non-axial dipole terms assuming zero-means (Supplementary Figure S2). This allowed us to construct a model where $\overline{g_1^0}$ is specified as an input from a prescribed distribution. For a specified $S_{VD}(\lambda = 0)$ and $\overline{g_1^0}$, β remains to be constrained, since α is dependent on $\overline{g_1^0}$ and β terms. Here, the β term which minimizes $\chi^2_{S_{VD}}$ was chosen for the *BB18* models (2.82, Table 1).

5.3 Estimating g_1^0 Variation

The standard deviation of g_1^0 is estimated through minimising D_{KS} across *BB18* models while varying the standard deviation of g_1^0 (Supplementary Figure S4). Here, we account for the contribution of non-axial-dipolar fields through approximating the variance of the non-g_1^0 terms through α and β, which can be estimated through comparison with PSV data. Uncertainty in PINT data is approximated by including a Gaussian-distributed noise term which approximates the median percent error of the PINT dataset ($\delta F\%n$ = 15%, the true standard deviation accounting for sample size, Paterson et al. (2010)).

The choice of σ_1^0 is dependent on the assumption of how much noise is present in the palaeointensity record. The use of a normally distributed noise term almost certainly under-predicts noise at lower field strengths. Here, we chose 15% noise based on the median percent error in the PINT dataset, however, it is conceivable that up to 20% noise is possible, which would reduce the model parameter σ_1^0 correspondingly (Supplementary Figure S4). This yielded a best-fitting σ_1^0 of $\sim$10.8 μT.

5.4 Zonal Non-zero-mean Terms

For defining *BB18.Z3*, which includes zonal terms with non-zero-mean values, a multi-objective genetic algorithm (Deb & Kalyanmoy, 2001) was employed to search for global minima in residuals. Here, three objective functions were independently defined: sum of squared error (SSE) between a given model and the PSV10 S_{VD} dataset, the SSE for the ΔI dataset, and spectral power (Lowes, 1974) at the core-mantle boundary for spherical harmonic degrees 2 through 10 (W), defined as follows:

$$SSE_{S_{VD}} = \sum_{i=1}^{N_b} \left(S_{VD,i} - S_{VD,i}^{PSV10}\right)^2 \tag{8}$$

$$SSE_{\Delta I} = \sum_{i=1}^{N_b} \left(\Delta I_i - \Delta I_i^{PSV10}\right)^2 \tag{9}$$

$$W = \sum_{l=2}^{10} \left(\frac{R_E}{R_c}\right)^{2(l+2)} \sum_{m=0}^{l} [(g_l^m)^2 + (h_l^m)^2] \tag{10}$$

where $S_{VD,i}$ (and ΔI_i) and $S_{VD,i}^{PSV10}$ (and ΔI_i^{PSV10}) are the estimates for the i^{th} of N_b latitude bins for the GGP model and PSV10, respectively. Previously determined model parameters for *BB18* ($\overline{g_1^0}$, σ_1^0, α and β) are retained. Because dipole mean and variance are not adjusted in minimising higher order terms, the misfit between modelled VDM distributions and the PINT dataset is not considered in this analysis.

The first two objectives (equations 8-9) represent our desired model predictions of PSV behaviour, while the third objective (equation 10) yields models consistent with a white-noise source at the CMB for degrees $l > 1$ (Figure 3). Since no single solution exists which minimizes all three defined objectives, a set of solutions can be found beyond which no further minimization in one objective can be achieved without increasing another objective (effectively, a trade-off "surface", Supplementary Figure S5). From this set of solutions, the pair of zonal terms which yields a minimum to the sum of misfit for S_{VD} and ΔI is chosen, i.e., the 'knee' of the trade-off relation between $SSE_{S_{VD}}$

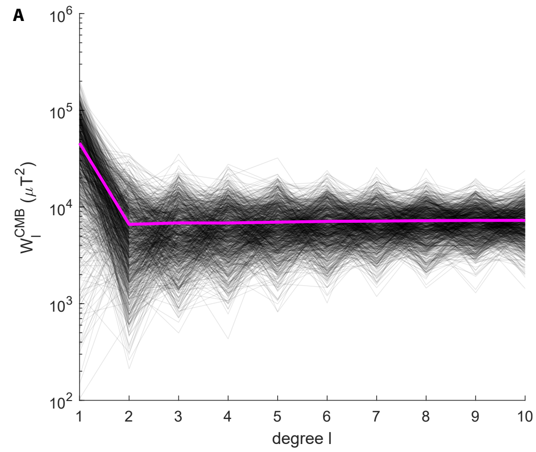


Figure 3. Power spectra at the core-mantle boundary (Lowes, 1974) of 1000 realisations of *BB18.Z3* (black lines). Magenta line shows the mean power spectrum for *BB18.Z3*.

and $SSE_{\Delta I}$ from the set of solutions where W has already been minimised. Solutions including zonal non-zero terms for spherical harmonic degrees 2 and 3 were explored.

6 New BB18 GGP Models and Methods

Presented here are two new GGP models named *BB18* and *BB18.Z3*. *BB18* assumes all non-axial dipole terms have a mean of zero, whereas *BB18.Z3* allows for non-zero means for the g_2^0 and g_3^0 Gauss coefficients. Both models introduce a covariance pattern (Σ) informed from dynamo simulations correlation matrix $\bar{\rho}$.

The resulting Gauss coefficients are drawn from a multivariate normal distribution with the probability density function (P):

$$P = \frac{1}{\sqrt{|\Sigma|(2\pi)^k}} \exp\left(-\frac{1}{2}(c_l^m - \overline{c_l^m})\Sigma^{-1}(c_l^m - \overline{c_l^m})^T\right) \tag{11}$$

where $k = 120$, the number of Gauss coefficients, and c_l^m and $\overline{c_l^m}$ are Gauss coefficients and their means. For spherical harmonic degrees 1-4, the observed covariance pattern from dynamo simulations introduced in the prior section is applied (Figure 2), and for higher degrees ($5 \leq l \leq 10$) no covariance is applied (i.e., independence). The specified model parameters are detailed in Table 1 with non-zero correlation coefficient terms reported in Table 3. Alternative *BB18* family GGP models without covariance were also explored (Supplementary Materials Section S2, Tables S3-S4 and Figure S6). While the

Table 3. Mean correlation coefficients ($\bar{\rho}$) for select terms determined from dynamo simulations

	(g_1^0, g_3^0)	(g_1^1, g_3^1)	(h_1^1, h_3^1)	(g_2^0, g_4^0)	(g_2^1, g_4^1)	(h_2^1, h_4^1)	(g_2^2, g_4^2)	(h_2^2, h_4^2)
$\bar{\rho}$	0.51	0.55	0.53	0.14	0.60	0.58	0.42	0.37

variant models yielded improved fits to PSV10 relative to extant GGP models, the addition of a covariance structure results in better fits overall, as we shall show in the following section.

7 Results

Considering all three metrics of resemblance between GGP models and the palaeomagnetic field for the past 10 million years, *BB18* models simultaneously achieve quantifiable improvements over prior GGP models. The *BB18* models are able to reproduce the median VDM and distribution observed in the PINT data, yielding a $p_{KS} \gg 0.05$, suggesting the null hypothesis that *BB18* and PINT sample the same underlying distribution cannot be rejected (Table 1, Figure 4A), which earlier GGP models do not. VGP dispersion (S_{VD}) predictions from the *BB18* models yield improved fits to the PSV10 dataset, as measured by equation 5 (Figure 4B, Table 1) and produce predictions with confidence intervals which contain all VGP dispersion estimates (and all but one inclination anomaly estimate, Figure 4). Between *BB18* and *BB18.Z3*, we see a small improvement in goodness of fit for *BB18.Z3*, likely due to the slight hemispheric asymmetries that the non-GAD zonal terms introduce.

Much of the improvement in fit in the *BB18* models, with respect to existing GGP models considered here, can be seen at the highest latitudes, which are less well sampled relative to lower latitudes (and thus do not contribute as much in the χ^2 metric). Prior GGP models yield S_{VD} curves with a prominent difference to Model G: an inflection point at some mid-latitude point which moves towards the equator as the difference in S_{VD} at the equator versus high latitudes increases, whereas Model G has no inflection point. Introducing a covariance matrix to the GGP models (i.e., *BB18*-family) reduces the effect of this inflection point while still yielding a latitude dependence in VGP dispersion. In the Supplementary Materials Section S2, variants of *BB18* without covariance are explored. While these models yield improved fits relative to existing GGP models, *BB18* models with covariance presented here have lower χ^2 values and visually improved fits at high latitudes to *BB18* models without covariance.

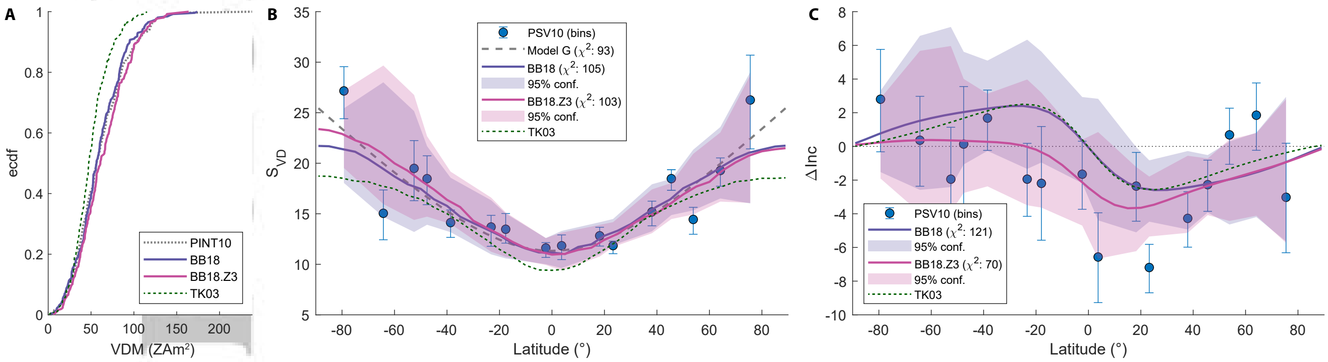


Figure 4. *BB18* model predictions, with *TK03* (Tauxe & Kent, 2004) for comparison, following the style presented in Figure 1.

BB18.Z3 (which includes non-zero-mean zonal degree 2 and 3 terms) yields comparable χ^2 values to *CP88* and *CJ98nz* when compared to the PSV10 (Cromwell et al., 2018) inclination anomaly estimates (Figure 4C, Table 1). The *BB18* model assumes a time-averaged GAD field and yields higher χ^2 values relative to existing GGP models. Similarly, *TK03* also assumes GAD, and yields a somewhat lower χ^2 than *BB18* (but substantially higher than GGP models with zonal terms, including *BB18.Z3*), however, *TK03* sacrifices goodness of fit for VGP dispersion.

8 Discussion

The *BB18* family of GGP-style models provides a flexible framework for the generation of statistical field models, which incorporates the correlation pattern observed in dynamo simulations to improve PSV predictions. Prior GGP models are able to reproduce some aspects of the palaeomagnetic field, but are unable to simultaneously reproduce all three metrics considered in this study (VDM distribution, VGP dispersion and inclination anomaly). While the *BB18* models, like prior GGP models considered here, are unable to satisfy the L^2 normalisation expectation, the specific models presented here, *BB18* and *BB18.Z3*, yield predictions which are in closer agreement with PSV data for the past 10 Myr as compiled by PSV10 than GGP models considered here while also reproducing the VDM distribution of the PINT dataset.

While other statistical properties beyond correlation of Gauss coefficients are available from dynamo simulations, here we chose to only incorporate the correlation pattern. To first order, *BB18* models reproduce the VGP dispersion of PSV10 better than dynamo simulations, however, it is worth acknowledging that simulations are able to reproduce some salient features of PSV behaviour (e.g., latitude-dependent VGP dispersion; Lhuillier & Gilder, 2013; Sprain et al., 2019). Furthermore, we are not aware of any dynamo simulation that reproduces the hemispheric asymmetry of inclination anomalies observed in PSV10 (albeit with some models reproducing the amplitude of the peak inclination anomaly observed). With respect to the mean values and standard deviations of Gauss coefficients, the appropriate scaling law to relate dimensionless simulation values to physical units remains an open question, with different scaling approaches yielding strengths which can vary substantially (Christensen & Wicht, 2015); this also precludes the determination of model parameter α.

Conceivably, the model parameter β could be determined directly from dynamo simulations, and indeed, in the suite of dynamo simulations considered, we do observe larger variances of $l - m$ odd terms relative to $l - m$ even terms of the same degree, yielding β values > 1. Estimated β terms from simulations were all < 2, below the β terms found to best fit PSV10 observations. Closer inspection of the ratio of odd to even Gauss coefficient variance within each degree suggests greater complexity than modelled in the GGP framework (i.e., differences in variance within a degree, violating the assumption of identical distributions of Gauss coefficients), however, the source of this complexity and whether this behaviour is found in Earth's magnetic field are beyond the scope of this study (Supplementary Figures S7–S8). Because of the assumptions and complexities associated with directly importing additional statistical behaviour from dynamo simulations, we have employed a conservative approach of modifying the GGP approach as little as possible while still capturing what we think are fundamental dynamo characteristics.

We note that earlier GGP models of Constable and Johnson (1999) are also able to yield high VGP dispersion predictions at high-latitudes, albeit with under-predictions of equatorial VGP dispersion. The high-latitude dispersion is due to the additional variance given to the g_2^1 and h_2^1 terms [an observation also made by Quidelleur and Courtillot (1996)]. By contrast, *BB18* models achieve increases in high-latitude VGP dispersion due to the positive correlation between g_2^1 and g_4^1 (h_2^1 and h_4^1) terms. As mentioned previously, Hulot and Gallet (1996) identified both the significance of order $m = 1$ terms to the latitude dependence of VGP dispersion, as well as the inability to distinguish between contributions from variance and covariance. While previous GGP models have improved fit to VGP dispersion through directly adjusting the σ_2^1 terms, the improved fit to data by *BB18* models is achieved through a process which is consistent with the observed behaviour of dynamos in numerical simulations (i.e., covariance).

BB18 models reproduce the distribution of VDMs observed for the past 10 million years without sacrificing fit to PSV measures, in contrast with existing GGP models considered in this study (Table 1). This outcome can be achieved by adjusting the mean and variance of the axial dipole. However, simply adjusting the model parameters of a *TK03*-style model is not sufficient (see Supplementary Materials Section S3, Figure S10). By reintroducing a separate model parameter for the variance of the axial dipole term, decoupling σ_1^0 from the variances of the other Gauss coefficients (which are determined by α and β), the observed VDM distribution can be reproduced. The increased variance of the axial dipole term in *BB18* models is consistent with the observations of Constable and Johnson (1999). We also find that, visually, *BB18* models are capable of reproducing the variation and mean trend observed in the PINT dataset of palaeointensity versus latitude (Supplementary Figure S9). We note that there are a few caveats to the assumptions made in determining model parameters with respect to VDM observations (Section 5.1, 5.3). In our efforts to estimate $\overline{g_1^0}$, we chose not to increase the complexity of our model by accounting for the potential bias when converting from VDMs, given the likelihood of additional, unaccounted for sources of error.

The third metric used in this study, the pattern and amplitude of inclination anomalies, requires additional consideration. In our study, inclination anomaly predictions from GGP models are treated in the same manner as palaeomagnetic data in the PSV10 dataset; specifically, inclinations are determined using unit vector magnetic directions and subsequently binned into 10° latitude groups. In the PSV10 dataset, two salient observations suggest that the observed inclination anomalies represent persistent non-GAD field contributions for the past 10 million years: there is a pronounced asymmetry between Northern and Southern hemisphere inclination anomaly estimates, and the maximum observed inclination anomaly is greater than 5°. These features are reproduced in early GGP models (Constable & Parker, 1988; Constable & Johnson, 1999), which used a different palaeomagnetic dataset than PSV10, assuming a small (∼1-2 μT) quadrupole contribution. Prior studies (Constable & Parker, 1988; Quidelleur & Courtillot, 1996; Constable & Johnson, 1999) empirically determined zonal terms g_2^0 on the order of 1-5% of g_1^0, with one analysis by Muxworthy (2017) suggesting an octupole contribution of $\sim$15%. We reproduce the observed inclination anomaly asymmetries through the contribution of small zonal quadrupole and octupole mean terms, <5% the strength of g_1^0, in *BB18.Z3*, which is our preferred model (Table 1).

When full vector magnetizations recording a GAD field are considered, no inclination anomalies are expected, however, due to the latitude dependence of both inclination and field strength, the treatment of magnetic directions as unit vectors results in small ($\sim$2-5°) inclination anomalies, anti-symmetric about the equator and peaking near $\sim$20-30° latitude. Therefore, some inclination anomalies are expected in a time-averaged GAD field when calculated from unit vector magnetizations. However, significant deviations from either zero inclination anomaly or the anti-symmetric anomaly arising from unit vector treatment may be due to persistent non-GAD contributions to the time averaged field. Alternative methods to calculate ΔI and additional data since PSV10, presented in Behar et al. (2019), suggest that the inclination anomaly estimates of PSV10 may be biased due to data selection and inclination anomaly calculation methods. If this is the case, then the persistent non-GAD contribution to the time-averaged field is likely to be negligible, and the *BB18* model is optimal.

9 Conclusions

The new GGP models presented in this study (*BB18* and *BB18.Z3*) both yield improved fits to the VGP dispersion estimates of PSV10 relative to existing GGP models, approaching what can be achieved with Model G-style fits of (Doubrovine et al., 2019), while also predicting field directions and intensities which cannot be done with Model G. Furthermore, *BB18* models are also able to reproduce the distribution of field strengths observed for the past 10 million years, which prior GGP models are unable to do. We find that the introduction of a covariance matrix allows for improved reproductions of the observed latitude dependence of VGP dispersion. This finding reinforces expected theoretical symmetry relationships of the field (Hulot & Gallet, 1996) and numerical dynamo simulations (Bouligand et al., 2005; Sanchez et al., 2019) which predict a covariance between Gauss coefficients. Generating accurate predictions of VGP dispersion at all latitudes is necessary to determine whether palaeomagnetic datasets sufficiently average secular variation and have properly excluded transitional directions and outliers. Identifying the precise physical processes which yield the observed covariance, what parameters control the amplitude of covariance, and further tests of the assumptions in GGP models (e.g. Hulot & Bouligand, 2005; Khokhlov & Hulot, 2017) are critical questions for future study.

The addition of zonal non-zero-mean terms yields an improved fit, relative to GAD field models, for VGP dispersion and inclination anomaly estimates from the PSV10 dataset. This supports previous assertions that the time-averaged field of the past 10 million years is not a perfect geocentric axial dipole, but one with a more complex mean field morphology. Field strength compilations (e.g. Biggin et al., 2015; Smirnov et al., 2016; Shcherbakova et al., 2017; Bono et al., 2019; Kulakov et al., 2019; Hawkins et al., 2019) demonstrated that earlier times record different VDM distributions from the past 10 million years. It is suspected that for other intervals further back in geologic time, VGP dispersion and other estimates of PSV behaviour are different than seen for this most recent interval (e.g., Tarduno et al., 2002; Biggin, Strik, & Langereis, 2008; Biggin, van Hinsbergen, et al., 2008; Biggin et al., 2009; Smirnov et al., 2011; de Oliveira et al., 2018; Doubrovine et al., 2019). Given the variation of field strength and morphology, new statistical field models based on the approach applied in this study are needed, which can reproduce the statistical properties of the time-averaged field and the validity of these assumptions during those intervals.

Acknowledgments

The authors declare no competing financial interests. Funding for R. K. B., A. J. B., R. H., and D. G. M. was provided by The Leverhulme Trust Research Leadership Award, RL-2016-080; for C. J. D., by NERC fellowship NE/L011328/1. A portion of the geodynamo simulations were performed on the UK National service ARCHER (via allocation through the Mineral Physics Consortium). We thank Johannes Wicht for sharing dynamo simulation outputs. We also thank Catherine Constable and Alexandre Fournier for constructive reviews, and Johannes Wicht, Greig Paterson and Courtney Sprain for helpful discussions. Code and model parameters are available at the EarthRef Digital Archive: `https://earthref.org/ERDA/2420`.

References

Behar, N., Shaar, R., Tauxe, L., Asefaw, H., Ebert, Y., Heimann, A., … Ron, H. (2019). Paleomagnetism and Paleosecular Variations From the Plio-Pleistocene Golan Heights Volcanic Plateau, Israel. *Geochemistry, Geophysics, Geosystems*, 2019GC008479. doi: 10.1029/2019GC008479

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

Biggin, A. J., Piispa, E. J., Pesonen, L. J., Holme, R., Paterson, G. A., Veikkolainen, T., & Tauxe, L. (2015). Palaeomagnetic field intensity variations suggest Mesoproterozoic inner-core nucleation. *Nature*, *526*(7572), 245-248. doi: 10.1038/nature15523

Biggin, A. J., Strik, G. H. M. A., & Langereis, C. G. (2008). Evidence for a very-long-term trend in geomagnetic secular variation. *Nature Geoscience*, *1*(6), 395-398. doi: 10.1038/ngeo181

Biggin, A. J., Strik, G. H. M. A., & Langereis, C. G. (2009). The intensity of the geomagnetic field in the late-Archaean: New measurements and an analysis of the updated IAGA palaeointensity database. *Earth, Planets and Space*, *61*(1), 9-22. doi: 10.1186/BF03352881

Biggin, A. J., van Hinsbergen, D. J., Langereis, C. G., Straathof, G. B., & Deenen, M. H. (2008). Geomagnetic secular variation in the Cretaceous Normal Superchron and in the Jurassic. *Physics of the Earth and Planetary Interiors*, *169*(1-4), 3-19. doi: 10.1016/j.pepi.2008.07.004

Bono, R. K., Tarduno, J. A., Nimmo, F., & Cottrell, R. D. (2019). Young inner core inferred from Ediacaran ultra-low geomagnetic field intensity. *Nature Geoscience*, 1. doi: 10.1038/s41561-018-0288-0

Bouligand, C., Hulot, G., Khokhlov, A., & Glatzmaier, G. A. (2005). Statistical palaeomagnetic field modelling and dynamo numerical simulation. *Geophysical Journal International*, *161*(3), 603-626. doi: 10.1111/j.1365-246X.2005.02613.x

Butler, R. F. (1992). *Paleomagnetism: Magnetic domains to geologic terranes.* Boston: Oxford: Blackwell Scientfic.

Christensen, U., & Aubert, J. (2006). Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields. *Geophysical Journal International*, *166*(1), 97–114. doi: 10.1111/j.1365-246X.2006.03009.x

Christensen, U., & Wicht, J. (2015). Numerical Dynamo Simulations. In *Treatise on Geophysics* (pp. 245–277). Elsevier. doi: 10.1016/B978-0-444-53802-4.00145-7

Constable, C. G., & Johnson, C. L. (1999). Anisotropic paleosecular variation models: Implications for geomagnetic field observables. *Physics of the Earth and Planetary Interiors*, *115*(1), 35-51. doi: 10.1016/S0031-9201(99)00065-5

Constable, C. G., & Parker, R. L. (1988). Statistics of the geomagnetic secular variation for the past 5 m.y. *Journal of Geophysical Research: Solid Earth*, *93*(B10), 11569-11581. doi: 10.1029/JB093iB10p11569

Cox, A. (1970). Latitude dependence of the angular dispersion of the geomagnetic field. *Geophysical Journal of the Royal Astronomical Society*, *20*(3), 253–269.

Cromwell, G., Johnson, C. L., Tauxe, L., Constable, C. G., & Jarboe, N. A. (2018). PSV10: A Global Data Set for 0–10 Ma Time-Averaged Field and Paleosecular Variation Studies. *Geochemistry, Geophysics, Geosystems*, *19*(5), 1533-1558. doi: 10.1002/2017GC007318

Davies, C. J., & Constable, C. G. (2014). Insights from geodynamo simulations into long-term geomagnetic field behaviour. *Earth and Planetary Science Letters*, *404*, 238-249. doi: 10.1016/j.epsl.2014.07.042

Davies, C. J., & Gubbins, D. (2011). A buoyancy profile for the Earth′s core. *Geophysical Journal International*, *187*(2), 549-563. doi: 10.1111/j.1365-246X.2011.05144.x

Deb, K., & Kalyanmoy, D. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms.* New York, NY, USA: John Wiley & Sons, Inc.

Dekkers, M. J., & Böhnel, H. N. (2006). Reliable absolute palaeointensities independent of magnetic domain state. *Earth and Planetary Science Letters*, *248*(1), 508-517. doi: 10.1016/j.epsl.2006.05.040

de Oliveira, W. P., Franco, D. R., Brandt, D., Ernesto, M., Neto, C. F. d. P., Zhao, X., . . . Martins, R. S. (2018). Behavior of the Paleosecular Variation During the Permian-Carboniferous Reversed Superchron and Comparisons to the Low Reversal Frequency Intervals Since Precambrian Times. *Geochemistry, Geophysics, Geosystems*, *19*(4), 1035-1048. doi: 10.1002/2017GC007262

Doubrovine, P. V., Veikkolainen, T., Pesonen, L. J., Piispa, E., Ots, S., Smirnov, A. V., ... Biggin, A. J. (2019). Latitude Dependence of Geomagnetic Paleosecular Variation and its Relation to the Frequency of Magnetic Reversals: Observations From the Cretaceous and Jurassic. *Geochemistry, Geophysics, Geosystems*, *0*(0). doi: 10.1029/2018GC007863

Dziewonski, A. M., Lekic, V., & Romanowicz, B. A. (2010). Mantle Anchor Structure: An argument for bottom up tectonics. *Earth and Planetary Science Letters*, *299*(1), 69–79. doi: 10.1016/j.epsl.2010.08.013

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap.* New York: Chapman & Hall.

Fisher, R. (1953). Dispersion on a Sphere. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *217*(1130), 295-305. doi: 10.1098/rspa.1953.0064

Gubbins, D., & Zhang, K. (1993). Symmetry properties of the dynamo equations for palaeomagnetism and geomagnetism. *Physics of the Earth and Planetary Interiors*, *75*(4), 225–241. doi: 10.1016/0031-9201(93)90003-R

Hawkins, L. M. A., Anwar, T., Shcherbakova, V. V., Biggin, A. J., Kravchinsky, V. A., Shatsillo, A. V., & Pavlov, V. E. (2019). An exceptionally weak Devonian geomagnetic field recorded by the Viluy Traps, Siberia. *Earth and Planetary Science Letters*, *506*, 134–145. doi: 10.1016/j.epsl.2018.10.035

Hulot, G., & Bouligand, C. (2005). Statistical palaeomagnetic field modelling and symmetry considerations. *Geophysical Journal International*, *161*(3), 591-602. doi: 10.1111/j.1365-246X.2005.02612.x

Hulot, G., & Gallet, Y. (1996). On the interpretation of virtual geomagnetic pole (VGP) scatter curves. *Physics of the Earth and Planetary Interiors*, *95*(1), 37-53. doi: 10.1016/0031-9201(95)03106-5

Johnson, C. L., & Constable, C. G. (1996). Palaeosecular variation recorded by lava flows over the past five million years. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *354*(1704), 89-141. doi: 10.1098/rsta.1996.0004

Johnson, C. L., & McFadden, P. (2015). The Time-Averaged Field and Paleosecular Variation. In *Treatise on Geophysics* (p. 385-417). Elsevier. doi: 10.1016/B978-0-444-53802-4.00105-6

Khokhlov, A., & Hulot, G. (2017). On the cause of the non-Gaussian distribution of residuals in geomagnetism. *Geophysical Journal International*, *209*(2), 1036-1047. doi: 10.1093/gji/ggx071

Konôpková, Z., McWilliams, R. S., Gómez-Pérez, N., & Goncharov, A. F. (2016). Direct measurement of thermal conductivity in solid iron at planetary core conditions. *Nature*, *534*(7605), 99–101. doi: 10.1038/nature18009

Kulakov, E., Sprain, C., Doubrovine, P., Smirnov, A., Paterson, G., Hawkins, L., ... Biggin, A. (2019). Analysis of an updated paleointensity database (Q_{PI}-PINT) for 65-200 Ma: Implications for the long-term history of dipole moment through the Mesozoic. *Journal of Geophysical Research: Solid Earth*, 2018JB017287. doi: 10.1029/2018JB017287

Lee, S. (1983). *A study of the time-averaged palaeomagnetic field for the last 195 million years.* (Ph.D. Thesis). Australian National University, Canberra.

Lhuillier, F., & Gilder, S. A. (2013). Quantifying paleosecular variation: Insights from numerical dynamo simulations. *Earth and Planetary Science Letters*, *382*, 87–97. doi: 10.1016/j.epsl.2013.08.048

Lowes, F. J. (1974). Spatial Power Spectrum of the Main Geomagnetic Field, and Extrapolation to the Core. *Geophysical Journal of the Royal Astronomical Society*, *36*(3), 717-730. doi: 10.1111/j.1365-246X.1974.tb00622.x

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, *46*(253), 68-78. doi: 10.1080/01621459.1951.10500769

Masters, T. G., Johnson, S., Laske, G., Bolton, H., Davies, J. H., Jephcoat, A. P., ... O'Nions, R. K. (1996). A shear - velocity model of the mantle. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *354*(1711), 1385–1411. doi: 10.1098/rsta.1996.0054

McElhinny, M. W., & McFadden, P. L. (1997). Palaeosecular variation over the past 5 Myr based on a new generalized database. *Geophysical Journal International*, *131*(2), 240-252. doi: 10.1111/j.1365-246X.1997.tb01219.x

McFadden, P. L., Merrill, R. T., & McElhinny, M. W. (1988). Dipole/quadrupole family modeling of paleosecular variation. *Journal of Geophysical Research: Solid Earth*, *93*(B10), 11583-11588. doi: 10.1029/JB093iB10p11583

Merrill, R. T., McElhinny, M. W., & McFadden, P. L. (1996). *The Magnetic Field of the Earth: Paleomagnetism, the Core, and the Deep Mantle* (Vol. 63). Academic Press.

Muxworthy, A. R. (2017). Considerations for Latitudinal Time-Averaged-Field Palaeointensity Analysis of the Last Five Million Years. *Frontiers in Earth Science*, *5*. doi: 10.3389/feart.2017.00079

Opdyke, N. D., Kent, D. V., Foster, D. A., & Huang, K. (2015). Paleomagnetism of Miocene volcanics on Sao Tome: Paleosecular variation at the Equator and a comparison to its latitudinal dependence over the last 5 Myr. *Geochemistry, Geophysics, Geosystems*, *16*(11), 3870-3882. doi: 10.1002/2015GC005901

Oruba, L., & Dormy, E. (2014). Transition between viscous dipolar and inertial multipolar dynamos. *Geophysical Research Letters*, *41*(20), 7115–7120. doi: 10.1002/2014GL062069

Parker, R. L. (1994). *Geophysical inverse theory.* Princeton University Press.

Paterson, G. A., Heslop, D., & Muxworthy, A. R. (2010). Deriving confidence in paleointensity estimates. *Geochemistry, Geophysics, Geosystems*, *11*(7). doi: 10.1029/2010GC003071

Pozzo, M., Davies, C., Gubbins, D., & Alfè, D. (2012). Thermal and electrical conductivity of iron at Earth's core conditions. *Nature*, *485*(7398), 355–358. doi: 10.1038/nature11031

Quidelleur, X., & Courtillot, V. (1996). On low-degree spherical harmonic models of paleosecular variation. *Physics of the Earth and Planetary Interiors*, *95*(1), 55-77. doi: 10.1016/0031-9201(95)03115-4

Quidelleur, X., Valet, J.-P., Courtillot, V., & Hulot, G. (1994). Long-term geometry of the geomagnetic field for the last five million years: An updated secular variation database. *Geophysical Research Letters*, *21*(15), 1639-1642. doi: 10.1029/94GL01105

Sanchez, S., Wicht, J., Bärenzung, J., & Holschneider, M. (2019). Sequential assimilation of geomagnetic observations: Perspectives for the reconstruction and prediction of core dynamics. *Geophysical Journal International*, *217*(2), 1434–1450. doi: 10.1093/gji/ggz090

Selkin, P. A., & Tauxe, L. (2000). Long-term variations in palaeointensity. *Philosophical Transactions of the Royal Society of London A: Mathe-*

matical, Physical and Engineering Sciences, *358*(1768), 1065-1088. doi: 10.1098/rsta.2000.0574

Shaw, J. (1974). A New Method of Determining the Magnitude of the Palaeomagnetic Field: Application to five historic lavas and five archaeological samples. *Geophysical Journal of the Royal Astronomical Society*, *39*(1), 133-141. doi: 10.1111/j.1365-246X.1974.tb05443.x

Shcherbakova, V. V., Biggin, A. J., Veselovskiy, R. V., Shatsillo, A. V., Hawkins, L. M. A., Shcherbakov, V. P., & Zhidkov, G. V. (2017). Was the Devonian geomagnetic field dipolar or multipolar? Palaeointensity studies of Devonian igneous rocks from the Minusa Basin (Siberia) and the Kola Peninsula dykes, Russia. *Geophysical Journal International*, *209*(2), 1265–1286. doi: 10.1093/gji/ggx085

Smirnov, A. V., Tarduno, J. A., & Evans, D. A. D. (2011). Evolving core conditions ca. 2 billion years ago detected by paleosecular variation. *Physics of the Earth and Planetary Interiors*, *187*(3–4), 225-231. doi: 10.1016/j.pepi.2011.05.003

Smirnov, A. V., Tarduno, J. A., Kulakov, E. V., McEnroe, S. A., & Bono, R. K. (2016). Palaeointensity, core thermal conductivity and the unknown age of the inner core. *Geophysical Journal International*, *205*(2), 1190–1195. doi: 10.1093/gji/ggw080

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

Tanaka, H., Kono, M., & Uchimura, H. (1995). Some global features of palaeointensity in geological time. *Geophysical Journal International*, *120*(1), 97–102. doi: 10.1111/j.1365-246X.1995.tb05913.x

Tarduno, J. A., Cottrell, R. D., & Smirnov, A. V. (2002). The Cretaceous superchron geodynamo: Observations near the tangent cylinder. *Proceedings of the National Academy of Sciences*, *99*(22), 14020-14025. doi: 10.1073/pnas.222373499

Tauxe, L., & Kent, D. V. (2004). A Simplified Statistical Model for the Geomagnetic Field and the Detection of Shallow Bias in Paleomagnetic Inclinations: Was the Ancient Magnetic Field Dipolar? In J. E. T. Channell, D. V. Kent,

W. Lowrie, & J. G. Meert (Eds.), *Timescales Of The Paleomagnetic Field* (p. 101-115). American Geophysical Union. doi: 10.1029/145GM08

Thellier, E., & Thellier, O. (1959). Sur l'intensité du champ magnétique terrestre dans le passé historique et géologique. *Annales de Geophysique*, *15*, 285.

Valet, J.-P., Brassart, J., Meur, I. L., Soler, V., Quidelleur, X., Tric, E., & Gillot, P.-Y. (1996). Absolute paleointensity and magnetomineralogical changes. *Journal of Geophysical Research: Solid Earth*, *101*(B11), 25029-25044. doi: 10.1029/96JB02115

Vandamme, D. (1994). A new method to determine paleosecular variation. *Physics of the Earth and Planetary Interiors*, *85*(1), 131-142. doi: 10.1016/0031-9201(94)90012-4

Wicht, J., & Meduri, D. G. (2016). A gaussian model for simulated geomagnetic field reversals. *Physics of the Earth and Planetary Interiors*, *259*, 45-60. doi: 10.1016/j.pepi.2016.07.007

Wicht, J., Stellmach, S., & Harder, H. (2015). Numerical dynamo simulations: From basic concepts to realistic models. In W. Freeden, M. Z. Nashed, & T. Sonar (Eds.), *Handbook of geomathematics* (pp. 779–834). Springer.

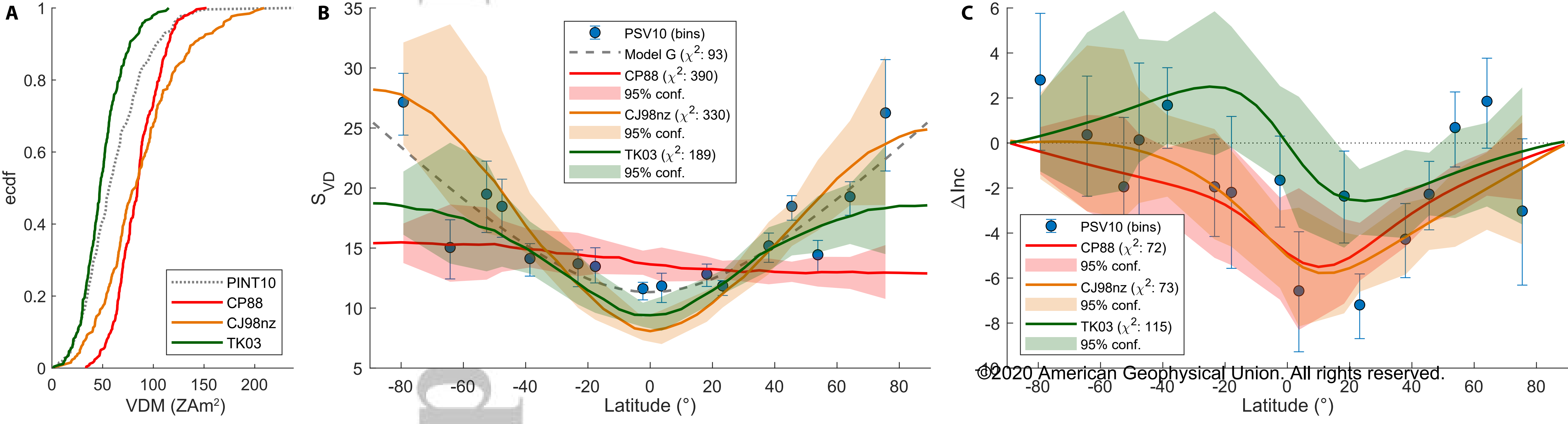
Wicht, J., & Tilgner, A. (2010). Theory and modeling of planetary dynamos. *Space Science Reviews*, *152*(1), 501–542. doi: 10.1007/s11214-010-9638-y

Wilson, R. L. (1961). The Thermal Demagnetization of Natural Magnetic Moments in Rocks. *Geophysical Journal International*, *5*(1), 45-58. doi: 10.1111/j.1365-246X.1961.tb02928.x

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. doi: 10.1111/j.1365-246X.2005.02651.x

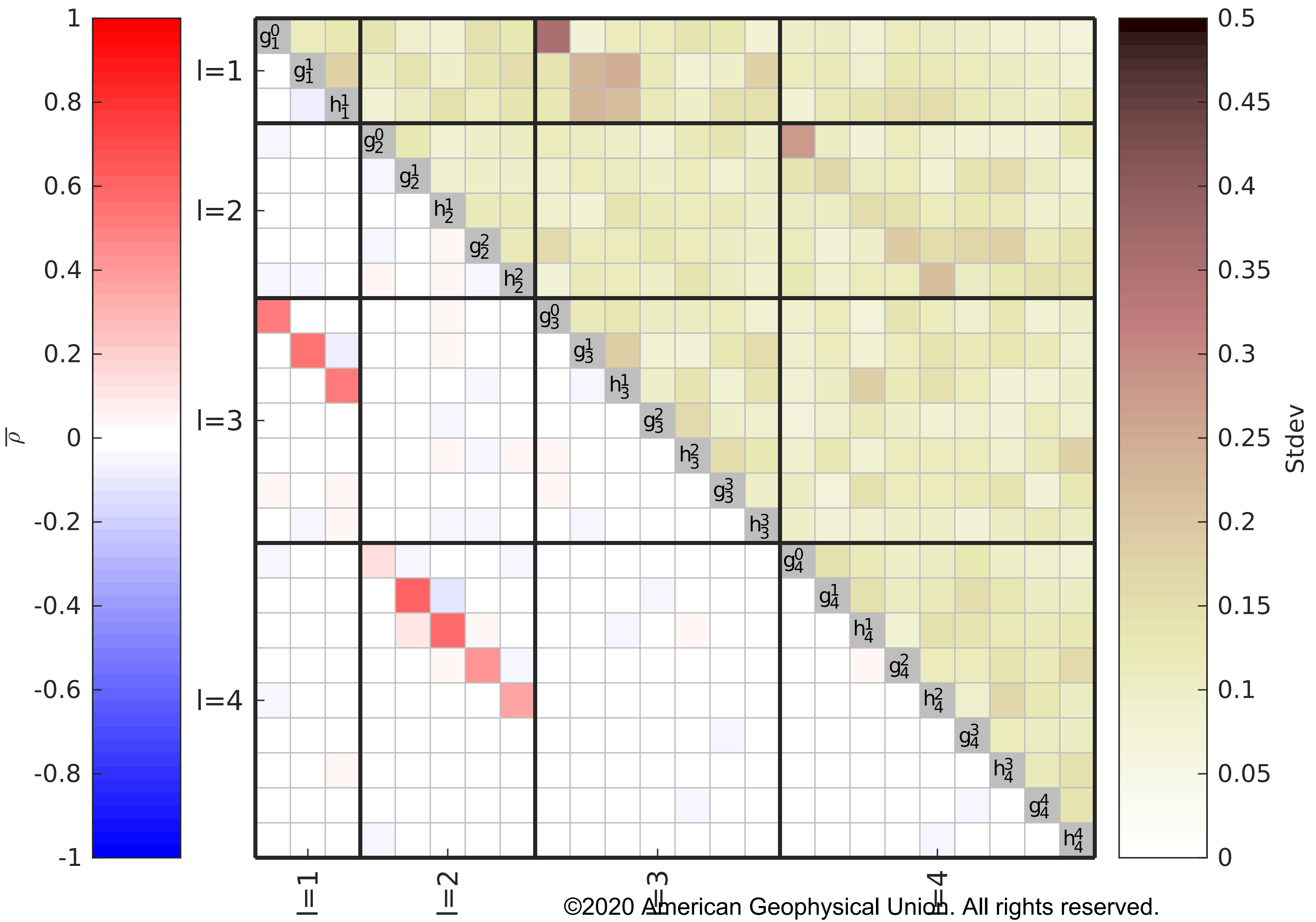
Yamamoto, Y., Tsunakawa, H., & Shibuya, H. (2003). Palaeointensity study of the Hawaiian 1960 lava: Implications for possible causes of erroneously high intensities. *Geophysical Journal International*, *153*(1), 263-276. doi: 10.1046/j.1365-246X.2003.01909.x

Figure 1.



A
ecdf
VDM (ZAm²)
PINT10
CP88
CJ98nz
TK03
B
S_VD
Latitude (°)
PSV10 (bins)
Model G (χ²: 93)
CP88 (χ²: 390)
95% conf.
CJ98nz (χ²: 330)
95% conf.
TK03 (χ²: 189)
95% conf.
C
ΔInc
Latitude (°)
PSV10 (bins)
CP88 (χ²: 72)
95% conf.
CJ98nz (χ²: 73)
95% conf.
TK03 (χ²: 115)
95% conf.

Figure 2.



$\overline{\rho}$
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
l=1
l=2
l=3
l=4
g_1^0 g_1^1 h_1^1
g_2^0 g_2^1 h_2^1 g_2^2 h_2^2
g_3^0 g_3^1 h_3^1 g_3^2 h_3^2 g_3^3 h_3^3
g_4^0 g_4^1 h_4^1 g_4^2 h_4^2 g_4^3 h_4^3 g_4^4 h_4^4
Stdev
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0

Figure 3.

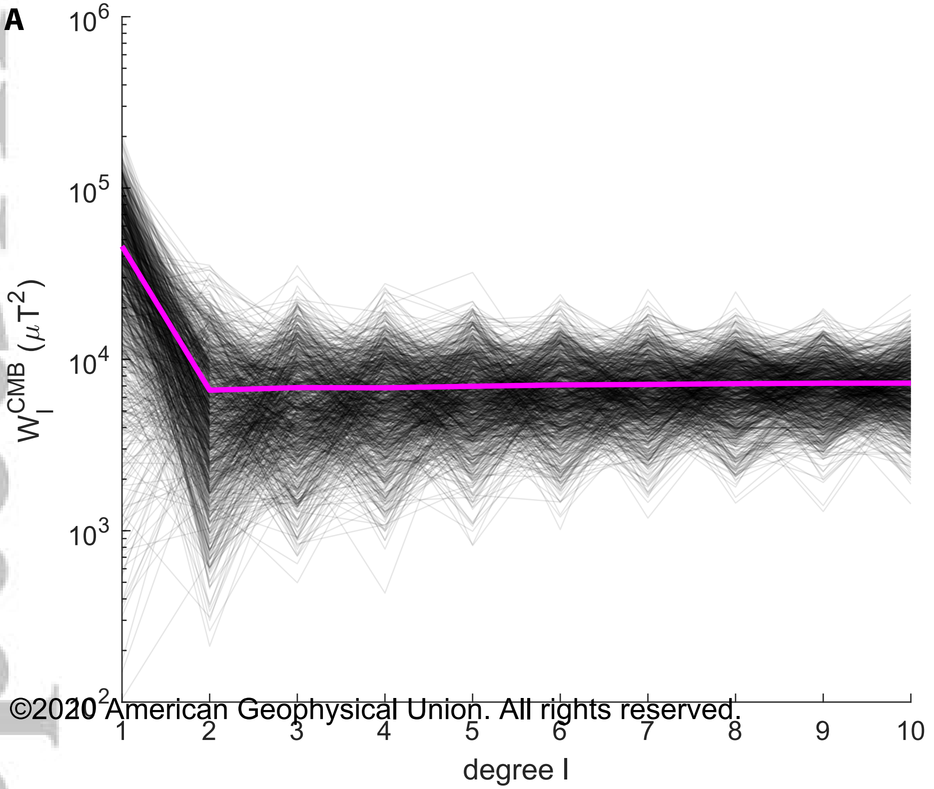
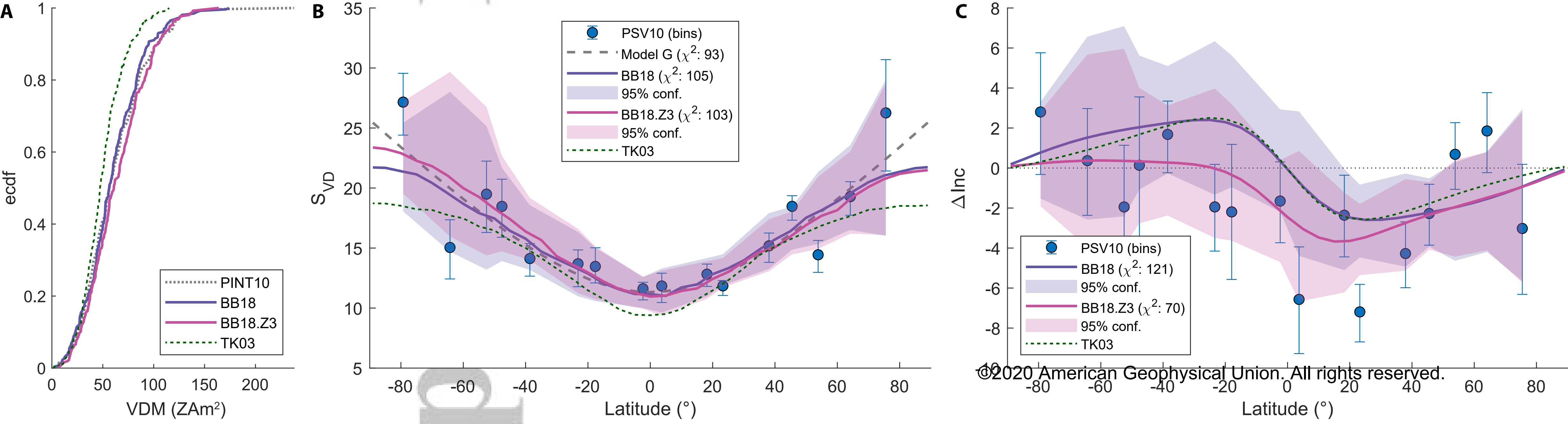

A
10^6
10^5
10^4
10^3
10^2
W_l^CMB (μT^2)
1
2
3
4
5
6
7
8
9
10
degree l

Figure 4.



A
ecdf
VDM (ZAm²)
PINT10
BB18
BB18.Z3
TK03
B
S_{VD}
Latitude (°)
PSV10 (bins)
Model G (χ^2: 93)
BB18 (χ^2: 105)
95% conf.
BB18.Z3 (χ^2: 103)
95% conf.
TK03
C
ΔInc
Latitude (°)
PSV10 (bins)
BB18 (χ^2: 121)
95% conf.
BB18.Z3 (χ^2: 70)
95% conf.
TK03