Geophysical Journal International

Geophys. J. Int. (2013) **195,** 86–97
Advance Access publication 2013 July 9

doi: 10.1093/gji/ggt218

A stochastic model for palaeomagnetic field variations

Bruce A. Buffett,[1] Leah Ziegler[2] and Cathy G. Constable[3]

[1] *Department of Earth & Planetary Science, University of California, Berkeley, CA* 94720*, USA. E-mail: bbuffett@berkeley.edu*
[2] *College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR* 97331-5503*, USA*
[3] *Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA* 92093-0225*, USA*

Accepted 2013 May 29. Received 2013 May 24; in original form 2013 February 11

SUMMARY

Regeneration of the Earth's magnetic field by convection in the liquid core produces a broad spectrum of time variation. Relative palaeointensity measurements in marine sediments provide a detailed record over the past 2 Myr, but an explicit reconstruction of the underlying dynamics is not feasible. A more practical alternative is to construct a stochastic model from estimates of the virtual axial dipole moment. The deterministic part of the model (drift term) describes time-averaged behaviour, whereas the random part (diffusion term) characterizes complex interactions over convective timescales. We recover estimates of the drift and diffusion terms from the SINT2000 model of Valet *et al.* and the PADM2M model of Ziegler *et al.* The results are used in numerical solutions of the Fokker–Planck equation to predict statistical properties of the palaeomagnetic field, including the average rates of magnetic reversals and excursions. A physical interpretation of the stochastic model suggests that the timescale for adjustments in the axial dipole moment is set by the dipole decay time τ_d. We obtain $\tau_d = 29$ kyr from the stochastic models, which falls within the expected range for the Earth's core. We also predict the amplitude of convective fluctuations in the core, and establish a physical connection to the rates of magnetic reversals and excursions. Chrons lasting longer than 10 Myr are unlikely under present-day conditions. However, long chrons become more likely if the diffusion term is reduced by a factor of 2. Such a change is accomplished by reducing the velocity fluctuations in the core by a factor of $\sqrt{2}$, which could be attributed to a shift in the spatial pattern of heat flux from the core or a reduction in the total core heat flow.

Key words: Probability distributions; Geomagnetic excursions; Magnetostratigraphy; Palaeointensity; Reversals: process, timescale, magnetostratigraphy.

1 INTRODUCTION

Variations in the geomagnetic field occur over a broad range of timescales (Constable & Johnson 2005). Historical and satellite observations (Jackson *et al.* 2000; Olsen & Mandea 2008) record variations over years to centuries, whereas palaeomagnetic data reveal long-term trends over timescales as long as 10^8 yr (Johnson *et al.* 1995). The broad spectrum of geomagnetic variation is thought to reflect processes in the core, possibly modulated by slow changes in the overlying mantle (Courtillot & Olson 2007). Historical observations (Jackson *et al.* 2000) and relatively recent palaeomagnetic data (Korte *et al.* 2005) have sufficient spatial and temporal resolution to represent the geomagnetic field in a spherical harmonic expansion (e.g. Korte & Constable 2005). Such detailed estimates have been used to draw inferences about flow at the top of the core (Jackson 1997; Pais *et al.* 2004; Duberry & Finlay 2007). On longer timescales, the data are not sufficient to constrain the spatial structure of time variations. Instead the geomagnetic field is represented as a dipole. Direction measurements define a virtual geomagnetic pole (VGP) and intensity measurements are interpreted as an equivalent or virtual axial dipole moment (VADM). Variations in both VGP and VADM contain valuable information about the geodynamo, but the connection to specific physical processes is not directly established.

Statistical models offer another means of expressing the variability of the geomagnetic field. Constable & Parker (1988) described the statistical properties of the magnetic field in terms of the coefficients of a spherical harmonic expansion. The coefficients of the non-dipole field are represented by independent Gaussian distributions with zero mean. An estimate of the correlation time for the non-dipole field is several hundred years, based on historical and satellite observations (Hulot & LeMouël 1994; Lhuillier *et al.* 2011). Variations in the axial dipole are described by a pair of Gaussian distributions with non-zero means, coinciding with the time-averaged axial dipole in the normal and reversed states. Refinements in the statistical model of Constable & Parker (1988)

have been proposed to improve the description of both direction and intensity variations (Quidelleur & Courtillot 1996; Constable & Johnson 1999).

An alternative approach is to treat the time dependence of the geomagnetic field as a stochastic process (e.g. Brendel *et al.* 2007). The general description of a stochastic process includes a deterministic component, as well as a random component to represent the complex interplay between convective fluctuations and the magnetic field. In this study, we develop a stochastic model for the palaeomagnetic field with the aim of extracting insights about the underlying dynamics. We focus on estimates of VADM because the available records are long enough to describe the behaviour of the geomagnetic field over several reversals. Currently, VADM stacks are based on a combination of relative palaeointensities from marine sediment cores and a smaller number of absolute intensities to calibrate the relative palaeointensities (Guyodo & Valet 1999; Valet *et al.* 2005; Ziegler *et al.* 2011). Sediment cores provide a long and continuous record of geomagnetic variation, although the acquisition of magnetization effectively averages the intensity over timescales of 10^3 yr (Lund & Keigwin 1994; Roberts & Winklhofer 2004). Averaging in time tends to reduce (although probably not eliminate) the contribution from fluctuations in the non-dipole field because these fluctuations occur over times comparable or shorter than 10^3 yr. In addition, spatial averaging of globally distributed records into a single composite stack reduces the contribution of regional-scale non-dipole features. Consequently, we can reasonably attribute most of the variation in VADM stacks to changes in the axial dipole, which is useful when interpreting the physical significance of the stochastic model.

Stochastic models have two main components. The first component is a deterministic part, which describes the time-averaged trend of the VADM as a function of the dipole intensity. This part of the stochastic model is often called the drift term because it reflects the tendency of the field to adjust towards an equilibrium state. (The equilibrium state will subsequently be defined as the state for which the drift term vanishes.) The second component is a random part, which represents the influence of convective fluctuations on the evolution of the dipole field. Both the deterministic and random components of the stochastic model can be recovered from estimates of the VADM. We demonstrate the approach using the models of Valet *et al.* (2005) and Ziegler *et al.* (2011). Once the random and deterministic components are established, a probabilistic description of the geomagnetic field is given by the Fokker–Planck equation (e.g. Risken 1989). Numerical and approximate analytical solutions to the Fokker–Planck equation are used to predict statistical properties of the geomagnetic field, including the mean rate of magnetic reversals and excursions. We also explore the physical significance of the stochastic model by examining the processes responsible for both the random and deterministic components of the model.

2 OVERVIEW OF STOCHASTIC MODELS

Stochastic models offer a simple description of the axial dipole moment over geological timescales. Fluctuations in the geomagnetic field on the timescale of a convective overturn in the core can be treated as a random component. A nominal estimate of the overturn time is R/v, where $R = 3.48 \times 10^6$ m is the radius of the core and $v \approx 5 \times 10^{-4}$ m s^{-1} is a representative value of the convective velocity (Hulot *et al.* 2002; Holme & Olsen 2006). The drift term represents the long-term tendency of the dipole to adjust towards the time-averaged dipole strength. We denote the axial dipole moment by $x(t)$ and describe its time evolution by the stochastic differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v(x) + \sqrt{D(x)}\Gamma(t)\,, \tag{1}$$

where $v(x)$ and $D(x)$ are, respectively, the drift and diffusion terms. The noise source, $\Gamma(t)$, is commonly assumed to be Gaussian with a vanishing time average

$$\langle \Gamma(t) \rangle = 0\,. \tag{2}$$

It is also customary to assume that the correlation time of the noise source is short compared with the sampling of $x(t)$. In this case, the autocorrelation function of $\Gamma(t)$ can be approximated by a Dirac delta function,

$$\langle \Gamma(t_1)\Gamma(t_2) \rangle = 2\delta(t_1 - t_2)\,, \tag{3}$$

where the factor of 2 is simply a convenient convention (e.g. Risken 1989). When the diffusion term varies with x (i.e. multiplicative noise) we must specify whether $D(x)$ is evaluated before or after the application of the impulsive noise source. For a continuously evolving system where (3) is an approximation of a process with a short correlation time, it is appropriate to adopt the Stratonovich convention and evaluate $D(x)$ using the mean value of x over a short time step (e.g. Van Kampen 1992). In this case, a gradient in $D(x)$ contributes to the drift term in a statistical description of the process (see later).

The probability distribution $P(x, t)$ of the stochastic process is governed by the Fokker–Planck equation

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[v_e(x)P(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[D(x)P(x,t)\right]\,, \tag{4}$$

where

$$v_e(x) = v(x) + \frac{1}{2}\frac{\mathrm{d}D(x)}{\mathrm{d}x} \tag{5}$$

defines the effective drift for a Stratonovich process. Equivalent information about the probability distribution of $x(t)$ could be inferred from a large number of realizations of (1), but the Fokker–Planck equation offers a powerful tool for describing the statistical properties of a stochastic process.

In principle, the drift and diffusion terms can be extracted from a realization of the stochastic process (e.g. Friedrich *et al.* 2011). Defining the nth moment of the process as

$$M^{(n)}(x,\tau) = \langle [x(t+\tau) - x(t)]^n \rangle \rfloor_{x(t)=x}\,, \tag{6}$$

the drift term is given by

$$v_e(x) = \lim_{\tau \to 0} \frac{1}{\tau} M^{(1)}(x,\tau) \tag{7}$$

and the diffusion term is

$$D(x) = \lim_{\tau \to 0} \frac{1}{2\tau} M^{(2)}(x,\tau)\,. \tag{8}$$

In practice, time-series are sampled at finite intervals (say Δt), which means that the moments cannot be evaluated in the limit of vanishing τ. Instead, the drift and diffusion terms are approximated using finite values of τ. Theoretical corrections have been derived for the drift and diffusion terms (Sura & Barsugli 2002; Gottschall & Peinke 2008), although it is more expedient to estimate the τ-dependence of $M^{(1)}(x, \tau)$ and $M^{(2)}(x, \tau)$ using a sequence of finite values for τ (e.g. $\tau = \Delta t, 2\Delta t$, etc.). A further complication arises if the noise source is correlated at small values of τ. We illustrate the general approach and address some of the complications in Section 3.

3 APPLICATION OF STOCHASTIC MODEL TO PALAEOMAGNETIC AXIAL DIPOLE FIELD

VADM models are constructed using relative palaeointensities from marine sediment cores and a smaller number of absolute intensities to provide the necessary calibration. The model of Valet *et al.* (2005), known as SINT2000, is obtained by stacking independent records of relative palaeointensity. Similarly, the model of Ziegler *et al.* (2011) is constructed from independent records, but the time dependence is parametrized using cubic B-splines. A maximum-likelihood estimate for the spline coefficients is regularized by imposing a penalty on the model roughness. The authors use a trade-off parameter to adjust the balance between minimizing model roughness and minimizing misfit to the observations. The preferred model of Ziegler *et al.* (2011) is called PADM2M. Both SINT2000 and PADM2M are evaluated at increments of 1000 yr over the past 2 Myr (see Fig. 1). A signed estimate of the VADM can be obtained by multiplying the VADM by the known polarity of the geomagnetic field (Cande & Kent 1995).

Moments $M^{(n)}(x, \tau)$ of the VADM models are calculated using a time difference of $\tau = 1$ kyr or larger. Larger values of τ are preferred if the noise source is correlated at the shortest time interval. In fact, we expect correlations in the random component of the stochastic model due to the finite time required to acquire a remanent magnetization in sediments. A correlation is also expected in the PADM2M model due to the regularization used in the inversion for the spline coefficients. In either case, we can estimate and correct for a correlation in the noise source by evaluating the drift, $v_e(x, \tau)$, and diffusion, $D(x, \tau)$, using series of finite values for τ.

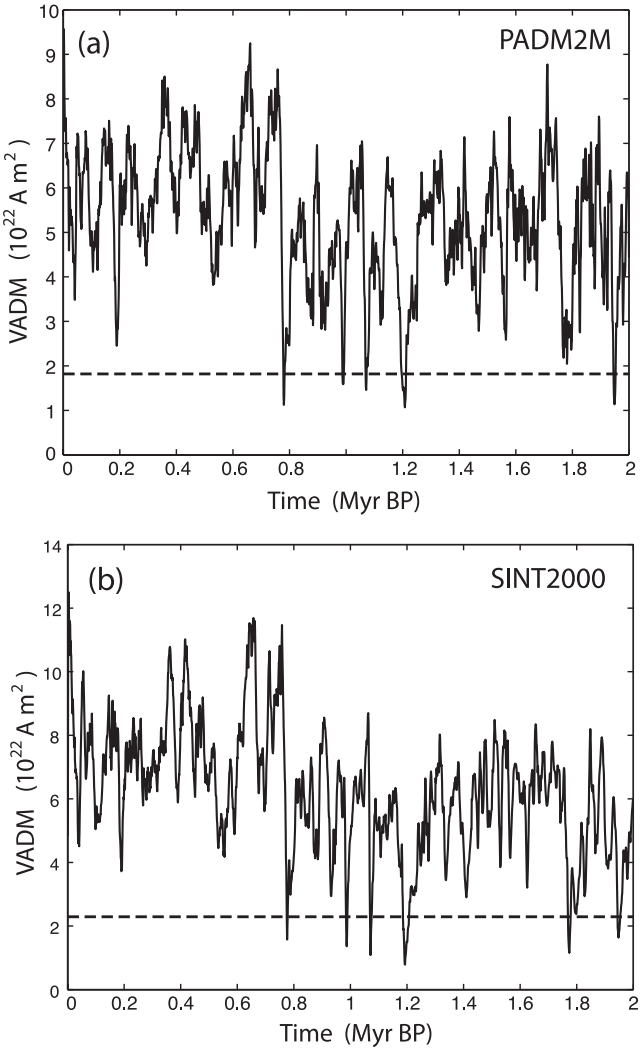

Figure 1. Virtual axial dipole moment over the past 2 Myr before present from the model of (a) Ziegler *et al.* (2011) and (b) Valet *et al.* (2005). Low values (dashed line) are predicted to occur at a rate of 3 Myr^{-1}.

We illustrate the approach by calculating $v_e(x, \tau)$ and $D(x, \tau)$ when the unsigned VADM is subdivided into eight bins. Each $x(t)$ in the time-series is used to calculate $[x(t + \tau) - x(t)]^n$ for a specific choice of τ. The value is assigned to a bin based on the amplitude of $x(t)$. Averages for $M^{(1)}(x, \tau)$ and $M^{(2)}(x, \tau)$ are evaluated in each bin, and the resulting drift and diffusion terms are computed using (7) and (8). Fig. 2 shows $v_e(x, \tau)$ and $D(x, \tau)$ as a function of τ for both PADM2M and SINT2000. Each line represents for the results for a single bin, identified by the value of x at the centre of the bin. No values are reported when the number of entries in a particular bin is less than 100.

A correlation in the random component reduces $x(t + \tau) - x(t)$ when τ is shorter than the correlation time of the noise. On the other hand, the difference $x(t + \tau) - x(t)$ should become larger when τ is greater than the correlation time of the noise. Reduced values for $v_e(x, \tau)$ and $D(x, \tau)$ at small τ in Fig. 2 probably reflect the influence of correlations in the time-series. Once the time difference exceeds $\tau \approx 4$–5 kyr, the drift and diffusion terms become relatively independent of τ. For the drift term this means that the average difference, $\langle x(t + \tau) - x(t) \rangle$, is well approximated by the average of a first-order Taylor expansion, $\langle \tau \, \partial x/\partial t \rangle$. From (7), it follows the drift term (for finite τ) is given by $v_e(x, \tau) \approx \langle \partial x/\partial t \rangle$. As long as the higher order terms in the Taylor expansion are small we expect $v_e(x, \tau)$ to be independent of τ. The diffusion term is computed from the mean-squared deviation $\langle [x(t + \tau) - x(t)]^2 \rangle$. When the time integral of the noise source is approximated by a random walk, we expect the mean-squared deviation to increase linearly with τ (Van Kampen 1992). In this case the diffusion term, $D(x, \tau)$, is independent of τ at the leading order. An additional contribution to the mean-squared deviation arises from the deterministic component, which can be approximated by $\langle [\tau \partial x/\partial t]^2 \rangle \approx \tau^2 v_e^2$). We see that the deterministic contribution to the mean-squared deviation occurs at the next order in τ, which vanishes in the limit $\tau \to 0$. Increasing τ beyond 10 kyr (not shown) causes more substantial variations in both $v_e(x, \tau)$ and $D(x, \tau)$, which implies that higher order terms in the Taylor expansion or deviations from a random walk become more important.

For the purposes of estimating the drift and diffusion terms, we confine our attention to values of $v_e(x, \tau)$ and $D(x, \tau)$ at $\tau > 4$ kyr. Small variations in $v_e(x, \tau)$ and $D(x, \tau)$ at larger τ imply that corrections to (7) and (8) for finite τ are small (Sura & Barsugli 2002). Consequently, we approximate $v_e(x)$ and $D(x)$ using the estimates at $\tau = 5$ kyr. Alternatively, we could use a pair of estimates at $\tau =$ 5 and 10 kyr to approximate the drift and diffusion terms in the limit $\tau \to 0$ using an approach described by Stanton (1997). However, the difference between these two approaches is small when the bin contains a large number of samples. Consequently, we use the difference between the two approaches as a rough measure of the uncertainty in the drift and diffusion coefficients. We also assess consistency of the results by comparing estimates computed using a different number of bins. Fig. 3 shows estimates for $v_e(x)$ and $D(x)$, evaluated at $\tau = 5$ kyr using 8, 9 and 10 bins in the unsigned VADM. We also show a smooth curve fit through the individual estimates (details are given below).

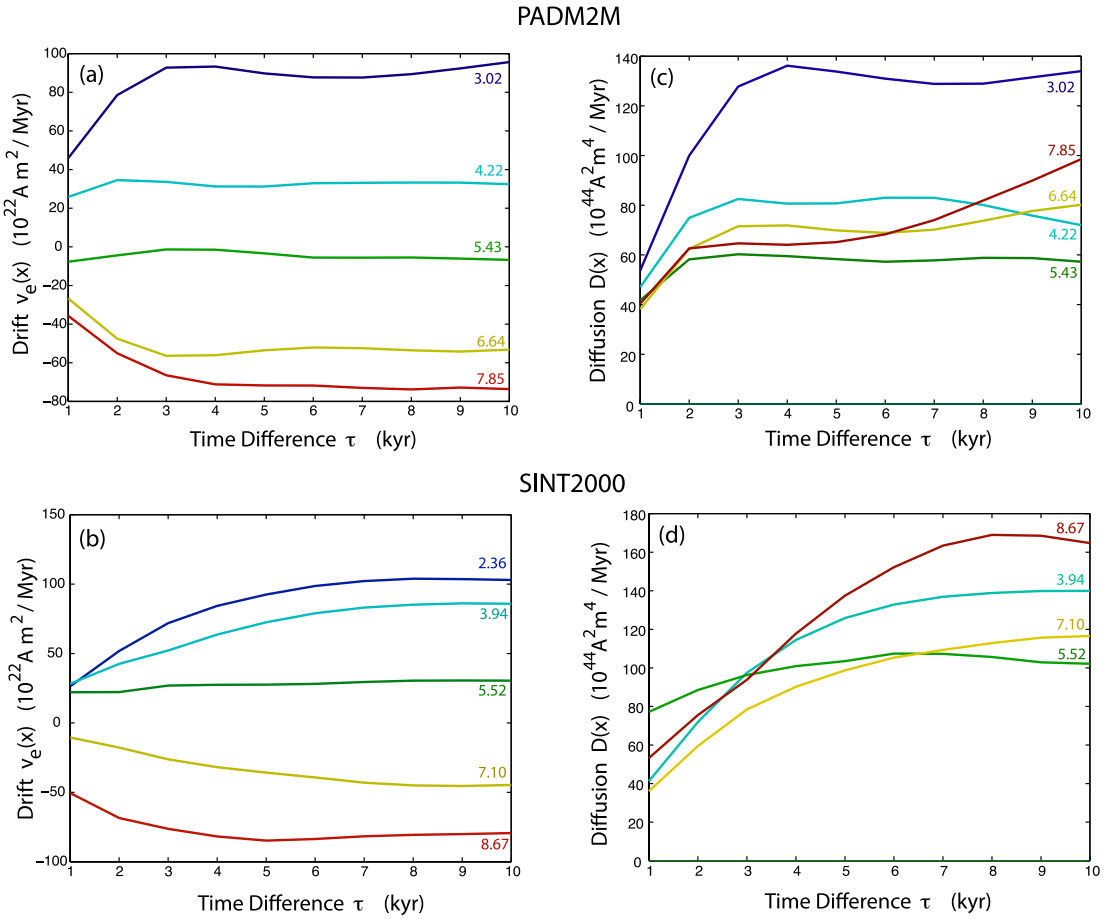

Figure 2. (a,b) Drift $v_e(x, \tau)$ and (c,d) diffusion $D(x, \tau)$ terms calculated using the PADM2M and SINT2000 models. Each line is identified by the value of x at the midpoint of the bin.

Both SINT2000 and PADM2M include relatively few instances of a weak dipole field. As a result, the sampling is insufficient to reliably determine $v_e(x)$ and $D(x)$ at low x. On the other hand, we can anticipate the behaviour of $v_e(x)$ near $x = 0$. A positive drift for $x < 5 \times 10^{22}$ A m^2 causes the intensity of the field to increase with time, whereas a negative drift for $x > 5 \times 10^{22}$ A m^2 causes the field to decrease. An equilibrium is expected when $v_e(x) = 0$, although the random component continually disturbs x away from the equilibrium value. When the dipole field reverses ($x < 0$), the drift term drives the dipole towards a new equilibrium in the reversed state, implying that $v_e(-x) = -v_e(x)$. Even if the expected symmetry between the reversed and normal states is not exactly satisfied, the drift term should change sign during a reversal. Consequently, we require $v_e(x)$ to vanish at $x = 0$ and fit a smoothing spline through the individual estimates of $v_e(x)$. By comparison, the behaviour of $D(x)$ near $x \approx 0$ is not known. However, we do expect $D(x)$ to be an even function of x. Consequently, we represent the diffusion term using a best-fitting low-degree polynomial in x (i.e. $D(x) = c_0 + c_2x^2 + c_4x^4$). We explore the consequences of other choices below.

Fig. 3 reveals that the drift term is nearly a linear function of x around the equilibrium state $x = x_{\rm eq}$. When the drift term is represented by

$$v_e(x) = -\gamma(x - x_{\rm eq}). \tag{9}$$

We recover $\gamma = 34$ Myr^{-1} and $x_{\rm eq} = 5.3 \times 10^{22}$ A m^2 for PADM2M and $\gamma = 35$ Myr^{-1} and $x_{\rm eq} = 6.1 \times 10^{22}$ A m^2 for SINT2000. Substituting the linear representation from (9) into the stochastic differential equation shows that small departures from equilibrium, $\epsilon = x - x_{\rm eq}$, are described by

$$\frac{{\rm d}\epsilon}{{\rm d}t} = -\gamma\,\epsilon + \sqrt{D(x_{\rm eq} + \epsilon)}\Gamma(t)\,. \tag{10}$$

The deterministic part of (10) relaxes towards equilibrium with a characteristic timescale of $\gamma^{-1} \approx 29$ kyr, which reflects the time required for the dynamo process to adjust to a disturbance in the dipole field. Thus the slope of the drift term near the equilibrium state contains important information about the dynamics. Over a broader range of x we can represent the drift term by $v_e(x) = -\nabla U(x)$, where $U(x)$ can be viewed as a potential for the dipole field. Inverting for $U(x)$ from the drift terms yields a bistable potential well (see Fig. 4). The drift term drives the dipole field towards an equilibrium at the base of one of the potential wells, but the random component continually displaces x from the equilibrium state. A reversal occurs when a sequence of random perturbations are sufficient to push the dipole field across the barrier at $x = 0$. Once the dipole field changes sign, the drift term drives the signed dipole moment, $x(t)$, into a new equilibrium.

4 NUMERICAL REALIZATIONS OF THE STOCHASTIC PROCESS

Realizations of the stochastic process are obtained by numerically integrating (1). We extend the model to yield a signed VADM by defining $v_e(-x) = -v_e(x)$ and $D(-x) = D(x)$, consistent with

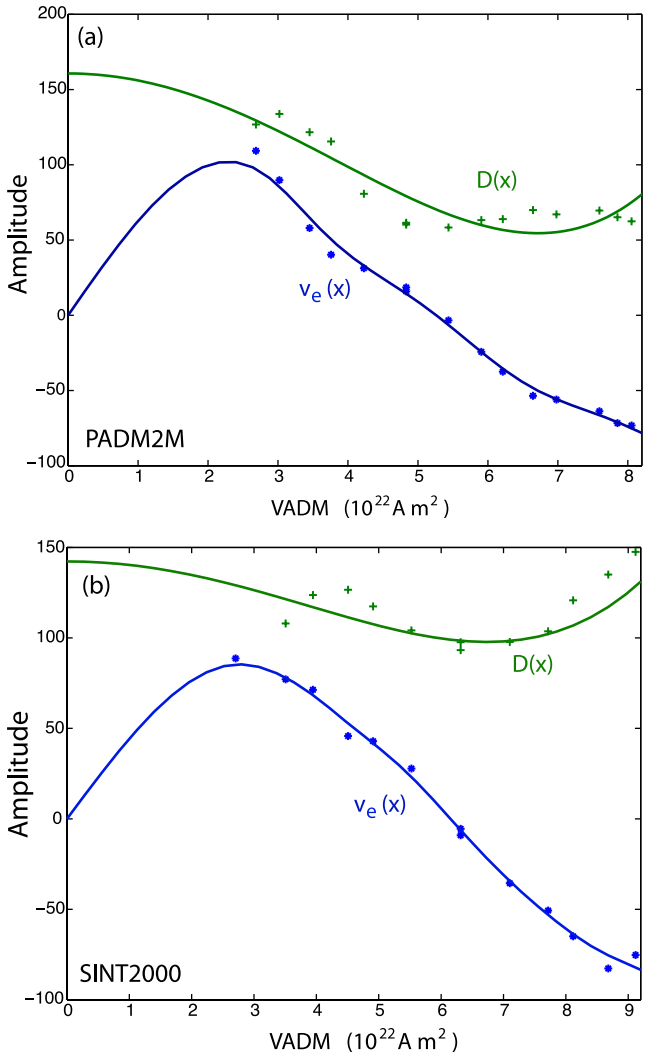

Figure 3. Drift, $v_e(x)$ and diffusion, $D(x)$, terms for the (a) PADM2M and (b) SINT2000 models. A smooth spline fit extrapolates the drift term to small values of VADM. A low-degree polynomial is used to extrapolate the diffusion term (see text for details).

expectations that magnetic induction is independent of the sign of the field (e.g. Jones 2011). Let t_n denote a set of discrete times separated by a constant time step Δt. When $x_n = x(t_n)$ is known, the solution at the next time step is approximated by (Risken 1989)

$$x_{n+1} = x_n + v_e(x_n)\Delta t + \sqrt{D(x_n)\Delta t}\, w_n, \tag{11}$$

where w_n is an independent Gaussian random variable with zero mean and variance 2 [i.e. the variance of the noise source in (3)]. Because $D(x)$ is evaluated at x_n rather than the mean value over the time step, we use the effective drift $v_e(x)$ to account for the effects of multiplicative noise. Each realization is distinct, but the mean properties of the solution converge when the integration time is sufficiently long.

Fig. 5 shows a 10-Myr realization using the parameters recovered from the PADM2M and SINT2000 models. Both solutions have several reversals and spends roughly equal time in the normal and reversed polarities. The mean of the unsigned VADM for the PADM2M numerical realization is 5.29×10^{22} A m^2 and the standard deviation is 1.47×10^{22} A m^2. By comparison, the mean and standard deviation of the PADM2M model is 5.32×10^{22} and 1.48×10^{22} A m^2, respectively. Similarly, the SINT2000 realization

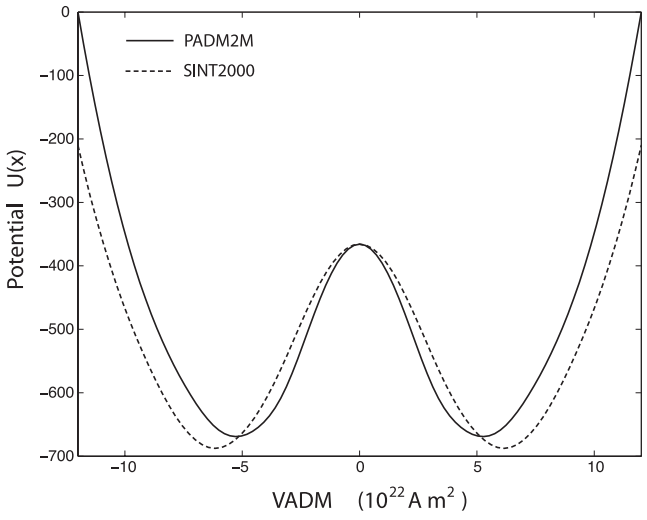

Figure 4. Potential well, $U(x)$, constructed from PADM2M and SINT2000 using symmetry about $x = 0$. Random perturbations displace the dipole moment from the equilibrium values at the base of the potential wells. The dipole field reverses polarity when perturbations are large enough to push $x(t)$ into the other potential well. The mean time between reversals is set by the shape of the potential well [or equivalently $v_e(x)$] and the amplitude of the drift term, $D(x)$.

yields a mean and standard deviation of 6.15×10^{22} and 1.91×10^{22} A m^2, whereas the SINT2000 model has a mean and standard deviation of 6.24×10^{22} and 1.97×10^{22} A m^2, respectively. Small differences are expected because each realization of the numerical model yields a different result. The uncertainty in the mean and standard deviation of the numerical solution can be inferred from multiple realizations. We find that the mean and standard deviation from independent realizations scatter around the mean and standard deviation of input models, suggesting that the stochastic models are consistent with the statistical properties of the VADM models. We also find reasonably good agreement in the histograms of the unsigned VADM (see Fig. 6). These histograms are normalized to give discrete probability distributions using the same number of bins employed in constructing $v_e(x)$ and $D(x)$. The smooth solid lines give the steady-state probability distributions obtained from a numerical solution of the Fokker–Planck equation (see Section 5).

5 PROBABILITY DISTRIBUTION OF THE AXIAL DIPOLE FIELD

A statistical description of the dipole field is given by the Fokker–Planck equation in (4). Numerical solutions are integrated using a finite volume approximation in x (see Appendix). The discrete equations for the probability distribution $P_i(t) = P(x_i, t)$ can be written in the form

$$\frac{\partial \mathbf{P}}{\partial t} = \mathbf{AP}, \tag{12}$$

where $\mathbf{P} = [P_1, P_2, \ldots, P_n]^T$ defines the probability distribution at a given time and matrix $\mathbf{A}$ approximates the operator on the right-hand side of (4). Boundary conditions on $P(x, t)$ require $P(\pm\infty, t) = 0$, although in practice it suffices to set $P_0 = P_{n+1} = 0$ at $x = x_0$ and x_{n+1} when the range of x is sufficiently large. An initial value problem also requires the probability distribution at $t = 0$. When the dipole moment at $t = 0$ is known, the probability distribution is effectively a Dirac

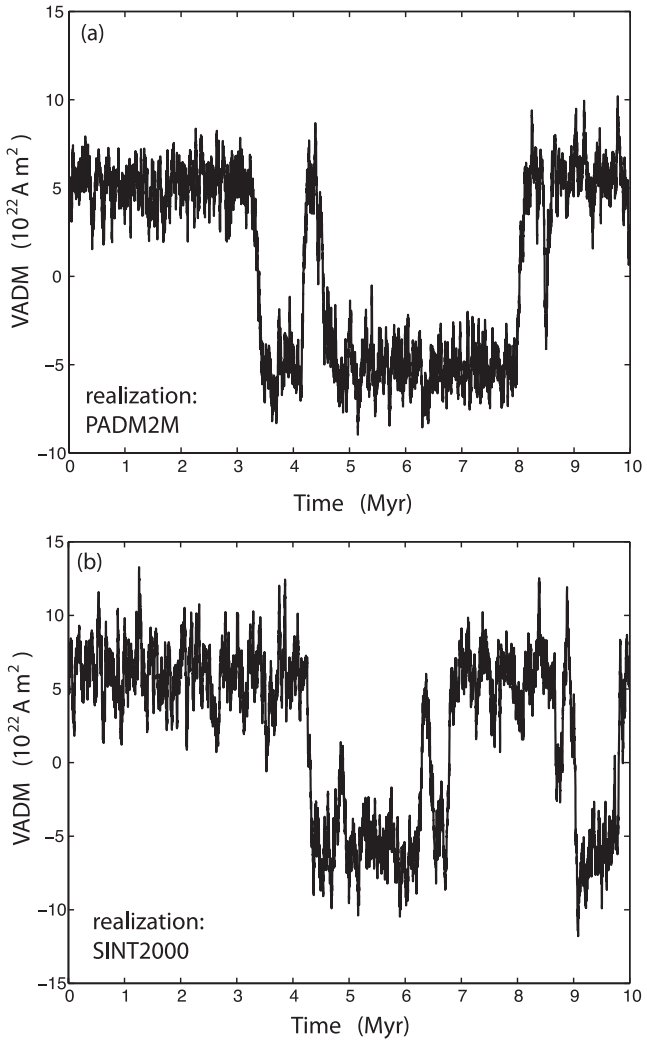

Figure 5. A numerical realizations of $x(t)$ over a 10-Myr interval using (a) the PADM2M-based model and (b) the SINT2000-based model.

delta function at the initial value of the dipole moment

$$P(x, 0) = \delta(x - x(0)). \tag{13}$$

In numerical calculations, the initial distribution is approximated by a narrow Gaussian distribution centred at $x = x(0)$. As time evolves the initial probability distribution drifts and broadens under the influence of the drift and diffusion terms in the Fokker–Planck equation. Useful insights into the process can be drawn from the first few eigenvalues of the discrete system in (12).

5.1 An eigenvalue problem

When the time dependence of the discretized probability distribution is represented by

$$P_i(t) = \tilde{P}_i \, e^{\lambda t}, \tag{14}$$

the system of equations in (12) reduces to

$$(\mathbf{A} - \lambda \mathbf{I})\tilde{\mathbf{P}} = 0, \tag{15}$$

where $\mathbf{I}$ is the identity matrix and λ are the eigenvalues of A. All of the eigenvalues are either negative or zero. The smallest $|\lambda|$ is $\lambda_0 = 0$, which corresponds to the steady-state probability distribution. Fig. 7(a) shows the steady-state distribution (or zeroth eigenfunction $\tilde{\mathbf{P}}_0$) for the PADM2M stochastic model, which closely approximates the histogram for PADM2M (see Fig. 6). The next smallest eigenvalue is $\lambda_1 = -0.97$ Myr^{-1}; the corresponding eigenfunction, $\tilde{\mathbf{P}}_1$, is shown in Fig. 7(b). Superposition of $\tilde{\mathbf{P}}_0$ and $\tilde{\mathbf{P}}_1$ yields another probability distribution in which the dipole moment is largely confined to one polarity (see Fig. 7c). In fact, the expected value of x in Fig. 7(c) defines the time average and closely coincides with the base of one of the potential wells. If we use $\tilde{\mathbf{P}}_0 + \tilde{\mathbf{P}}_1$ as an initial condition in the Fokker–Planck equation, the first eigenfunction decays with time while the zeroth eigenfunction persists. In other words, the initial probability distribution in Fig. 7(c) relaxes towards a steady-state distribution with an e-folding time of $|\lambda_1^{-1}| = 1.03$ Myr. A similar calculation for the SINT2000 model yields $|\lambda_1^{-1}| = 1.24$ Myr. We show below that this timescale gives a good estimate of the mean time between reversals.

5.2 Steady-state probability distribution

A useful approximation to the steady-state distribution can be found with a few simplifications to the drift and diffusion terms. First, we use the small-departure approximation for $v_e(x)$ from (9), which introduces the slope, γ, of the drift term around the equilibrium value, x_{eq}. Secondly, we let $D(x) \approx D(x_{eq})$. With these two simplifications in (4), the steady-state Fokker–Planck equation can be integrated directly for $P(x, \infty)$. The resulting normal distribution

$$P(x, \infty) = \frac{1}{2}\sqrt{\frac{\gamma}{2\pi D_{eq}}} e^{-\gamma(x - x_{eq})^2/2D_{eq}} \tag{16}$$

has a normalization factor included such that total probability under a single peak at $x = x_{eq}$ is 1/2. An identical normal distribution is assumed in the region around $x = -x_{eq}$. These local distributions give a reasonable good description of the steady-state probability in Fig. 7(a). The standard deviation of the dipole moment about x_{eq} is approximately $\sigma \approx \sqrt{D_{eq}/\gamma}$. Using $\gamma = 34$ Myr^{-1} and $D_{eq} = 69 \times 10^{44}$ A^{2}m^4 Myr^{-1} gives $\sigma = 1.42 \times 10^{22}$ A m^2 for the PADM2M-based model. Recall that the standard deviation of the PADM2M model is 1.48×10^{22} A m^2. For SINT2000, we obtain $\sigma = 1.68 \times 10^{22}$ A m^2 using $\gamma = 35$ Myr^{-1} and $D_{eq} = 99 \times 10^{44}$ A^2 m^4, compared with $\sigma = 1.97 \times 10^{22}$ A m^2 from the SINT2000 stack. Much of the disagreement in σ is probably due to the approximations used in (16) because the standard deviation from long-time realizations yield much better agreement with the VADM models. Still, the normal distribution in (16) is a useful approximation for the steady-state distribution. The broad agreement suggests that the stochastic models reproduce the variability of the VADM models. On the other hand, the mean time between reversals (as inferred from eigenvalue λ_1) hints at an inconsistency with the observed rate of reversals (Merrill & McFadden 1994). We explore this question in Section 6.

6 MEAN REVERSAL AND EXCURSION RATES

The mean time for an event (like a reversal) is often described in literature as the mean first-passage time. Similar problems arise in a wide range of applications (e.g. Risken 1989). The calculation of a mean first-passage time relies on a solution of the Fokker–Planck equation with boundary conditions that define the event of interest. (Typically an event occurs when $x(t)$ first moves outside a specific range of values.) Suppose we start with an initial dipole

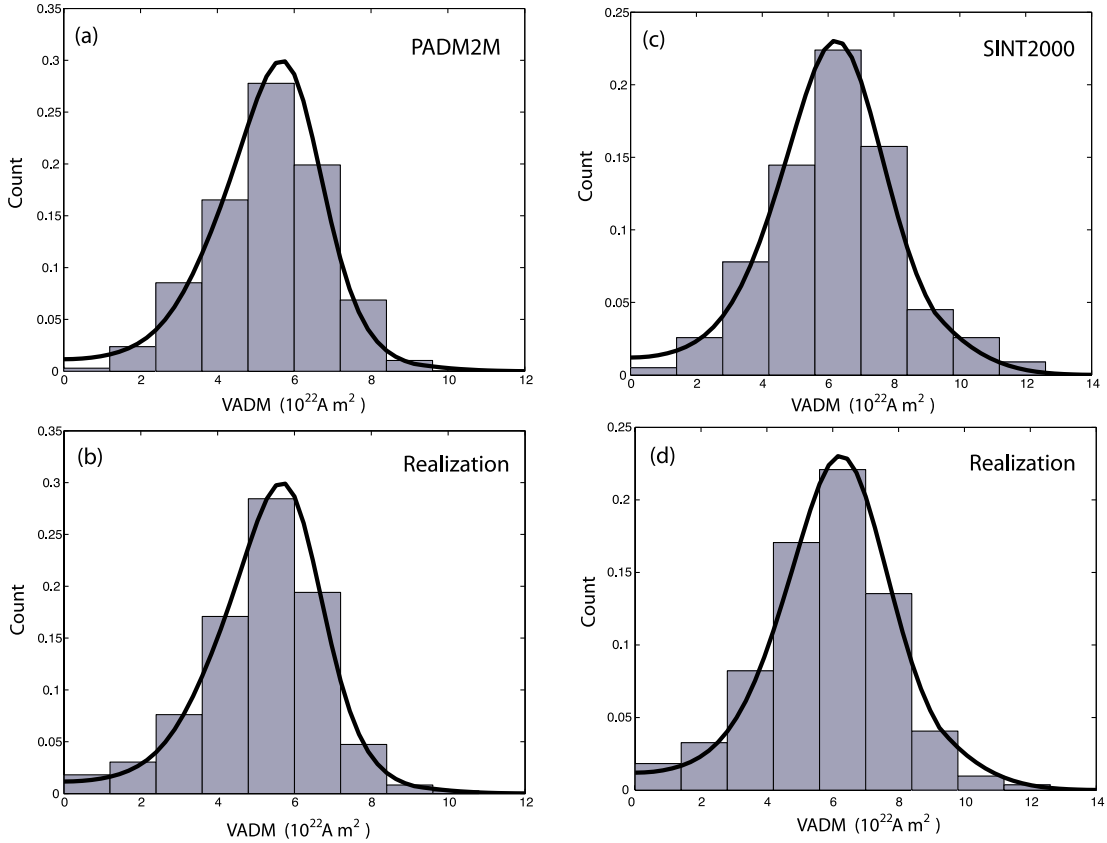

Figure 6. (a) Histogram of the PADM2M model and (b) a numerical realization. (c) Histogram of SINT2000 model and (d) a numerical realization. The solid lines indicate the steady-state probability distribution computed from the Fokker–Planck equation.

moment $x = x(0)$ in the range of positive polarity ($0 < x(0) < \infty$). The initial probability distribution, $P(x, 0) = \delta(x - x(0))$, expresses the certainty of the initial value of the VADM. As time advances the initial probability distribution evolves according to (4) to include a finite probability that $x(t)$ occupies the opposite polarity. The mean first-passage time is defined in terms of the time-dependent probability that $x(t)$ leaves the initial range $0 < x(0) < \infty$. Similarly, we could define the mean time for an excursion in terms the time required for $x(t)$ to leave a more restricted initial range of $x_{ex} < x(0) < \infty$, where the lower bound is set the value of x at the time of the Laschamp excursion in the PADM2M model. In other words, the mean excursion time is defined as the average time required for $x(t)$ to drop below $x_{ex} = 3.48 \times 10^{22}$ A m^2. Such a drop appears to be sufficient to cause a large tilt of the dipole from the geographic axis in the presence of an equatorial dipole and non-dipole field (Merrill & McFadden 1994).

Consider a large number of realizations that begin with an initial (signed) dipole moment $x(0)$ in the range $0 < x(0) < \infty$. Each realization will eventually leave the interval $0 < x < \infty$, although the time T required will vary between realizations. We seek the average time, $\bar{T}$, for realizations to leave the interval $0 < x < \infty$ or, equivalently, the average time for a reversal. Let $P(x, t|x(0))$ denote the conditional probability distribution of realizations that begin at $x(0)$ but have not reversed at time t. The distribution $P(x, t|x(0))$ satisfies the Fokker–Planck equation on the interval $0 < x < \infty$ with an initial condition $P(x, 0|x(0)) = \delta(x - x(0))$. Boundary conditions at $x = 0$ and $x = \infty$,

$$P(0, t|x(0)) = P(\infty, t|x(0)) = 0\,, \tag{17}$$

ensure that the probability distribution vanishes outside the range of positive polarities. In effect, realizations that leave the interval $0 < x < \infty$ are removed from further consideration. A numerical solution of this initial value problem is used to define the mean reversal time.

The number of realizations that survive in the interval $0 < x(t) < \infty$ progressively decreases with time. Each loss of a realization implies a polarity reversal, so the probability of reversal during the time interval $(t, t + dt)$ is

$$f(t)\,\mathrm{d}t = \left[-\int_0^\infty \dot{P}(x, t|x(0))\,\mathrm{d}x \right] \mathrm{d}t, \tag{18}$$

where the dot denotes the time derivative. It follows that the mean first-passage time is given by

$$\bar{T} = \int_0^\infty T f(T)\,\mathrm{d}T. \tag{19}$$

A direct determination of $f(t)$ can be computed from (18) using a numerical solution of the Fokker–Planck equation. At each time step, the right-hand side of the Fokker–Planck equation in (4) is evaluated and the resulting time derivative, $\dot{P}(x, t)$, is integrated over x; the limits of the integral are based on the definition of the event (i.e. $0 < x < \infty$ for the case of a reversal). Fig. 8 shows the reversal time distribution, $f(t)$, for the PADM2M model using an initial condition $x(0) = 5.3 \times 10^{22}$ A m^2. After an initial transient $f(t)$ decreases exponentially with time. The mean reversal time from (19) is $\bar{T} = 1.05$ Myr, which is close to the value inferred from eigenvalue λ_1. The mean reversal time for the SINT2000 model is $\bar{T} = 1.27$ Myr, which is also close to the value inferred from the corresponding λ_1.

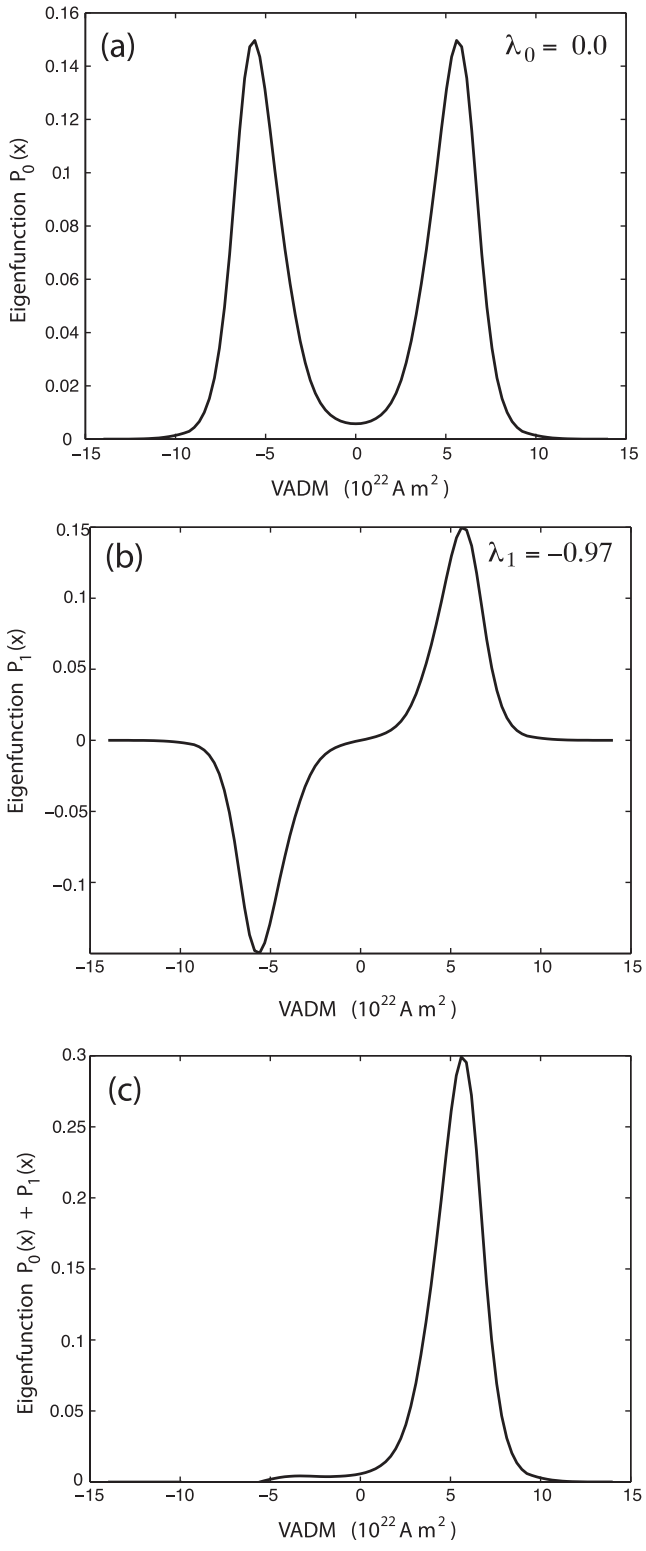

Figure 7. Eigenfunctions computed from (15) using the PADM2M model. (a) Steady-state probability distribution corresponds to the zeroth eigenvalue $\lambda_0 = 0$; (b) Probability distribution associated with the first eigenvalue λ_1; (c) Superposition of the probability distributions in (a) and (b). The resulting distribution relaxes to the steady-state distribution on the timescale λ_1^{-1}.

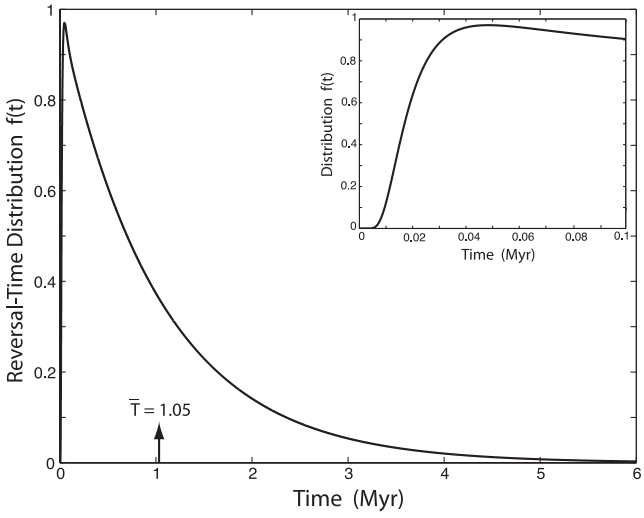

Figure 8. Reversal time distribution, $f(t)$, computed from a numerical solution of the Fokker–Planck equation with an initial condition $x(0) = 5.3 \times 10^{22}$ A m^2. After a short transient (inset) the probability distribution decreases exponentially with t.

The connection to the eigenvalue problem is straightforward. The first eigenfunction

$$P_1(x, t) = \tilde{P}_1(x)\mathrm{e}^{\lambda_1 t} \tag{20}$$

obeys the Fokker–Planck equation and satisfies the same boundary conditions imposed on $P(x, t|x(0))$. Thus the solutions for $P_1(x, t)$ and $P(x, t|x(0))$ differ only in the initial condition. Recall that $P(x, 0|x(0))$ is equal to a delta function, whereas $P_1(x, 0)$ is set by the eigenfunction. Both $P_1(x, t)$ and $P(x, t|x(0))$ decay with time as 'realizations' leave the interval $0 < x < \infty$. The integral of $P_1(x, t)$ over the interval $0 < x < \infty$ has an exponential dependence on time, whereas the integral of $P(x, t|x(0))$ has a nearly exponential dependence on time after a short transient. In fact, the transient period corresponds to the adjustment of $P(x, t|x(0))$ from the initial delta function to a distribution that closely approximates the shape of the eigenfunction. Subsequent decreases in $P(x, t|x(0))$ occur with an e-folding time of roughly λ_1^{-1}. According to (19), the mean reversal time is set by λ_1^{-1} plus a short transient period.

The mean time between excursions is calculated by redefining the event to represent a realization leaving the interval $x_{ex} < x < \infty$. Strictly speaking, the mean excursion time defines the time required for the dipole field to drop below x_{ex} when started at the equilibrium value. Such an event is likely to cause a large angular deviation in the direction of the VGP if the non-dipole and equatorial dipole parts of the field are comparable to the present-day values. On the other hand, our definition of an excursion using $x(t) < x_{ex}$ is indirect and assumes that the Laschamp event is representative of other excursions. Numerical calculations for PADM2M give a mean excursion time of $\bar{T} = 0.119$ Myr or about 8.4 excursions Myr^{-1}. For SINT2000, we obtain $\bar{T} = 0.121$ Myr, or about 8.3 excursion Myr^{-1}.

Independent palaeomagnetic observations suggest that reversals, and possibly excursions, occur more frequently (Merrill & McFadden 1994; Roberts 2008). One way to reconcile the predictions and observations is to increase the overall amplitude of noise [i.e. $D(x)$], so the stochastic model jumps more frequently between the two stable potential wells in Fig. 4. Increasing $D(x)$ by a factor of 2 increases the reversal rate of the PADM2M-based model to roughly 4 Myr^{-1}, which is more consistent with observations. Unfortunately, this revision has the consequence of producing a mismatch between

the observed and predicted variations in the dipole moment. Recall that the standard deviation of $x(t)$ in the stochastic model is approximated by $\sqrt{D_{\rm eq}/\gamma}$. If we increase $D_{\rm eq}$ by a factor of 2, the resulting standard deviation becomes $\sigma = 2.0 \times 10^{22}$ A m^2, compared with $\sigma = 1.48 \times 10^{22}$ A m^2 for the PADM2M model. In this case, the stochastic model does not reproduce the statistical behaviour of the input model.

An alternate way to reproduce the observed reversal rate is to adjust the values of $v_e(x)$ and $D(x)$ at low x, where observations provide few direct constraints on the stochastic model. To explore this idea, we define an event in the stochastic model to correspond to a specific recurrence time. For example, the value of $x(t)$ for the PADM2M-based stochastic model is expected to drop below 1.82×10^{22} A m^2 three times in 1 Myr. This prediction agrees very well with the actual occurrence of low dipole moment in the PADM2M model. In Fig. 1, we observe six events below 1.82×10^{22} A m^2 over the past 2 Myr, including the Cobb Mountain subchron at 1.2 Myr BP. An identical recurrence rate of 3 Myr^{-1} is predicted by the SINT2000-based stochastic model if the threshold for the event is set at $x = 2.29 \times 10^{22}$ A m^2. The SINT-2000 stack in Fig. 1 shows seven events in the past 2 Myr, which agrees reasonably well with predictions. The point is that the stochastic model successfully predicts the occurrence of low dipole moments, but it does not correctly predict the reversal rate. Interestingly, each of the low dipole events in Fig. 1 corresponds to a reversal, so once $x\,(t)$ drops below $\approx 2 \times 10^{22}$ A m^2 there appear to be no impediment to reversing polarity.

This behaviour can be interpreted several ways. For example, we might flatten or truncate the barrier between the potential wells over the interval $|x| < 2 \times 10^{22}$ A m^2. Once the solution reaches $x \approx 2 \times 10^{22}$ A m^2, there is no further obstacle to overcome. Such a potential well requires $v_e(x)$ to vanish for $|x| < 2 \times 10^{22}$ A m^2, meaning that persistent dipole generation does not begin until $|x|$ exceeds a threshold. Alternatively, we could increase $D(x)$ at low x, ensuring that the random component is more effective in jumping the solution over the barrier between the potential wells at low x. It is also possible that the VADM from relative palaeointensity measurements is adversely affected by non-dipole and equatorial dipole components at low x. As a result, the recovery of drift and diffusion terms at low x may not solely reflect the axial dipole moment. It appears that independent estimates of reversal rates offer valuable insights into the reversal process.

We conclude this section with a brief comment on the probability of long chrons. To a good approximate, the reversal-time distribution can be represented by an exponential distribution

$$f(t) = \frac{1}{\bar{T}} \mathrm{e}^{-t/\bar{T}}. \tag{21}$$

The same probability distribution is often used to describe chron lengths (Lowrie & Kent 2004), although here the independent variable is time (measured from the initial condition) rather than chron length. On the other hand, equivalent inferences can be drawn about chron lengths. For example, the probability of a chron lasting longer than t_c is

$$p(t > t_c) = \int_{t_c}^{\infty} f(t)\,\mathrm{d}t = \mathrm{e}^{-t_c/\bar{T}}. \tag{22}$$

A chron lasting longer than 2 Myr has a probability of $p = 0.15$ when $\bar{T} = 1.05$ Myr. The true probability is likely be much lower with a more realistic value of $\bar{T}$. The Cretaceous superchron ($t_c = 37$ Myr) has a vanishingly small probability of $p = 5 \times 10^{-16}$ when $\bar{T} = 1.05$ Myr. To bring the duration of the Cretaceous superchron into line with the mean reversal time of the stochastic model, we need to reduce the diffusion term. Reducing $D(x)$ in the PADM2M-based model by a factor of 2 increases the mean reversal time to $\bar{T} = 12.3$ Myr. In this case, the probability of the Cretaceous superchron is about 5 per cent. The precise reduction in $D(x)$ is not uniquely set by the duration of the superchron, but it does appear that some reduction is required to make the Cretaceous superchron feasible. We interpret the significance of this change in Section 7.

7 PHYSICAL INTERPRETATION OF DRIFT AND DIFFUSION TERMS

A physical interpretation of the stochastic model requires an understanding of the processes that regenerate the dipole field. While many details are still poorly understood, the results of numerical geodynamo models offer some insights. The axial or z-component of the dipole moment is defined by

$$x(t) = \frac{1}{2} \int_V \hat{\mathbf{z}} \cdot [\mathbf{r} \times \mathbf{J}(\mathbf{r}, \mathbf{t})]\,\mathrm{d}V, \tag{23}$$

where $\mathbf{J}$ is the electric current density, $\mathbf{r}$ is the position vector and V is the volume of the Earth's core. When there are no currents outside V (i.e. no external sources), the axial moment can be expressed in terms of the magnetic field $\mathbf{B}$ (Davidson 2001) as

$$x(t) = \frac{3}{2\mu} \int_V \hat{\mathbf{z}} \cdot \mathbf{B}\,\mathrm{d}V, \tag{24}$$

where μ is the magnetic permeability. Only the dipole part of the magnetic field contributes to the integral in (24) because the other spherical harmonic components with degrees $n > 1$ vanish when integrated over the spherical volume. To make the role of the dipole field more explicit, we introduce the notation $\overline{\mathbf{B}}$ to denote the projection of the total magnetic field onto the $n = 1$ term of a vector spherical harmonic expansion (e.g. Chandrasekhar 1981).

Numerical geodynamo models suggest that the z-component of the magnetic field is generated by helical convection inside the core (Olson *et al.* 1999). These flows lift and twist an initially azimuthal field (denoted by $\mathbf{B}_\phi$), causing electric currents in the azimuthal direction (Parker 1955). The time evolution of $\overline{\mathbf{B}}$ is governed by the induction equation

$$\frac{\partial \overline{\mathbf{B}}}{\partial t} = \nabla \times \overline{(\mathbf{v} \times \mathbf{B})} + \eta \nabla^2 \overline{\mathbf{B}}, \tag{25}$$

where the magnetic diffusivity $\eta = 1/(\mu\sigma_{\rm e})$ is defined in terms of the electrical conductivity $\sigma_{\rm e}$. The characteristic length scale of helical flow in the core is probably small compared with the typical length scale of the B_ϕ field because the viscosity of the liquid core is much smaller than the magnetic diffusivity (Sakuraba & Robert 2009). Consequently, an order of magnitude estimate for the induction term is

$$\overline{\mathbf{v} \times \mathbf{B}} \approx \epsilon B_\phi v^2 l/\eta, \tag{26}$$

where v is the typical convective velocity and l is the width of a helical convection column (Moffatt 1970). The parameter ϵ is introduced to represent the fraction of magnetic induction that projects into the $n = 1$ toroidal component. Physically, a helical flow produces both positive and negative B_z, so there is considerable cancellation in the integral for the dipole moment (see eq. 24). Numerical models suggest that the main contribution to the dipole moment occurs when the upward and downward-directed field lines are separated by flow in the core. Expulsion of one of the field pair from the core volume produces a net change in the dipole moment (Olson

et al. 1999). Consequently, we can view ϵ as the fraction of field loops that have one sign of B_z ejected from the core.

The diffusion term in (25) can be approximated by

$$\eta\nabla^2\overline{\mathbf{B}} \approx \overline{\mathbf{B}}/\tau_\mathrm{d}, \tag{27}$$

where $\tau_\mathrm{d} = R^2/(\pi^2\eta)$ is the dipole diffusion time and $R = 3480$ km is the radius of the core (e.g. Gubbins & Roberts 1987). This approximation is exact when $\overline{\mathbf{B}}$ has the structure of the slowest dipole decay mode. More generally, $\overline{\mathbf{B}}$ can be represented by a linear superposition of all dipole decay modes. The resulting value for τ_d would be a suitable average of the corresponding diffusion times.

Differentiating (24) with respect to time defines an equation for the evolution of the dipole moment. The rate of change of $\overline{\mathbf{B}}$ is governed by the induction equation in (25), so the evolution of $x(t)$ is described by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = S(t) - \frac{x(t)}{\tau_\mathrm{d}}, \tag{28}$$

where the source term $S(t)$ is approximated by

$$S(t) \approx \frac{3\epsilon}{2\mu}\left(\frac{B_\phi v^2 l}{\eta L}\right) V\,. \tag{29}$$

Here, $L \approx R$ characterizes the length scale of variations in $\overline{\mathbf{v}\times\mathbf{B}}$ because we are dealing exclusively with the $n = 1$ component of the source. Taking the time average of (28) yields

$$\langle S\rangle = \frac{\langle x\rangle}{\tau_\mathrm{d}}, \tag{30}$$

which means that the time evolution of the dipole moment can be written as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{1}{\tau_\mathrm{d}}(x(t) - \langle x\rangle) + \Delta S(t)\,. \tag{31}$$

The first term on the right-hand side of (31) corresponds to the drift term of the stochastic model and $\Delta S = S(t) - \langle S\rangle$ represents a random fluctuation in the source term about the time average.

We are now in a position to interpret the drift and diffusion terms. Our estimate of the drift term was approximated by

$$v_e \approx -\gamma(x - x_\mathrm{eq}), \tag{32}$$

where the best-fitting slope for the PADM2M model was $\gamma = 34$ Myr^{-1}. Comparison with the drift term in (31) indicates that γ is set by the dipole decay time. The stochastic models gives $\tau_\mathrm{d} = \gamma^{-1} = 29$ kyr. By comparison, the expected dipole decay time is $\tau_\mathrm{d} = 24.5$–49 kyr for an electrical conductivity of $\sigma_\mathrm{e} = 0.5 \times 10^6$ to 10^6 S m^{-1} and a magnetic permeability of $\mu = 4\pi \times 10^{-7}$ in SI units. While the decay time from the stochastic model lies within the expected range, the best-fitting conductivity, $\sigma_\mathrm{e} = 0.6 \times 10^6$, is close to the low end of the range (Braginsky & Roberts 1995). Higher values from more recent studies (de Koker *et al.* 2012; Pozzo *et al.* 2012) can be reconciled with the stochastic model if the radial structure of $\overline{\mathbf{B}}$ is more complicated than that of the slowest decay mode. On the other hand, $\overline{\mathbf{B}}$ cannot depart dramatically from the slowest decay mode. The next fastest decay mode has a decay time of $\tau_\mathrm{d} = R^2/(4\pi^2\eta)$, which yields 12.5 kyr for the higher conductivity, and even shorter times for the lower conductivity. To reconcile the value of γ^{-1} with reasonable values for the electrical conductivity, it seems that most of the energy in $\overline{\mathbf{B}}$ must accumulate in the slowest decay mode.

Insights into the efficiency of dipole generation can be inferred from the time average of $S(t)$ in (30). Using $\langle x\rangle = 5.3 \times 10^{22}$ A m^2 and $\tau_\mathrm{d} = 29$ kyr gives $\langle S\rangle = 5.8 \times 10^{10}$ A m^2 s^{-1}. This value can be compared with the approximate estimate for $S(t)$ in (29). Using $B_\phi = 1$ mT, $l = 100$ km and $v = 10^{-4}$ m s^{-1} gives $\epsilon = 1.5 \times 10^{-3}$. This low efficiency suggests that a small fraction of the total magnetic induction contributes to the dipole field. Equivalently, a small fraction of the field loops produced by helical flow have one sign of B_z ejected from the core; the vast majority of the field loops remain trapped inside the core. Interestingly, the value of $\langle x\rangle$ implies a volume-averaged axial field of $B_z = 0.25$ mT, according to (24). As noted previously, this part of the field is associated with the dipole component, so the higher degree components in a spherical harmonic expansion are not represented in the volume average. A small magnetic field at degree $n = 1$ is compatible with the idea that most of the energy in the magnetic field is contained in small-scale components.

The significance of the diffusion term, $D(x)$, is connected with fluctuations in $x(t)$ over convective timescales (say $\tau_c = 1$ kyr). The magnitude of a random fluctuation is given by

$$\Delta x(\tau_c) = \int_0^{\tau_c} \Delta S(t)\,\mathrm{d}t = \sqrt{D\tau_c}\,w_n, \tag{33}$$

where w_n is a random variable with a standard deviation of $\sqrt{2}$. Consequently, the rms variation in $\Delta x(\tau_c)$ is

$$\sigma_{\Delta x} = \sqrt{2D\,\tau_c} \tag{34}$$

or, equivalently, about 7 per cent of the time average $\langle x\rangle$ in both the PADM2M and SINT2000 models. By comparison, the duration of the Cretaceous superchron can plausibly be explained by reducing D by a factor of 2. In this case, the rms variation in $x(t)$ over τ_c is about 5 per cent of the time average. Such a decrease in the rms variation could be achieved by reducing the rms velocity fluctuations by a factor of $\sqrt{2}$, given the velocity dependence of $S(t)$ in (29). Changes in the spatial pattern of heat flux at the core–mantle boundary have been proposed as one mechanism to alter reversal rates (Glatzmaier *et al.* 1999; Olson *et al.* 2010). Alternatively, a change in velocity fluctuations may be associated with a change in the total core heat flow (Driscoll & Olson 2012). Our results suggest that a modest change in the velocity fluctuations can have a large influence on average reversal rate.

8 CONCLUSIONS

Fluctuations in VADM can be described quantitatively using a simple stochastic model. We construct two stochastic models using estimates of the drift and diffusion terms recovered from PADM2M (Ziegler *et al.* 2011) and SINT2000 (Valet *et al.* 2005). The results are used in numerical solutions of the Fokker–Planck equation to predict statistical properties of the axial dipole, including the mean time between reversals and excursions. The mean reversal time is 1.05 Myr for the PADM2M-based model and 1.27 Myr for the SINT2000-based model. Evidence for more frequent reversals in the recent geological past (Merrill & McFadden 1994) can be reconciled with predictions from the stochastic models by adjusting the drift and diffusion terms at low values of VADM (denoted by x). Interestingly, the stochastic models are able to reproduce the infrequent occurrence of low x in the models of Valet *et al.* (2005) and Ziegler *et al.* (2011). For example, we expect the VADM to drop below $x \approx 2 \times 10^{22}$ A m^2 at a rate of 3 Myr^{-1}, which is consistent with both VADM models. Each of these events coincides with a reversal, suggesting a threshold $|x| < 2 \times 10^{22}$ A m^2 for the occurrence of reversals. Decreasing the amplitude of the drift term in the interval $|x| < 2 \times 10^{22}$ A m^2 promotes reversals by effectively lowering the

barrier between stable polarities. Physically, a lower barrier means that persistent dipole generation does not occur until the threshold value of x is exceeded. Alternatively, we can increase the amplitude of the diffusion term in the interval $|x| < 2 \times 10^{22}$ A m^2. A larger diffusion term increases the random fluctuations, shortening the time required to jump between stable polarities. It is also possible that estimates of VADM are more affected non-axial dipole field at low x, so the recovery of drift and diffusion terms is less reliable in this region.

Geomagnetic excursions are approximated by the condition $x < x_{ex}$, where $x_{ex} = 3.48 \times 10^{22}$ A m^2 is set by the value of x in PADM2M at the time of the Laschamp excursion (Laj *et al.* 2000). Assuming that other excursions occur when x drops below x_{ex}, we obtain an average time between excursions of 0.119 Myr using the PADM2M-based model. The corresponding excursion rate is 8.4 excursions Myr^{-1}. The SINT200-based stochastic model yields a mean excursion time of 0.121 Myr, or 8.3 excursions Myr^{-1}. These results are in rough agreement with previous estimates (Laj & Channell 2007; Roberts 2008), although a direct correspondence between VADM and excursions is not strictly correct because angular deviations of the VGP from the geographic axis also depend on the non-axial dipole part of the field.

A physical interpretation of the drift and diffusion terms offers insights into the dynamics of the core. The timescale for adjustments in the axial dipole moment, $x(t)$, is controlled by the slope of the drift term $v_e(x)$ near the average value x_{eq}. We argue that the slope of the drift term is primarily determined by the dipole diffusion time τ_d. The drift terms from the VADM time models yield $\tau_d = 29$ kyr, which is consistent with current estimates of the diffusion time, although the best-fitting electrical conductivity, $\sigma_e = 6 \times 10^5$, is on the lower end of the current range. We also predict that dipole generation is relatively inefficient; only about 0.1 per cent to the total magnetic induction contributes to the axial dipole. Most of the magnetic energy in the dipole field appears to accumulate in the slowest decay mode.

The diffusion term is used to assess the amplitude of convective fluctuations. The rms variations in $x(t)$ on the timescale of a convective overturn is about 7 per cent of the average value x_{eq}. Assuming that the dipole is generated by an ensemble of several hundred independent convection columns, we expect the rms variation in the field generation of a single column to be comparable to the mean generation of a single column. Long chrons lasting tens of millions of years are most easily explained by a modest decrease in the diffusion term. For example, the Cretaceous superchron becomes likely when the present-day estimate of $D(x)$ is decreased by a factor of 2. The mean time between reversals increases to 12.3 Myr and the rms variation in $x(t)$ for a convective overturn time is about 5 per cent of the time average. Such a change in D is achieved by altering the rms fluctuations in the core velocity field by a factor of $\sqrt{2}$. This change in velocity is probably connected to the magnitude and/or spatial structure of the heat flux at the core–mantle boundary.

ACKNOWLEDGEMENTS

We thank Gauthier Hulot and an anonymous reviewer for many helpful comments and suggestions. This work is partially support by the National Science Foundation; EAR-1045277 (to BB), EAR-1049579 (to LZ) and EAR-1246826 (to CC).

REFERENCES

Braginsky, S.I. & Roberts, P.H., 1995. Equations governing convection in Earth's core, *Geophys. astrophys. Fluid Dyn.,* **79,** 1–97.

Brendel, K., Kuipers, J., Barkema, G.T. & Hoyng, P., 2007. An analysis of the fluctuations of the geomagnetic dipole, *Phys. Earth planet. Inter.,* **162,** 249–255.

Cande, S. & Kent, D., 1995. Revised calibration of the geomagnetic polarity timescale for the late Cretaceous and Cenozoic, *J. geophys. Res.,* **100,** 6093–6095.

Chandrasekhar, S., 1981. *Hydrodynamic and Magnetohydrodynamic Stability,* Dover Press, New York.

Constable, C.G. & Johnson, C.L., 1999. Anisotropic paleosecular variation models: implications for geomagnetic field observables, *Phys. Earth planet. Inter.,* **115,** 35–51.

Constable, C.G. & Johnson, C.L., 2005. A paleomagnetic power spectrum, *Phys. Earth planet. Inter.,* **153,** 61–73.

Constable, C.G. & Parker, R.L., 1988. Statistics of the geomagnetic secular variation for the past 5 m.y., *J. geophys. Res.,* **93,** 11 569–11 581.

Courtillot, V. & Olson, P., 2007. Mantle plumes link magnetic superchrons to phanerozoic mass depletion events, *Earth planet. Sci. Lett.,* **260,** 495–504.

Davidson, P.A., 2001. *An Introduction to Magnetohydrodynamics,* Cambridge University Press, Cambridge, UK.

de Koker, N., Steinle-Neumann, G. & Vicek, V., 2012. Electrical resistance and thermal conductivity of liquid Re alloys at high P and T, and heat flux in Earth's core, *Proc. Natl. Acad. Sci.,* **11,** 4070–4073.

Driscoll, P. & Olson, P., 2012. Superchron cycles driven by variable core heat flow, *Geophys. Res. Lett.,* **38,** L09304, doi:10.1029/2011GL046808.

Dumberry, M. & Finlay, C., 2007. The eastward and westward drift of the Earth's magnetic field the last three millenia, *Earth planet. Sci. Lett.,* **254,** 147–157.

Friedrich, R., Peinke, J., Sahimi, M. & Reza-Rahimi-Tabar, M., 2011. Approaching complexity by stochastic methods: from biological systems to turbulence, *Phys. Rep.,* **506,** 87–162.

Glatzmaier, G.A., Coe, R.S., Hongre, L. & Roberts, P.H., 1999. The role of the Earth's mantle in controlling the frequency of geomagnetic reversals, *Nature,* **401,** 885–890.

Gottschall, J. & Peinke, J., 2008. On the definition and handling of different drift and diffusion estimates, *New J. Physics,* **10,** 083034, doi:10.1088/1367-2630/10/8/083034.

Gubbins, D. & Roberts, P.H., 1987. Magnetohydrodynamics of the Earth's core, in *Geogmagnetism,* Vol. 2, ed. Jacobs, J.A., Academic Press, New York.

Guyodo, Y. & Valet, J.P., 1999. Global changes in geomagnetic intensity during the past 800 thousand years, *Nature,* **399,** 315–319.

Holme, R. & Olsen, N., 2006. Core surface flow modeling from high-resolution secular variation, *Geophys. J. Int.,* **166,** 518–528.

Hulot, G. & LeMouël, J.-L., 1994. A statistical approach to the Earth's main magnetic field, *Phys. Earth planet. Inter.,* **82,** 167–183.

Hulot, G., Eymin, C., Langlais, B., Mandea, M. & Olsen, N., 2002, Small-scale structure of the geodynamo inferred from Oersted and Magsat satellite data, *Nature*, **416**, 620–623.

Jackson, A., 1997. Time-dependency of tangentially geostrophic core surface motion, *Phys. Earth planet. Inter.,* **103,** 293–311.

Jackson, A., Jonkers, A.R.T. & Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records, *Phil. Trans. R. Soc. Lond., A,* **358,** 293–311.

Johnson, H.P., Van Patten, D., Tivey, M. & Sager, W.W., 1995. Geomagnetic polarity reversal rate for the Phanerozoic, *Geophys. Res. Lett.,* **22,** 231–234.

Jones, C.A., 2011. Planetary magnetic fields and fluid dynamo, *Ann. Rev. Fluid Mech.,* **43,** 583–614.

Korte, M. & Constable, C.G., 2005. Continuous geomagnetic models for the past 7 millenia II: CALS7K, *Geochem. Geophys. Geosyst.,* **6,** Q02H16, doi:10.1029/2004GC000801.

Korte, M., Genevey, A., Constable, C.G., Frank, U. & Schnepp, E., 2005. Continuous geomagnetic field models for the past 7 millennia, 1. A

new global data compilation, *Geochem. Geophys. Geosyst.,* **6,** Q02H15, doi:10.1029/2004GC000800.
Laj, C & Channell, J.E.T., 2007. Geomagnetic excursions, in *Treatise on Geophysics,* Vol. 5, pp. 373–316, Geomagnetism, ed. Kono, M., Elsevier, Amsterdam.
Laj, C., Kisset, C., Mazaud, A., Channell, J.E.T. & Beer, J., 2000. North Atlantic paleointensity stack since 75 ka (NAPIS-75) and the duration of the Laschamp event, *Phil. Trans. R. Soc. A,* **358,** 1009–1025.
Lhuillier, F., Fornier, A., Hulot, G. & Aubert, J., 2011. The geomagnetic secular-variation timescale in observations and numerical dynamo models, *Geophys. Res. Lett.,* **38,** L09306, doi:10.1029/2011GL047356.
Lowrie, W. & Kent, D.V., 2004. Geomagnetic polarity timescales and reversal frequency regimes, in *Timescales of the Paleomagnetic Field,* Geophysical Monograph Series, 145, American Geophysical Union, Washington, DC.
Lund, S.P. & Keigwin, L., 1994. Measurement of the degree of smoothing in sediment paleomagnetic secular variation records: an example from late Quaternary deep-sea sediments of the Bermuda Rise, western North Atlantic Ocean, *Earth planet. Sci. Lett.,* **122,** 317–330.
Merrill, R.T. & McFadden, P.L., 1994. Geomagnetic field stability: reversal events and excursions, *Earth planet. Sci. Lett.,* **121,** 57–69.
Moffatt, H.K., 1970. Turbulent dynamo action at low magnetic Reynolds number, *J. Fluid Mech.,* **41,** 432–452.
Olsen, N. & Mandea, M., 2008. Rapidly changing flows in the Earth's core, *Nature Geosci.,* **1,** 390–394.
Olson, P., Christensen, U. & Glatzmaier, G.A., 1999. Numerical modeling of the geodynamo: mechanisms of field generation and equilibration, *J. geophys. Res.,* **104,** 10 383–10 404.
Olson, P., Coe, R.S., Driscoll, P.E., Glatzmaier, G.A. & Roberts, P.H., 2010. Geodynamo reversal frequency and heterogeneous core-mantle boundary heat flow, *Phys. Earth planet. Inter.,* **180,** 66–79.
Pais, M.A., Oliveria, O. & Hogueira, F., 2004. Nonuniqueness of inverted core-mantle boundary flows and deviations from tangential geostropy, *J. geophys. Res.,* **109,** B08105, doi:10.1029/2004JB003012.
Parker, E.N., 1955. Hydromagnetic dynamo models, *Astrophys. J.,* **122,** 293–314.
Pozzo, M., Davies, C., Gubbins, D. & Alfe, D., 2012. Thermal and electrical conductivity of iron at Earth's core conditions, *Nature,* **485,** 355–358.
Quidelleur, X. & Courtillot, V., 1996. On low degree spherical harmonic models of paleosecular variation, *Phys. Earth planet. Inter.,* **95,** 55–77.
Risken, H., 1989. *The Fokker-Planck Equation,* Springer, New York.
Roberts, A.P., 2008. Geomagnetic excursions: knowns and unknowns, *Geophys. Res. Lett.,* **35,** L17307, doi:10.1029/2008GL034719.
Roberts, A.P. & Winklhofer, M., 2004. Why are geomagnetic excursions not always recorded in sediments? constraints from post-depositional remanent magnetization lock-in modelling, *Earth planet. Sci. Lett.,* **227,** 345–359.
Stanton, R., 1997. A nonparametric model of term structure dynamics and the market price of interest rate risk, *J. Finance,* **52,** 1973–2002.
Sura, P. & Barsugli, J., 2002. A note on estimating drift and diffusion parameters from time series, *Phys. Lett. A,* **305,** 304–311.
Van Kampen, N.G., 1992. *Stochastic Methods in Physics and Chemistry,* North-Holland, Amsterdam.
Valet, J.P., Meynadier, L. & Guyodo, Y., 2005. Geomagnetic field strength and reversal rate over the past 2 million years, *Nature,* **435,** 802–805.
Ziegler, L.B., Constable, C.G., Johnson, C.L. & Tauxe, L., 2011. PADM2M: a penalized maximum likelihood model of the 0–2 Ma paleomagnetic axial dipole moment, *Geophys. J. Int.,* **184,** 1069–1089.

APPENDIX: NUMERICAL APPROXIMATION

A numerical solution of the Fokker–Planck equation in (4) is obtained using a finite volume approximation . Discrete equations are defined by integrating the Fokker–Planck equation over a finite interval Δx. Let x_n denote the midpoint of the finite interval and let $x_\mathrm{l} = x_n - \Delta x/2$ and $x_\mathrm{u} = x_n + \Delta x/2$ be the lower and upper limits of the interval. The rate of change of the probability distribution, $P(x, t)$, is approximated by

$$\int_{x_\mathrm{l}}^{x_\mathrm{u}} \frac{\partial P}{\partial t}\,\mathrm{d}x \approx \dot{P}_n \Delta x, \tag{A1}$$

where $\dot{P}_n$ is the time derivative of P at $x = x_n$. The next term in the Fokker–Planck equation involves the drift coefficient, $v_e(x)$, which is conveniently represented in terms of a potential $U(x)$ using

$$v_e(x) = -\nabla U(x). \tag{A2}$$

Integrating over Δx yields

$$-\int_{x_\mathrm{l}}^{x_\mathrm{u}} \frac{\partial}{\partial x}\,(v_e P)\,\mathrm{d}x \approx a_{n-1}P_{n-1} + a_n P_n + a_{n+1}P_{n+1}, \tag{A3}$$

where

$$a_{n-1} = -(U_n - U_{n-1})/\Delta x, \tag{A4}$$

$$a_n = (U_{n+1} - 2U_n + U_{n-1})/\Delta x, \tag{A5}$$

$$a_{n+1} = (U_{n+1} - U_n)/\Delta x. \tag{A6}$$

The final diffusion term is

$$\int_{x_\mathrm{l}}^{x_\mathrm{u}} \frac{\partial^2}{\partial x^2}\,(DP)\,\mathrm{d}x \approx b_{n-1}P_{n-1} + b_n P_n + b_{n+1}P_{n+1}, \tag{A7}$$

where

$$b_{n-1} = D_{n-1}/\Delta x, \tag{A8}$$

$$b_n = -2D_n/\Delta x, \tag{A9}$$

$$b_{n+1} = D_{n+1}/\Delta x. \tag{A10}$$

Time stepping of the discrete system is done using the Crank–Nicolson method.