5.05 Geomagnetic Secular Variation and Its Applications to the Core

A. Jackson and C. C. Finlay, ETH Zürich, Zurich, Switzerland

5.05.1 Introduction

The purpose of this chapter is to review the origins of our current knowledge of the secular variation of the Earth's magnetic field, that is, the slow changes that occur on timescales of years to centuries. There is clearly overlap with the description of the present geomagnetic field (Chapters 5.02 and 5.09), which treats changes in the field from centuries to millennia.

The source of our knowledge on the timescales we are concerned with is primarily direct historical observations of the field; we review the available data, followed by the treatment of the data to generate mathematical models of the field in space and time. We then discuss interpretations of these models in terms of some of the physical processes occurring at the core surface. We stop short of describing the actual calculation of models of fluid flow at the core surface, as this is covered in detail in Volume 8, but we lay the groundwork by developing an exposition of the governing equations and the approximations that are frequently used.

5.05.1.1 Historical Background

This section gives a very brief overview of the development of geomagnetism, and does not purport to be comprehensive. Fuller treatments of the history can be found in various places; for example, relevant chapters of Merrill *et al.* (1998) or Chapman and Bartels (1940), Malin (1987), or Stern (2002). A detailed account of geomagnetism up to 1500 can be found in Crichton Mitchell (1932, 1937, 1939); recent articles on nineteenth century geomagnetism are those of Good (1985, 1988). Excellent discussions of geomagnetic instruments can be found in McConnell (1980) and Multhauf and Good (1987). An authoritative source on virtually every aspect of geomagnetic history is the epistle by Jonkers (2000).

It is generally acknowledged that the Chinese were the first to discover the directive property of lodestone, almost certainly by the first century AD. Its development as a primitive navigational device was slow, though the declination had almost certainly been discovered by the ninth century and compasses were certainly in use by the eleventh century; early observations of declination are given by Needham (1962) and Smith and Needham (1967). The first recorded observation of declination in Europe was by George Hartmann in 1510; inclination was discovered by Robert Norman in 1576. The fact that the field underwent slow changes with time (the secular variation) was not discovered until 1635: by comparing a series of records taken at London previously, Henry Gellibrand showed that secular variation was a real effect. Relative intensities of the field were measured at the end of the eighteenth century by La Perouse, D'Entrecasteaux, and Humboldt, by comparing the periods of oscillation of a magnetic needle at different places. Measurements of the absolute intensity of the field were not made until a method was devised by Gauss in 1832 (see, e.g., Malin, 1982). The method was published in Gauss (1833a); an English translation of the abstract of a paper read in Göttingen in December 1832 can be found in Gauss (1833b).

While early observations of the field are extremely valuable, some problems do exist. For example, before the discovery of secular variation, some observations are undated as the need to record the date was not apparent. The accuracy with which an observer's position was known is also a source of error. Although the measurement of latitude was precise even by late fifteenth century (e.g., an accuracy of 10 min of arc was claimed by 1484 (John II's commission, 1509)), the measurement of longitude at sea remained a problem until approximately 1770 with the introduction of accurate chronometers by Harrison. The result of this poor knowledge of longitude led to the practice of 'running down the parallel', or sailing to the correct line of latitude before sailing due east or west along that parallel to the desired location. Although this practice meant that the ship's company often arrived at their desired destination, it does mean that large navigation errors could occur in the quoted positions of magnetic observations. To a large extent these errors can be alleviated by examination of the original ship's log and plotting the positions on a modern chart. This procedure has been performed for sixteenth, seventeenth, and eighteenth century data by Bloxham (1985, 1986a, 1986b), Hutcheson (1990) (see also Hutcheson and Gubbins, 1990), and Barraclough (1985), and Jackson *et al.* (2000) (hereinafter JJW2000); in addition, the latter authors also developed a statistical theory for accounting for imprecision in longitude.

The Greenwich meridian was adopted as an international longitude standard only in 1884, and some national conventions remained in use later than that date. Consequently, care must be taken as to which of the particular national conventions of Paris, the observatory at Pulkova (Leningrad), Washington, or San Fernando was being used. One example of French marine data measuring longitude from Paris until at least 1895 has been given by

Jackson (1989); this difference of 2° 13′ of longitude between Paris and Greenwich is small, but extremely significant.

5.05.1.2 Early Theories of the Secular Variation

Beginning with the seminal works *Epistola de Magnete* by Peregrinus (1269) and the better-known *De Magnete* by Gilbert (1600), various authors have sought to explain the Earth's magnetic field by models, some physical, some mathematical. Though Gilbert's model explained a considerable part of the static field, after the discovery of the secular variation a whole new dimension was opened up, requiring explanation. It is not our purpose to adumbrate the numerous models created over time to explain the temporal variation of the field. However, Jonkers (2000) has provided just such a list, comprising a remarkable compilation of theories of the field up to 1800, starting with Peregrinus (1269) and ending with Churchman's (1794) petitions to the English Board of Longitude, requesting acceptance of his theories for use in determining longitude. A shorter description can also be found in Jonkers (2003).

5.05.2 Data

We refer the reader to Chapter 5.04 for information on how measurements of the field are taken, and for definitions of the quantities that are typically reported: the declination *D*, the inclination *I*, the horizontal and total intensities (*H* and *F*, respectively), and the Cartesian components *X*, *Y*, and *Z* in the easterly, northerly, and downward directions. The availability of different data types varies as a function of time, chiefly as a result of the needs for navigation, followed by the drive of scientific curiosity in the eighteenth and nineteenth centuries. It should be noted that until Gauss' invention of a method for the determination of absolute intensity in 1832, only the morphology of the field can be determined from direct measurements.

5.05.2.1 Catalogs and Compilations of Data

The earliest catalogs, of Stevin (1599), Kircher (1641), and Wright (1657), are deficient in that they contain undated observations. Around 1705, the French hydrographer Guillaume Delisle compiled some 10 000 observations (mostly of declination) in his notebooks, trying to establish regularity in secular acceleration; these were never published but still exist in the Archives Nationales in Paris. The next important compilation of magnetic data was made by Mountaine and Dodson (1757) who claimed to have based their tables of declination at different points on the Earth on over 50 000 original observations of the field. The original observations of this enormous collection were never recorded and are lost: the work merely printed grids of averaged data with no reference to sources, numbers of data, etc. This claim regarding the number of data involved has attracted some skepticism; however, the work of Jonkers *et al.* (2003) indicates that it is undoubtedly the case that the authors' claim for the number of data is true. The early work of Mountaine and Dodson should almost certainly receive more prominence than it does, representing probably the first large-scale attempt to describe the morphology of the field. Maps based on the data were subsequently produced.

The main era of printed compilations of geomagnetic data was the nineteenth century, featuring the work of Hansteen, Becquerel, Sabine, and Van Bemmelen. In 1819, the Norwegian astronomer and physicist Christopher Hansteen published *Untersuchungen über den Magnetismus der Erde*, which listed data from land surveys and 73 nautical voyages from 1589 to 1816. His collection includes many of the great scientific expeditions during the latter half of the eighteenth century, including Cook's three voyages, contributing over 6500 declination and 1200 inclination observations. Another valuable addition was made by A. C. Becquerel's *Traité Expérimental de l' Électricité et du Magnétisme* (1840), which contains the only comprehensive collection of relative intensities. Several Philosophical Transactions papers by astronomer Edward Sabine span the period 1818-70 with exceptionally good coverage, although, as various authors have noted, they are far from comprehensive. Finally, in the 1890s, Dutch physicist Willem van Bemmelen processed 165 nautical sources prior to 1741 (Bemmelen, 1899).

Finally, one of the largest single compilations, featuring over 28 000 data points of all types, is the *Catalogue of Magnetic Determinations in USSR and in Adjacent Countries from 1556 to 1926* in three volumes. It was compiled and published by Russian physicist B. P. Veinberg in 1929–33 (Veinberg, 1929–33) and contained original data from Russia and neighboring states, mostly obtained in the first decades of the twentieth century. A review of the previous compilations of magnetic data that have been produced over time can be found in Barraclough (1982).

Another category of sources comprise time series for specific locations, normally major cities where investigators have set up permanent instruments, for instance, at national astronomical observatories. Past observers include Graham, London clock maker and the discoverer of diurnal variation, and Gilpin in England, academics Celsius and Hiörter in Sweden (who studied the correlation of needle disturbance with the occurrence of auroras), MacDonald on Sumatra (eighteenth century), and, in France, many scholars and astronomers summarized in Alexandrescu *et al.* (1996). But despite their achievements, a mere handful of cities can boast a series of more or less regular observations spanning over a century prior to 1800. A review of recent efforts to make these data series available to a modern audience is given by Alexandrescu *et al.* (1996), who also list all early geomagnetic observations made in Paris (1541-1883, based in part on earlier work by Raulin (1867) and Rayet (1876)); see **Figure 1**. Other capitals with a sustained tradition of geomagnetic observations include London (Malin and Bullard, 1981, Barraclough *et al.*, 2000; see **Figure 18**), Rome (Cafarella *et al.*, 1992), and Edinburgh (Barraclough, 1995).

It should be noted that only when observations in compliations can be confidently ascribed to individual observations (with well-defined times and locations and not derived by interpolation) have they included in the database of historical observations of Jonkers *et al.* (2003) (hereinafter JJM2003), and used historical field models such as *gufm1* described in JJW2000.

5.05.2.2 Surveys, Repeat Stations, and Marine Data

Recent interest in historical secular variation has led to original observations being compiled for other time periods; a comprehensive review of available data has been given recently by JJM2003. It is impossible to detail all characteristics of this data set. Suffice it to say that the largest part of the data set originates in marine observations of the declination, typically taken for the purposes of navigation. It has been possible to characterize the accuracy with which observers measured the declination at sea; it is better than half a degree for the 17th and 18th centuries taken as a whole. When one compares this accuracy with modern measurements, it transpires that the old measurements have a signal/noise ratio which is not too bad compared to modern measurements, for the simple reason that both types of measurements suffer the contaminating effect of the crustal magnetic field. This contributes $\sim 60\sigma/H$ degrees of error in declination where σ is the rms horizontal crustal field, and H the local horizontal field strength; for $\sigma \sim 200$ nT and $H = 20\,000$–$40\,000$ nT the error contribution is 0.3–0.6°, and thus commensurate with the observational error.

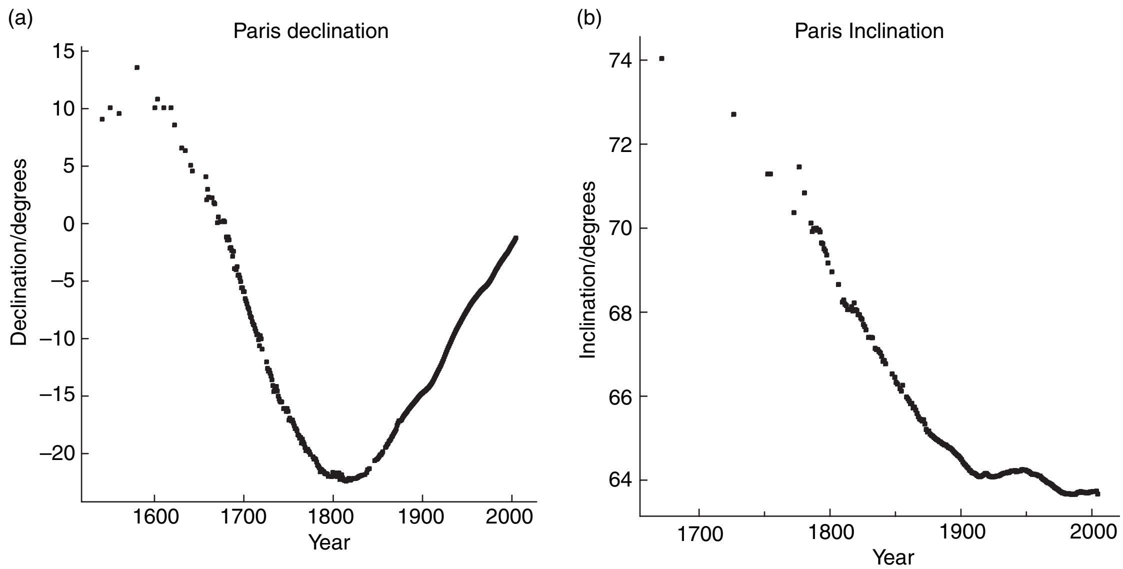

Figure 1 Declination (*D*) and Inclination (*I*) in Paris during historical times, based on Alexandrescu *et al.* (1996) and more recent observatory data. The data have been reduced to the current site of Chambon-la-Forêt; early observations are made by many different observers at several different sites.

An accurate position is of course a crucial part of any magnetic measurement, and thus latitude and longitude need to be known. This poses no problems on land, but at sea the determination of position can be challenging. From the sixteenth century onward, the backstaff provided a method for latitude determination, often said to be accurate to 10 min of arc; the empirical findings of JJW2000 agree with this. For early data, a well-known difficulty is the imprecision in longitude prior to the invention of the marine chronometer. A very detailed study of this was undertaken in JJW2000, who showed that navigational error generally generated a Brownian motion-type effect, such that the errors increased with the square root of voyage duration. Empirically, the data suggest that a typical 25 day voyage might accumulate 2° of error (though voyages often achieved much better than this). The effect can be ameliorated in most voyages by using the fact the voyage arrived at a known location — thus the whole voyage can be retrospectively corrected for the accumulated errors, giving a lower error. The appropriate model for the errors becomes the so-called Brownian bridge, and the error at the mid-point of the journey is reduced to exactly half what it would have been using the simple Brownian motion model — hence typically it is 1° for a 25 day journey. The effect of this imprecision in location is to increase the error budget by an amount depending on the gradient of declination with respect to longitude — a fairly representative figure is less than 1° of declination change per degree of longitude. Hence one can see that even early data have contributions to their error budget (from observational imprecision, crustal magnetic fields, and longitudinal inaccuracy) that are not too dissimilar. Inclination was initially measured on land at London (1576); the next extant measurement was taken shortly after at Uranienborg (1584) by Brahe. The first example of a measurement made on an expedition was by Weymouth in Frobisher Bay, Canada, in 1602. For the first inclination observation south of the equator, one had to wait until 1680 when one Benjamin Harry took observations on board the Berkeley Castle en route to the far east.

The nineteenth century saw burgeoning scientific expeditions on land, which included the measurement of intensity as well as D and I — thus giving the first vector measurements of the field. De Rossel's measurements of the oscillation time of a dip needle in 1791 provides the first measurements of relative intensity between several places on the Earth. Humboldt and Erman also provide well-known relative intensity measurements prior to Gauss' (1833a) invention of a method to determine absolute intensity (see Malin, 1982). The net result of the collation of the known historical observations as described in JJM2003 is a data set which is summarized in **Table 1**. The available data are summarized in geographical plots in **Figures 2–10**.

A vast source of magnetic field data for the twentieth century was prepared by the US Coast and Geodetic Survey for the 1965 world charts (Hendricks and Cain, 1963), and is accessible in machine-readable form from the world data centers. The cutoff date for the collection was arbitrarily set at 1900. This data set has been used by many authors over the years, and has been the basis for many International Geomagnetic Reference Field models; we shall not dwell on a description of this data set, as it has been described numerous times (e.g., Bloxham *et al.*, 1989, Sabaka *et al.*, 1997). Its temporal distribution dominates the temporal distribution of the data used to create *gufm1* during the twentieth century, which is shown in **Figure 11**. Twentieth century data is characterized by a constant improvement in measurement accuracy (*see* Chapter 5.02). Marine surveys continued in the twentieth century, the most notable being the voyages of the nonmagnetic surveying ship, the *Carnegie*. A new type of data emerged with the advent of aeromagnetic surveys, most notably Project Magnet; for details, see Langel (1993).

Table 1 Temporal distribution of databased geomagnetic measurements; a single record may contain a land sighting and/or up to three types of measurement (D, I, and intensity (H or F))

Period	Records	D	I	H	F	Total
1510–1589	162	160	2	0	0	162
1590–1699	13 673	12 001	53	0	0	12 054
1700–1799	85 070	68 076	1 747	0	36	69 859
1800–1930	78 162	71 323	17 723	11 404	4 779	105 229
Total	177 067	151 560	19 525	11 404	4 815	187 304

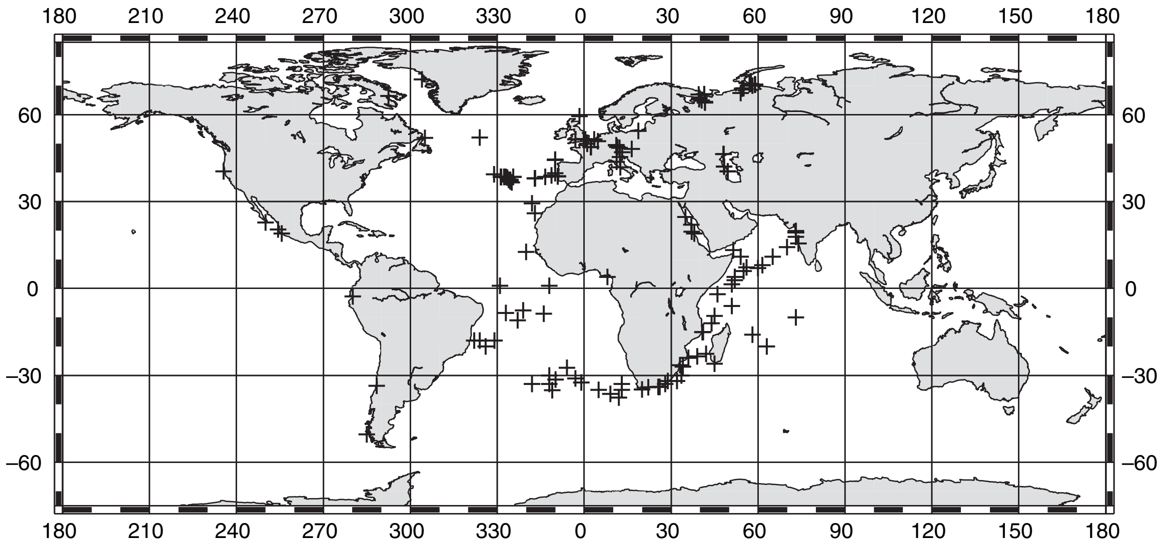

Figure 2 Geographical data distribution of declination observations made before 1590; $n = 160$; some points may overlap; cylindrical equidistant projection.

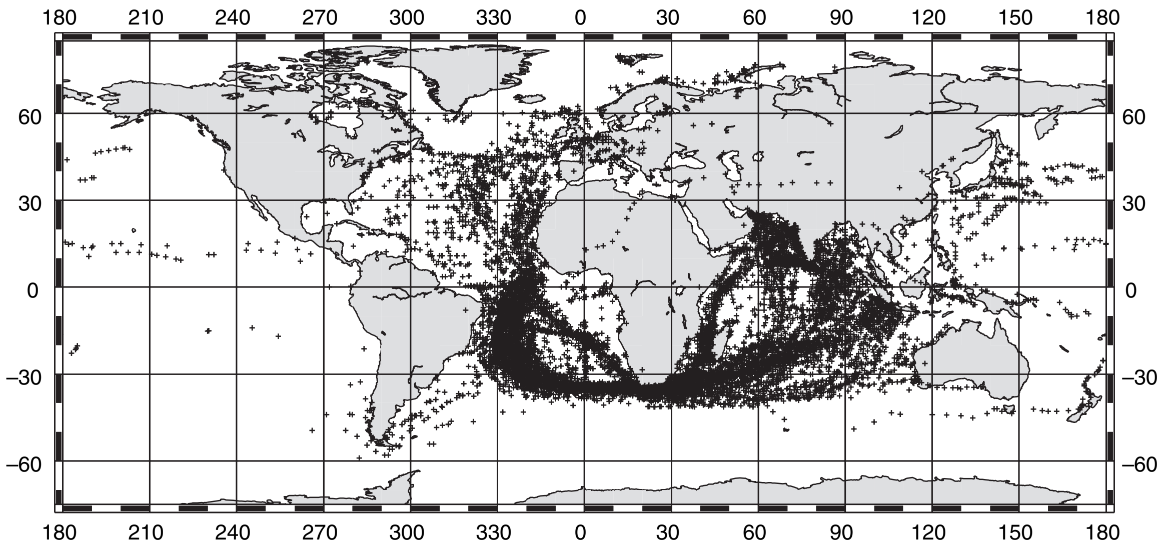

Figure 3 Geographical data distribution of declination observations made in 1590–1699; $n = 12\,001$; some points may overlap; cylindrical equidistant projection.

5.05.2.3 Observatory Data

The establishment of the Göttingen Magnetic Union (Magnetische Verein) in 1834 by Gauss and Weber heralded the establishment of an observatory network at sites around the world where observations of the magnetic field would be made with regularity. With the adoption of the 'magnetic crusade' of Sabine, Herschel, and Lloyd by the British learned bodies in 1838, Germany and Britain took a lead in driving forward observational geomagnetism (Cawood, 1977, 1979). The number of observatories gradually grew and their distribution increased toward the distribution of today (**Figure 12**), although some former observatories have closed due to a multitude of factors. Some of the history of the growth of the observatory network can be found in Chapman and Bartels (1940).

From the point of view of studies of the secular variation due to the motions of the liquid outer core, the most important product derived from the

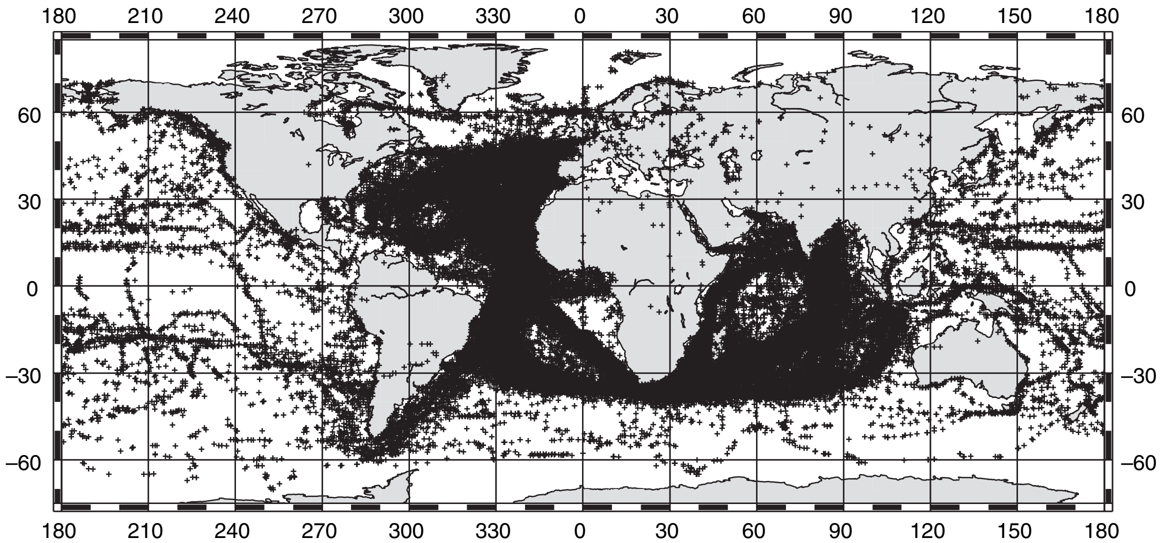

Figure 4 Geographical data distribution of declination observations made in 1700–99; $n = 68\,076$; some points may overlap; cylindrical equidistant projection.

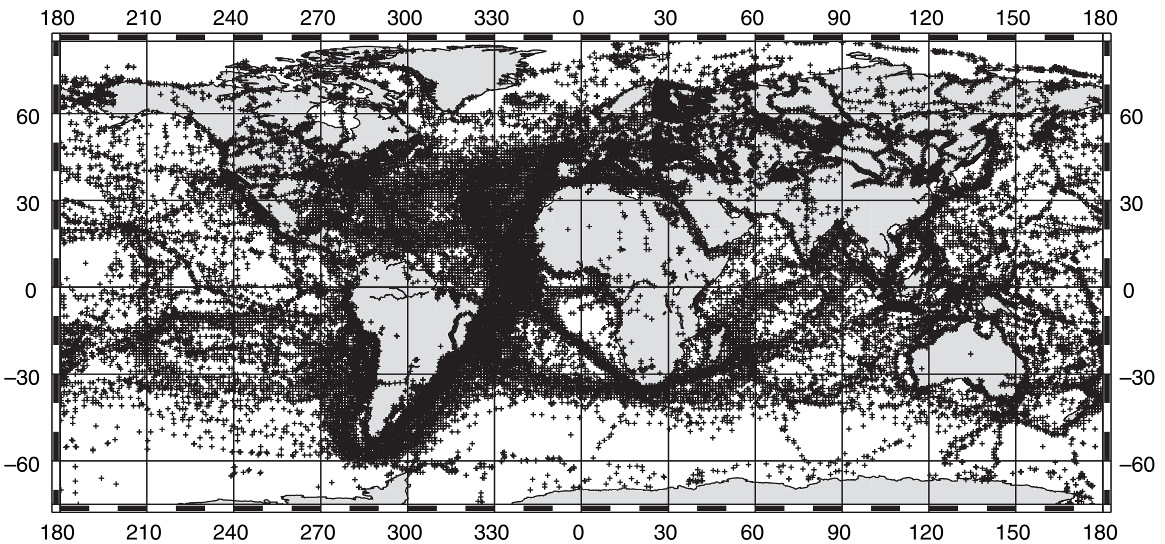

Figure 5 Geographical data distribution of declination observations made in 1800–1930; $n = 71\,323$; some points may overlap; cylindrical equidistant projection.

continuous monitoring of the observatories is the so-called annual mean, representing the yearly averaged value of the geomagnetic elements. Although the current definition of an annual mean is a mean over all data, there has, in the past, been some variability in exactly what is reported as an annual mean. For example, occasionally data reported from the five International Quiet Days every month has been used as an annual mean. Table 2 shows the frequency of the different types of data that are included in the definitive annual means data file, held by the World Data Centre for Geomagnetism at the British Geological Survey (Edinburgh). This inconsistency, which cannot be corrected retrospectively (since the original data no longer exist), leads to inevitable difficulties in treating the data, because the data contain different amounts of external magnetic field contribution. Compromises are always required in treating historical magnetic data, and so far these observatory data have been treated as if they were homogeneously recorded; perhaps it will be possible to treat them in a way that recognizes their different characteristics in

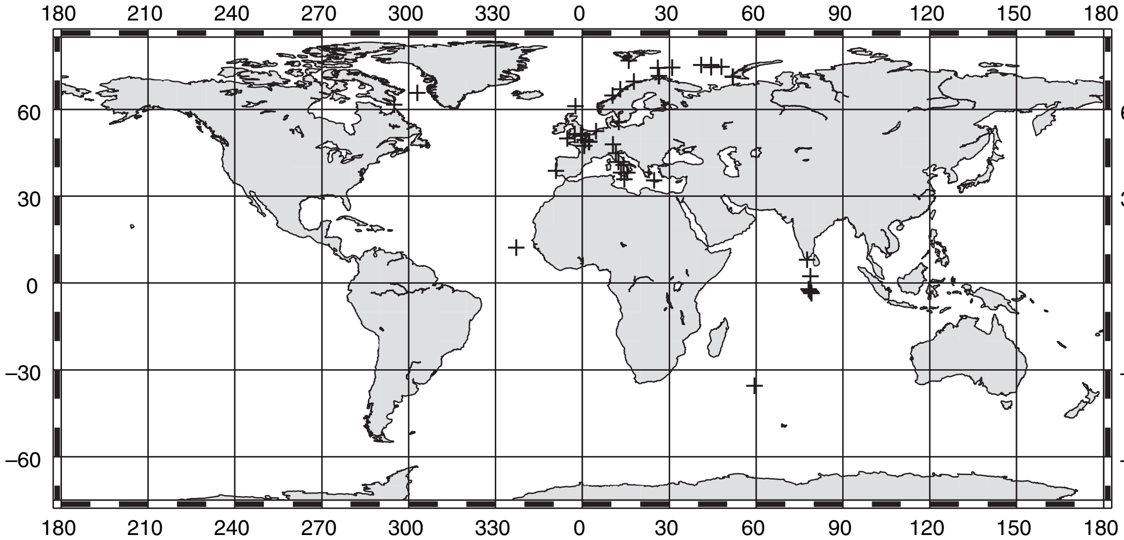

Figure 6 Geographical data distribution of inclination observations made in 1590–1699; $n = 53$; some points may overlap; cylindrical equidistant projection.

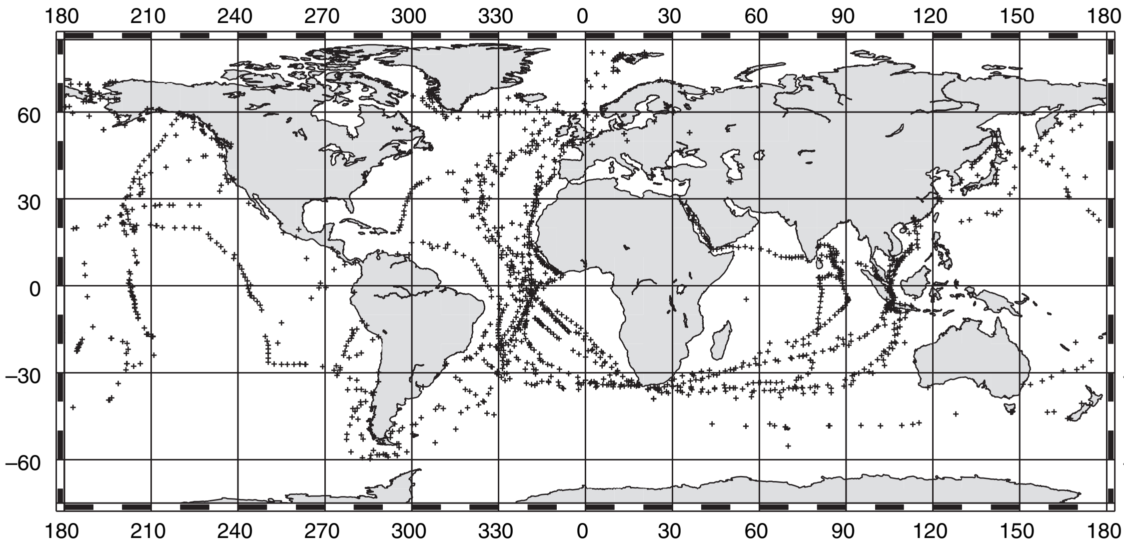

Figure 7 Geographical data distribution of inclination observations made in 1700–99; $n = 1747$; some points may overlap; cylindrical equidistant projection.

the future. Figure 13 shows the distribution of observatory annual mean data through time, from the first observations originating with the formation of the Göttingen Magnetic Union, to the present day.

It is straightforward to treat single observations of the field (such as those made by surveys or satellites) as being independent measurements that can be fitted simultaneously in a least-squares process. Some words are in order regarding the treatment of observatory data in time-dependent field modeling. Observatories obviously supply critical data on the secular variation, and indeed the accuracy of many of the modern field models, rests on the observatory time series. A problem that must be recognized, however, is the fact that the observatories are subject to a (quasi-) constant field associated with the magnetization of the crust in the region that they are located. If observatory data are mixed with other types of data (survey, satellite data), this so-called observatory bias must be recognized; otherwise, it will bias the solution for the main field because an observatory time series essentially records it many times. Two approaches have

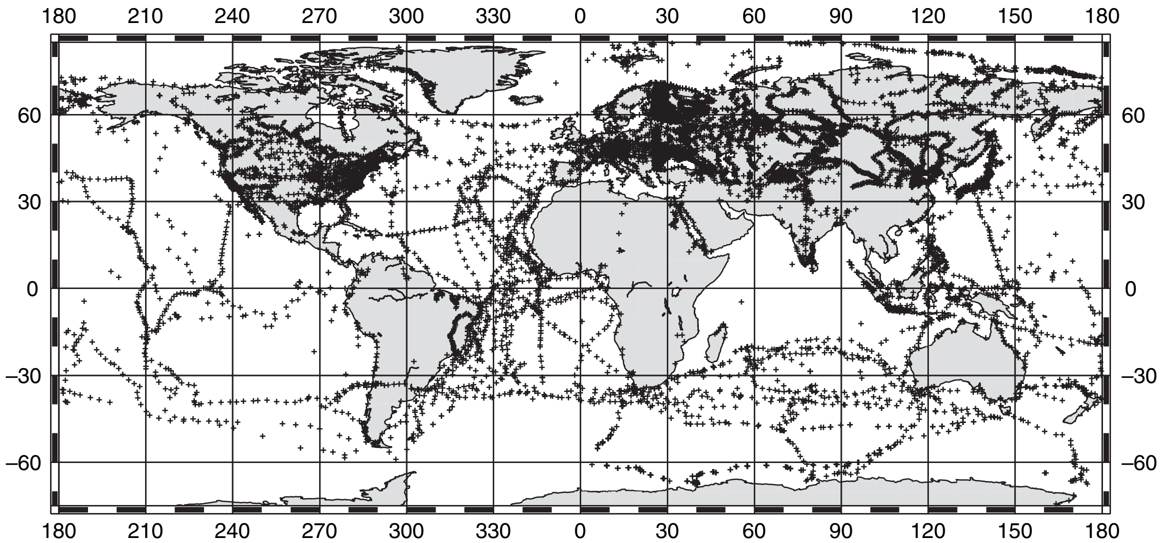

Figure 8 Geographical data distribution of inclination observations made in 1800–1930; $n = 17\,723$; some points may overlap; cylindrical equidistant projection.

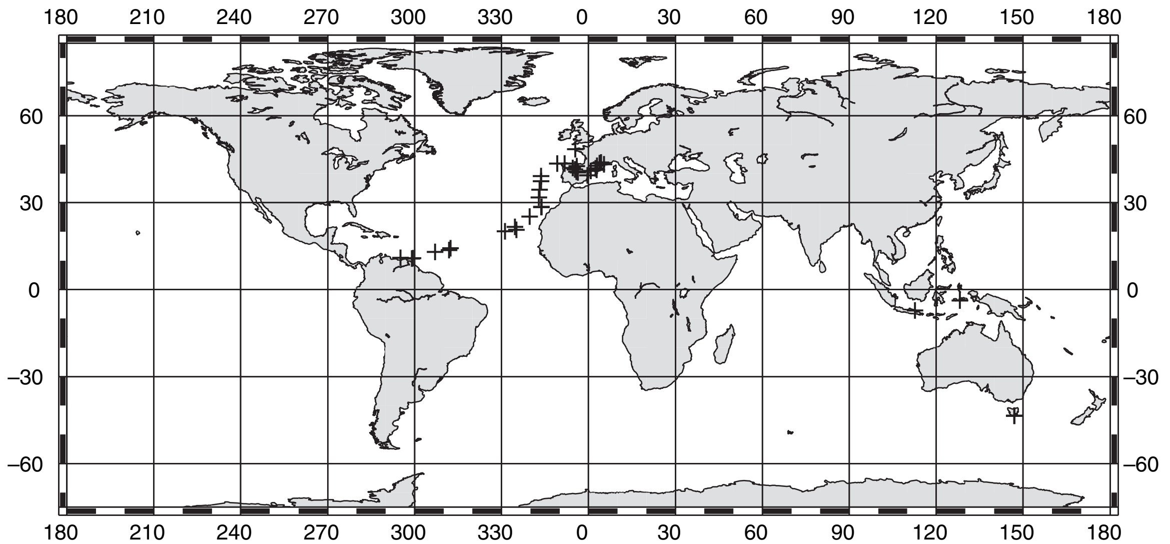

Figure 9 Geographical data distribution of intensity observations made in 1700–99; $n = 36$; some points may overlap; cylindrical equidistant projection.

developed for dealing with this. The first, developed by Langel *et al.* (1982), is to solve for the observatory biases (three per observatory in the *X*-, *Y*-, and *Z*-directions) as unknowns at the same time as solving for the magnetic field. This technique continues to be adopted in the comprehensive series of field models (see below), and works very effectively. The second approach is to desensitize the observatory data to the presence of the bias (see, e.g., Bloxham and Jackson, 1992). An effective way of doing this is to work with the rate of change of the field from the observatory, and hence first differences of observatory data are used in the *ufm* and *gufm* series of models (see below). There appears to be very little difference in the results of the two approaches.

5.05.2.4 Satellite Data

Satellite data play a crucial role in determining a detailed global picture of the secular variation. An extensive discussion of the special character of satellite data can be found in Chapter 5.02, and we shall

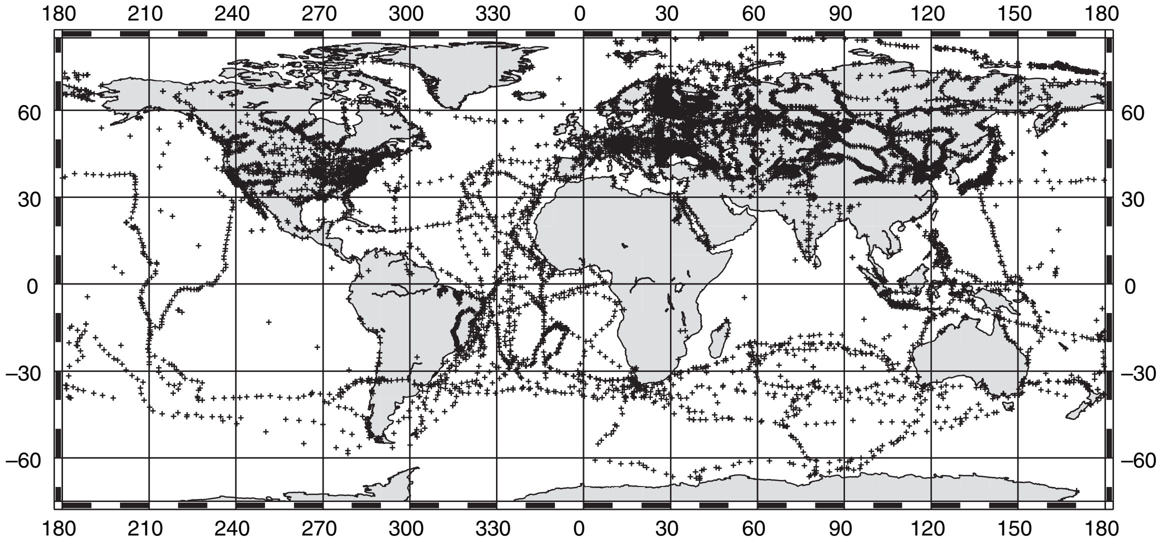

Figure 10 Geographical data distribution of intensity observations made in 1800–1930; $n = 16\,183$; some points may overlap; cylindrical equidistant projection.

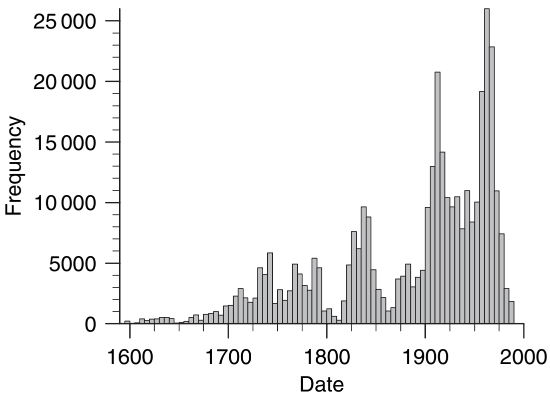

Figure 11 The overall number historical data (as described in JJM2003) together with observatory data, twentieth century survey and repeat station data, and satellite data used in the construction of *gufm1* (Jackson *et al.* 2000). Note that this depicts a subset of data available, as some data selection has taken place, based on criteria designed to avoid the effect of the correlation in errors due to the crust.

not duplicate that here. Nevertheless, in **Table 3**, we list some of the satellites that have been used for magnetic field determination over time, and their different characteristics.

5.05.3 Time-Dependent Models of the Main Field

We now turn to the use that is made of the data sets that have been described in the previous section. The tool that has been most commonly applied has been spatial spherical harmonic analysis, first applied by Gauss (1839). His analysis demonstrated the predominantly internal origin of the field.

The principles of spherical harmonic analysis are described in Chapter 5.02. They were applied by many authors in the nineteenth and twentieth centuries, with various amendments in order to deal with the fact that primarily nonlinear functions of the Gauss coefficients were being measured, namely D, I, H, and F; such developments are described fully in Barraclough (1978). As one such example, consider how to treat measurements of declination in the spherical harmonic inverse problem. We have that

$$D = \tan^{-1}\frac{Y}{X} \tag{1}$$

and the northerly (X) and easterly (Y) components are linearly related to the Gauss coefficients $\{g_l^m; h_l^m\}$ forming the model vector $\mathbf{m}$. Let us write these relations as $X = \mathsf{A}_x^{\mathrm{T}}\mathbf{m}$ and $Y = \mathsf{A}_y^{\mathrm{T}}\mathbf{m}$. If we rearrange [1] into the form,

$$X \sin D = Y \cos D \tag{2}$$

one can form a linear constraint on $\mathbf{m}$ of the form

$$\left[(\sin D)\mathsf{A}_x - (\cos D)\mathsf{A}_y\right]^{\mathrm{T}}\mathbf{m} = 0 \tag{3}$$

This can be fit in a least-squares sense, but note that the data enter in defining the linear relation, rather

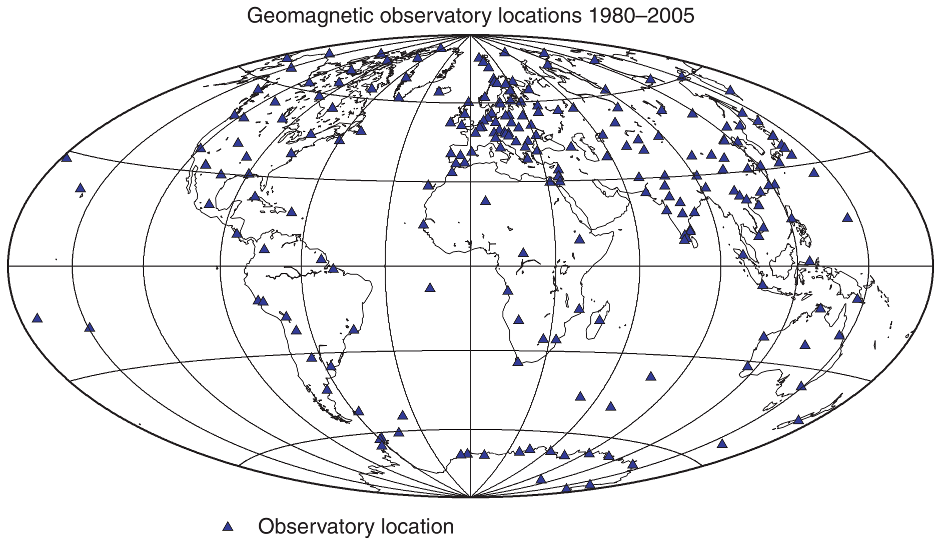

Figure 12 The distribution of observatories operating at some point in the last 25 years.

than as a target for the prediction. Numerous other schemes for dealing with nonlinear data are described in Barraclough (1978). With the advent of significant computer power, the need to deal with nonlinear data in such a way has diminished and iterative schemes, as described in Chapter 5.02, are more commonplace.

In the years following the early applications of Gauss' method, the technique was applied to the field at different epochs, the interest being primarily in the evolution of global averages such as the dipole moment. Being before the advent of modern computers, it was impossible to deal with true measurements of the field without some preliminary reduction of the data — thus the source for the spherical harmonic analyses were field values at regular intervals read from charts which had been constructed by interpolating the original data by hand. Useful descriptions of these types of model can be found in Barraclough (1978) or Langel (1987).

5.05.3.1 Methodologies

Chapter 5.02 discusses the determination of static models of the field. In this section, we will review a selection of the most widely used time-dependent field models, and the techniques used to derive them. We restrict attention to models that have been produced specifically as time dependent; only passing reference is made to models designed to describe either the static magnetic field or its rate of change at a particular point in time; for models of this type, *see* Chapter 5.02.

Our description focuses on models of the magnetic field **B** which are simultaneously models of its spatial ((r, θ, ϕ) in spherical coordinates) dependence and the temporal dependence (t denotes time). The standard technique which is common to all the analyses we will describe is to employ the spherical harmonic expansion of the field in terms of Gauss coefficients $\{g_l^m; h_l^m\}$ for the internal field; some of the most recent models also incorporate coefficients representing the external field. All the models will

Table 2 Types of observations classed as annual means and used for secular variation modeling

Flag	Annual means derived from	Percentage
1	Data for all days	72.4
2	Data for quiet days	4.0
3	Preliminary data	1.9
4	Absolute observations only	9.0
5	Incomplete data (<12 months, but $\geq$6 months)	3.4
6	Very incomplete data (<6 months)	1.4
7	Limited absolute control (introduced in 1996)	0.4
0	Unknown	7.6

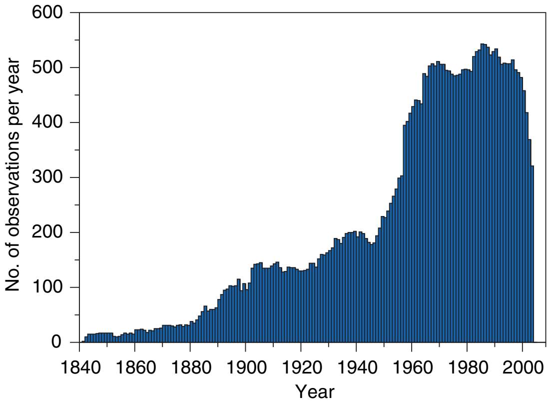

Figure 13 Distribution of observatory annual means through time, reflecting availability as of October 2006. The falloff in recent times is due to the delay in observatories reporting definitive data to the world data centers.

employ the Schmidt quasi-normalization common in geomagnetism.

A time-dependent model of the field necessarily must be built using a data set spanning a period of time, denoted herein $[t_s, t_e]$. In order that a spherical harmonic analysis can be performed, a parametrization is required for the temporal variation of the field. The unifying idea, common to all analyses, is to use an expansion for the Gauss coefficients of the form

$$g_l^m(t) = \sum_i {}^i g_l^m \phi_i(t) \qquad [4]$$

where ϕ_i are a set of basis functions and the ${}^i g_l^m$ are a set of unknown coefficients. (A similar expansion is of course used for h_l^m.) The different models that have been produced over the last few decades differ in their choice of the $\phi_i(t)$. With an expansion of the form [4], the unknown coefficients $\{{}^i g_l^m; {}^i h_l^m\}$ are denoted as a model vector $\mathbf{m}$, and when linear data such as the elements (X, Y, Z) are required to be synthesized (denoted by vector $\mathbf{d}$) the resulting forward problem is linear and of the form

$$\mathbf{d} = \mathsf{A}\mathbf{m} \qquad [5]$$

where A is often termed the equations of condition or design matrix and describes how model parameters are combined to give predictions that can be compared to the data.

The inverse problem of finding the coefficients $\mathbf{m}$ is generally solved by finding a model minimizing the least-squares difference between the model predictions and the data, sometimes together with a measure of the field complexity to help resolve the issue of nonuniqueness (*see* Chapter 5.02). More generally, when I, D, F, and H data are involved so that the relation between the model parameters and the data is a nonlinear function (which we write as $\mathbf{d} = \mathbf{f}(\mathbf{m})$), the model must be found iteratively. If $[\mathsf{A}]_{ij} = \partial f_i / \partial m_j$, and if C_e is the data covariance matrix, then the model solution is sought by an iterative scheme, such as the quasi-Newton method,

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \left(\mathsf{A}^T \mathsf{C}_e^{-1} \mathsf{A}\right)^{-1} \left[\mathsf{A}^T \mathsf{C}_e^{-1} (\mathbf{d} - \mathbf{f}(\mathbf{m}_i))\right] \qquad [6]$$

In [6], $\mathbf{m}_i$ stands for the model at the ith iterate, and in principle the matrix A should be recomputed at every iterate. Such methods converge very rapidly since the effect of the nonlinearity is very mild.

Table 3 Satellite missions of relevance for measurement of the core secular variation. Accuracies refer to the intrinsic accuracies of the instrumentation, combined with the positional and orientation accuracy

Name	Inclination	Dates	Altitude (km)	Accuracy (nT)	Remarks
Cosmos 49	50°	1964	261–488	22	Scalar
OGO-2	87°	1965–67	413–1510	6	Scalar
OGO-4	86°	1967–69	412–908	6	Scalar
OGO-6	82°	1969–71	397–1098	6	Scalar
Magsat	97°	1979–80	325–550	6	Vector
DE-1	90°	1981–1991	568–23 290	?	Vector (spinning)
DE-2	90°	1981–83	309–1012	~30(*F*)/100	Low accuracy vector
POGS	90°	1990–93	639–769	?	Scalar, timing problems
UARS	57°	1991–94	560	?	Vector (spinning)
Ørsted	97°	1999–	600–850	*a*	Vector
CHAMP	87°	2000–	350–460	*a*	Vector
SAC-C	98°	2000–04	702–709	*a*	Scalar

[a]For accuracies of the present missions, *see* Chapter 5.02.
Inclination is measured as the angle at which the satellite crosses the equator while passing from the Southern Hemisphere to the Northern Hemisphere.

5.05.3.1.1 Taylor series models

The earliest time-dependent models used a Taylor series expansion for the Gauss coefficients of the form

$$g_l^m(t) = g_l^m(t_0) + \dot{g}_m^m(t_0)(t - t_0) + \ddot{g}_l^m(t_0)\frac{(t - t_0)^2}{2!} + \cdots \qquad [7]$$

about some central epoch here denoted t_0. This expansion is of the form [4], with the identification $\phi_n(t) = (t - t_0)^n/n!$ and ${}^n g_l^m = (\partial_t)^n g_l^m(t_0)$, the nth time derivative at the central epoch.

In the case of the Taylor expansion, A is a dense matrix. The first models to be produced this way were those of Cain *et al.* (1965) and Cain *et al.* (1967), who produced models GSFC(4/64) and GSFC(12/66) with temporal expansions truncated at first derivative and second derivative terms, respectively. The truncation level was subsequently raised to third derivative terms in the model GSFC(9/80) of Langel *et al.* (1982). More recently, Taylor series expansion techniques have been used to provide time-dependent models of satellite data, covering the only short intervals of a few years. For example, Olsen (2002) used a first-order expansion and Maus *et al.* (2005) used a second-order expansion; for such models of satellite data covering only a few years, Taylor series expansion models are reasonable.

When one wishes to produce a model of the field spanning a long time period, it is clear that a large number of terms will be required in [7], and it no longer remains an attractive method because of numerical instabilities and lack of flexibility of the parametrization.

5.05.3.1.2 Two-step models

A variety of models have been made by a two-step process of first making a series of spatial models at particular epochs, followed by some form of interpolation. For example, the International Geomagnetic Reference Fields and Definitive Geomagnetic Reference Fields (DGRFs) are strictly snapshot models of the field for particular epochs, but they can be used to calculate the magnetic field at times intermediate between two epochs by linear interpolation between the models. As a result it is possible to evaluate the DGRFs at any point in time between 1900 and the present day, though from a purist point of view they are not strictly time-dependent models of the magnetic field. The stepping stone between such two-step models and the more sophisticated approach of using a spline representation of temporal behaviour (see Section 5.05.3.1.3) was the pioneering paper of Langel *et al.* (1986). These authors used a spline temporal basis to interpolate between single epoch secular variation models.

5.05.3.1.3 Time-dependent models based on cubic B-splines

After the mid-1980s, more flexible representations of the time dependency were introduced. Beginning with Bloxham (1987), who used Legendre polynomials, a variety of functions have been employed. The most commonly used and referenced time-dependent field models along with their time span and modeling approach are summarized in **Table 4**.

The methods employed by different workers have gradually converged toward the use of cubic B-splines as temporal basis functions following the example of Bloxham and Jackson (1992), who were heavily

Table 4 Characteristics of widely used models of the time-varying magnetic field

Model	L	N	Time period	Expansion	Regularized	Author
GSFC(4/64)	5	2	1940–63	Taylor	No	Cain *et al.* (1965)
GSFC(12/66)	10	3	1900–66	Taylor	No	Cain *et al.* (1967)
GSFC(9/80)	13	4	1960–80	Taylor	No	Langel *et al.* (1982)
MFSV/1900/1980/OBS	14	8	1900–80	Legendre	Yes	Bloxham (1987)
	14	10	1820–1900, 1900–80	Chebyshev	Yes	Bloxham and Jackson (1989)
ufm1, ufm2	14	63	1690–1840, 1840–1990	B-spline	Yes	Bloxham and Jackson (1992)
gufm1	14	163	1690–1990	B-spline	Yes	Jackson *et al.* (2000)
CM3	13	14	1960–85	B-spline integrals	Yes	Sabaka *et al.* (2002)
CM4	13	24	1960–2002.5	B-spline integrals	Yes	Sabaka *et al.* (2004)
CHAOS	18	10	1999–2006	B-spline and Taylor	Yes	Olsen *et al.* (2006)

L is the maximum degree of the internal secular variation. *N* is the number of temporal basis functions used for each Gauss coefficient. The CHAOS model uses B-splines up to degree 14 and a first-order Taylor expansion between 15 and 18. For other satellite-derived models, *see* Chapter 5.02.

influenced by the approach of Langel *et al.* (1986). There are two reasons for the popularity of the cubic B-spline method. First, when global basis functions such as Legendre or Chebyshev expansions are used (see, e.g., Bloxham (1987) or Bloxham and Jackson (1989)), the design matrix remains dense and requires considerable memory for its storage, whereas a B-spline basis is a 'local' basis, meaning that the basis functions are zero outside a small range (see **Figure 14**). This fact leads to a design matrix which is sparse (in fact it is banded), and storage requirements are minimized. Second, the B-splines provide a flexible basis for smoothly varying descriptions of data. One can show that of all the interpolators passing through a time series of points (say $f(t_i), i = 1, \ldots, N$), an expansion in cubic B-splines of order 4 ($\hat{f}(t)$ say) is the unique interpolator which minimizes a particular measure of roughness $\mathcal{R}$ (see, e.g., De Boor, 2002),

$$\mathcal{R} = \int_{t_s}^{t_e} \left[\frac{\partial^2 \hat{f}(t)}{\partial t^2}\right]^2 \mathrm{d}t \qquad [8]$$

The idea of attempting to construct a smooth representation in time is an application of 'Occam's razor', that there should be no extra detail in the representation than that truly demanded by the data. This idea of 'regularization' has been employed in many of the models of **Table 4** from that of Bloxham (1987) onward. Those models that employ regularization typically minimize a combination of norms $\mathcal{N}$ on the core–mantle boundary (CMB) of the form

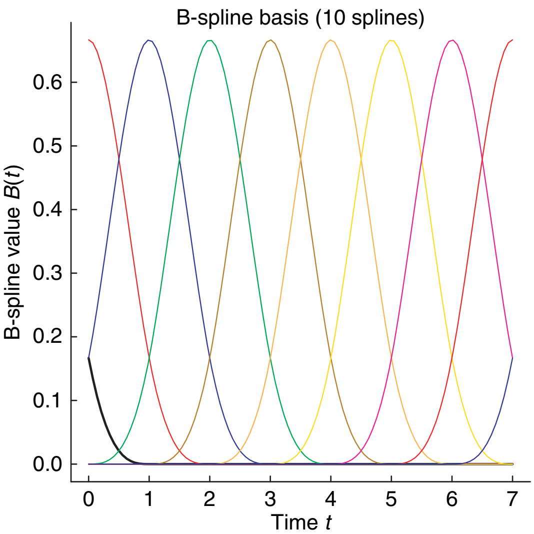

Figure 14 B-splines of order 4 (cubic B-splines). Local temporal basis of cubic B-splines used in the construction of time-dependent geomagnetic field models.

$$\mathcal{N} = \int_{t_s}^{t_e} \left[\nabla_h^{(n_1)} \partial_h^{(n_2)} B_r\right]^2 \mathrm{d}\Omega\, \mathrm{d}t \qquad [9]$$

where B_r is the radial field on the CMB. The models produced by Bloxham, Jackson, and co-workers use $n_1 = 0$ and $n_2 = 2$ in one norm, and $n_1 = 1$ and $n_2 = 0$ (approximately, to be precise the ohmic heating norm of Gubbins (1975) is used) in a second norm; this is slightly different to the choices made by Sabaka, Olsen, and co-workers in their comprehensive models and the CHAOS time-dependent model of satellite data (see below). A rather different form of regularization was recently proposed by Jackson (2003) that involves maximizing the entropy of the field model rather than penalising spatial or temporal gradients. This new method has so far been used to produce single epoch models but it could also be applied to provide both spatial and temporal regularization of time-dependent models; such models are currently under development.

Regularized field models are found by minimizing an objective function consisting of a measure (often the L_2 least squares norm) of the misfit of the time-dependent model to the data along with spatial and temporal norms measuring the field complexity. The relative weights of the spatial and temporal norms are scaled by the sizes of so-called damping parameters λ_S and λ_T. The choice of the damping parameters are made by trading off the desire that the data be fit within their estimated errors, the desire that the spatial complexity of the time-dependent model at the core surface be compatible with accurate single-epoch models, and the requirement that no unnecessary temporal oscillations be introduced. The models such as *ufm* and *gufm* satisfy each of these criteria.

5.05.3.1.3.(i) The ufm1, ufm2, and gufm1 models The *ufm1*/*ufm2* and *gufm1* field models share a common aim, namely to model the long-term secular variation at the core surface as accurately as possible over the past few centuries. They were built using the cubic B-spline basis with knots every 2.5 years, and from the largest data sets possible at the time: *ufm1*/*ufm2* used over 250 000 data originating from old ships' logs, survey data, observatories and satellite missions; a description of the oldest data can be found in Bloxham (1986b) and Bloxham *et al.* (1989). The *gufm1* model was built from similar data from the twentieth century, but a vastly expanded historical data set, described in Jonkers *et al.* (2003) – the model contains over 365 000 data and consists of 36 512 parameters. **Figure 11** shows the time distribution of the data used in *gufm1*. No account is explicitly taken of external fields in these models.

5.05.3.1.2.(ii) The comprehensive models An effort began in the early 1990s to build a comprehensive series of field models which took account of many effects which are recorded in geomagnetic data in addition to the core secular variation. The first model was reported by Sabaka and Baldwin (1993); Sabaka *et al.* (2002) described the most recent model formulation in detail while Sabaka *et al.* (2004) discuss its extension to include *Ørsted* and *CHAMP* satellite data. We will specifically report on the CM4 model of Sabaka *et al.* (2004). In general terms, the model includes representations of the main field, its secular variation, and both local-time (Sun-synchronous) and seasonal modes of the magnetospheric and ionospheric fields, as well as describing ring-current variations through the Dst index and internal fields induced by time-varying external fields. The data used in creating the model consists of *POGO*, *Magsat*, *Ørsted*, and *CHAMP* satellite data (totalling over 1.6 million observations) and over 500 000 observatory data; the latter consist of either a 1.00 a.m. observation (actually an hourly mean) on the quietest day of the month during the 1960–2002.5 period, plus observations every 2 h on quiet days during the *POGO* and *Magsat* missions.

Comprehensive models take into account not only the time-varying core magnetic field (out to degree 13) but also the static crustal field from degree 14 to degree 65. Because a model of the lithospheric field to this degree captures only a small proportion of the total lithospheric signal, it is necessary to also solve for 1635 observatory biases, generally three components at each observatory. The novel features of CM4 arise in its very sophisticated treatment of the external magnetic fields, and we will discuss these in some detail.

The ionospheric field is modeled as currents flowing in a thin shell at an altitude of 110 km. This leads to magnetic fields which are derived from potentials below and above this layer, which influence the observatory and satellite data, respectively (since all the satellites fly above this layer). In quasi-dipole coordinates, the currents are allowed to vary with 24 h, 12 h, 8 h, and 6 h periods, as well as annually and semi-annually. Induced fields are accounted for by assuming that the conductivity distribution of the Earth varies only in radius, which means that an external spherical harmonic can only excite its corresponding internal spherical harmonic. The magnetospheric field is also parametrized in a similar way, with both daily and seasonal periodicities, but also a modulation is allowed based on the Dst index. In order to take into account the poloidal F-region currents through which the satellites fly, a parametrization is made in terms of a toroidal magnetic field, which also has periodic time variations.

The model is estimated by a iteratively reweighted least-squares method, using Huber weights, and the core contribution is regularised as in [9] $n_1 = 2$ and $n_2 = 1$ in one norm and $n_1 = 0$ and $n_2 = 2$ in another. This difference from the *ufm*/*gufm* method simply represents a different approach; the fundamental quantity in the comprehensive models is the secular variation $\partial_t B_r$, which has an expansion in B-splines, and the main field B_r is found as the integral of this using the 1980 value as the offset or integration constant. All the other parameters are regularized in a similar way, by smoothing on spheres at different altitudes, representing the physical locations of the sources. In total, CM4 consists of 25 243 free parameters.

5.05.3.1.2.(iii) CHAOS field model of recent satellite data The satellites *Ørsted*, *CHAMP*, and *SAC-C* are currently providing unprecedented coverage of Earth's magnetic field. Olsen *et al.* (2006) have recently produced a cubic B-spline-based time-dependent field model spanning the interval 1999–2006 covered by data from these satellites. This model is not fully comprehensive in the sense that only a relatively simple external field model is co-estimated and toroidal fields and coupling currents in the ionosphere and magnetosphere are not explicitly considered. For spherical harmonic degrees 1–14 of the internal field, a time-dependent cubic B-spline representation is employed with knot points every year, for degrees 15–18 a static field and the linear secular variation (the first order term in a Taylor series expansion) is solved for, while for degrees 19–50 only the static field is solved for. The near-magnetospheric external fields are represented by a degree 2 spherical harmonic expansion in solar magnetospheric co-ordinates (including Dst-dependence), while a degree 2 zonal expansion in geocentric solar magnetospheric co-ordinates is used to parametrize fields with a far magnetosphere origin. Perhaps the most revolutionary aspect of this model is that the Euler angles required to transform measurements from the satellite magnetometer frame to the geocentric frame are co-estimated along with the model parameters which avoids the inconsistency of alignment using a pre-existing field model (see Olsen *et al.* (2006) for further details). Iteratively reweighted least-squares estimation involving Huber weights in the covariance matrix (see, e.g., Olsen, 2002) is used to find the preferred field model, with regularization in time only by minimizing the mean squared amplitude of the second time derivative of **B** integrated over Earth's surface and averaged over time. The CHAOS model represents the state-of-the-art in terms of determining the secular variation in the twenty-first century; future efforts will aim to incorporate this accurate information into longer time span models of the field evolution.

5.05.3.1.2.(iv) Comparison between CM3, CM4, and gufm1 To illustrate the fidelity with which the present field models are able to model observatory data, we show in Figure 15 a comparison of model *gufm1*'s predictions with some observatory annual mean data sets.

To show CM4's performance on very short timescales, Figure 16 compares the model to hourly mean values for the month of April 1990, data that were not used in deriving the model. It is clear that the model is capable of predicting variations rather well, though with more difficulty at the Antarctic station SBA (Scott Base).

Table 5 compares the performance of models *gufm1* and CM3 against observatory data, showing almost identical performance. This comes about principally because of the large intrinsic variance of the data at some observatories, which neither field model is able to capture.

Figure 17 shows a comparison of the model predictions for the variation in the first six Gauss coefficients over century and decade timescales. Although small differences exist, particularly in estimates of the instantaneous secular variation, it is apparent that modeling has reached a stage where there is considerable consensus between the models.

In the next section, we move on to describe the characteristics of secular variation as observed on Earth's surface and inferred at the CMB. We will ultimately (Section 5.05.5) describe possible underlying physical mechanisms in terms of core hydromagnetics. Most of the results shown in the next section (unless explicitly stated otherwise) are derived from the *gufm1* field model of Jackson *et al.* (2000), which, as we have shown in this section, provides a good representation of the historical field evolution.

5.05.4 Historical Field Evolution – Long-Term Secular Variation

5.05.4.1 Field Evolution at the Earth's Surface

The magnetic field at Earth's surface has changed significantly over the past 400 years. This can clearly be seen, for example, in the long times series of

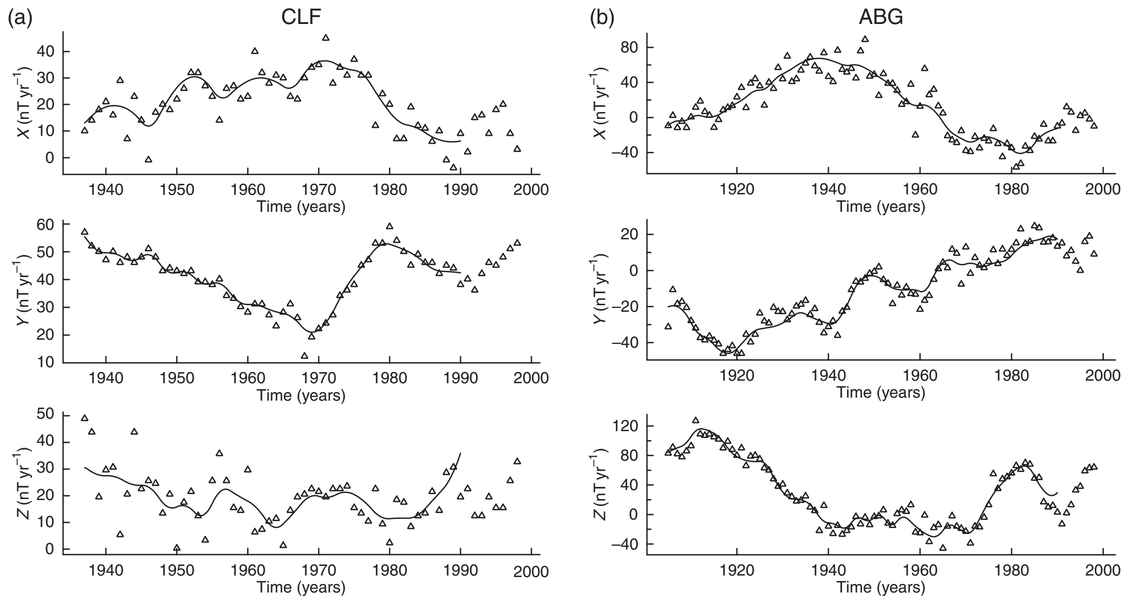

Figure 15 Comparison of secular variation models and first differences of annual means. Two observatories (a) Chambon-la-Foret, France, and b) Alibag, India, are shown comparing observed field rate of change with predictions from the model *gufm1* (solid lines). The symbols show the rate of change of the field, as obtained from first differences of annual means. The *X*-, *Y*-, and Z-components are in the northerly, easterly, and downward directions, respectively. Because the post-1990 data were not used in the creation of *gufm1*, there is a small mismatch at the end of the data series — this shows the difficulty in predicting the secular variation.

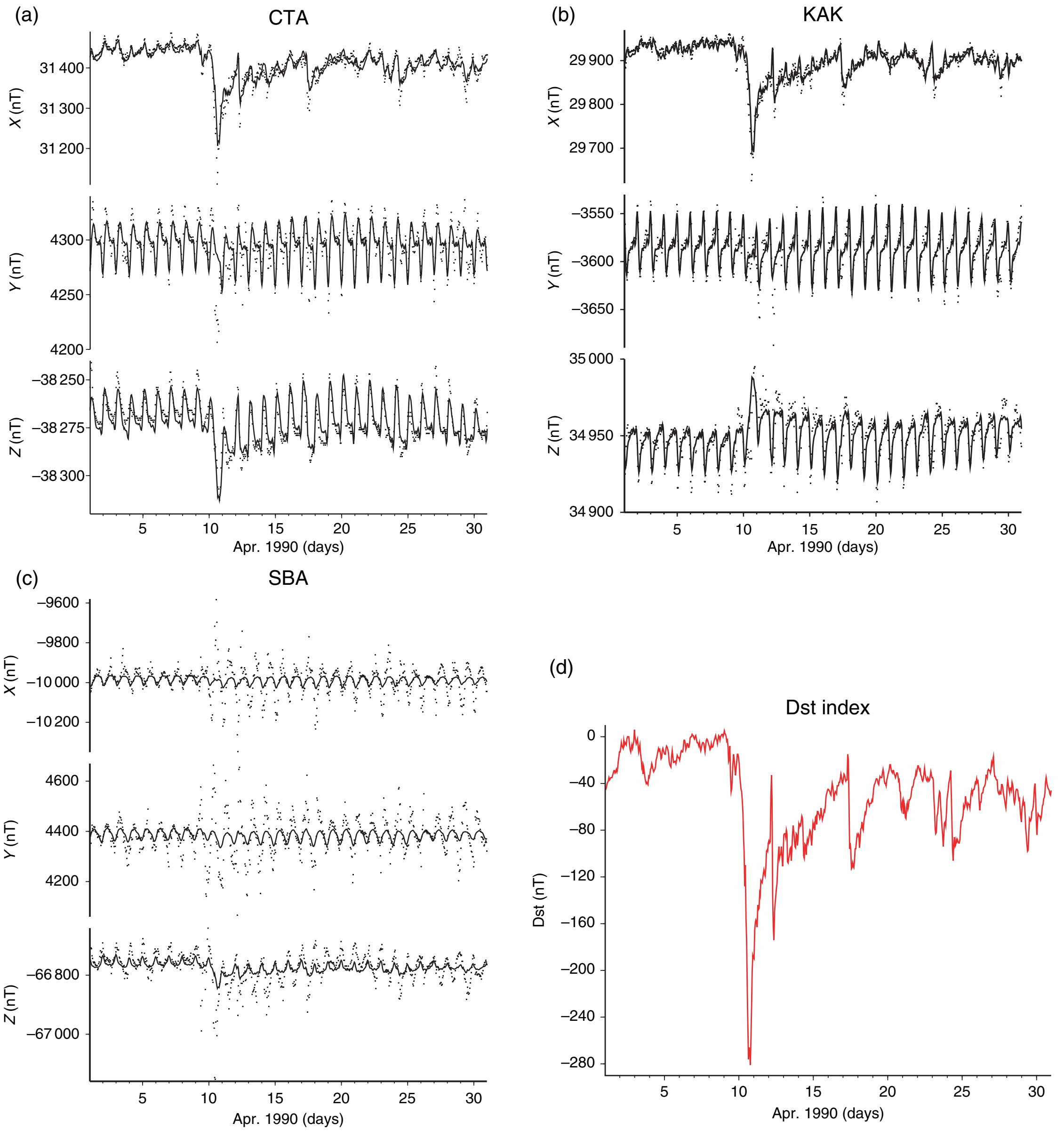

Figure 16 Comparison of one month (Apr. 1990) of hourly mean data. Observed *X*-, *Y*-, and *Z*-, components (dots) from selected observatories and predictions (solid line) from CM4. (a) Charters Towers; (b) Kakioka; (c) Scott Base; (d) the Dst index for April. Note the commencement of a magnetic storm on the 10th day. The Dst index is used in the synthesis of predictions at individual observatories; this is particularly noticeable in the predictions of the *X* component in (a) and (b).

measurements cataloged by Malin and Bullard (1981) (**Figure 18**). In fact, it was such measurements of changes in declination by Henry Gellibrand in 1634 that first indicated the existence of geomagnetic secular variation. Note that in this figure, since declination and inclination are nonlinearly related to the model **m**, it is not possible to account for observatory crustal biases.

The best way to appreciate global field changes (i.e., secular variation) is for the reader to study contour maps of different field components and compare how they have evolved. In **Figure 19**, the declination (D) at Earth's surface is shown in AD 1590 and 1990, while **Figure 20** shows the inclination anomaly (I_a) (defined as the difference between the observed inclination and that of a geocentric dipole) at the same

Table 5 Comparison of rms differences (in nanotesla) between observatory annual means and predictions from the models *gufm1* and CM3, the latter with or without its external contribution

Component	No. of data	gufm1	CM3 (all)	CM3 (no external)
X	4047	17.71	17.48	18.09
Y	4047	21.27	21.45	21.47
Z	4047	24.55	24.49	24.53

epochs. **Figures 21**, **22**, and **23** catalog the evolution of the vertical component of the field (which is much larger than the horizontal components except at low latitudes) at AD 1590, 1690, 1790, 1890 and 1990.

5.05.4.1.1 The westward drift

Perhaps the most striking aspect of the geomagnetic secular variation over the past 400 years is the westward motion of the field at Earth's surface. This phenomenon has been recognized since the time of Halley (1683, 1692) and was first analyzed in detail by Bullard *et al.* (1950), who concluded that the non-dipole part of the field had moved westward at a rate of 0.18° per year during the first half of the twentieth century. Bullard *et al.* (1950) and later Yukutake (1962) suggested that the westward drift was not globally constant but rather depended on latitude; subsequently, Yukutake and Tachinaka (1969) realised that it could be better explained by separating the field into standing and drifting parts. The latitudinal dependence of the drift rate was conclusively demonstrated by Jault *et al.* (1988).

The westward motion of the field is most easily seen by following the motion of the agonic lines (where $D=0$) in **Figure 19**. It can be seen that in 1590 one agonic line bisected the African continent, running through the Cape of Good Hope (which at this time was named Cape Agulhas ('needle cape') by sailors due the coincidence of the directions of true and magnetic north there); fast-forwarding 400 years to AD 1990 we find that the same agonic line has now moved westward, so that it now bisects southern America. The maxima and minima of inclination anomalies centered on low latitudes can also be tracked westward, for example, the inclination anomaly high that was present over Africa in 1590 now lies on the western edge of south America. Contour maps of the vertical component of the magnetic field are dominated by the axial dipole component of the field which is unchanged by westward motion due to its axisymmetric nature; however, the westward motion of nonaxial dipole parts of the field can still be discerned in the maps of **Figures 21**, **22** and **23**, especially by following long-lived distortions in the magnetic equator. A southwest-to-northeast trending element of the magnetic equator can be followed from its initial location at the Indian ocean in 1590, through to Africa in 1790 and the Atlantic in 1890, to the eastern edge of southern America in 1990.

5.05.4.1.2 Hemispherical asymmetry

The description of westward motion of field features in the previous section focused on high-amplitude features moving across the Atlantic hemisphere (longitude 90E to 90W). In contrast, the field evolution in the Pacific hemisphere is characterized by lower-amplitude features and a lack of systematic secular variation. This asymmetry between the hemispheres was first discussed by Fisk (1931), and it has been suggested that this could be a consequence of the influence of lower-mantle inhomogeneities on the dynamo in the core (Doell and Cox, 1971). This interpretation remains controversial, as it is known that asymmetric field morphologies are transiently possible during highly supercritical core convection (see Volume 8).

5.05.4.1.3 Axial dipole decay

The westward drift is only part of the observed secular variation. The largest contribution to present-day secular variation comes from the decay of the axial dipole part of field. The axial dipole has decayed rapidly at an average rate of 5% per century since the first direct measurements of intensity (Barraclough, 1974). Gubbins *et al.* (2006) have recently used paleointensity measurements from the database of Korte *et al.* (2005) along with estimates of field directions from *gufm1* to infer g_1^0 for the interval 1590–1840 and found that field decay rate was much slower (almost constant) during this earlier interval. **Figure 24** shows both the extrapolation of Barraclough (1974) used by Jackson *et al.* (2000) and the result of Gubbins *et al.* (2006). The variability in the rate of change of the axial dipole over past four

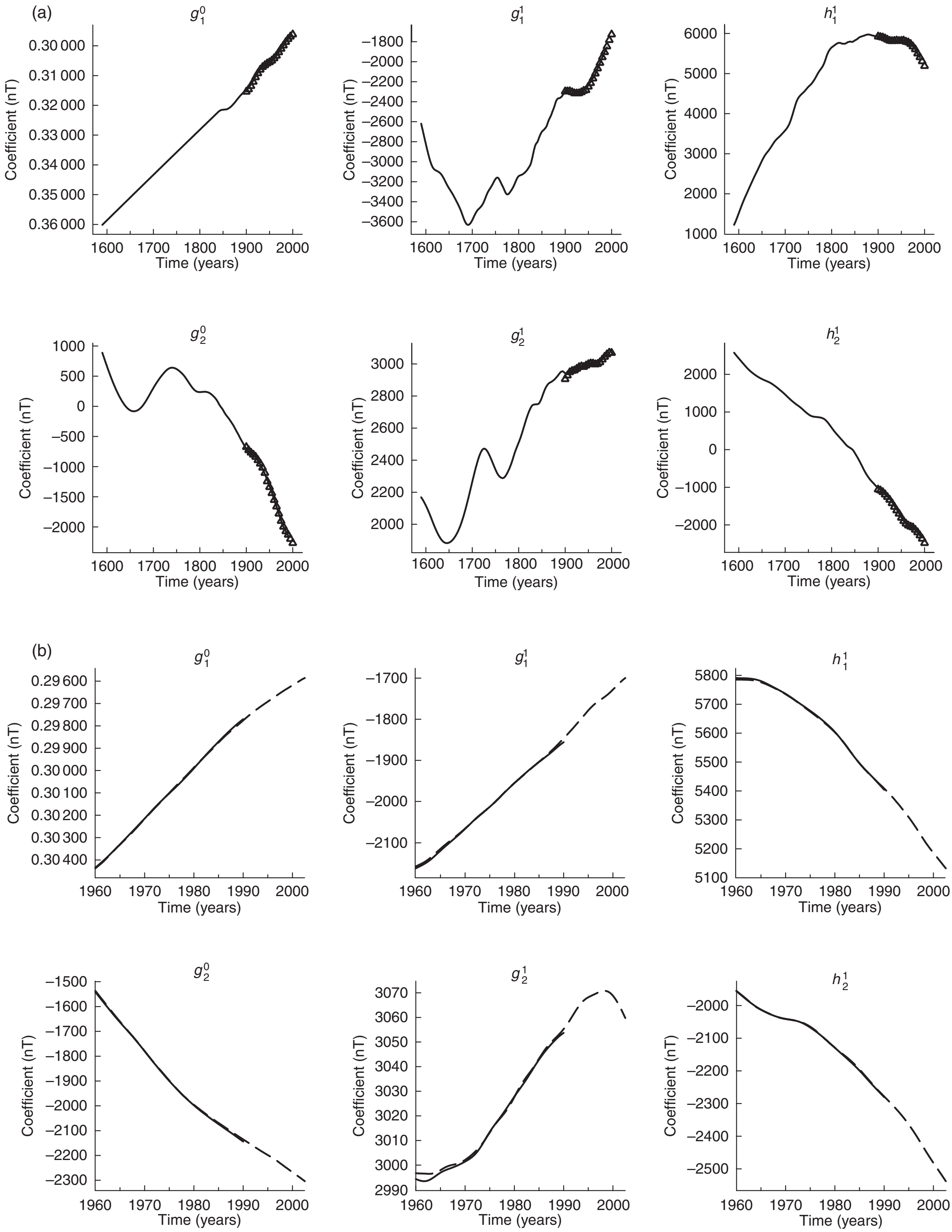

Figure 17 Comparison of model values for first six Gauss coefficients. (a) 1590–1990; (b) 1960–2002.5. Solid is *gufm1*, dashed is CM4, and the triangles are DGRFs. In (a), g_1^0 has been fixed to decrease at a rate of $15\,\mathrm{nT\,yr^{-1}}$ prior to 1840; in the absence of intensity data, it is necessary to fix the amplitude of the solutions.

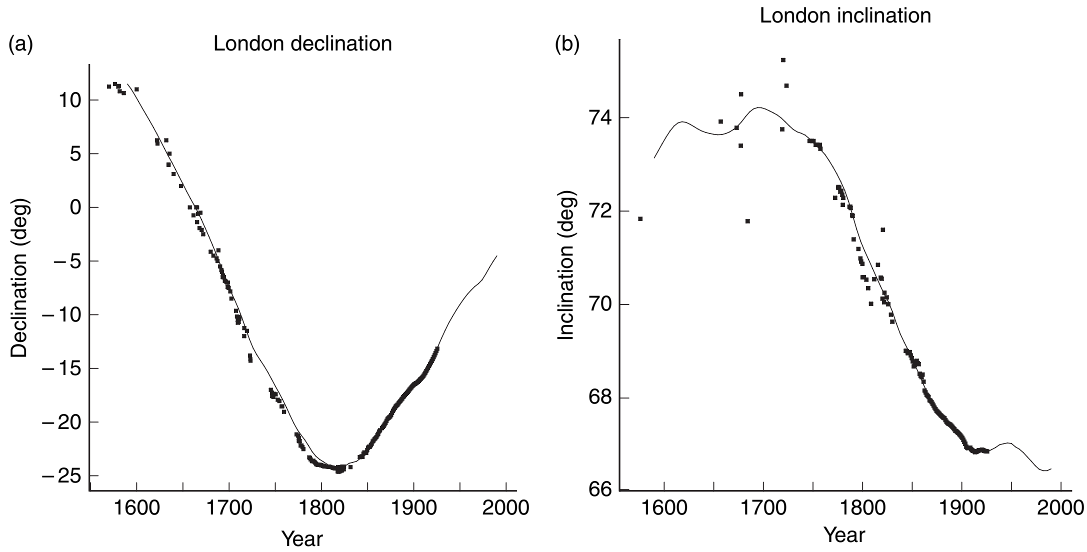

Figure 18 Declination (*D*) and inclination (*I*) in London during historical times (Malin and Bullard, 1981), and the fit by *gufm1* of Jackson *et al.* (2000) (line). The data were not used in the construction of *gufm1*, and provide an independent check of its fidelity. Note that the inclination prior to 1700 is very poorly constrained, and the model *gufm1* differs from some archeomagnetic measurements of inclination in Europe (Le Goff, personal communication, 2005).

centuries suggests that the dynamo process generating the main field is not steady, but is continuously fluctuating in strength.

5.05.4.1.4 Timescale associated with different wavelengths (spherical harmonic degrees)

A useful statistical estimate of how changes in Earth's magnetic field at the surface depend on the length scale under consideration is the reorganization (or correlation) time $\tau(l)$ introduced by Hulot and Le Mouël (1994),

$$\tau(l) = \sqrt{\frac{\sum_m \left(g_l^m\right)^2 + \left(h_l^m\right)^2}{\sum_m \left(\dot{g}_l^m\right)^2 + \left(\dot{h}_l^m\right)^2}} \qquad [10]$$

This quantity is a measure of how long it takes for power at spherical harmonic degree l to be completely changed (altered by an amount equal to its current value) given its present rate of change. Physically, this corresponds to the time taken to completely reorganize field features of a particular size. In order to calculate τ (l) one requires only a model of the main field and its time derivative at a given time. In **Figure 25**, $\tau(l)$ derived from the CHAOS model (Olsen *et al.*, 2006) is presented.

The CHAOS model contains global data sets from the *Ørsted*, *CHAMP*, and *SAC-C* satellites (see earlier description) and is the most accurate short-time-span field model yet derived, thus giving the best picture currently available of τ, to the highest possible degree l. Olsen *et al.* (2006) found that a power law $\tau(l) = 890 l^{-1.35}$ yr provided a good fit to the calculated values of $\tau(l)$ so that for $l = 1$, $\tau \sim 900$ yr while for $l = 16$, $\tau \sim 20$ yr. These results illustrate that secular variation processes span a wide range of timescales but that the small length scales that can now be accurately monitored are evolving on timescales short enough to allow detailed study in the coming decades. Holme and Olsen (2006), analyzing an earlier model (CO2003), found $\tau \sim 1000 l^{-1.45}$; they have argued that the simple power-law functional form of $\tau(l)$ suggests a consistent (l-independent) form of advection–diffusion balance producing the secular variation. For further discussion on the core processes underlying secular variation, *see* Chapter 8.04.

5.05.4.1.5 Evolution of integrated rate of change of vertical field at Earth's surface

It is also interesting to consider the evolution of a global measure of the amplitude of the instantaneous rate of change of vertical field at Earth's surface. A suitable measure is the root mean square (rms) of $\dot{B}_r$ integrated over Earth's surface. This quantity is plotted in **Figure 26** from 1840 to 1990, for the *ufm1* model.

Dramatic changes in the integrated instantaneous rate of B_r are observed to have occurred. These changes are thought to be robust as *ufm1* is a good

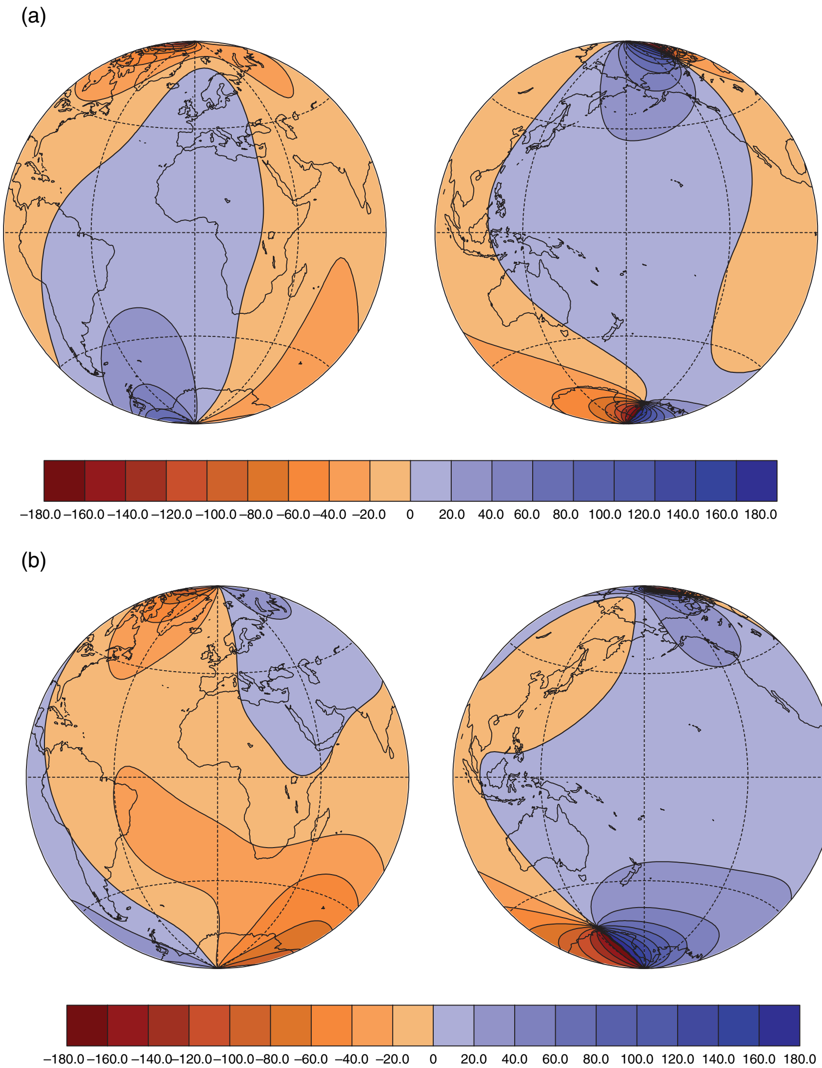

Figure 19 Declination D at Earth's surface in (a) AD 1590 and (b) AD 1990 from the model *gufm1* of Jackson *et al.* (2000). Plots are Lambert equal area projections of the Atlantic and Pacific hemi-spheres. Color bars are at 20° intervals, red being negative and blue positive. Note the westward displacement of the agonic lines where $D = 0$.

representation of the globally averaged field evolution at the surface, as is evident in comparisons of the model prediction with observed secular variation. Particularly dramatic is the 20% increase in the amplitude of the secular variation at the start of the twentieth century. This is thought to be associated with an increase in the rms core surface flow velocity; Hulot *et al.* (1993) inferred by inversion of the observed secular variation (*see* Chapter 8.10) that the rms flow speed increased at this time, and it appears to be the case that zonal (axisymmetric) core flow speeds altered precisely in the required way for observed decadal length-of-day changes to be explained by geostrophic core motions (Jackson, 1997). It is also remarkable that there are a number of local maxima and minima in **Figure 26** occuring throughout the twentieth century. These extrema seemingly mark reorganizations of the global secular variation, and at least some of them appear coincident with so-called geomagnetic jerks that are discussed in the next section.

5.05.4.1.6 Geomagnetic jerks

Geomagnetic jerks or secular variation impulses are abrupt changes in the second time derivative of the geomagnetic field at Earth's surface

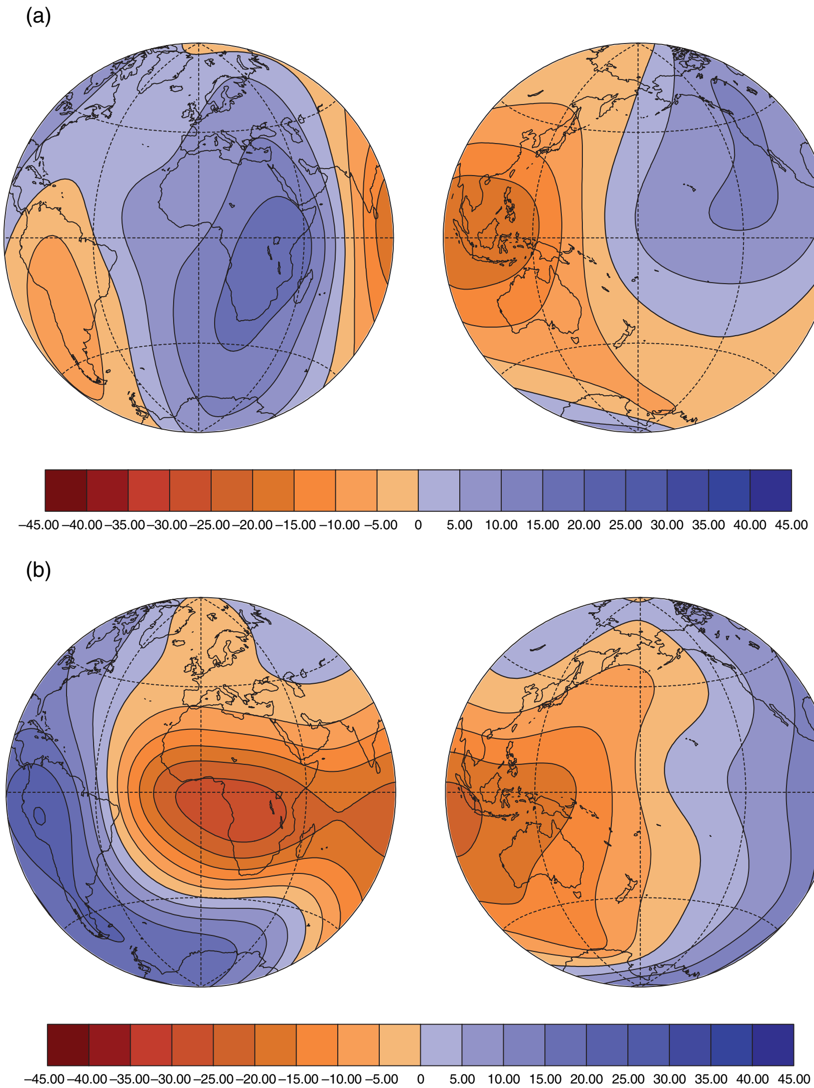

Figure 20 Inclination anomaly I_a at Earth's surface in (a) AD 1590 and (b) AD 1990 from the model *gufm1* of Jackson *et al.* (2000). Plots are Lambert equal area projections of the Atlantic and Pacific hemispheres. Inclination anomaly is the inclination of the field minus that expected for a geocentric dipole. Color bars are at 5° intervals, red being negative and blue positive.

(see, e.g., Courtillot and Le Mouël, 1984). During the twentieth century, they were found to separate intervals of linearly changing secular variation and have been unambiguously identified as having occurred in 1901, 1913, 1925, 1969, 1978, 1991, and 1999 (Alexandrescu *et al.*, 1995; Macmillan, 1996; Mandea *et al.*, 2000). The signature of jerks can be seen particularly clearly at European observatories, for example, in **Figure 27**, which shows the evolution of the linear rate of change of the eastward component of the geomagnetic field ($\dot{Y}$) in Niemegk.

A 12 month running average filter has been applied to the central differences of monthly mean data to produce this time series, following the methodology of Mandea *et al.* (2000). The jerk events are captured (at least in a smoothed manner) by the internal field representation of global models such as CM4 and *gufm1* — this can been seen, for example, in **Figure 15**, which compares model results to observatory annual means. Jerks are not always observed at all locations and those that are observed are not simultaneous; Alexandrescu *et al.* (1996) noted that, for example

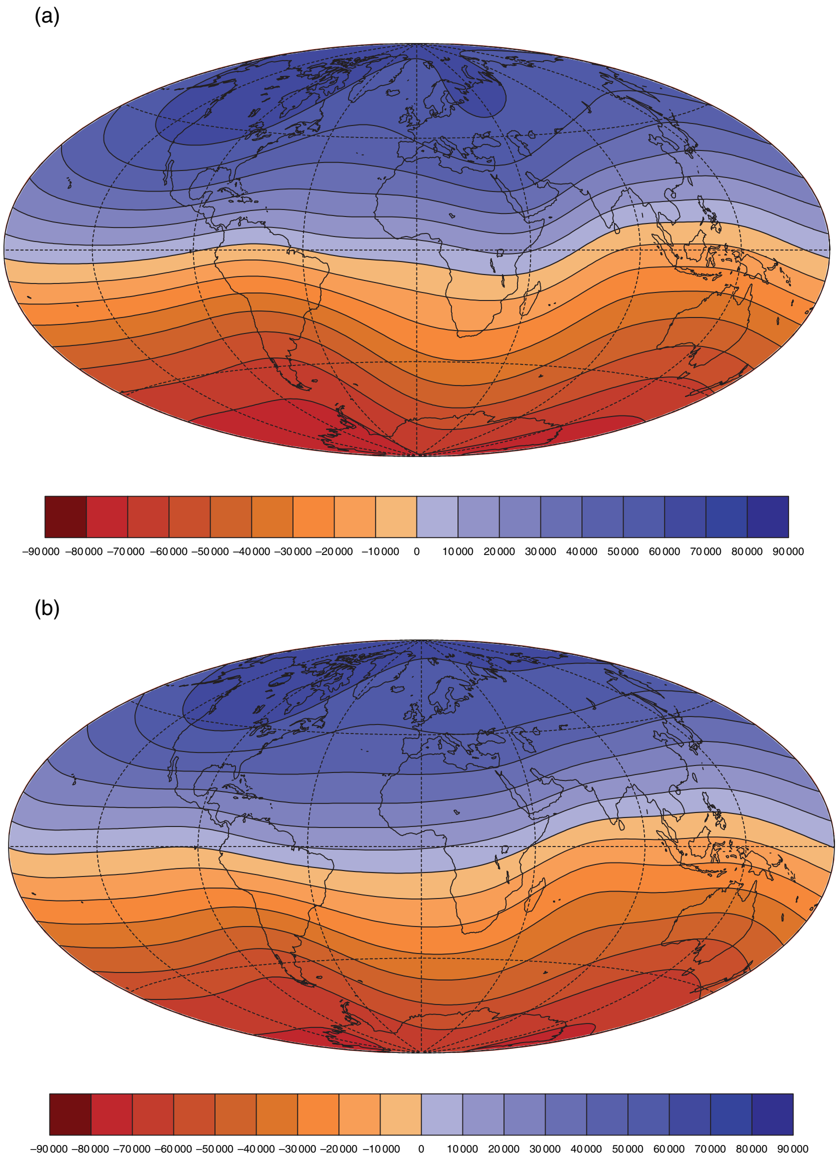

Figure 21 Vertical magnetic field B_Z at Earth's surface in (a) AD 1590 and (b) AD 1690 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection; each color interval represents a 10 000 nT increment.

in 1969, the signature of the jerk tended to be observed later in the Southern Hemisphere (see **Figure 28**).

The physical process causing jerks, as well as geographic variations in their detectability and time delays in their observation, are not well understood. Bloxham *et al.* (2002) suggested that jerks might be the surface manifestation of a superposition of torsional oscillations (a special class of axisymmetric, geostrophic, hydromagnetic waves likely to be present in Earth's outer core – see Volume 8), and that variations in their detectability might be the result of variation in the field morphology at the core surface. Alexandrescu *et al.* (1999) and Nagao *et al.* (2003) have suggested that variations in mantle conductivity could explain the observed delays in jerk observations. Much work remains to be carried out in understanding the physical mechanisms involved and in testing the various hypotheses.

Variations in the main geomagnetic field have their origin in Earth's core. Most insight into the

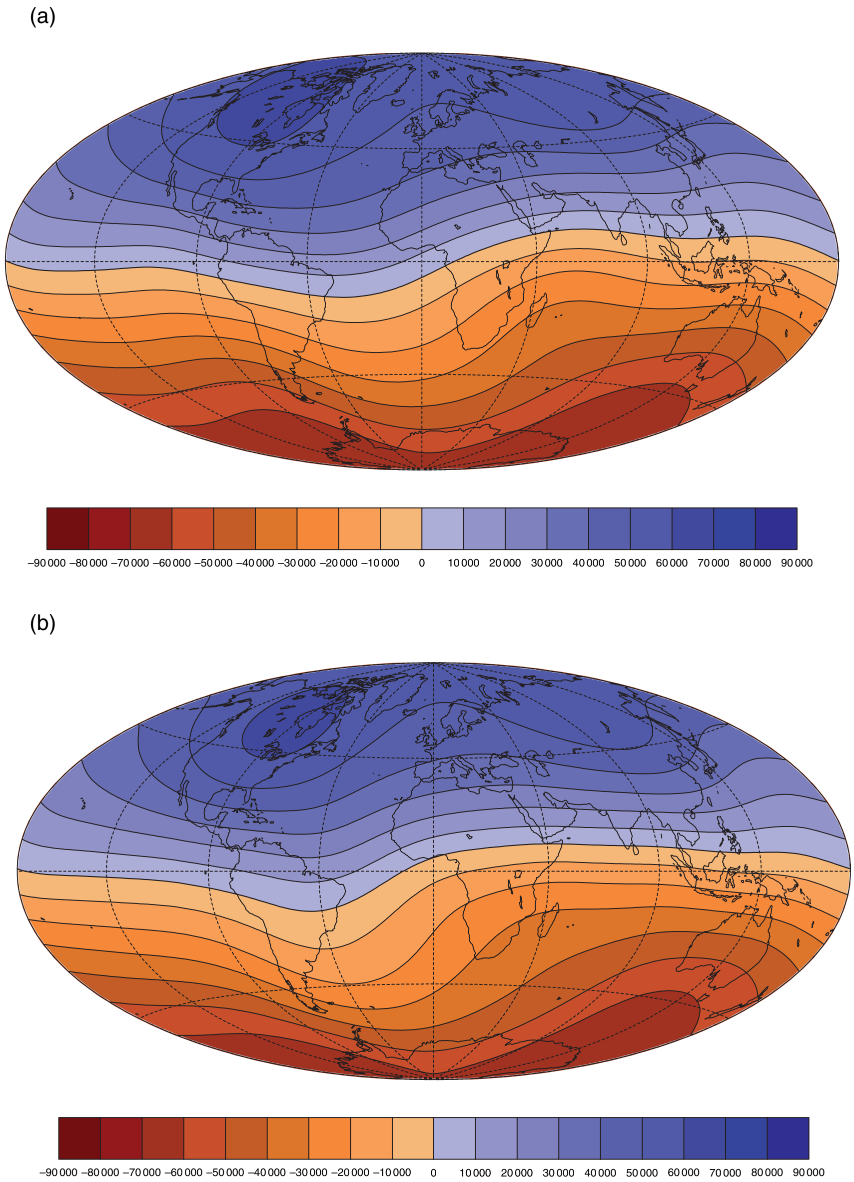

Figure 22 Vertical magnetic field B_Z at Earth's surface in (a) AD 1790 and (b) AD 1890 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection; each color bar represents a 10 000 nT increment.

physical mechanisms causing the field evolution can therefore be obtained by examining the patterns of field evolution at the core surface. To determine the core field we adopt the approximation of treating the mantle as a perfect insulator. This approximation has been studied by Benton and Whaler (1983), who show that when variations are considered whose periods are longer than annual, the error introduced is small when the mantle has an electrical conductivity structure as currently believed (i.e., from 10^{-2} to $10\,\mathrm{S\,m^{-1}}$). In the next section, the patterns of field evolution that result from such an approach are discussed in detail.

5.05.4.2 Field Evolution at the Core Surface

The evolution of the geomagnetic field at the core surface over the past few centuries was first described in detail by Bloxham and Gubbins (1985) and Bloxham *et al.* (1989) by considering a series of

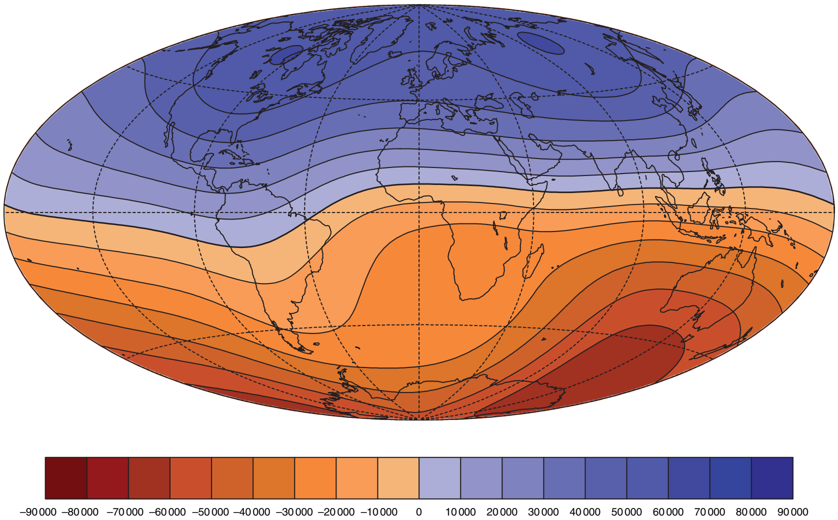

Figure 23 Vertical magnetic field B_Z at Earth's surface in AD 1990 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection; each color bar represents a 10 000 nT increment.

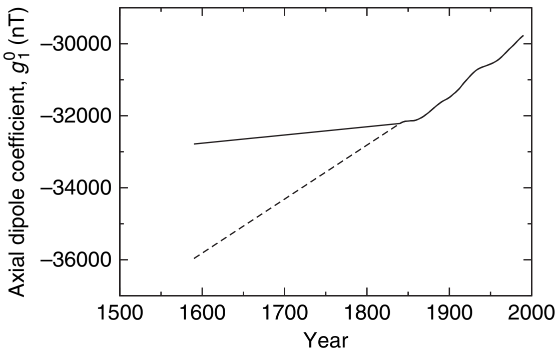

Figure 24 Evolution of g_0^1 in nanotesla since direct measurements began in 1590; absolute intensity measurements began sporadically only in 1832; the two slopes shown before 1840 are (a) the extrapolation based on the average rate of fall since 1840 and (b) derived from indirect paleointensity measurements by Gubbins *et al.* (2006). The solid line indicates g_1^0 from *gufm1* with pre-1840 decay of 2.28 nT and dashed line g_1^0 from *gufm1* with pre-1840 decay of 15 nT. Adapted from Gubbins D, Jones AL, and Finlay CC (2006). Fall in Earth's magnetic field is erratic. *Science* 312: 900–902.

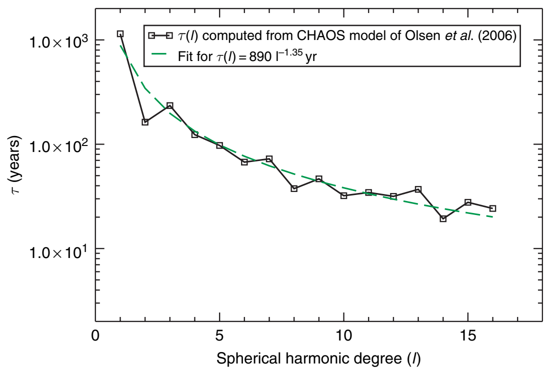

Figure 25 Estimated reorganization timescales τ. The reorganization (or correlation) time $\tau(l) = \sqrt{\sum_m (g_l^m)^2 + (h_l^m)^2 / \sum_m (\dot{g}_l^m)^2 + (\dot{h}_l^m)^2}$ (Hulot and Le Mouël, 1994) is derived from the CHAOS field model (Olsen *et al.*, 2006), giving an instantaneous estimate of the time taken for power at spherical harmonic degree l of the field to be completely changed or renewed. Solid line shows the CHAOS estimates; dashed line shows a two-parameter exponential fit to the data.

single-epoch models. The picture they described has been borne out by the more recent time-dependent field models *ufm1* (Bloxham and Jackson, 1992) and *gufm1* (Jackson *et al.*, 2000), so we shall reiterate their findings here before discussing more recent developments. Contour plots of the historical evolution of the vertical field at the core surface are found in **Figures 29, 30, and 31**.

The structure of the vertical field at the core surface is considerably more complicated than at the surface, because higher degree spherical harmonics are amplified more (by a factor $(a/c)^{(l+2)}$ where a is the radius of the Earth, c is the core radius, and l is the spherical harmonic degree) during the downward continuation procedure. This is one reason why it is preferable to downward continue regularized field models rather

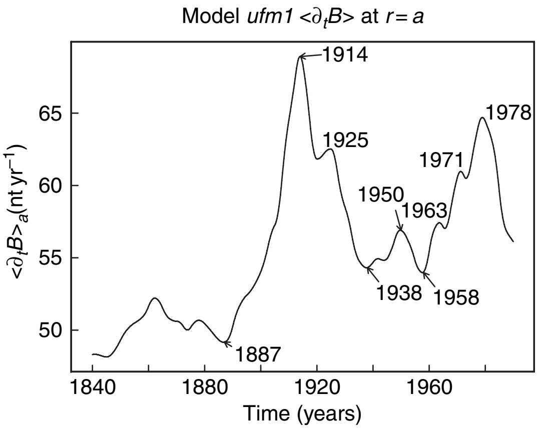

Figure 26 The RMS B_r integrated over Earth's surface from the *ufm1* time-dependent field model.

than those that have been simply truncated, and may contain noise contributions in the higher degree spherical harmonics. Downward continuing truncated field models also unfortunately introduces the possibility of unwanted Gibbs ringing effects due to the sharp cutoff in spectral space (see e.g., Whaler and Gubbins, 1981; Shure *et al.*, 1982; Gubbins, 1983). The *gufm1* model results presented and discussed here have been regularized so the power spectrum for the model has reached negligible values by the time the nominal cutoff at spherical harmonic degree $l = 14$ is reached.

5.05.4.2.1 High latitude, approximately stationary flux lobes

Probably the most prominent feature in the maps of the vertical field at the core surface are the high-intensity flux lobes (by which we mean the areas of flux maxima, of either sign) under Arctic Canada, Siberia, and under the eastern and western edges of Antarctica; they can be seen particularly clearly in **Figure 31**. These lobes are responsible for the predominantly axial dipole field structure observed at the surface and have remained approximately stationary (wobbling slightly about a mean position) over the past four centuries. Gubbins and Bloxham (1987) identified these high-latitude flux lobes as the signature of columnar convection rolls in the core (Busse, 1975) which are thought to be a major ingredient in the geodynamo process (Kono and Roberts, 2002). They proposed that flow convergence associated with downwelling in the convection rolls is responsible for producing the observed field concentrations. Bloxham and Gubbins (1987) ascribed the relative stationarity of these flux lobes to the influence of heat flow inhomogeneities at the CMB associated with the structure of mantle convection. More detailed studies using geodynamo simulations (Olson and Christensen, 2002; Bloxham, 2002) have confirmed the feasibility of this mechanism.

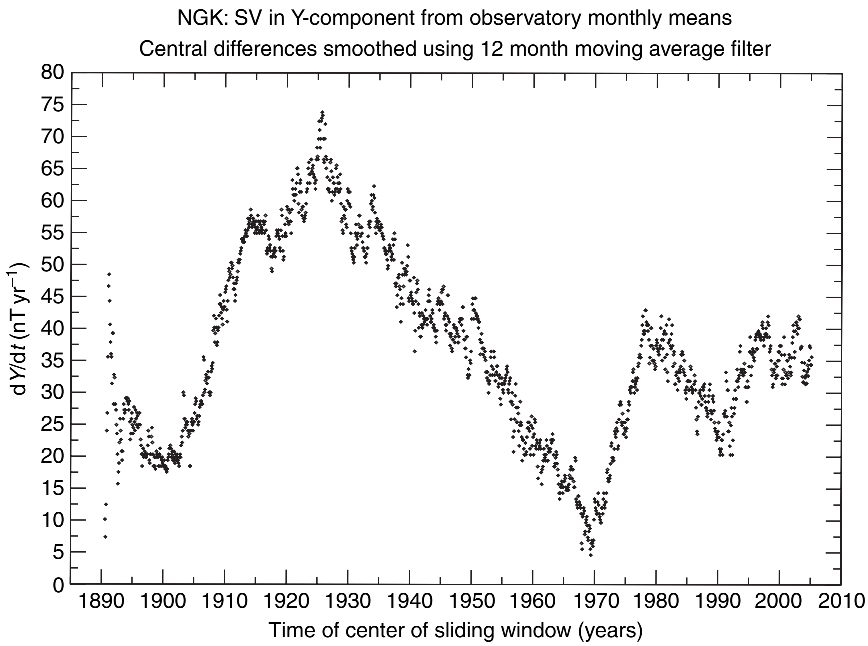

Figure 27 Central differences of monthly means of $\dot{Y}$ at the Niemengk observatory processed using a 12 month moving average filter.

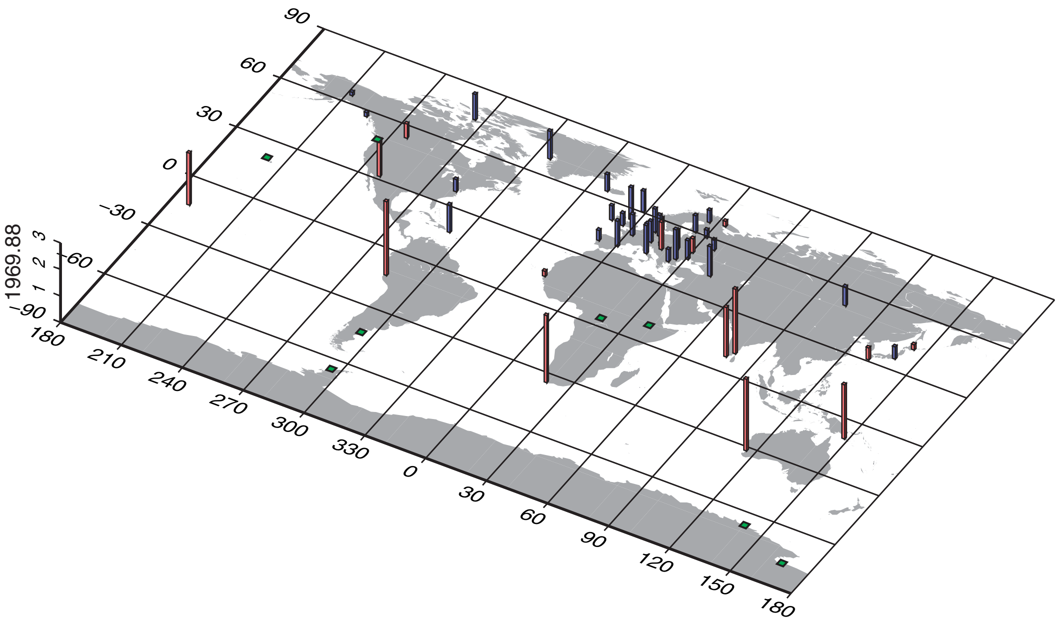

Figure 28 Geographical distribution of the times of occurence of the 1969 jerk measured by Alexandrescu *et al.* (1996). A linear combination of *X*- and *Y*-field components was analyzed using wavelet ridge functions and the jerk onset time estimated. Blue bars represent negative delays relative to the mean occurence time (1969.88) and correspond to earlier jerks, whereas red bars represent positive delays relative to the mean (later jerks). The scale bar varies from 0 to 3 years. Green squares represent locations where jerks were not detected.

5.05.4.2.2 Reversed flux patches

The presence of reversed flux features at the core surface is a major difference to the field structure observed at Earth's surface. Most prominent of these are the patch that is found close to the geographical North Pole throughout most of the past 400 years, and the large feature that extends from under southern Africa across to under southern America that has been formed by the coalescence of two earlier patches. Gubbins (1987) and Gubbins *et al.* (2006) have linked the growth and migration of the South Atlantic patch to the rapid decay of the axial dipole field observed since 1840. The significance of the changes in the flux through these patches is discussed in Section 5.05.5.4. If taken at face value, the growth of the South Atlantic patch implies a failure of a particularly attractive approximation for the core, the so-called frozen flux hypothesis, which consequently means that it is very difficult to retrieve fluid motions at the core surface. It is important to recognize that the increase in quality, quantity, and distribution of data throughout time leads to increased complexity in the field models, and it is very difficult to disentangle this effect from true diffusional effects; we refer the reader to the discussion in Section 5.05.5.4.

5.05.4.2.3 Low-latitude, westward-drifting field features

Bloxham and Gubbins (1985) pointed out the presence of a number of rapidly westward-moving field concentrations at low and mid-latitudes, especially clear in the Atlantic hemisphere. Bloxham *et al.* (1989) further noted that beneath Europe and the Atlantic Ocean during the twentieth century there was a westward-moving sequence of field highs and lows, and referred to this as a mid-latitude polar wave. They suggested this could be the signature of a wave with azimuthal wave number between $m=5$ and $m=9$. Jackson (2003) examined very high resolution images of the field at the core surface in 1980 and 2000, constructed using high-quality satellite data and utilizing a maximum entropy regularization technique. He showed that the wave-like feature identified in the Northern Hemisphere by Bloxham *et al.* (1989) has a counterpart at low latitude on the other side of the geomagnetic equator with amplitude considerably higher than was evident in previous studies.

Since these wave-like features appear to move essentially east–west, their motion can be tracked using plots of field amplitude as a function of time and longitude (TL plots). TL plots of the vertical component of the field at the core surface between

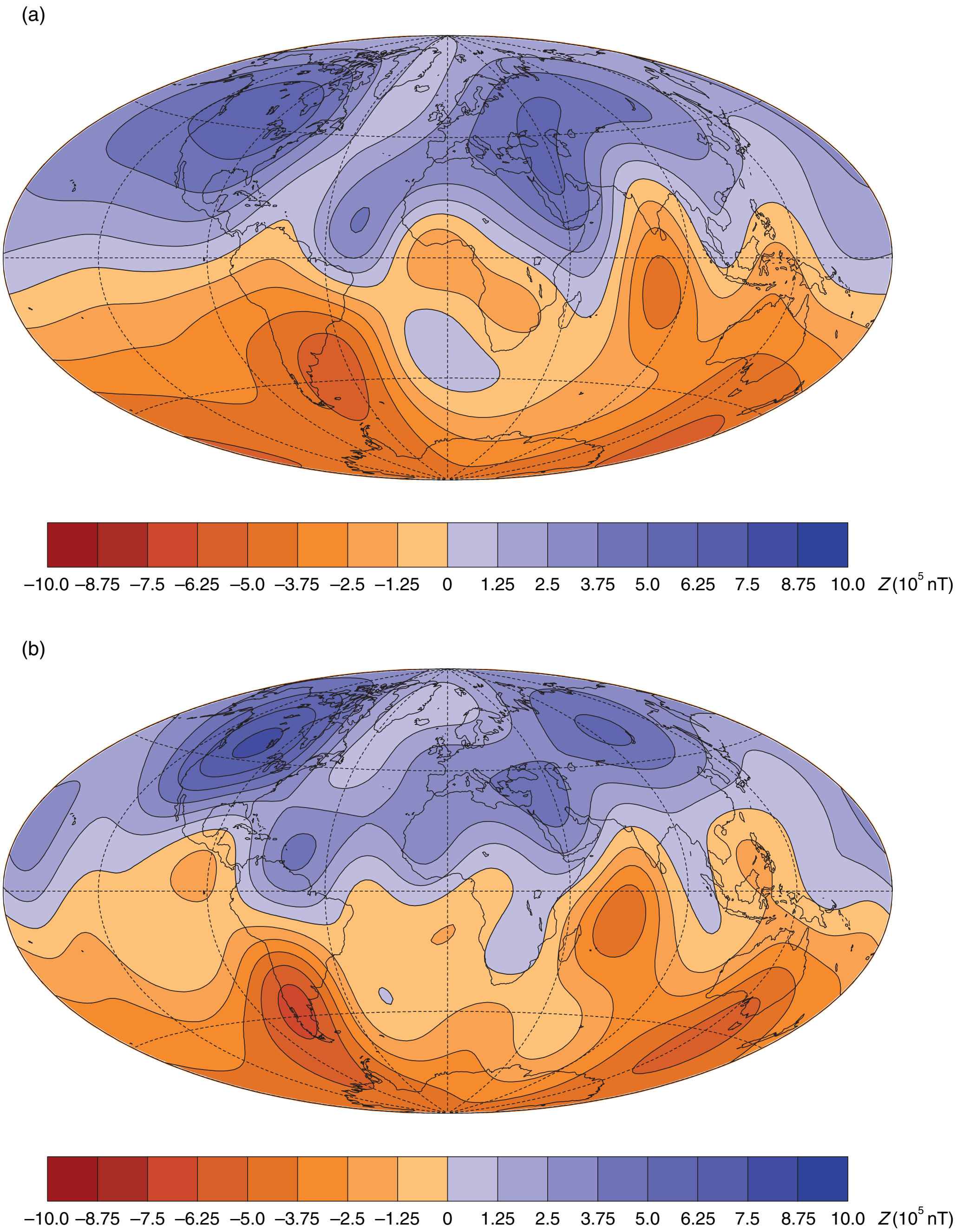

Figure 29 Vertical magnetic field B_Z at the core surface in (a) AD 1590 and (b) AD 1690 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection, units are nanotesla.

1590 and 1990 (from the *gufm1* model) at latitudes 60° N, 40° N, 20° N, at the equator, at 20° S, and at 60° S are presented in **Figure 32**.

In TL plots, linear features of high field amplitude running up-down the page indicate stationary flux features such as the high-latitude flux features (see, e.g., at latitudes 60° N between longitudes − 120° and − 90° as well as between + 90° and + 120°, similarly at 60° S, near longitudes − 90° and + 120°). At lower latitudes, for example at 20° N after 1900, at the equator between longitudes 0 and + 90° and at 20° S there are some hints of diagonal lines of high field intensity that represent azimuthally (east–west) moving field features. Unfortunately, it is rather difficult to analyze these

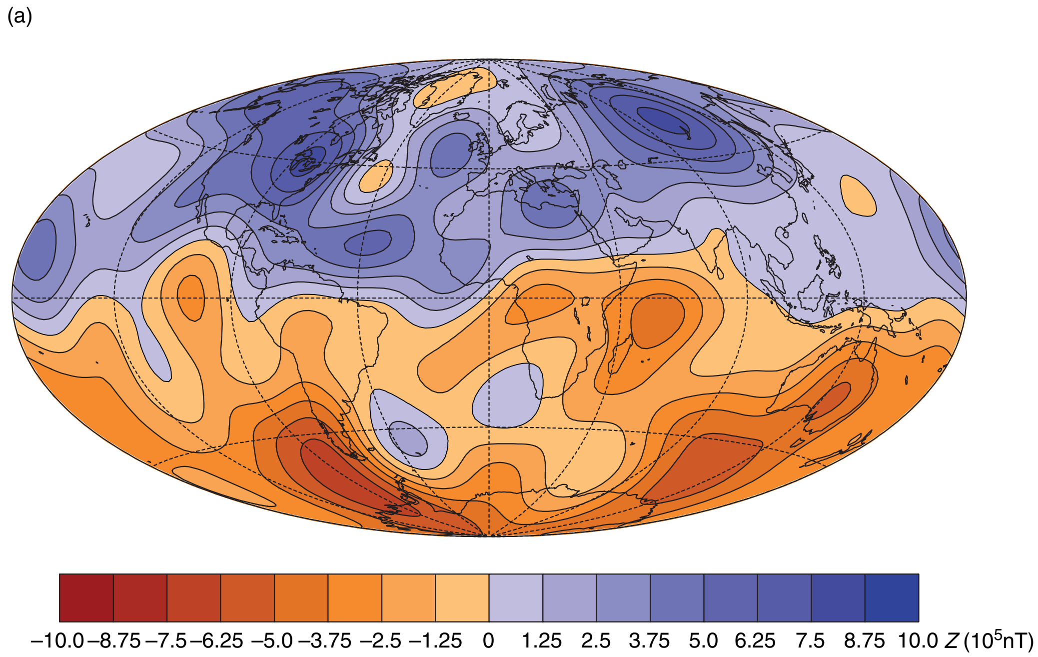

Figure 30 Vertical magnetic field B_Z in (a) AD 1790 and (b) AD 1890 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection, units are nanotesla.

features because they are swamped by stationary features that are not of interest in this context. To get round this problem, Finlay and Jackson (2003) high-pass-filtered the vertical field from the *gufm1* model, removing the time-averaged axisymmetric field and all field components varying on timescales longer than the 400 years to obtain a field which they denoted by $\tilde{B}_r$. The result of this processing is shown in TL plots at the same latitudes as before in **Figure 33**. Note that the first and last 40 years of the record have been disregarded to eliminate filter warm-up effects, namely the fact that the edges of the time series affect the filtered output.

The filtering reveals clear westward-moving, wave-like, signals at low latitudes (between 20° N

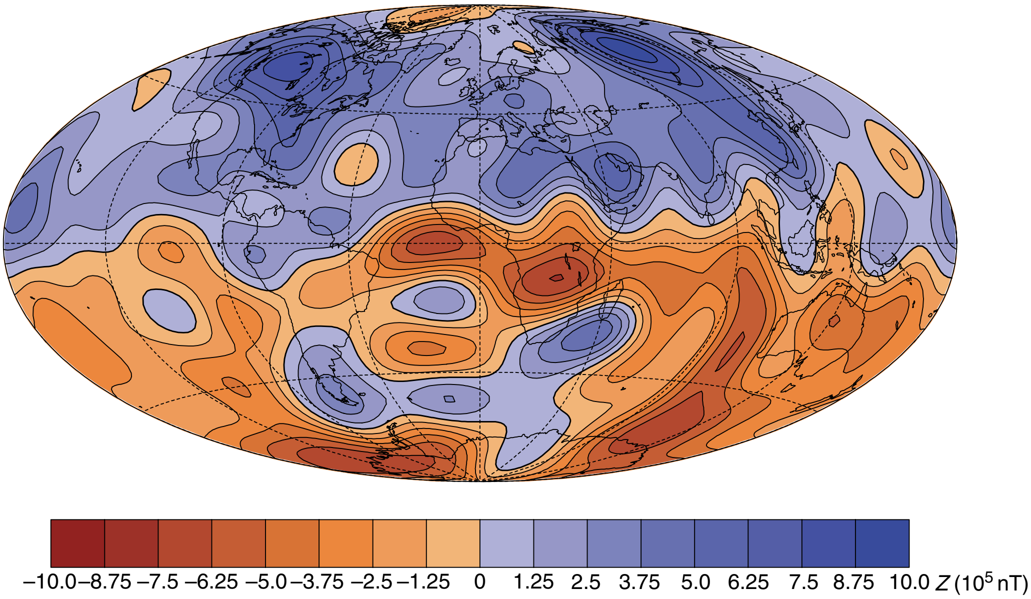

Figure 31 Vertical magnetic field B_Z at the core surface in AD 1990 from the model *gufm1* of Jackson *et al.* (2000). Plots are Molleweide projection, units are nanotesla.

and S and particularly striking at the equator). No clear wave-like motions were found at higher latitudes, indicating that such patterns of secular variation are confined to low latitudes on timescales shorter than 400 years. By measuring the power traveling at different angles in the TL plots at all latitudes, Finlay (2005) showed that it is possible to construct latitude-azimuthal speed (LAS) plots, which summarize the relative strength, location, and rate of azimuthal secular variation processes. Such plots, constructed when the vertical field at the core surface from *gufm1* is high-pass-filtered with thresholds of 2500, 600, and 200 years, are shown in **Figure 34**.

The LAS power plots suggest that three distinct types of azimuthal secular variation have been operating during the past four centuries. At mid- to high latitudes in the Northern Hemisphere, there are weak signals probably associated with the wobbles of the high-latitude flux lobes. These motions are both eastward and westward and appear most clearly when long timescale field variations are retained in **Figures 34(a)** and **34(b)**. Next, there is the strong equatorially confined signal with speed of approximately 17 km yr^{-1} westward, as described by Finlay and Jackson (2003). This is the dominant signal when only field variations with timescales shorter than 400 years are considered, and looks in TL plots to have the form of a wave-like disturbance. Finally, on all timescales, there is a strong westward signal in the Southern Hemisphere, which is particularly clear when the filter threshold is much longer than the record length. It seems to be associated with the westward motion of reversed flux features in the South Atlantic and is particularly strong in the twentieth century.

No in-depth study of meridional motions of field features at the core surface has yet been carried out. Such a study would be of interest, especially considering the possible links between meridional motions and proposed reversal mechanisms (Gubbins, 1987; Wicht and Olson, 2004).

5.05.5 Interpretation in Terms of Core Processes

The observed evolution of the internally generated part of Earth's magnetic field is a consequence of the motions in the liquid metal outer core. In order to understand and model the mechanisms underlying these changes, we must employ the mathematical framework of magnetohydrodynamics – the marriage of Maxwell's laws of electromagnetism and the principles of hydrodynamics or fluid mechanics. In this section, equations describing the evolution of the core magnetic field and the generation of core fluid motions are derived, and useful approximations are discussed; we stop short of describing attempts to invert field observations for core fluid motions at the CMB, which is the territory of Volume 8.

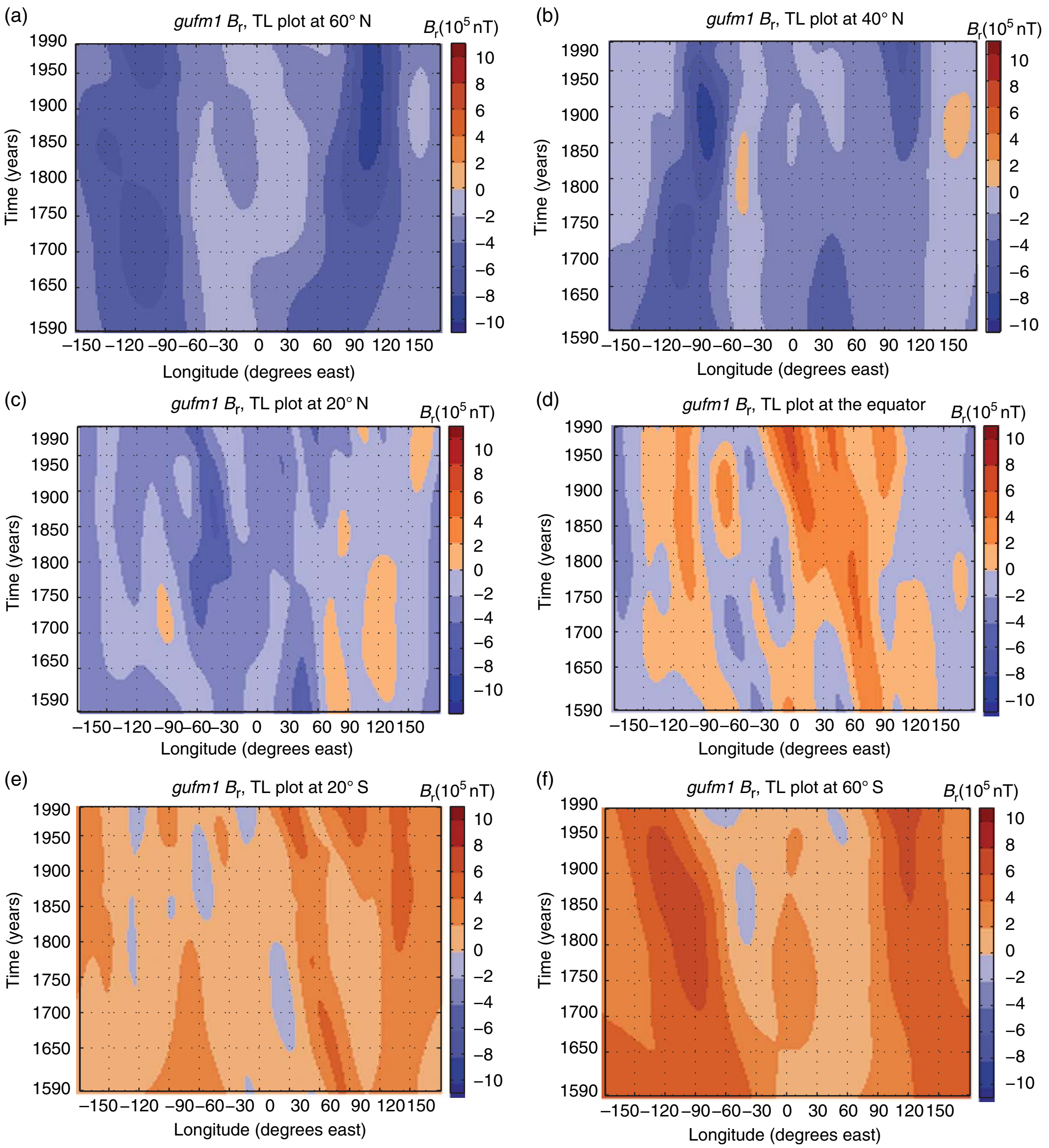

Figure 32 Time–longitude (TL) plots of the unfiltered vertical magnetic field at the core surface from the field model *gufm1* at latitudes 60° N in (a), at 40° N in (b), at 20° N in (c), at the equator in (d), at 20° S in (e), and at 60° S in (f).

5.05.5.1 Maxwell's Equations and Moving Frames

Maxwell's equations for an electrically conducting fluid moving with a velocity $\mathbf{u}$ in the presence of a magnetic field $\mathbf{B}$, an electric field $\mathbf{E}$, and an electric current density $\mathbf{J}$ are

$$\nabla \cdot \mathbf{B} = 0 \quad \text{Absence of free magnetic monopoles} \quad [11]$$

$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} \quad \text{Faraday's law of magnetic induction} \quad [12]$$

$$\frac{1}{\mu_0}(\nabla \times \mathbf{B}) = \mathbf{J} \quad \text{Ampere's law of magnetostatics} \quad [13]$$

where μ_0 is the magnetic permeability of free space which is applicable to nonferromagnetic fluids. It should be noticed that these equations are somewhat simpler than the usual, most general form of Maxwell's equations described in, for example,

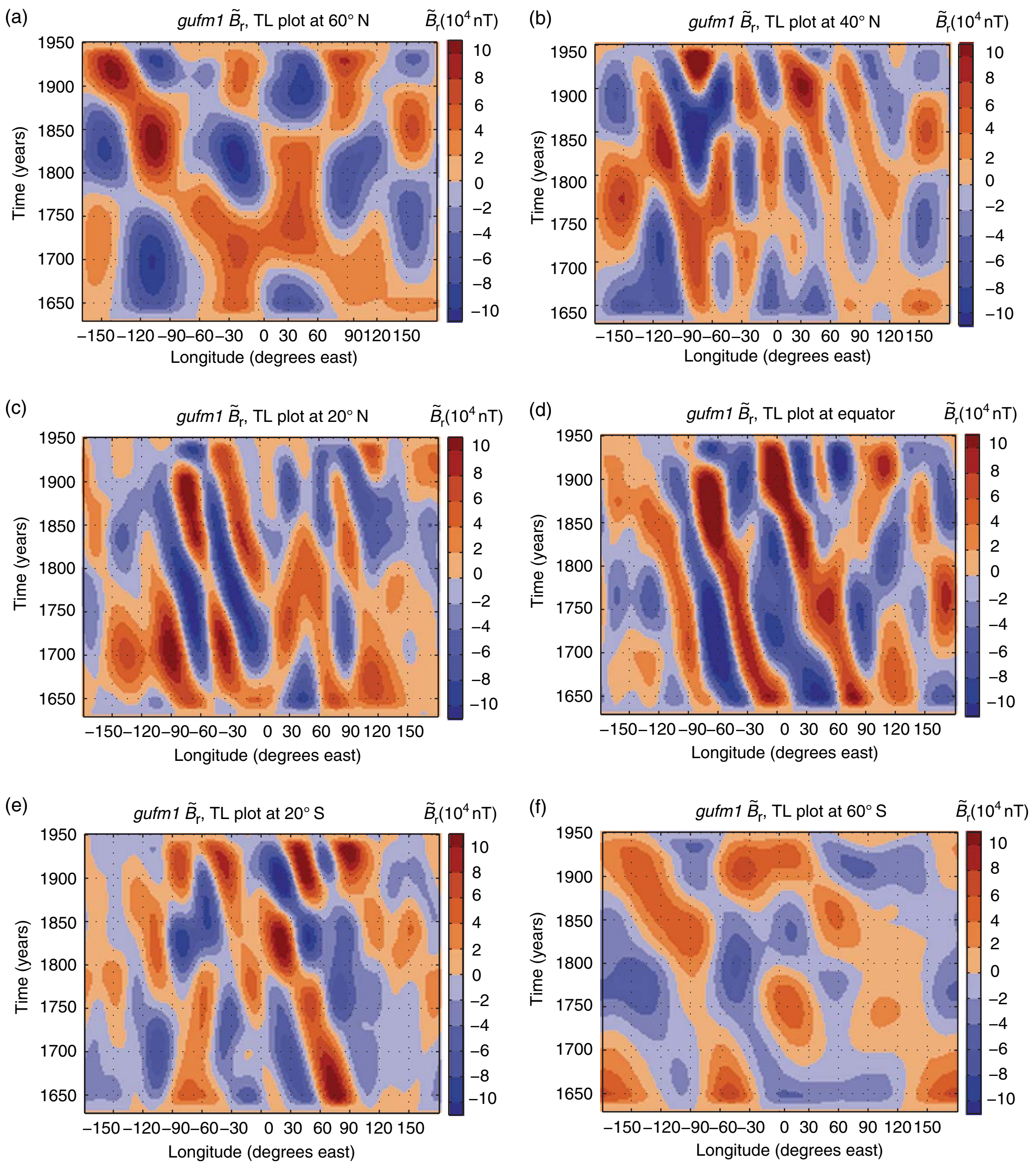

Figure 33 TL plots of the processed vertical magnetic field at the core surface with time-averaged axisymmetric component subtracted and high-pass-filtered with cutoff period 400 years, from the field model *gufm1* at latitudes 60° N in (a), at 40° N in (b), at 20° N in (c), at the equator in (d), at 20° S in (e), and at 60° S in (f).

Jackson (1999) or Backus *et al.* (1996). The fact that the liquid metal flows we are interested in have speeds $|\mathbf{u}| \ll c$ (the speed of light) has enabled the well-known displacement current term in Ampere's law to be neglected and allowed the (decoupled) Gauss law of electrostatics to be dispensed with. This powerful simplification is known as the magnetohydrodynamic or MHD approximation. In this scenario, Ohm's law for the electrically conducting and moving fluid takes the form,

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \qquad [14]$$

where σ is the electrical conductivity of the fluid. The mathematical formalism can be further

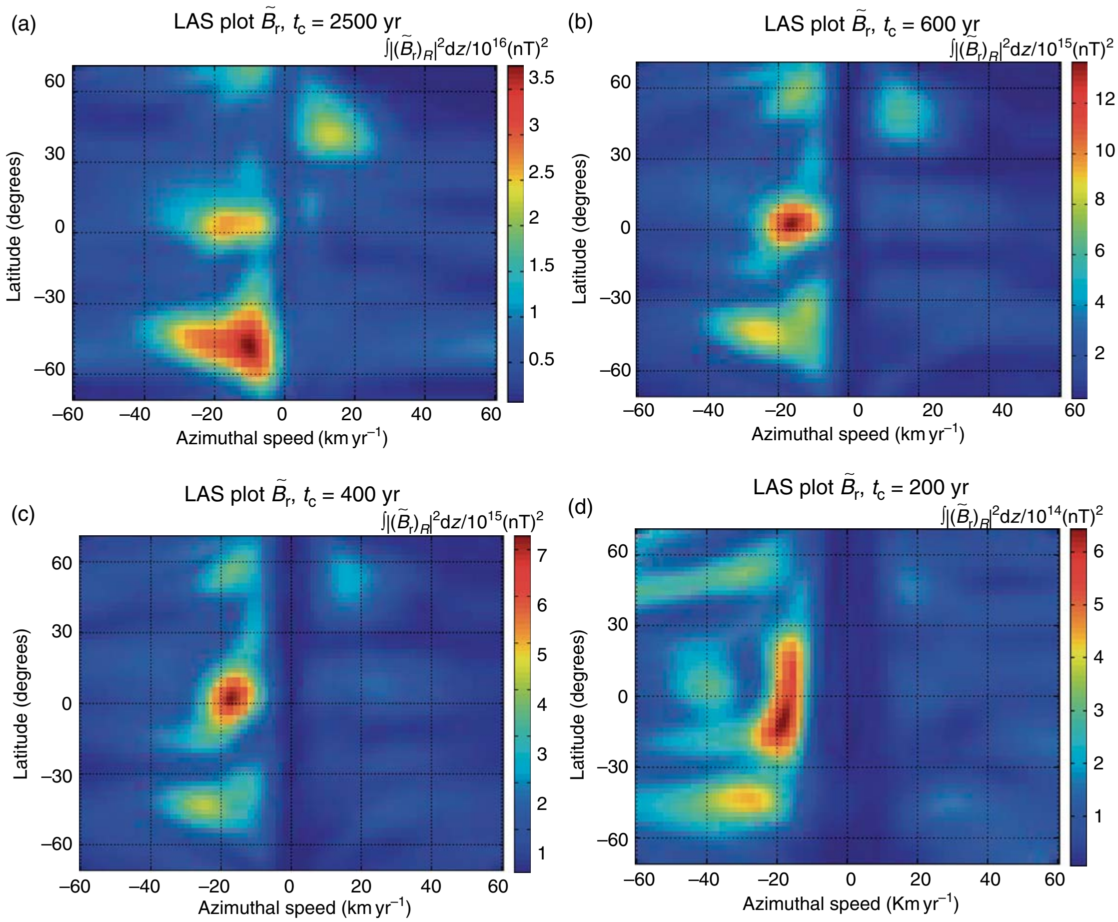

Figure 34 Summing the power traveling at different angles in TL plots using a Radon transform method, LAS power plots are constructed for the high-pass-filtered vertical magnetic field at the core surface from *gufm1*. (a) Result when the high-pass filter threshold is 2500 years, (b) 600 years, (c) 400 years (the case shown for the TL plots in **Figure 33**), and (d) 200 years.

compacted by realising that eqns [11]–[14] can be combined to yield a single prognostic equation governing the evolution of magnetic fields. Substituting from eqn [13] into eqn [14],

$$\nabla \times \mathbf{B} = \mu_0\sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \qquad [15]$$

Taking the curl of this,

$$\nabla \times (\nabla \times \mathbf{B}) = \mu_0\sigma(\nabla \times \mathbf{E} + \nabla \times (\mathbf{u} \times \mathbf{B})) \qquad [16]$$

Substituting from eqn [12] and using the vector identity that $\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2\mathbf{B}$, this becomes

$$\nabla(\nabla \cdot \mathbf{B}) - \nabla^2\mathbf{B} = \mu_0\sigma\left(-\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{u} \times \mathbf{B})\right) \qquad [17]$$

Using eqn [11], rearranging and defining the magnetic diffusivity $\eta = 1/(\mu_0\sigma)$ yields

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta\nabla^2\mathbf{B} \qquad [18]$$

This is known as the magnetic induction equation. The first term on the right-hand side represents the magnetic induction effects whereby electrical currents are generated when an electrical conductor moves through a magnetic field. The second term represents the changes in magnetic fields due to dissipative Joule heating effects that are the consequence of the flow of electric currents in a material with finite resistivity.

5.05.5.2 The Induction Equation in a Spherical Earth

Focusing our attention on the MHD of Earth's core, with a view to interpreting secular variation processes, in this section we discuss the induction equation at the core–mantle boundary. In particular, we discuss the approximations commonly made and the consequences of the present uncertainty regarding the electrical conductivity of the lower mantle.

We begin with the induction equation [18] governing changes with time in the magnetic field due to the effects of magnetoadvection and magnetic diffusion. We will assume for the moment that the flow $\mathbf{u}$ is given and examine its effects on the field, the so-called 'kinematic problem;' the question of the dynamics will be examined below. It is useful to introduce the toroidal–poloidal, or Mie (Backus, 1986; Backus *et al.*, 1996), representation:

$$\mathbf{B} = \mathbf{B}_{\mathrm{T}} + \mathbf{B}_{\mathrm{P}} = \nabla \times (T\hat{\mathbf{r}}) + \nabla \times \nabla \times (P\hat{\mathbf{r}}) \quad [19]$$

where $T(r, \theta, \phi)$ and $P(r, \theta, \phi)$ are the toroidal and poloidal scalars, respectively, defining the toroidal $\mathbf{B}_{\mathrm{T}}$ and poloidal $\mathbf{B}_{\mathrm{P}}$ ingredients of $\mathbf{B}$. This representation is valid for any solenoidal vector field satisfying

$$\nabla \cdot \mathbf{B} = 0 \quad [20]$$

The sphericity of the core suggests that we continue to work in spherical polar coordinates, though note that it is the case that the core is ellipsoidal with equatorial radius greater by approximately 1 part in 400 than the polar radius. In terms of approximating the surface for the purposes of plotting fields or fluid motion, the ignorance of the oblate spheroidal nature of the core can be seen to introduce negligible errors; note, however, that the ellipticity might be important in a dynamical context, because of the coupling that it can generate between mantle and core. In spherical polar coordinates (r, θ, ϕ), the toroidal and poloidal ingredients may be written in terms of the toroidal and poloidal scalars in the form,

$$\mathbf{B}_{\mathrm{T}} = \left(0, \frac{1}{r\sin\theta}\frac{\partial T}{\partial \phi}, -\frac{1}{r}\frac{\partial T}{\partial \theta}\right) \quad [21]$$

and

$$\mathbf{B}_{\mathrm{P}} = \left(\frac{\mathcal{L}^2 P}{r^2}, \frac{1}{r}\frac{\partial^2 P}{\partial r \partial \theta}, \frac{1}{r\sin\theta}\frac{\partial^2 P}{\partial r \partial \phi}\right) \quad [22]$$

where $\mathcal{L}$ is the angular momentum operator of quantum mechanics, defined by

$$\mathcal{L}^2 = -\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\} \quad [23]$$

We can use the same representation for the fluid flow $\mathbf{u}$, if we assume incompressibility. While this is not strictly true (the density increase with pressure across the outer core is approximately 20%), it is a fair approximation.

We now briefly consider the consequences of a mantle with finite conductivity for the spherical induction equation (most studies take no account of the conductivity of the mantle and treat it as a simple insulator). If the mantle has a spherically symmetric distribution of electrical conductivity which is of moderate amplitude, then an insulating mantle is likely to be a reasonable approximation. However, at the present time very little is known about the three-dimensional heterogeneity in conductivity, and in the deep mantle even the spherically symmetric part is poorly known. Thus, the insulating mantle assumption may be in considerable error and it is worth dwelling for a moment on whether this leads to difficulties in interpreting secular variation. In a conducting mantle, revise [18] (having first set $\mathbf{u}$ to zero) to account for the fact that η in the mantle may be variable; this gives the diffusion equation for a heterogeneous η:

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\eta \nabla \times \mathbf{B}) \quad [24]$$

When we analyze its toroidal and poloidal ingredients, we find that in general there can be poloidal-to-toroidal conversion and vice versa; however, in the case that there is spherical symmetry ($\eta = \eta(r)$), a remarkable simplification occurs: the toroidal and poloidal ingredients of $\mathbf{B}$ obey

$$\frac{\partial T}{\partial t} = \eta \nabla^2 T + \frac{1}{r}\frac{\mathrm{d}\eta}{\mathrm{d}r}\frac{\partial}{\partial r}(rT); \; \frac{\partial P}{\partial t} = \eta \nabla^2 P \quad [25]$$

There is thus complete decoupling between the poloidal and toroidal ingredients of the field, even when the mantle conductivity is radially varying. Inasmuch as the spherically symmetric assumption is valid, this simplifies the task of understanding the secular variation.

At the Earth's surface we have electrically insulating boundary conditions and thus the toroidal magnetic field must vanish; this arises because Ampere's law relates fields $\mathbf{B}$ to currents $\mathbf{J}$ thus:

$$\begin{aligned} \mu_0 \mathbf{J} &= \nabla \times \mathbf{B} = \nabla \times \mathbf{B}_{\mathrm{T}} + \nabla \times \mathbf{B}_{\mathrm{P}} \\ &= \nabla \times \nabla \times (T\hat{\mathbf{r}}) + \nabla \times \nabla \times \nabla \times (P\hat{\mathbf{r}}) \\ &= \nabla \times \nabla \times (T\hat{\mathbf{r}}) + \nabla \times \left[\left(-\nabla^2 P\right)\hat{\mathbf{r}}\right] \end{aligned} \quad [26]$$

Comparing this expression for the current with our definition of the toroidal–poloidal decomposition, it is clear that poloidal field results from toroidal currents, and toroidal field from poloidal currents. If $\mathbf{J} = 0$, as in an insulator, then $\mathbf{B}_{\mathrm{T}} = 0$, while $\mathbf{B}_{\mathrm{P}}$ does not necessarily vanish, although P must satisfy $\nabla^2 P = 0$ (in which case we call it a potential field).

In principle, we can find the toroidal field within the mantle from measurements of the electric field at the sea bottom (e.g., Runcorn, 1955; Lanzerotti *et al.*, 1993; Shimizu *et al.*, 1998; Shimizu and Utada, 2004), but the electric field from the toroidal field from the core is likely to be small compared with the field from other sources, especially that from ocean currents.

We are therefore only able to monitor the poloidal part of the magnetic field at the Earth's surface. Turning now to the CMB, it is extremely unlikely that the toroidal magnetic field vanishes there on account of the finite (and possibly large) conductivity there. Several factors come to our rescue to ameliorate what otherwise would seem like a hopeless situation. We temporarily assume that the CMB is a free-slip boundary, so that $\mathbf{u} \cdot \hat{\mathbf{r}} = 0$, but $\mathbf{u} = \mathbf{u}_\mathrm{h} \neq 0$). Then, following Bullard and Gellman (1954), we can show that

$$[\nabla \times (\mathbf{u}_\mathrm{h} \times \mathbf{B})]_\mathrm{P} = [\nabla \times (\mathbf{u}_\mathrm{h} \times \mathbf{B}_\mathrm{P})]_\mathrm{P} \qquad [27]$$

from which we obtain the poloidal induction equation at the CMB,

$$\frac{\partial \mathbf{B}_\mathrm{P}}{\partial t} = [\nabla \times (\mathbf{u}_\mathrm{h} \times \mathbf{B}_\mathrm{P})]_\mathrm{P} + \eta \nabla^2 \mathbf{B}_\mathrm{P} \qquad [28]$$

In the above, $[\ ]_P$ stands for the poloidal part of the equation. We discover, somewhat counterintuitively, that the poloidal secular variation depends only on the poloidal magnetic field! This result hangs entirely on the fact that radial motions are zero at the CMB. The result would not be true elsewhere in the core, where radial motions are crucial for the production of poloidal field from toroidal magnetic field, in order for the dynamo to operate. Note also that the toroidal part of the induction equation does not separate so easily, and that the rate of change of toroidal field depends on both the toroidal field and the poloidal field. Horizontal flow in this case shears both poloidal and toroidal magnetic field in order to create toroidal secular variation.

It is a fortuitous fact that when one analyzes the different components of the poloidal induction equation, the radial part has a particularly simple form, and gives an equation which will be central to much of our discussion. The radial induction equation reads

$$\partial_t B_\mathrm{r} + \mathbf{u}_\mathrm{h} \cdot \nabla_\mathrm{h} B_\mathrm{r} + B_\mathrm{r} \nabla_\mathrm{h} \cdot \mathbf{u}_\mathrm{h} = \frac{\eta}{r} \nabla^2 (r B_\mathrm{r}) \qquad [29]$$

demonstrating that only radial fields, and their derivatives, need to be known for radial secular variation to be calculable when a particular flow is prescribed.

It is incumbent on us to realise the shortcomings in the above analysis. We applied the condition $\mathbf{u} \cdot \hat{\mathbf{r}} = 0$ rather than the true nonslip condition $\mathbf{u} = 0$ at the CMB. In reality, there is a boundary layer over which the flow adjusts to the nonslip boundary condition, and we really apply the induction equation at the top of the free stream, the bottom of the boundary layer. We need to know the difference in the values of $\mathbf{B}$ across this boundary layer, denoted $[\mathbf{B}]$.

Various analyses of the boundary layer have been carried out, and no consensus has been reached (e.g., Backus, 1968; Hide and Stewartson, 1973; Jault and Le Mouël, 1991). The important issue for our purposes is that the radial derivatives in all three components of $\mathbf{B}$ are expected to be much bigger than the horizontal derivatives. When one uses this fact along with the divergence-free constraint on the field, one finds

$$\frac{\partial B_\mathrm{r}}{\partial r} + \nabla_\mathrm{h} \cdot \mathbf{B}_\mathrm{h} \sim \frac{\partial B_\mathrm{r}}{\partial r} = 0 \qquad [30]$$

This leads to the conclusion that $[B_\mathrm{r}] = 0$, so maps of the radial component of magnetic field immediately above the CMB also represent the radial component of magnetic field at the top of the free stream. The same cannot be said to be true for the horizontal components, but we omit a discussion on the possible jumps in $\mathbf{B}_h$ mainly because the induction equation for $\mathbf{B}_h$ involves the toroidal field, about which we have no direct knowledge.

5.05.5.3 The Navier–Stokes Equation

Moving on to consider the fluid dynamics of the core, we adopt the hypothesis that on macroscopic length scales core fluid can be well approximated as a continuum (see, e.g., Batchelor, 1967), suppose that it is to first approximation incompressible, obeys Newtonian laws of viscosity, and is uniformly rotating. Then, in a frame of reference rotating with the fluid, the conservation of momentum is encapsulated in the Navier–Stokes equation, which under the Boussinesq approximation (e.g., Gubbins and Roberts, 1987) reads

$$\rho_0 \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + 2\boldsymbol{\Omega} \times \mathbf{u} \right) = -\nabla p + \rho' \mathbf{g} + \mathbf{J} \times \mathbf{B} + \rho_0 \nu \nabla^2 \mathbf{u} \qquad [31]$$

where ρ_0 and ρ' are the hydrostatic density and departure from hydrostatic density, respectively, $\boldsymbol{\Omega}$ is the Earth's rotation vector, p is the nonhydrostatic part of the pressure, $\mathbf{g}$ the acceleration due to gravity, ν the kinematic viscosity, and $\mathbf{J}$ the current density. The Boussinesq approximation is a simplification frequently adopted for the core, and ignores variations in density except those that are responsible for thermal buoyancy through the term $\rho'\mathbf{g}$; its applicability to the core is under current scrutiny in the field of numerical simulation of core dynamics, since the compressibility of the core does cause a change of approximately 20% in the core density between the inner- and outer-core boundaries. It suffices as an approximation for our purposes, as we shall predominantly be limiting our discussion to the surface of the core.

Considerable simplification can be made if one analyzes the likely sizes of the terms in the equation, concentrating on the flow in the main body of the core (outside the boundary layers). First, the Rossby number,

$$R_o = \frac{U}{\Omega L} \simeq 4 \times 10^{-6} \qquad [32]$$

compares the nonlinear advective term on the left-hand side of [31] with the Coriolis term. We take $L \sim 3 \times 10^6$ m as a characteristic length scale for the core. Then the estimate above is based on values for U (roughly half a millimeter per second) gleaned from analysis of the secular variation (e.g., Volume 8 and Section 5.05.4.2.3), and hence (as in much of our analysis) there is a slight sense of circularity. Similarly, the Ekman number (the ratio of viscous forces to the Coriolis force) is given by

$$E = \frac{\nu}{\Omega L^2} \qquad [33]$$

where Ω is the rotation rate. If we take L as before and $\nu \sim 10^{-6}\,\mathrm{m\,s^{-2}}$, we find the classic value of $E \sim 10^{-15}$, indicating that viscous effects are negligible in the main body of the core if a laminar value for ν is adopted. The inertial term is somewhat more difficult – there is a mode of oscillation in the core that can occur on decade timescales, the so-called torsional oscillation (*see* Chapter 8.10) which may not be negligible. It is easily excited and it is inappropriate to compare it to the Coriolis force because it is unaffected by it. We will neglect the inertial term on the grounds that it is only significant when the period approaches the diurnal period, except in the force balance when averaged over cylinders coaxial with the rotation axis. This leads us to a very useful approximation in core studies, the so-called magnetostrophic approximation,

$$\rho_0(2\boldsymbol{\Omega} \times \mathbf{u}) = -\nabla p + \rho'\mathbf{g} + \mathbf{J} \times \mathbf{B} \qquad [34]$$

The difficult term in this equation is the last term on the right-hand- side, the Lorentz force $\mathbf{L}$. We write it in the form,

$$\begin{aligned} \mathbf{L} = \mathbf{J} \times \mathbf{B} &= \frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B} \\ &= \frac{1}{\mu_0}\left[-\frac{1}{2}\nabla B^2 + (\mathbf{B}\cdot\nabla)\mathbf{B}\right] \end{aligned} \qquad [35]$$

An approximation called the tangential geostrophy approximation, proposed independently by Hills (1979) and Le Mouël (1984), would neglect the horizontal components of this term when compared to all others; to aid our development, we write the horizontal and radial components of $\mathbf{L}$ as $\mathbf{L}_\mathrm{h}$ and $\mathbf{L}_\mathrm{r}$, respectively. For tangential geostrophy, we require that

$$M = \frac{|\mathbf{L}_\mathrm{h}|}{2\rho_0\Omega U} = \frac{|(\mathbf{B}\cdot\nabla)\mathbf{B}_\mathrm{h}|}{2\mu_0\rho_0\Omega U} \ll 1 \qquad [36]$$

There are two contributions to $(\mathbf{B}\cdot\nabla)\mathbf{B}_\mathrm{h}$:

$$B_\mathrm{r}\frac{\partial \mathbf{B}_\mathrm{h}}{\partial r} \quad \text{and} \quad (\mathbf{B}_\mathrm{h}\cdot\nabla_\mathrm{h})\mathbf{B}_\mathrm{h} \qquad [37]$$

We need to estimate $|\mathbf{B_h}|$ and $|\partial\mathbf{B}_\mathrm{h}/\partial_r|$. Le Mouël (1984) argues that if the toroidal field is small at the CMB (if the mantle is a perfect insulator it must vanish) and its radial gradient is small, then these terms are of order B_P^2/L, where B_P is the size of the poloidal field at the CMB ($\simeq 5 \times 10^{-4}$ T), giving $M \simeq 10^{-3}$.

If we adopt this approximation, we have

$$2\rho_0(\boldsymbol{\Omega} \times \mathbf{u}) = -\nabla p + \rho'\mathbf{g} + \mathbf{L}_\mathrm{r} \qquad [38]$$

Curling this equation, we obtain

$$2\rho_0(\boldsymbol{\Omega}\cdot\nabla)\mathbf{u} = \mathbf{g} \times \nabla\rho' + \nabla \times \mathbf{L}_\mathrm{r} \qquad [39]$$

which, in the case that $\mathbf{L}_\mathrm{r} = 0$, is the thermal wind equation, of great importance in meteorology. The radial component gives the so-called tangentially geostrophic constraint,

$$\nabla_\mathrm{h}\cdot(\mathbf{u}_\mathrm{h}\cos\theta) = 0 \qquad [40]$$

which remains true regardless of whether $\mathbf{L}_\mathrm{r} \neq 0$. Given these assumptions, we therefore have a strong constraint on the types of allowed fluid motions at the CMB. The interested reader can see in Volume 8 that output from self-consistent geodynamo simulations tend to suggest that the tangential geostrophy approximation is reasonably well obeyed. The significance of the constraint is that it vastly reduces the types of allowable fluid motions when the inverse problem for $\mathbf{u}$ is solved (for further details, *see* Chapter 8.04).

5.05.5.4 The Frozen Flux Approximation

5.05.5.4.1 The neglect of magnetic diffusion and its physical consequences

Interpretation of secular variation using the induction equation [18] is often simplified by neglecting the contribution of magnetic diffusion. In the limit of a perfectly electrically conducting fluid (zero magnetic diffusivity), the induction equation becomes

$$(\partial_t + \mathbf{u}\cdot\nabla)\mathbf{B} = (\mathbf{B}\cdot\nabla)\mathbf{u} \qquad [41]$$

The left-hand side of this equation is the advective derivative describing how the magnetic field changes as one moves along with the fluid, while the right-hand side tells us that such changes occur through the stretching of the magnetic field by fluid motions.

Further intuition follows if we think about how a velocity field $\mathbf{u}$ would advect a material line element d$\boldsymbol{l}$. We imagine the line element being drawn in the fluid at some instant and subsequently moved along and stretched by the fluid motions. The total rate of change of d$\boldsymbol{l}$ is then $\mathbf{u}(\boldsymbol{r}+\mathrm{d}\boldsymbol{l})-\mathbf{u}(\boldsymbol{r})$, where $\boldsymbol{r}$ and $\boldsymbol{r}+\mathrm{d}\boldsymbol{l}$ are position vectors at the two ends of d$\boldsymbol{l}$. The equation describing the evolution of d$\boldsymbol{l}$ therefore has the form,

$$(\partial_t + \mathbf{u}\cdot\nabla)\mathrm{d}\boldsymbol{l} = \mathbf{u}(\mathbf{r}+\mathrm{d}\boldsymbol{l}) - \mathbf{u}(\mathbf{r}) = (\mathrm{d}\boldsymbol{l}\cdot\nabla)\mathbf{u} \qquad [42]$$

Inspection of eqns [41] and [42] reveals they have precisely the same form. This simple example demonstrates that because magnetic fields evolve in an identical manner to material line elements in a fluid if magnetic diffusion is negligible, a field line found on a particular fluid element at some initial instant must continue to lie on that element at all subsequent times. The magnetic field effectively appears to be frozen into the fluid as it moves. This result is known as Alfvén's theorem (part I) after Hannes Alfvén who first derived it; we shall discuss the second part of the theorem in a moment. Neglect of magnetic diffusion in the induction equation and its consequences are most commonly referred to as the frozen flux approximation and we shall use the latter terminology.

Another important property that results from assuming that a fluid is a perfect electrical conductor can be demonstrated by returning to Faraday's law of magnetic induction (eqn [12]). Earlier, this was stated in its differential form. The integral form when applied to material curves of an electrically conducting fluid in motion (see, e.g., Davidson, 2001) takes the form,

$$\oint_C \mathbf{E}'\cdot\mathrm{d}\boldsymbol{l} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_S \mathbf{B}\cdot\mathrm{d}\boldsymbol{S} \qquad [43]$$

where $\mathbf{E}'=\mathbf{E}+\mathbf{u}\times\mathbf{B}$ is the total electric field in a reference frame moving along with d$\boldsymbol{l}$ at velocity $\mathbf{u}$, C is a closed material curve composed of line elements d$\boldsymbol{l}$, and S is any surface that spans C. Now, from Ohm's law, $\mathbf{J}=\sigma\mathbf{E}'$, so

$$\frac{1}{\sigma}\oint_C \mathbf{J}\cdot\mathrm{d}\boldsymbol{l} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_S \mathbf{B}\cdot\mathrm{d}\boldsymbol{S} \qquad [44]$$

but under the frozen flux approximation $\sigma\rightarrow\infty$; therefore,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_S \mathbf{B}\cdot\mathrm{d}\boldsymbol{S} = 0 \qquad [45]$$

In a perfect electrical conductor, the integrated magnetic field (or magnetic flux) through any material surface is thus always preserved. This is known as Alfvén's theorem (part II).

5.05.5.4.2 Application of the frozen flux hypothesis to the generation of secular variation at the core surface

Roberts and Scott (1965) were the first to suggest that the frozen flux approximation could be applied to the problem of modeling secular variation. Many authors (see, e.g., Backus *et al.*, 1996) refer to this as the frozen flux hypothesis. Roberts and Scott argued that there are two distinct timescales associated with the induction equation. Considering a length scale $\mathcal{L}_B$ over which the magnetic field changes, a characteristic flow speed $\mathcal{U}$, and the magnetic diffusivity η, these timescales are defined as

$$\tau_{\mathrm{adv}} = \frac{\mathcal{L}_B}{\mathcal{U}} \quad \text{Advection time scale} \qquad [46]$$

and

$$\tau_{\mathrm{dif}} = \frac{\mathcal{L}_B^2}{\eta} \quad \text{Magnetic diffusion time scale} \qquad [47]$$

The ratio of these timescales $R_{\mathrm{m}} = \tau_{\mathrm{dif}}/\tau_{\mathrm{adv}} = \mathcal{U}\mathcal{L}_B/\eta$ is known as the magnetic Reynolds number and gives a crude measure of the relative strength of advection to magnetic diffusion. Taking estimates of $\mathcal{L}_B = 10^6$ m (the approximate scale of the outer-core container and of large-scale features in Earth's magnetic field), $\mathcal{U} = 5\times10^{-4}\,\mathrm{ms}^{-1}$ (speed of observed westward field motions, thought to be caused by core flow), and $\eta = 2\,\mathrm{m}^2\,\mathrm{s}^{-1}$ (from estimates of liquid iron electrical conductivity at core pressures and temperatures (Braginsky and Roberts, 1995)) gives estimates of $\tau_{\mathrm{adv}} = 65$ yr and $\tau_{\mathrm{dif}} = 1.6\times10^4$ yr in the core (Roberts and Glatzmaier, 2000). On this basis, Roberts and Scott suggested that making the frozen flux assumption was a reasonable (though imperfect) approximation when modeling the core motions causing large-scale secular variation. This theoretical argument has been the subject of much comment and debate over the past 40 years; we will return later to the question of its validity. First, we will give details on its formal consequences and review attempts to determine whether these are compatible with observed secular variation.

5.05.5.4.3 Consequences of frozen flux approximation at the core surface

Backus (1968) described the conditions for the main field morphology and secular variation to be consistent with a frozen flux theory of core motions. Neglecting magnetic diffusion, he showed that the radial part of the induction equation reduces to (c.f. eqn [29])

$$\partial_t B_r + \nabla_h \cdot (\mathbf{u}_h B_r) = 0 \qquad [48]$$

He then deduced that this implies a set of conditions on null-flux points and curves (where $B_r = 0$). The most important of these are

$$\int_S \partial_t B_r \mathrm{d}S = 0 \qquad [49]$$

where S is surface bounded by a null-flux curve C and

$$\partial_t B_r = 0 \qquad [50]$$

where two null-flux curves C_1 and C_2 intersect. From the first condition follows a condition closely related to eqn [45], which states that the integrated radial magnetic field through a null flux curve C (an example of a material curve) is preserved. The simple derivation is as follows. We know that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_S B_r \mathrm{d}S = \int_S \partial_t B_r \mathrm{d}S + \int_C B_r u_n \mathrm{d}l = 0 \qquad [51]$$

where $\mathrm{d}l$ is a line element along the null flux curve C and u_n is the normal component of its velocity. Using eqn [49], the first term on the right-hand side is zero and, since C is a null-flux curve, the second term is also zero, giving

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_S B_r \mathrm{d}S = 0 \qquad [52]$$

From eqn [52], it further follows that the sum of the unsigned flux over all null-flux curves must also be conserved,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{S'} |B_r| \mathrm{d}S = 0 \qquad [53]$$

with the integration now over the entire core surface S'. It should be noted that eqn [53] is a weaker constraint than eqn [52] on the signed flux through individual flux patches because contributions from small patches will be swamped by those from the larger Northern and Southern Hemisphere patches. The unsigned flux condition will therefore only be violated if there are large amounts of magnetic diffusion occurring on a global scale; it could still be approximately obeyed even if magnetic diffusion was occurring locally. On the other hand, condition [53] is less likely to be adversely affected by errors in the field models; so, it is a reasonable and observational test of the global applicability of the frozen flux approximation.

Backus' results additionally require that the topology of the field must be invariant so that null-flux curves cannot split or coalesce. In practice, this is a rather difficult condition to satisfy as it requires only a small amount of diffusion in order to be violated. It therefore seems unlikely that this condition would be satisfied by the magnetic field at the core surface, where frozen flux is at best a useful approximation, so we shall not discuss this condition further.

5.05.5.4.4 Attempts to test the frozen flux approximation using geomagnetic observations

The conditions described in the previous section (which will be referred to collectively as the Backus conditions) have enabled workers to test the validity of the frozen flux approximation using models of the main field and secular variation constructed from geomagnetic observations. These tests fall into two main categories: (1) attempts to estimate whether the Backus conditions are violated in field models constructed without any constraints on field evolution; (2) attempts to build field models that are constrained to obey the Backus conditions but that also satisfactorily fit the observations.

The first category involves investigating whether observations are sufficiently accurate to observe diffusion. Booker (1969) was the first to investigate this issue. He attempted to estimate the integral of $\partial_t B_r$ through the null flux curves represented by the magnetic equator and found that the dipole part of the time derivative was nonzero, possibly violating eqn [49]. However, the field models he used lacked the necessary resolution at high spherical harmonic degrees to definitively calculate the full secular variation integral and he was thus unable to definitely identify the presence of diffusion. Gubbins (1983) repeated the calculation and arrived at a similar conclusion, again being limited by the accuracy of observations.

Hide and Malin (1981) and later Voorhies and Benton (1982) used the criteria for invariance of the unsigned flux at the core surface (eqn [53]) as a method for calculating the outer-core radius and found values that agreed well with the seismologically determined value. These results indicated that the unsigned flux is not changing rapidly at the core surface and was evidence in support in the application of the frozen flux approximation on large length scales. Benton and Voorhies (1987) later extended these analyses to show that there had been little change in the unsigned flux over the interval 1945–85. Bloxham *et al.* (1989) considered the evolution of the unsigned flux over a much longer interval from 1715 to 1980 and found that it remained approximately constant over that interval. They concluded that this global requirement for frozen flux was satisfied by their field models but pointed out that changes in the flux of smaller null-flux patches would not be seen by this method. Similar results for the unsigned flux were recently obtained by Holme and Olsen (2006) using a high-quality, satellite-derived field model (CO2003).

The second question of possible changes in the flux through individual null flux curves was addressed in a series of papers by Bloxham and Gubbins in the mid- to late 1980s including Bloxham and Gubbins (1985), Gubbins and Bloxham (1985), Bloxham and Gubbins (1986), Bloxham (1986a), and in the landmark study of Bloxham *et al.* (1989) where the definitive results were reported. In the latter paper, a sequence of single epoch, regularized, field models spanning 1715–1980 were studied and changes of the flux though null-flux curves at the core surface were calculated. Major flux changes through some curves were found; in particular, a patch that moved from the Indian Ocean to southern Africa was found to increase its flux dramatically. Bloxham and Gubbins regarded this as conclusive evidence for the violation of frozen flux; however, others have expressed doubts over the rigor of their arguments. In order for frozen flux to have been demonstrably violated, the changes in the flux must be larger than possible changes in the flux due to errors in the field models. Backus (1988) has argued that the error estimates of Gubbins and Bloxham (1985) and Bloxham and Gubbins (1986) are rather optimistic. He suggests that their models do not fully solve the uniqueness problem because the damping (regularization) parameter is arbitrarily chosen; their error estimates depend crucially on the value of this parameter and are likely to actually be much larger than those quoted. Furthermore, O'Brien (1996) has demonstrated that even the number of null-flux patches in field models based on excellent satellite data is difficult to definitively determine. It therefore seems that inferences based on the changes in flux through individual null-flux patches in the past (when there were significant variations in the distribution and accuracy of observations) should be viewed with a healthy degree of skepticism.

The compatibility of observations with the Backus conditions for frozen flux has recently become apparent thanks to the construction of field models that are constrained to obey these conditions. Gubbins (1984) used a Lagrangian constraint approach to enforce zero radial secular variation through null-flux curves (eqn [49]) and found that it was possible to do this over the interval 1959–74. Bloxham and Gubbins (1986) and Bloxham *et al.* (1989) produced field models that satisfied the condition that the flux of B_r through null-flux curves be conserved (eqn [52]). They implemented this by imposing an additional penalty during the inversion process; models with flux integrals differing from a predefined reference model were heavily penalized. They reported that it was possible to find models that satisfied these flux constraints, and that their imposition often improved field models where the data quality was poor. However, they also reported that the constrained models had a slightly higher misfit than the unconstrained models, using this to support their contention that frozen flux was in fact violated. This argument is subject to similar caveats regarding errors in field models as the argument concerning changes in the flux integrals in unconstrained models. Benton *et al.* (1987) constructed field models with constrained flux through null-flux curves covering the interval 1977.5–1982.5 and were able to demonstrate that the secular variation predicted by these models for times outside the span of input data was an improvement on the predictions of unconstrained models. They took this to be evidence in favor of the frozen flux hypothesis.

Most recently, Constable *et al.* (1993), O'Brien *et al.* (1997), and Jackson *et al.* (2007) have attacked the same problem but used a different parametrization of the field based on a spherical triangle tessellation at the core surface. Constable *et al.*

(1993) constructed field models satisfying the flux constraints of eqn [52] for the epochs 1945.5 and 1980.0, O'Brien *et al.* (1997) did the same for the epochs 1915 and 1980, while Jackson *et al.* (2007) managed to do so for the epochs 1882, 1915, 1945, 1980, and 2000. In all cases, it was found that reasonable misfit levels to the observations could be achieved, and it was noted that in order to reject the frozen flux hypothesis it would be necessary to demonstrate that no such models could be found.

We remark here that the published time-dependent field models of Bloxham and Jackson (1992) and Jackson *et al.* (2000) both show an obvious growth in the intensity of reversed flux patches in the Southern Hemisphere in the twentieth century. Gubbins (1987) has pointed out that a simple and physically appealing explanation of this phenomenon is the expulsion of toroidal flux by upwelling core fluid (Allan and Bullard, 1966; Bloxham, 1986a; Drew, 1993). It is therefore unfortunate that the present field observations seem incapable of constraining field models sufficiently to discriminate between this mechanism and a mechanism involving frozen flux advection.

Although our discussion has focused on the radial magnetic field and its secular variation, a limited amount of work has been carried out on whether the horizontal components of the magnetic field are consistent with the frozen flux hypothesis. Assuming continuity of the horizontal magnetic field components across the magnetic field boundary layer close to the core surface (note that this assumption is questionable – see, e.g., Jault and Le Mouël (1991)), a further set of consistency conditions can be constructed (Backus, 1968; Gubbins and Roberts, 1987). An attempt to test these conditions has been made by Barraclough *et al.* (1989); they found no evidence that the constraints are violated, but confessed that the field models used were not yet sufficiently accurate enough to allow a stringent test.

5.05.5.4.5 Theoretical issues concerning the frozen flux hypothesis

The simple scaling arguments of Roberts and Scott regarding the plausibility of using the frozen flux assumption to model secular variation has been the subject of some debate, especially since the claims in the 1980s that the signature of magnetic diffusion had been detected.

Gubbins and Kelly (1996) have pointed out that, for the special case of steady flows, the frozen flux hypothesis is invalid because the balance in the induction equation must be between secular variation and diffusion, with frozen flux effects being negligible. Although undoubtedly true, this observation appears to be of little relevance when trying to model changes in Earth's magnetic field that are time dependent across a wide variety of timescales (even the dynamo process appears to be fundamentally time dependent as seems apparent from considering the frequency of geomagnetic excursions over the past 700 000 years (Gubbins, 1999)). A more worrisome objection raised by Gubbins and Kelly is that the frozen flux approximation is a singular limit of the induction equation. It lowers the differential order of the system from second to first order, raising the possibility that physically relevant solutions might be filtered out.

Concerns over how frozen flux approximation arises as a limit of a magnetic advection–diffusion process have been addressed in detail by Gubbins (1996). Gubbins has developed a novel formalism for determining core motions that takes magnetic diffusion into account and shows that the frozen flux approximation can be formally retrieved in the limit when $(\eta/\omega)^{1/2}$ tends to zero, where ω is the frequency of the secular variation. This implies that we should only expect the frozen flux approximation to work well when field variations are rapid and the concomitant fluid motions highly time dependent. Gubbins used his formalism to derive an estimate of the toroidal field gradient close to the core surface necessary to explain the amount of flux change through the South Atlantic reversed flux patch between 1905.5 and 1965.5 as estimated by Bloxham and Gubbins (1985). He arrived at a plausible scenario by considering a 1.2 mT toroidal field at a depth of 60 km in the core (diminishing at a rate of $20\,\mathrm{nT\,m^{-1}}$ toward the surface) with a horizontal length scale of 10^6 m. When this field was acted on by an upwelling flow with a radial rate of change of $0.02\,\mathrm{yr^{-1}}$ in the presence of a magnetic diffusivity of $1.6\,\mathrm{m^2\,s^{-1}}$, it was found that it was possible to generate the required flux change of 500 MWb.

This estimate, together with the earlier forward calculations of Allan and Bullard (1966), Bloxham (1986a), and Drew (1993), have established the plausibility of toroidal flux expulsion as a mechanism that could cause localized growth of pairs of reverse and normal flux patches. Such arguments however fail to convince on the need to apply a diffusive formalism globally. Rather, they serve as a caution against interpreting core flow inferred using the frozen flux approximation without first considering whether the specific local patterns of secular variation being explained might be produced by a diffusive mechanism. For example, it might be very reasonable to use the frozen flux approximation to understand the rapid azimuthal motion of flux patches but not to understand rapid growth and decay of field concentrations. Such reasoning was used by Dumberry and Bloxham (2006) in their study of global azimuthal core flows on millennial timescales, derived from archeomagnetic field models.

Love (1999) raised an objection to the frozen flux hypothesis in the case of a nearly steady dynamo. Although clearly correct under the conditions considered, the applicability of his examples to the Earth is questionable since, as noted earlier, the geodynamo does not appear to be steady on any known timescales. Recent feasibility tests using more dynamically plausible dynamo models (Roberts and Glatzmaier, 2000; Rau *et al.*, 2000) have, on the other hand, demonstrated that the frozen flux approximation is a useful construct. These important tests will be described in more detail below. The important point to take away from the arguments of Love (1999) is that the original scaling argument of Roberts and Scott was overly simplistic – in real secular variation there is not just one length scale and timescale of interest (rather both magnetic and velocity fields will contain power over a range of length- and timescales). To address whether the frozen flux hypothesis is a useful approximation really requires experiments studying the induction effects of Earth-like velocity fields acting on Earth-like magnetic fields. Unfortunately, it is still beyond our ability to accurately simulate all aspects of the dynamics of Earth's core (Glatzmaier, 2002), though present geodynamo models already capture many important aspects of Earth's magnetic field and its evolution (Kuang and Bloxham, 1997; Christensen *et al.*, 1998) and are certainly useful tools when considering the validity of the frozen flux hypothesis.

Roberts and Glatzmaier (2000) were the first to use a numerical geodynamo model in an attempt to evaluate the frozen flux hypothesis. They calculated the variations in the unsigned flux from a high resolution simulation and found that it varied by $\sim 3\%$ over timescales estimated to be equivalent to 150 years. This suggested that for plausible magnetic and velocity fields the frozen flux hypothesis is a good approximation on the global scale. Unfortunately, no attempt was made to investigate how much individual flux integrals varied over the same interval. The authors further remark that for frozen flux approximation to be useful it is not required to be strictly true; rather, it must only involve errors smaller than those associated with our incomplete knowledge of the field at the core surface. They point out that the errors in the unsigned flux from truncation of the main field at degree 12 in their model (only one aspect of the error present in observationally based field models) are much larger than the amount by which the unsigned flux is varying. Again, a more stringent assessment would involve studying such issues for individual null-flux patches.

Rau *et al.* (2000) have performed a large number of tests on a different suite of geodynamo models. The main focus of their study was to take the time-dependent magnetic field output from the simulation, carry out inversions for the underlying core surface flow on the basis of the frozen flux hypothesis, and compare the results with the known simulation flow. We are most interested in their preliminary test that aimed to determine how well the assumption of frozen flux was satisfied in the models. They observed that the known secular variation and the secular variation predicted on the basis of the frozen flux hypothesis using the known flow were similar, finding high correlation coefficients in the range 0.7–0.8. They observed that the diffusive contribution is significant in only a few isolated locations and that the deviations from the frozen flux assumption are not so large as to preclude its use in the determination of core motions from the magnetic data. It should be noted that their models have R_m in the range 118–320; R_m for the Earth is expected to be rather larger than this (around 500 – see Roberts and Glatzmaier (2000)), and hence the frozen flux approximation might perform even better in reality. Perhaps the most important finding of this study is that even though the frozen flux assumption is not perfectly satisfied (the formal necessary conditions would be violated to an extent), useful information can be extracted concerning fluid motions at the core surface and substantial parts of the secular variation pattern can be explained using the frozen flux approximation. Errors in the field models (due to limits on the range of core field spherical harmonic coefficients that can be determined by observations) and in the oversimplified dynamical assumptions can apparently lead to more serious problems in the inversion than violation of the frozen flux approximation.

One final remark should be made concerning the inevitable failure of the frozen flux hypothesis (Backus and Le Mouël, 1986). We recollect the radial component of the induction equation in its exact form:

$$\partial_t B_r + \nabla_h \cdot (\mathbf{u}_h B_r) = \eta \frac{1}{r} \nabla^2 (r B_r) \quad [54]$$

It is clear that the frozen flux hypothesis fails in all locations where $|\nabla_h \cdot (\mathbf{u}_h B_r)| \leq |\eta r^{-1} \nabla^2 (r B_r)|$. In particular, the approximation always fails on the curve S where

$$\nabla_h \cdot (\mathbf{u}_h B_r) = 0 \quad [55]$$

Backus and Le Mouël call these curves 'leaky curves', and the region around them where the approximation fails the 'leaky belt'. The problem, of course, is that one cannot locate these places *a priori*, because one does not know $\mathbf{u}$, and much depends on the relative scales of variation of B_r, $\mathbf{u}$, including the radial variation of B_r.

One thing is certain: any point where B_r and $\nabla_h B_r$ vanish are inevitably on the leaky curve. Backus called such places 'touch points'; it is at these places that null-flux curves can appear or disappear. Of course, it is incredibly difficult, if not impossible, to determine accurately the positions of such points on the CMB (Backus, 1988), and hence very little effort has been expended on such activity.

5.05.5.4.6 Summary of applicability of the frozen flux approximation in core studies

After 40 years of study using observations, theoretical arguments, and numerical tests, the worth and limitations of using the frozen flux approximation as an explanation of secular variation are now apparent. The approximate invariance of the unsigned flux over intervals of several hundred years in both observationally based field models and numerical simulations demonstrates that flux is well conserved globally on such timescales. On the other hand, field models spanning over 100 years that satisfy flux conservation conditions and satisfactorily fit observations tell us that geomagnetic data is not yet capable of unambiguously detecting the presence of diffusion. Frozen flux is therefore undoubtedly a useful tool in understanding the advective motions that cause much of the observed secular variation. It should however always be borne in mind that the frozen flux assumption is only an approximation and that magnetic diffusion will inevitable be present at some level, causing violation of flux constraints over long time intervals or where field gradients are very large.

5.05.5.5 Other Invariants

We have concentrated on the frozen flux approximation, and its observable (and sometimes testable) consequences. Depending on the additional approximations one is willing to make, there are other invariants that can sometimes be testable. When one makes assumptions about the type of fluid flow that may be occurring at the CMB, a number of other invariants arise – see, for example, those reviewed in the cases of tangentially geostrophic or toroidal fluid flow in Bloxham and Jackson (1991), and additional constraints described in Jackson and Hide (1996), Chulliat and Hulot (2001), and Chulliat (2004). The possibility of using deviations from exact satisfaction of the constraints by field models is discussed in Gubbins (1996) and Hulot and Chulliat (2003).

Here we briefly describe an invariant that arises from perhaps the most innocuous of assumptions additional to that of frozen flux, that of a poorly conducting (or to be strict, a perfectly insulating) mantle. The invariant arises from a consideration of eqn [34], namely the Navier–Stokes equation in the magnetostrophic limit. We consider a null-flux curve (on which $B_r = 0$) and examine the form of the Lorentz force $\mathbf{J} \times \mathbf{B}$; we can write the horizontal part of it as

$$[\mathbf{J} \times \mathbf{B}]_h = \mathbf{B}_h \mathcal{J}_r - \mathbf{J}_h B_r \quad [56]$$

If the mantle is a sufficiently poor conductor that the radial current $\mathcal{J}_r$ is negligible, then all terms on the right-hand side vanish (because $B_r = 0$ also), and we discover that the horizontal Lorentz force vanishes on null-flux curves. Null-flux curves must therefore move as if they are governed by a tangentially geostrophic force balance, even if the true force balance over the core is one of magnetostrophic equilibrium. The repercussion of this is the following: because null-flux curves are material curves (section 5.05.5.4.1), they must obey Kelvin's celebrated theorem (e.g., Gill, 1982), and conserve their planetary vorticity (we have ignored relative vorticity from the outset by dropping the nonlinear advection term in the Navier–Stokes equation). We therefore conclude that the following integral must hold:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_S \cos\theta \, \mathrm{d}S = 0 \quad [57]$$

over any patch S which is a null flux patch, where θ is colatitude. For more details, see Gubbins (1991) and Jackson (1996). The interesting aspect of these constraints is that null flux patches were previously free to shrink *in situ*, while conserving their flux (by concentrating the flux into a smaller area while increasing the amplitude of B_r). This new constraint places demands on their area, so that they cannot shrink *in situ*; they must instead move part of their area to a different latitude, either toward the pole (for shrinking) or away from the pole (for enlargement). These constraints have been imposed on field models by Jackson *et al.* (2007), who find that they can fit 100 years of historical data adequately even when the constraints are imposed.

5.05.6 Summary

We have described the basis of our knowledge of the secular variation of the magnetic field, starting from the fundamental data sets that are available. We subsequently moved through the mathematical techniques that are used for representing the four-dimensional vector $\mathbf{B}(\mathbf{r}, t)$, and discussed the underlying fluid dynamics and electrodynamics of the core. We ended with a discussion of some of the controversies surrounding the approximations employed, and what can be gleaned about the physical state at the top of the core. We have stopped short of describing the industry of computing models of core velocities, and the interesting conclusions that can be drawn from them. For example, models of flow at the core surface appear to have certain components linked to motions deeper within the core: these are the axisymmetric zonal flows that apparently are linked to certain modes of oscillation in the core, often termed torsional oscillations, that are most easily viewed as motions of nested cylinders (coaxial with the rotation axis) filling the core. These oscillations are almost certainly linked to changes in the rotation rate of the Earth, often coined changes in the length of day. This observation in itself is remarkable, but what is more exciting is the possibility of gleaning something about the state of the magnetic field within the core from exactly these oscillations, since they are strongly affected by the component of the interior magnetic field perpendicular to the rotation axis. For a discussion of this, see Volume 8.

The reader who has consulted Chapter 5.02 will see that satellites are beginning to provide models of secular variation with unprecedented accuracy; there is a tradeoff between accuracy, and the short observing time span. This means that historical models still have a role to play, since the accumulated effects of slow processes can be seen over long times even when the signal/noise level is lower. A number of challenges lay ahead, not least the question of how to optimally incorporate the newest satellite, repeat station, and observatory data in a homogeneous way with the older data described herein.

Acknowledgments

The authors thank Mioara Mandea for supplying data for some of the figures, and Terry Sabaka and Katia Pinheiro for supplying **Figures 16** and **28**, respectively. The advice of Susan Macmillan regarding observatory data is gratefully acknowledged. Nils Olsen has constantly given early access to his models, for which we are grateful.

References

Alexandrescu M, Courtillot V, and Le Mouël JL (1996) Geomagnetic field direction in Paris since the mid-sixteenth century. *Physics of the Earth and Planetary Interiors* 98: 321–360.

Alexandrescu M, Gibert D, Hulot G, Le Mouël JL, and Saracco G (1995) Detection of geomagnetic jerks using wavelet analysis. *Journal of Geophysical Research* 100(B10): 12557–12572.

Alexandrescu M, Gibert D, Hulot G, Le Mouël JL, and Saracco G (1996) Worldwide wavelet analysis of geomagnetic jerks. *Journal of Geophysical Research* 101(B10): 21975–21994.

Alexandrescu M, Gibert D, Le Mouël JL, Hulot G, and Saracco G (1999) An estimate of average lower mantle conductivity by wavelet analysis of geomagnetic jerks. *Journal of Geophysical Research* 104: 17735–17746.

Allan DW and Bullard EC (1966) The secular variation of the Earth's magnetic field. *Proceedings of the Cambridge Philosophical Society* 62(3): 783–809.

Backus G (1968) Kinematics of geomagnetic secular variation in a perfectly conducting core. *Philosophical Transactions of the Royal Society of London, Series A* 263: 239–266.

Backus GE (1986) Poloidal and toroidal fields in geomagnetic field modeling. *Reviews of Geophysics* 24: 75–109.

Backus G (1988) Bayesian inference in geomagnetism. *Geophysical Journal of the Royal Astronomical Society* 92: 125–142.

Backus GE and Le Mouël JL (1986) The region on the core-mantle boundary where a geostrophic velocity field can be determined from frozen-flux geomagnetic data. *Geophysical Journal of the Royal Astronomical Society* 85: 617–628.

Backus G, Parker R, and Constable C (1996) *Foundations of Geomagnetism*. Cambridge: Cambridge University Press.

Barraclough D (1974) Spherical harmonic analysis of the geomagnetic field for eight epochs between 1600 and 1910. *Geophysical Journal of the Royal Astronomical Society* 36: 497–513.

Barraclough DR (1978) Spherical harmonic analysis of the geomagnetic field, No.8: *Geomagnetic Bulletin of the Institute of Geological Sciences*. London: HMSO.

Barraclough DR (1982) Historical observations of the geomagnetic field. *Philosophical Transactions of the Royal Society of London, Series A* 306: 71–78.

Barraclough DR (1985) Halley's Atlantic magnetic surveys. In: Schröder W (ed.) *Historical Events and People in Geosciences*. New York: Peter Lang.

Barraclough DR (1995) Observations of the Earth's magnetic field in Edinburgh from 1670 to the present day. *Transactions of the Royal Society of Edinburgh: Earth Sciences* 85: 239–252.

Barraclough D, Gubbins D, and Kerridge D (1989) On the use of horizontal components of magnetic field in determining core motions. *Geophysical Journal International* 98: 293–299.

Barraclough DR, Carrigan JG, and Malin SRC (2000) Observed geomagnetic field intensity in London since 1820. *Geophysical Journal International* 141: 83–99.

Batchelor GK (1967) *An Introduction to Fluid Mechanics*. Cambridge: Cambridge University Press.

Becquerel AC (1840) *Traité Experimental de l'Electricité et du Magnétisme*. Paris: Fermin Didot Frères.

Bemmelen W van (1899) *Die Abweichung der Magnetnadel: Beobachtungen*, *Säcular-Variation*, *Wert- und Isogonensysteme bis zur Mitte des XVIIIten Jahrhunderts*, supplement to *Observations Made at the Royal Magnetic and Meteorological Observatory*, *Batavia*, 21: 1–109.

Benton ER, Estes RH, and Langel RA (1987) Geomagnetic field modeling incorporating constraints from frozen-flux electromagnetism. *Physics of the Earth and Planetary Interiors* 48: 241–264.

Benton ER and Voorhies CV (1987) Testing recent geomagnetic field models via magnetic flux conservation at the core–mantle boundary. *Physics of the Earth and Planetary Interiors* 48: 350–357.

Benton ER and Whaler KA (1983) Rapid diffusion of poloidal geomagnetic field through the weakly conducting mantle: a perturbation solution. *Geophysical Journal of the Royal Astronomical Society* 75: 77–100.

Bloxham J (1985) *Geomagnetic Secular Variation*. Unpublished Ph.D. Thesis, University of Cambridge.

Bloxham J (1986a) The expulsion of magnetic flux from the Earth's core. *Geophysical Journal of the Royal Astronomical Society* 87: 669–678.

Bloxham J (1986b) Models of the magnetic field at the core–mantle boundary for 1715, 1777, and 1842. *Journal of Geophysical Research* 91: 13954–13966.

Bloxham J (1987) Simultaneous stochastic inversion for geomagnetic main field and secular variation 1. A large scale inverse problem. *Journal of Geophysical Research* 92: 11597–11608.

Bloxham J (2002) Time-independent and time-dependent behaviour of high-latitude flux bundles at the core-mantle boundary. *Geophysical Research Letters* 29(18): 1854–1858.

Bloxham J and Gubbins D (1985) The secular variation of Earth's magnetic field. *Nature* 317: 777–781.

Bloxham J and Gubbins D (1986) Geomagnetic field analysis IV – Testing the frozen flux hypothesis. *Geophysical Journal of the Royal Astronomical Society* 84: 139–152.

Bloxham J and Gubbins D (1987) Thermal core–mantle interactions. *Nature* 325: 511–513.

Bloxham J, Gubbins D, and Jackson A (1989) Geomagnetic secular variation. *Philosophical Transactions of the Royal Society of London, Series A* 329(1606): 415–502.

Bloxham J and Jackson A (1989) Simultaneous stochastic inversion for geomagnetic main field and secular variation 2. 1820–1980. *Journal of Geophysical Research* 94: 15753–15769.

Bloxham J and Jackson A (1991) Fluid flow near the surface of the Earth's outer core. *Reviews of Geophysics* 29: 97–120.

Bloxham J and Jackson A (1992) Time-dependent mapping of the magnetic field at the core-mantle boundary. *Journal of Geophysical Research* 97(B13): 19537–19563.

Bloxham J, Zatman S, and Dumberry M (2002) The origin of geomagnetic jerks. *Nature* 420: 685–687.

Booker JR (1969) Geomagnetic data and core motions. *Proceedings of the Royal Society of London Series A - Mathematical Physical and Engineering Sciences* 309: 27–40.

Braginsky SI and Roberts PH (1995) Equations governing convection in Earth's core and the geodynamo. *Geophysical and Astrophysical Fluid Dynamics* 79: 1–97.

Bullard EC, Freedman C, Gellman H, and Nixon J (1950) The westward drift of the Earth's magnetic field. *Philosophical Transactions of the Royal Society of London, Series A* 243: 67–92.

Bullard EC and Gellman H (1954) Homogeneous dynamos and terrestrial magnetism. *Philosophical Transactions of the Royal Society of London, Series A* 247: 213–278.

Busse FH (1975) A necessary condition for the geodynamo. *Journal of Geophysical Research* 80: 278–280.

Cafarella L, De Santis A, and Meloni A (1992) Secular variation in Italy from historical geomagnetic field measurements. *Physics of the Earth and Planetary Interiors* 73: 206–221.

Cain JC, Daniels WE, Hendricks SJ, and Jensen DC (1965) An evaluation of the main geomagnetic field, 1940–1962. *Journal of Geophysical Research* 70: 3647–3674.

Cain JC, Hendricks SJ, Langel RA, and Hudson WV (1967) A proposed model for the International Geomagnetic Reference Field – 1965. *Journal of Geomagnetism and Geoelectricity* 19: 335–355.

Cawood J (1977) Terrestrial magnetism and the development of international collaboration in the early nineteenth century. *Annals of Science* 34: 551–587.

Cawood J (1979) The magnetic crusade: Science and politics in early Victorian Britain. *Isis* 70(no.254): 493–518.

Chapman S and Bartels J (1940) *Geomagnetism*, vol. II, 1049 pp. London: Oxford University Press.

Christensen UR, Olson P, and Glatzmaier G (1998) A dynamo model interpretation of geomagnetic field structures. *Geophysical Research Letters* 25: 1565–1568.

Chulliat A (2004) Geomagnetic secular variation generated by a tangentially geostrophic flow under the frozen-flux assumption – II. sufficient conditions. *Geophysical Journal International* 157(2): 537–552.

Chulliat A and Hulot G (2001) Geomagnetic secular variation generated by a tangentially geostrophic flow under the frozen-flux assumption – I. necessary conditions. *Geophysical Journal International* 147(2): 237–246.

Churchman J (1794) *The magnetic atlas or variation charts of the whole terraqueous globe; comprising a system of the variation and dip of the needle, by which, the observations being truly made, the longitude may be ascertained 2nd edn. London.*

Constable CG, Parker RL, and Stark PB (1993) Geomagnetic field models incorporating frozen-flux constraints. *Geophysical Journal International* 113: 419–433.

Courtillot V and Le Mouël JL (1984) Geomagnetic secular variation impulses. *Nature* 311: 709–716.

Crichton Mitchell A (1932) Chapters in the history of terrestrial magnetism I. *Terrestrial Magnetism and Atmospheric Electricity* 37: 105–146.

Crichton Mitchell A (1937) Chapters in the history of terrestrial magnetism II. *Terrestrial Magnetism and Atmospheric Electricity* 42: 241–280.

Crichton Mitchell A (1939) Chapters in the history of terrestrial magnetism III. *Terrestrial Magnetism and Atmospheric Electricity* 44: 77–80.

Davidson P (2001) *An Introduction to Magnetohydrodynamics*. Cambridge: Cambridge University Press.

De Boor C (2002) *A Practical Guide to Splines*. New York: Springer.

Doell RR and Cox A (1971) Pacific Geomagnetic Secular Variation. *Science* 171: 248–254.
Drew S (1993) Magnetic field line expulsion into a conducting mantle. *Geophysical Journal International* 115: 303–312.
Dumberry M and Bloxham J (2006) Azimuthal flows in the Earth's core and changes in the length of day on millennial timescales. *Geophysical Journal International* 165: 32–46.
Finlay CC (2005) *Hydromagnetic Waves in Earth's Core and Their Influence on Geomagnetic Secular Variation*. Unpublished PhD Thesis. University of Leeds.
Finlay CC and Jackson A (2003) Equatorially dominated magnetic field change at the surface of Earth's core. *Science* 300: 2084–2086.
Fisk HW (1931) Isopors and isoporic movements. *Intern. Geodet. Geophys. Union, Section Terrest. Magnet. Elec. Bull. Stockholm* 8: 280–292.
Gauss CF (1833a) *Intensitas vis magneticae terrestris ad mensuram absolutam revocata*. Gottingae: Sumtibus Dieterichianis.
Gauss CF (1833b) Intensitas vis magneticae terrestris ad mensuram absolutam revocata. *Abstracts of the Papers Printed in the Philosophical Transactions of the Royal Society of London* 3: 166–174.
Gauss CF (1839) Allgemeine Theorie des Erdmagnetismus: Resultate aus den Beobachtungen des magnetischen Verein im Jahre 1838. Leipzig.
Gilbert W (1600) *De magnete*. London: P. Short (English translation, Chiswick Press, London, 1900).
Gill AE (1982) *Atmosphere-Ocean dynamics*. Cambridge: Cambridge University Press.
Glatzmaier GA (2002) How realistic and geodynamo simulations?. *Annual Review of Earth and Planetary Sciences* 30: 237–257.
Good GA (1985) Geomagnetics and scientific institutions in 19th century America. *Transactions of the American Geophysical Union* 66: 521–526.
Good GA (1988) The study of geomagnetism in the late 19th century. *Transactions of the American Geophysical Union* 69: 218–232.
Gubbins D (1975) Can the Earth's magnetic field be sustained by core oscillations? *Geophysical Research Letters* 2(9): 409–512.
Gubbins D (1983) Geomagnetic field analysis – I. Stochastic inversion. *Geophysical Journal of the Royal Astronomical Society* 73: 641–652.
Gubbins D (1984) Geomagnetic field analysis – II. Secular variation consistent with a perfectly conducting core. *Geophysical Journal of the Royal Astronomical Society* 77: 753–766.
Gubbins D (1987) Mechanism for geomagnetic polarity reversals. *Nature* 326: 167–169.
Gubbins D (1991) Dynamics of the Secular Variation. *Physics of the Earth and Planetary Interiors* 68: 170–182.
Gubbins D (1996) A formalism for the inversion of geomagnetic data for core motions with diffusion. *Physics of the Earth and Planetary Interiors* 98: 193–206.
Gubbins D (1999) The distinction between geomagnetic excursions and reversals. *Geophysical Journal International* 173: F1–F3.
Gubbins D and Bloxham J (1985) Geomagnetic field analysis – III. Magnetic fields on the core-mantle boundary. *Geophysical Journal of the Royal Astronomical Society* 80: 695–713.
Gubbins D and Bloxham J (1987) Morphology of the geomagnetic field and implications for the geodynamo. *Nature* 325: 509–511.
Gubbins D, Jones AL, and Finlay CC (2006) Fall in Earth's magnetic field is erratic. *Science* 312: 900–902.
Gubbins D and Kelly P (1996) A difficulty with using the frozen flux hypothesis to find steady core motions. *Geophysical Research Letters* 23: 1825–1828.
Gubbins D and Roberts PH (1987) Magnetohydrodynamics of the Earth's core. In: Jacobs JA (ed.) *Geomagnetism*, Volume II, pp. 1–183. London: Academic Press.
Halley E (1683) A theory of the variation of the magnetic compass. *Philosophical Transactions of the Royal Society of London* 13: 208–221.
Halley E (1692) On the cause of the change in the variation of the magnetic needle, with a hypothesis of the structure of the internal parts of the Earth. *Philosophical Transactions of the Royal Society of London* 17: 470–478.
Hendricks SJ and Cain JC (1963) World magnetic survey data. *NASA Goddard Space Flight Centre*, *technical report*, *August 1963*.
Hide R and Malin SRC (1981) On the determination of the size of the Earth's core from observations of geomagnetic secular variation. *Proceedings of the Royal Society of London Series A - Mathematical Physical and Engineering Sciences* 374: 15–33.
Hide R and Stewartson K (1973) Hydromagnetic oscillations of the Earth's core. *Reviews of Geophysics* 10(2): 579–598.
Hills RG (1979) Convection in the Earth's mantle due to viscous shear at the core–mantle interface and due to large-scale buoyancy, *PhD Thesis*, New Mexico State University, Las Cruces, New Mexico.
Holme R and Olsen N (2006) Core-surface flow modelling from high resolution secular variation. *Geophysical Journal International* 166: 518–528.
Hulot G and Chulliat A (2003) On the possibility of quantifying diffusion and horizontal Lorentz forces at the earth's core surface. *Physics of the Earth and Planetary Interiors* 135(1): 47–54.
Hulot G, Le Huy M, and Le Mouël J-L (1993) Secousses (jerks) de la variation séculaire et mouvements dans le noyau terrestre. *Comples Rendus de l'Académie des Sciences Paris* 317: 333–341.
Hulot G and Le Mouël JL (1994) A statistical approach to earth's main magnetic field. *Physics of the Earth and Planetary Interiors* 82: 167–183.
Hutcheson KA (1990) *Geomagnetic Field Modelling*. Unpublished Ph.D. Thesis, University of Cambridge.
Hutcheson KA and Gubbins D (1990) Earth's magnetic field in the seventeenth century. *Journal of Geophysical Research* 95: 10769–10781.
Jackson A (1989). *The Earth's Magnetic Field at the Core–Mantle Boundary*. Unpublished Ph.D. Thesis, University of Cambridge.
Jackson A (1996) Kelvin's theorem applied to earth's core. *Proceedings of the Royal Society of London* 452: 2195–2201.
Jackson A (1997) Time-dependency of tangentially geostrophic core surface motions. *Physics of the Earth and Planetary Interiors* 103: 293–311.
Jackson A (2003) Intense equatorial flux spots on the surface of Earth's core. *Nature* 424: 760–763.
Jackson A, Constable CG, Walker MR, and Parker R (2007) Models of Earth's main magnetic field incorporating flux and radial vorticity constraints. *Geophysical Journal International* (in press).
Jackson A and Hide R (1996) Invariants in toroidal and tangentially geostrophic frozen-flux velocity fields. *Geophysical Journal International* 125: 925–927.
Jackson A, Jonkers ART, and Walker MR (2000) Four centuries of geomagnetic secular variation from historical records. *Philosophical Transactions of the Royal Society of London, Series A* 358: 957–990.
Jackson JD (1999) *Classical Electrodynamics, 3rd edn. New York: Wiley.*
Jault D, Gire C, and Le Mouël JL (1988) Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature* 333: 353–356.

Jault D and Le Mouël JL (1991) Physical properties at the top of the core and core surface motions. *Physics of the Earth and Planetary Interiors* 68: 76–84.

John II's commission (1509) *Regimento do astrolabe e do quadrante,* Lisbon.

Jonkers ART (2000) *North by Northwest. Seafaring, Science, and the Earth's Magnetic Field (1600–1800)*. Göttingen: Cuvillier Verlag.

Jonkers ART (2003) *Earth's Magnetism in the Age of Sail*. Baltimore, MD: Johns Hopkins University Press.

Jonkers ART, Jackson A, and Murray A (2003) Four centuries of geomagnetic data from historical records. *Reviews of Geophysics* 41: 1006 (doi:10.1029/2002RG000115).

Kircher A (1654) *Magnes sive de Arte Magnetica* (. . .), 3rd ed. Rome.

Kono M and Roberts PH (2002) Recent geodynamo simulations and observations of the geomagnetic field. *Reviews of Geophysics* 40: 1013 (doi:10.1029/2000RG00102).

Korte M, Genevey A, Constable C, Frank U, and Schnepp E (2005) Continuous geomagnetic field models for the 7 millennia I: A new global data compilation. *Geochemistry, Geophysics, Geosystems* 6(2): Q02H15 (doi:10.1029/ 2004GC000800).

Kuang W and Bloxham J (1997) An Earth-like numerical dynamo model. *Nature* 389: 371–374.

Langel RA (1987) The Main Field. In: Jacobs JA (ed.) *Geomagnetism, Volume I*, pp. 249–512. London: Academic Press.

Langel RA (1993) *Types and characteristics of data for geomagnetic field modeling*, NASA Conference Publication 3153. Washington, DC: NASA.

Langel RA, Estes RH, and Mead GD (1982) Some new methods in geomagnetic field modelling applied to the 1960–1980 epoch. *Journal of Geomagnetism and Geoelectricity* 34: 327–349.

Langel RA, Kerridge DJ, Barraclough DR, and Malin SRC (1986) Geomagnetic temporal change: 1903–1982, A spline representation. *Journal of Geomagnetism and Geoelectricity* 38: 573–597.

Lanzerotti LJ, Chave AD, Sayres CH, Medford LV, and Maclennan CG (1993) Large-scale electric-field measurements on the Earth's surface – A review. *Journal of Geophysical Research-Planets* 98(E12): 23525–23534.

Le Mouël JL (1984) Outer core geostrophic flow and secular variation of Earth's magnetic field. *Nature* 311: 734–735.

Love JJ (1999) A critique of frozen-flux inverse modelling of a nearly steady geodynamo. *Geophysical Journal International* 138: 353–365.

Macmillan S (1996) A geomagnetic jerk for the early 1990's. *Earth and Planetary Science Letters* 137: 189–192.

Malin SRC (1982) Sesquicentenary of Gauss's first measurement of the absolute value of magnetic intensity. *Philosophical Transactions of the Royal Society A* 306: 5–8.

Malin SRC (1987) Historical introduction to geomagnetism. In: Jacobs JA (ed.) *Geomagnetism*, Volume I, pp. 1–49. London: Academic Press.

Malin SRC and Bullard EC (1981) The direction of the Earth's magnetic field at London, 1570–1975. *Philosophical Transactions of the Royal Society of London, Series A* 299: 357–423.

Mandea M, Bellanger E, and LeMouel J-L (2000) A geomagnetic jerk for the end of the 20th century. *Earth and Planetary Science Letters* 183: 369–373.

Maus S, Lühr H, Balasis G, Rother M, and Mandea M (2005) Introducing POMME the postdam magnetic model of the Earth. In: Reigber C, Luhr H, and Schwintzer P (eds.) *CHAMP Mission Results II*, pp. 347–352. Berlin: Springer.

McConnell A (1980) *Geomagnetic instruments before 1900; an illustrated account of their construction and use*. London: Harriet Wynter.

Merrill RT, McElhinny MW, and McFadden PL (1998) *The Magnetic Field of the Earth*. London: Academic Press.

Mountaine W and Dodson J (1757) A letter to the Right Honourable the Earl of Macclesfield, President of the Council and Fellow of the Royal Society, concerning the variation of the magnetic needle; with a set of tables annexed, which exhibit the results of upwards of fifty thousand observations, in six periodic reviews, from the year 1700 to the year 1756, both inclusive; and are adapted to every 5 degrees of latitude and longitude in the more frequented oceans. *Philosophical Transactions of the Royal Society of London* 50: 329–350.

Multhauf RP and Good G (1987) *Smithsonian Studies in History and Technology*, Volume 48: *A brief history of geomagnetism and a catalogue of the collections of the National Museum of American History*, 87 pp. Washington: Smithsonian Institution Press.

Nagao H, Iyemori T, Higuchi T, and Araki T (2003) Lower mantle conductivity anomalies estimated from geomagnetic jerks. *Journal of Geophysical Research* 108: (doi 10.1029/ 2002JB001786).

Needham J (1962) *Science and civilisation in China, vol. 4, part 1. Cambridge: Cambridge University Press.*

O'Brien MS (1996) Resolving magnetic flux patches at the surface of the core. *Geophysical Research Letters* 23: 3071–3074.

O'Brien MS, Constable CG, and Parker RL (1997) Frozen flux modelling for epochs 1915 and 1980. *Geophysical Journal International* 128: 434–450.

Olsen N (2002) A model of the geomagnetic field and its secular variation estimated for epoch 2000. *Geophysical Journal International* 149: 454–462.

Olsen N, Lühr H, Sabaka TJ, *et al.* (2006) CHAOS – A model of Earth's magnetic field derived from CHAMP, Ørsted and SAC-C magnetic satellite data. *Geophysical Journal International* 166: 67–75.

Olson P and Christensen UR (2002) The time-averaged magnetic field in numerical dynamos with nonuniform boundary heat flow. *Geophysical Journal International* 151: 809–823.

Peregrinus P (1269) *Epistola de magnete*.

Rau S, Christensen U, Jackson A, and Wicht J (2000) Core flow inversion tested with numerical dynamo models. *Geophysical Journal International* 141: 484–497.

Raulin V (1867) Sur les variations seculaires du magnetisme terrestre. *Actes de la Société Linnéenne de Bordeaux* 26: 1–92.

Rayet MG (1876) Recherches sur les observations magnétiques faites à l'observatoire de Paris de 1667 à 1872. *Annales de l'Observatoire de Paris Memoires* 13.

Roberts PH and Glatzmaier GA (2000) A test of the frozen flux approximation using a new geodynamo model. *Philosophical Transactions of the Royal Society of London, Series A* 358: 1109–1121.

Roberts PH and Scott S (1965) On the analysis of the secular variation. 1. A hydromagnetic constraint: theory. *Journal of Geomagnetism and Geoelectricity* 17(2): 137–151.

Runcorn SK (1955) The electrical conductivity of the Earth's mantle. *Transactions of the American Geophysical Union* 36: 191–198.

Sabaka TJ and Baldwin RT (1993) Modeling the Sq magnetic field from POGO and Magsat satellite and contemporaneous hourly observatory data. *HSTX/G&G-9302*, *Hughes STX Corp.*, *7701 Greenbelt Road*, *Greenbelt*, *MD*.

Sabaka TJ, Langel RA, Baldwin RT, *et al.* (1997) The geomagnetic field 1900–1995, including the large-scale field from magnetospheric sources, and the NASA candidate

models for the 1995 revision of the igrf. *Journal of Geomagnetism and Geoelectricity* 49: 157–206.
Sabaka TJ, Olsen N, and Langel RA (2002) A comprehensive model of the quiet-time, near-Earth magnetic field: Phase 3. *Geophysical Journal International* 151: 32–68.
Sabaka TJ, Olsen N, and Purucker ME (2004) Extending comprehensive models of the Earth's magnetic field with Oersted and CHAMP data. *Geophysical Journal International* 159: 521–547.
Shimizu H, Koyama T, and Utada H (1998) An observational constraint on the strength of the toroidal magnetic field at the cmb by time variation of submarine cable voltages. *Geophysical Research Letters* 25(21): 4023–4026.
Shimizu H and Utada H (2004) The feasibility of using decadal changes in the geoelectric field to probe earth's core. *Physics of the Earth and Planetary Interiors* 142(3–4): 297–319.
Shure L, Parker RL, and Backus GE (1982) Harmonic splines for geomagnetic modelling. *Physics of the Earth and Planetary Interiors* 28: 215–229.
Smith PJ and Needham J (1967) Magnetic declination in medieval China. *Nature* 214: 1213–1214.
Stern DP (2002) A millennium of geomagnetism. *Reviews of Geophysics* 40: 1007 (doi:10.1029/2000RG000097).
Stevin S (1599) *De Havenvinding*. Leyden: Plantijn.
Veinberg BP (1929–33) *Catalogue of Magnetic Determinations in USSR and in Adjacent Countries from 1556 to 1926, 3 vols.* Leningrad: *Central Geophysical Observatory.*
Voorhies CV and Benton ER (1982) Pole-strength from MAGSAT and the magnetic determination of the core radius. *Geophysical Research Letters* 9: 258–261.
Whaler K and Gubbins D (1981) Spherical harmonic analysis of the geomagnetic field: An example of a linear inverse problem. *Geophysical Journal of the Royal Astronomical Society* 65: 645–693.
Wicht J and Olson P (2004) A detailed study of the polarity reversal mechanism in a numerical dynamo model. *Geochemistry, Geophysics, Geosystems* 5: Q03H10 (doi:10.1029/2003GC000602).
Wright E (1657) *Certaine errors of navigation detected and corrected, with many additions that were not in the former editions. 3rd edn. London: Joseph Moxton.*
Yukutake T (1962) The westward drift of the magnetic field of the Earth. *Bulletin of the Earthquake Research Institute, University of Tokyo* 40: 1–65.
Yukutake T and Tachinaka H (1969) Separation of the Earth's magnetic field into the drifting and standing parts. *Bulletin of the Earthquake Research Institute, University of Tokyo* 47: 65–97.