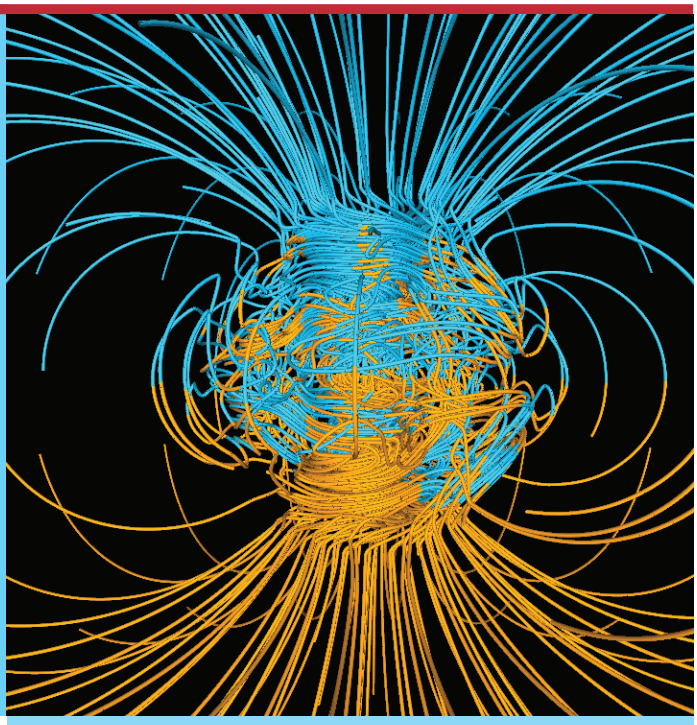
Modelling Changes of the Earth's Core Magnetic Field

Bachelor's Thesis



Ali Jabbar Hussein Muhammad
June 2017

Supervisor:
Chris Finlay

DTU Space
National Space Institute

Abstract

In an era when data of the Earth's magnetic field is delivered in high temporal and spatial resolution from satellite measurements, some associated challenges need to be addressed in order to make accurate field models. Unmodelled signals from ionospheric and magnetospheric currents pose a challenge in modelling fields of internal origin such as the core field. Today modellers use regularization to cope with these unmodelled signals by penalizing the square of second or third time-derivatives of the field, risking to enforce unwanted smoothing effects on the core surface and hence restricting inference on core dynamics. In this project we use 68077 data points from CHAMP in a span of 8.7 years to obtain robustly estimated model solutions of the field where an alternative approach of regularization is used. This method is less restrictive in minimizing model roughness and hence contains unmodelled signals that in some Gauss coefficients are manifested with a periodic behaviour. The model solutions are utilized to investigate the behaviour of the signals in temporal and geographical distribution on the Earth's surface, and we show that the signals have higher amplitudes at polar latitudes which we argue is a strong indication that the source of the noise is high-latitude polar/auroral current systems. Finally we attempt to address the question of whether and to which degree it is possible to contain the signals by means of data selection based on the interplanetary magnetic field and recent satellite-based auroral electrojet indices in our regularization framework. It is shown that such data selections have a positive effect on model characteristics, albeit to a rather limited extent. The performed data selections and the disturbance characterizations strongly indicate that the observed disturbances are solar driven currents with the most prominent amplitudes in the vicinity of polar regions.

Front page figure: three-dimensional numerical simulation of the geodynamo, simulated with the Glatzmaier-Roberts geodynamo model. Source: `https://websites.pmc.ucsc.edu/~glatz/geodynamo.html`

Contents

1 Introduction

1.1 The Aim and Structure of the Project

In this project we build main field models using spherical harmonic modelling and inverse problem methodology. The data used in building the models are data from CHAMP spanning approximately 8 years in the period 2002-2010. The objectives of the project include the assembly of the modelling framework that allows the determination of model solutions to the field with a spatial and temporal dependence and the adoption of known improvements on the models. This yields a foundation for the main objective of the course: analyzing the effects of data selections to cope with some of the challenges caused by ionospheric and magnetospheric current systems in field modelling. This will be a rather modest attempt that will involve a relatively small number of data, but it will allow us to characterize some aspects of the noise problem in satellite data and to investigate the effects of data selections on the models to a certain degree.

The project is structured as follows:

1. An introduction to the geomagnetic field, its contributions and the modelling process is given. Furthermore we give a short description of satellite missions and the data they deliver.

2. The modelling machinery that will give a framework for obtaining model parameters is introduced. With model parameters at disposal one may find the field components at specified time and location. This includes an important discussion on regularization.

3. The data used in this project will be discussed, together with prior selections and selections based on the interplanetary magnetic field and satellite-based auroral indices.

4. A section on the results and discussion together with comments on the outcomes and suggestions for further work.

1.2 Electrodynamics and the Geomagnetic Field

The classical theory of electromagnetism revolutionized our understanding of the electromagnetic interaction, one of the four fundamental forces of nature. The development of the theory culminated with a set of four equations that unified electric and magnetic phenomena together with a wave description of light as a manifestation of the same force, a unification which was brought together by J. C. Maxwell. In vacuum the equations are given by:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0}, & \boldsymbol{\nabla} \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \boldsymbol{\nabla} \cdot \mathbf{B} &= 0, & \boldsymbol{\nabla} \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \end{aligned} \tag{1}$$

Where the fields $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$ are dynamical quantities that depend on spacial coordinates and time and are determined by the sources, i.e. charges and currents.

One of a celestial body's key features is its magnetic field. Stars with their huge reservoir of dynamical plasma produce magnetic fields, and some planets such as the Earth and Jupiter also exhibit magnetic fields with varying intensities and complexities. Modern theories of the cause of the Earth's magnetic field (which from now own will be denoted by the geomagnetic field) strongly indicate that the field is mainly due to a so-called geodynamo process, in which effects of rotation and convection of a conducting fluid produces the field. From a geophysical point of view, the geomagnetic field is a very important quantity due to its link to other branches in geophsyics such as plate tectonics. Tracking the field to the past reveals significant physical insights about the evolution of our planet, and a deep

understanding of its mechanism is important to make future predictions, e.g. geomagnetic reversals and explaining core-dynamical events such as geomagnetic jerks, where such an understanding relies heavily on the accuracy of the available mathematical models of the field. Since the biggest (internal) source of the magnetic field is due to the geodynamo, the behavior of the field is strongly related to the complex motion of the fluid in the molten outer core.

1.3 Modelling the Field

An understanding of the behavior of the geomagnetic field is dependant on the models available that can describe the field accurately. The ability to build accurate models has been enhanced since satellite missions started to take high-resolution measurements of the field, whereas measurements before were restricted to observatories with an uneven geographical distribution.
The equations for determining a model for geomagnetic field are Maxwell's equations for the magnetic field $\mathbf{B}$. As will be shown later, the solution to the field is a solution to Laplace equation taking form of a series expansion of a potential determined uniquely by the expansion coefficients. The potential then determines the field, hence the expansion coefficients are the model parameters that must be determined to have a geomagnetic field model.
The modelling methodology in geomagnetism is by no way different from any other generic geophysical problem. In a so-called forward problem, a given model predicts data (measurements) of a physical system (e.g. the Earth's magnetic field), while in an inverse problem, models describing a physical system are constructed using available measurements (i.e. data). As mentioned earlier, the core is hypothesized to be a complex system of a dynamo process whose behavior is poorly understood, thus no geomagnetic model can be inferred directly. Knowing the relation between the measurements and the model parameters as prescribed by a physical theory will then allow us to determine the model parameters by applying inverse problem methods.

1.4 Satellite Geomagnetism

Measurements taken by a satellite offer a great range of advantages. A full geographical coverage is unquestionably one of them, in contrast to the limitation of distributing ground observatories evenly around the the Earth. Apart from a polar gap[1], a satellite mission spanning some years would provide a full coverage of the field in latitude and longitude [Olsen and Stolle, 2012]. This abundance of data is crucial for modern geomagnetic modelling, and satellites deliver data both in high temporal and spatial resolutions, both needed in making accurate models. Ground observatories can then be used to supplement satellite-made models, e.g. by testing the models against their measurements of the field at specific geographical locations or by actively taking part in the modelling process [Finlay et al., 2016].
However, satellite measurements also pose a challenge in field modelling. One such challenge is the fact that satellites are in motion with a speed of approximately 8 km s^{-1} at the low Earth orbit (LEO) in contrast to stationary stations whose measured variation in the field are solely related to the temporal change in the magnetic field. Hence methods to separate the spatial and the temporal dependences must be utilized in order to give correct models that incorporate both dependences. The biggest challenge in using satellite measurements is unquestionably the altitude of the satellite. Being in the LEO region where highly dynamical ionospheric and magnetospheric currents have considerable contributions to the measured field by the satellite's magnetometer. This is indeed the challenge that comprises the main concern of this project.
Satellite geomagnetism has a history back in the end 1950s and beginning 1960s, but in the previous

[1]Which results from the inclination of the satellite and span the circle of radius $|90°-i|$, where i is the satellite inclination. In the case of CHAMP, $i = 87°$.

two decades Ørested, CHAMP and SWARM have delivered data with high precision of absolute measurements of the geomagnetic field. In this project a number of data has been selected from CAHMP (described below) to construct geomagnetic models using spherical harmonic modelling.
CHAMP (**CHA**llenging **M**inisatellite **P**ayload) was a German satellite designed for geoscientific, atmospheric and ionospheric research and was managed by the German Research Centre for Geosciences (GFZ). The satellite operated in the period July 2000 - Sept. 2010 with an inclination of 87°, providing data in scalar and vector quantities. The satellite operated at altitude in the range of 260 - 450 km in the low Earth orbit (LEO).

1.5 The Contributions to the Field

The measured magnetic field by the satellite is a superposition of several contributions. The contributions to the measured field can be separated into two classes: those of internal origin and those of an external origin. The dynamo process of the Earth is thought to be generating and sustaining the core field, also denoted the main field. At the Earth's surface the core field comprises about 95% of the total field. Lithospheric[2] magnetization comprise most of the remaining internal field, denoted the crustal field. The fields of external origin include electric currents in the ionosphere and the magnetosphere.
The fact that the satellite's magnetometer measures a superposition of fields caused by different mechanisms poses a natural interest in the ability of separating the different parts of the field. In this project we consider core field models, where inference on core dynamics require that the data has been cleaned to a satisfactory degree of any unrelated mechanism. In particular, currents in the atmosphere (ionosphere and magnetosphere) have been studied thoroughly by modellers in order to exclude data that are thought to contain external effects in internal field modelling [Finlay et al., 2017]. These currents and their effect on core-field modelling are of great importance to this project and we shall give a more detailed discussion of them in later sections.

1.5.1 External Fields and Geomagnetic Indices

The geomagnetic field itself includes the field produced by all sources within and beyond the solid planet extending to the magnetopause after which the solar contribution dominates in what is called the interplanetary magnetic field (IMF). This field is carried along with solar plasma (solar wind) from the solar corona and varies accordingly depending on the solar activity. The external contribution to the field is located above the neutral atmosphere at the ionosphere and magnetosphere. These regions of the atmosphere are home to current systems that have been studied extensively, and they are of great importance when implementing data selections. In what follows some of these currents are introduced without delving into deep details.
The ionosphere extends from 60 km to 1000 km in altitude and is made of ionized atoms and free electrons, primarily due to the UV exposure from the Sun. The ionization depends on the solar activity and varies accordingly. The variations may diurnal or seasonal and additionally it depends on the Solar cycle. In the region 85-200 km in altitude where the ionospheric plasma is electrically conducting, a dynamo process takes place due to differential solar heating and lunar gravitational effects. This region is called the ionospheric dynamo region with several current systems including the equatorial electrojet (EEJ) and the mid-latitude solar quiet (Sq). An important fact in this regard is that due to the placement of satellites in LEO the ionospheric currents will be an "internal" source, hence it is important to carefully select data that are believed to be least disturbed.
The magnetosphere is the region in space that determines the magnetic dominance by a celestial ob-

[2]The geomagnetic field is either produced by electric currents such as those in the core or by magnetized media, for which the lithospheric contribution is the only source. This magnetization might be permanent or induced. The quasistatic approximation is valid in this connection and hence there is no significant contribution from the $\partial \mathbf{E}/\partial t$ term in (1).

ject i.e. by confining electrical particles. The Earth's magnetosphere interacts with the IMF, creating the so-called magnetopause on the dayside. Magnetospheric currents include the Chapman-Ferraro currents, tail currents and the ring currents. The ring current can reach values of hundreds of nT in disturbed conditions (i.e. geomagnetic storms).
Geomagnetic indices have been developed in order to quantify and characterize ionospheric and magnetospheric contributions. The indices are numerous and are used in various connections in geomagnetism [Kauristie et al., 2017]. Here we mention the ones that are utilized in this project.

- **Kp:** Kp is an index of global geomagnetic activity. The derivation of the index aims at characterizing geomangetic field disturbances due to particles from solar radiation. It is derived from 13 subauroral stations.
- **RC:** The RC index measures the strength of the magnetospheric ring currents. This index is an alternative to the Dst index that has a time-dependant base-line and thus is difficult to use in modelling. The index is derived in [Olsen et al., 2014] and is used in the CHAOS models.

1.5.2 Core Field and Dynamo Theory

The modern theory for generating and sustaining the geomagnetic field is the so-called dynamo theory[3]. At the core no permanent magnetization can be sustained due to the high temperatures that exceed Curie's temperature. Hence this rotating, convecting and electrically conducting fluid is thought to be the source of a dynamo mechanism that sustains the field. The induction of the magnetic field is caused by the convective outer core (molten iron), supported by kinetic energy from the planet's rotation. The Coriolis force in the neighborhood of the core creates fluid motion of electric currents in the form of columns aligned with the rotation axis of the Earth (Taylor columns). An induction equation can be determined from Maxwell's equations (1) and Lorentz force law together with Ohm's law:

$$\begin{aligned} \mathbf{F} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \qquad \text{(Lorentz force law)} \\ \mathbf{J} &= \sigma \mathbf{f} \qquad\qquad\quad \text{(Ohm's law)} \end{aligned} \tag{2}$$

Where σ is the electrical conductivity, $\mathbf{F}$ is the force and $\mathbf{f} = \mathbf{F}/q$. Using the identity $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot) - \nabla^2$ on $\mathbf{J}$ one obtains the induction equation:

$$\frac{\partial \mathbf{B}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \tag{3}$$

Where $\eta = 1/\mu\sigma$ is called the magnetic diffusivity. The first term in the above equation is the inductive part while the second is diffusive part. In the limit of $\mathbf{v} = \mathbf{0}$ the induction equation reduces to the diffusion equation. Due to the size of the Earth the magnetic field will not disappear (diffuse away) immediately but will decays away in an e-folding period[4] of estimated 20 ka. The geomagnetic field has existed for much longer period than this decay period as shown in paleomagnetic records, so for a self-sustaining dynamo an energy source must exist. Indeed core motion means that $\mathbf{v}$ is nonzero. On the other hand, when conductivity tends to infinity (i.e. $\eta \to 0$) the second term in the induction equation vanishes, the perfect conductor limit and also called the frozen flux approximation since the flux does not change (frozen-in field lines). Finally, the ratio between the inductive and the deffusive part is denoted the magnetic Reynolds number R_m, which can be estimated to be R_m 250, indicating that a geodynamo process is indeed possible.
The theory of geodynamo is rich and involves numerical solutions of complex nonlinear equations in the field of magnetohydrodynamics (MHD). These theories are of no relevance to this project and

[3] The introduction made here is following [Gubbins and Herrero-Bervera, 2007].
[4] e-folding period is the time for the field amplitude to drop by a factor of e.

hence shall not be considered further, but since geodynamo theories are the strongest available ones to satisfactory explain a planetary (celestial) magnetic field then they are of great interest in attempting to bridge modelling work with the geodynamo-theoretic work.

2 Modelling

In this section global geomagnetic modelling of the core field by a spherical harmonic (SH) representation is discussed. There are several methods in geomagnetic modelling, but in the case of dense geographical coverage (such as those provided by satellite data) the SH representation is advantageous and describes the field with satisfactory accuracy. Additionally, in this representation the process of separating internal and external contributions is relatively simple. SH modelling is thus widely applied in field modelling where models are characterized by several factors such as the expansion order, the incorporation of the temporal dependence, techniques in appropriately dealing with unmodeled fields and the use of regularization.
We shall construct a linear model through a SH expansion (as commonly done), discuss time-dependence through spline fitting and then construct the matrix that determines the model parameters through least square (LS) estimation. Improving the LS estimation by reiterative LS estimation will also be discussed and at the end regularization is introduced where its importance and related issues will be analyzed. The *general* methodology here follows that of the CHAOS models, with the most recent one being CHAOS-6 [Finlay et al., 2016], in particular, concerning the use of SH modelling, spline modelling and reiterative LS.

2.1 Spherical Harmonic Representaion

As described in the introduction, a model for the geomagnetic field is obtained by solving Maxwell's equation for $\mathbf{B}$. A global approach by using spherical harmonic expansions is obtained as follows. The core field is assumed to be measured in current-free regions on a surface of a sphere, so that $\mathbf{J} = \mathbf{0}$, and there are no displacement currents so that $\partial\mathbf{E}/\partial t = \mathbf{0}$. In this case the curl and the divergence of $\mathbf{B}$ vanish. A conservative vector field can be written as a gradient of a scalar potential $\mathbf{B} = -\boldsymbol{\nabla} V$, and together with $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$ the problem is reduced to solving Laplace equation for the scalar potential V:

$$\nabla^2 V = 0 \tag{4}$$

Since we are dealing with a spherical geometry, it is natural to solve the equation in spherical coordinates (r, θ, ϕ) in a geocentric system. Separation of variables then leads to solving the radial, polar and azimuthal equations. The field is furthermore divided into two potential fields, one of internal origin and the other of external origin. The potential of the internal field is given by the spherical harmonic expansion:

$$V(r,\theta,\phi) = a\sum_{n=1}^{N}\sum_{m=0}^{n}(g_n^m\cos m\phi + h_n^m \sin m\phi)\Big(\frac{a}{r}\Big)^{n+1} P_n^m(\cos\theta) \tag{5}$$

Where $a = 6371.2$ km is the mean Earth's radius, P_n^m is the associated Legendre polynomials and $\{g_n^m, h_n^m\}$ are the Gauss coefficients for the field of internal origin. It is noted that for the potential of external origin the expansion is similar to equation (5) with another radial dependence and a new set of Gauss coefficients. We remind again that *internal* is with respect to the satellite, hence internal field in this connection may contain contributions from e.g. ionospheric currents.
The field is hence obtained by

$$\mathbf{B} = -\boldsymbol{\nabla} V = -\left(\frac{\partial}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial}{\partial\theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\hat{\boldsymbol{\phi}}\right)V \tag{6}$$

2.2 The Linear Model and LS Estimation

The above equations show that the components of $\mathbf{B}$ depend linearly on the Gauss coefficients $\{g_n^m, h_n^m\}$, hence we can write equation (6) as a matrix equation:

$$\mathbf{d} = \mathbf{G}\mathbf{m} \tag{7}$$

Where $\mathbf{d}$ is the data vector containing the components of $\mathbf{B} = (\mathbf{B}_r, \mathbf{B}_\theta, \mathbf{B}_\phi)$ at some given locations, while $\mathbf{m}$ is a vector of the Gauss coefficients whose length depends on the choice of expansion degree N (denoted the spherical harmonic degree). More specifically its length will be given by $\aleph \equiv N(N+2)$, where e.g. $N = 1$ yields a dipolar field determined by the coefficients $\{g_1^0, g_1^1, h_1^1\}$. Thus a direct relation is established between the field (physical quantity) and the Gauss coefficients (model parameters) by a linear operator $\mathbf{G}$ that maps model parameters to model predictions which in our finite and discrete case is merely a matrix. The matrix $\mathbf{G}$ reflects the geometric dependence between $\mathbf{B}$ and $\{g_n^m, h_n^m\}$ and is a function of the spatial coordinates $\mathbf{r} = (r, \theta, \phi)$. The matrix $\mathbf{G}$ is determined using (6) once a degree N is chosen, and it has dimensions of $(3d \times \aleph)$ where d is the number of data. The matrices for the three dependences are then given by:

$$\mathbf{G}_r = -a\frac{\partial}{\partial r}\begin{pmatrix} \cos(\phi)P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin(N\phi)P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \\ \vdots & \ddots & \vdots \\ \cos(\phi)P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin(N\phi)P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \end{pmatrix}\Bigg|_{\mathbf{r}=\mathbf{d}} \quad \mathbf{G}_\theta = -\frac{a}{r}\frac{\partial}{\partial \theta}\begin{pmatrix} \cos(\phi)P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin(N\phi)P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \\ \vdots & \ddots & \vdots \\ \cos(\phi)P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin(N\phi)P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \end{pmatrix}\Bigg|_{\mathbf{r}=\mathbf{d}}$$
$$\mathbf{G}_\phi = -\frac{a}{r\sin\theta}\frac{\partial}{\partial \phi}\begin{pmatrix} \cos\phi P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin N\phi P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \\ \vdots & \ddots & \vdots \\ \cos(\phi)P_1^0(\theta,\phi)\left(\frac{a}{r}\right) & \dots & \sin(N\phi)P_N^N(\theta,\phi)\left(\frac{a}{r}\right)^{N+1} \end{pmatrix}\Bigg|_{\mathbf{r}=\mathbf{d}} \tag{8}$$

It is remarked that $\mathbf{G} = (\mathbf{G}_r, \mathbf{G}_\theta, \mathbf{G}_\phi)$[5], corresponding to the three components of the field. To make this clear, equation (7) is rewritten as:

$$\mathbf{d} = \begin{pmatrix} \mathbf{B}_r \\ \mathbf{B}_\theta \\ \mathbf{B}_\phi \end{pmatrix} = \begin{pmatrix} \mathbf{G}_r \\ \mathbf{G}_\theta \\ \mathbf{G}_\phi \end{pmatrix}\mathbf{m} \tag{9}$$

Knowing the model parameters $\{g_n^m, h_n^m\}$ will then allow a prediction of the field components at any location by applying (7).
This defines a so-called inverse problem: given data measurements that are assumed to originate from the core together with errors with a certain distribution, how do we determine the *best fitting* model[6]? In other words, the geomagnetic field model is determined *inversely* by utilizing available data and regression techniques. In particular, we may solve equation (7) by means of LS estimation, i.e. by minimizing the sum of the squared residuals[7]. More explicitly, we want to minimize the function f given by:

$$f(\mathbf{m}) = \sum_{i=1}^{d}\left|d_i - \sum_{j=1}^{\aleph} G_{ij}m_j\right|^2 = \|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 = (\mathbf{d} - \mathbf{G}\mathbf{m})^{\mathrm{T}}(\mathbf{d} - \mathbf{G}\mathbf{m}) \tag{10}$$

By equating the gradient of $f(\mathbf{m})$ to zero a unique solution for $\mathbf{m}$ is obtained and is given by the famous least square solution:

$$\mathbf{m} = \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{d} \tag{11}$$

2.3 Time-Dependence Through Spline Modeling

Models constructed by (11) at disposal allow to predict the field components at any desired location. However, since the core field *varies* with time, the temporal dependence must also be incorporated in the model. The time variation of the field is indeed the property that is wished to be investigated in

[5]When the matrices are written in components the notation $(G_{ij}^r, G_{ij}^\theta, G_{ij}^\phi)$ will be used.
[6]More accurately, a model that will satisfy some condition, such as a minimization of a cost function
[7]Where the residual is defined as the difference between the value of the data and the model prediction of that data

this project, so a systematic method to include time is needed and is done as follows. For a dynamic *time-dependent field* the solution of the Laplacian equation (4) is also given by the general form (5) where the time-dependence is carried by the coefficients $\{g_n^m(t), h_n^m(t)\}$[8]. Assuming that we have a time-series of a certain Gauss coefficient g_n^m (or h_n^m) for a given location, can we determine a model that describes the time-dependence smoothly[9] of g_n^m in terms of some coefficients a_n^m? This defines an inverse problem of the time-dependence in the model as described below.
One way of establishing such a relation is by using spline modelling. Splines are functions defined piecewisely by polynomials with a certain order. More formally, a spline is a piecewise polynomial real function $S : [a, b] \to \mathbb{R}$, where the interval $[a, b]$ is subdivided in the following manner:

$$a = t_0 < t_1 < t_2 < \ldots < t_{k-1} < t_k = b.$$

Then S is defined by its restirctions to the subintervals as $P_i : [t_{i-1}, t_i] \to \mathbb{R}$:

$$\begin{aligned} S(t) &= P_0(t), \quad t_0 \le t \le t_1, \\ S(t) &= P_1(t), \quad t_1 \le t \le t_2, \\ &\vdots \\ S(t) &= P_{k-1}(t), \quad t_{j-1} \le t \le t_j. \end{aligned} \tag{12}$$

Where the points t_k are called the knots of the spline. The highest order of P_k is called the order of the spline. We shall only work with cubic splines, i.e. fourth order splines.
As we have seen, splines are represented piecewisely by a set of polynomials. However, splines may also be represented uniquely by using the B-splines (basis spline) for a given set of knots $\{t_k\}$. This is done by representing a spline S in terms of a linear combination of B-splines $M(t, t_k)$[10]:

$$S(t) = \sum_k \alpha_k M(t, t_k) \tag{13}$$

For some set of coefficients α_k. It is remarked that the number k is related to the knots, hence the number of the coefficients α_k is determined by the knots spacing, which in our case is chosen to be equal. The functions $M(t, t_k)$ can be determined by recursive formulas [De Boor et al., 1978], but we shall not need these formulas explicitly since they shall be constructed numerically with MATLAB.

2.4 Merging Temporal and Spatial Dependence

Returning to our original problem, we are now in a position of expressing the Gauss coefficients $\{g_n^m(t), h_n^m(t)\}$ in terms of B-splines that will incorporate time dependence into the geomagnetic model in a smooth manner such that derivatives[11] can be determined. As formulated earlier, given time series of some model paramter for a time vector $\mathbf{t}$, we may expand the parameter into a linear combination in terms of some coefficients that will be denoted $\{(\alpha_n^m)_k, (\beta_n^m)_k\}$:

$$\begin{aligned} g_n^m(\mathbf{t}) &= \sum_k (\alpha_n^m)_k M(\mathbf{t}, t_k), \\ h_n^m(\mathbf{t}) &= \sum_k (\beta_n^m)_k M(\mathbf{t}, t_k). \end{aligned} \tag{14}$$

[8] The differential operator ∇^2 does not involve a time dependence.
[9] Smoothness is necessary as the aim is to find the secular variation, i.e. time-derivative of the field.
[10] Here $M(t, t_k)$ is used as an alternative for the more reasonable notation of $B(t, t_k)$ to avoid confusions with the magnetic field
[11] We shall only be considering the first derivative in this project, i.e. the secular variation.

Which can be expressed as the matrix equations:

$$\mathbf{g} = \mathbf{M}\boldsymbol{\alpha}, \quad \mathbf{h} = \mathbf{M}\boldsymbol{\beta} \tag{15}$$

This means that the new set of coefficients determine the model parameters for a geomagnetic model, since the expressions in (14) may be inserted in (5) and hence $\{(\alpha_n^m)_k, (\beta_n^m)_k\}$ determine the field uniquely for any given time and location. Since the components of $\mathbf{B}$ depend linearly on the Gauss coefficients which in return depend linearly on the set of coefficients $\{(\alpha_n^m)_k, (\beta_n^m)_k\}$, a matrix $\mathbf{\Lambda}$ can be constructed for the determination of model parameters. Denoting the matrix elements of $\mathbf{G}$ and $\mathbf{M}$ by G_{ij} and M_{dk} where the index d is for the time vector $\mathbf{t} = (t_1, \ldots, t_d)$ and k is for the knots, the vector $\mathbf{d}$ may now be written explicitly as:

$$\mathbf{d} = \begin{pmatrix} \mathbf{B}_r \\ \mathbf{B}_\theta \\ \mathbf{B}_\phi \end{pmatrix} = \begin{pmatrix} B_r(\mathbf{r}_1, t_1) \\ \vdots \\ B_r(\mathbf{r}_d, t_d) \\ B_\theta(\mathbf{r}_1, t_1) \\ \vdots \\ B_\theta(\mathbf{r}_d, t_d) \\ B_\phi(\mathbf{r}_1, t_1) \\ \vdots \\ B_\phi(\mathbf{r}_d, t_d) \end{pmatrix} = \begin{pmatrix} G^r_{11}\sum_k(\alpha^0_1)_k M(t_1,t_k) + \ldots + G^r_{1\aleph}\sum_k(\beta^N_N)_k M(t_1,t_k) \\ \vdots \\ G^r_{d1}\sum_k(\alpha^0_1)_k M(t_d,t_k) + \ldots + G^r_{d\aleph}\sum_k(\beta^N_N)_k M(t_d,t_k) \\ G^\theta_{11}\sum_k(\alpha^0_1)_k M(t_1,t_k) + \ldots + G^\theta_{1\aleph}\sum_k(\beta^N_N)_k M(t_1,t_k) \\ \vdots \\ G^\theta_{d1}\sum_k(\alpha^0_1)_k M(t_d,t_k) + \ldots + G^\theta_{d\aleph}\sum_k(\beta^N_N)_k M(t_d,t_k) \\ G^\phi_{11}\sum_k(\alpha^0_1)_k M(t_1,t_k) + \ldots + G^\phi_{1\aleph}\sum_k(\beta^N_N)_k M(t_1,t_k) \\ \vdots \\ G^\phi_{d1}\sum_k(\alpha^0_1)_k M(t_d,t_k) + \ldots + G^\phi_{d\aleph}\sum_k(\beta^N_N)_k M(t_d,t_k) \end{pmatrix} \tag{16}$$

Now that the explicit dependence of the field is known in terms of the unknown model parameters $\{(\alpha_n^m)_k, (\beta_n^m)_k\}$, the matrix $\mathbf{\Lambda}$ relating the field to these parameters is needed in order to inversely obtain a model that accounts for both the spatial and the temporal dependence of the field, i.e. we need to construct $\mathbf{\Lambda}$ such that:

$$\begin{pmatrix} B_r(\mathbf{r}_1, t_1) \\ \vdots \\ B_\phi(\mathbf{r}_d, t_d) \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & \ldots & \Lambda_{1(\aleph\kappa)} \\ \vdots & \ddots & \vdots \\ \Lambda_{(3d)1} & \ldots & \Lambda_{(3d)(\aleph\kappa)} \end{pmatrix} \begin{pmatrix} (\alpha^0_1)_1 \\ \vdots \\ (\beta^N_N)_\kappa \end{pmatrix} \tag{17a}$$

$$\mathbf{d} = \mathbf{\Lambda}\mathbf{m} \tag{17b}$$

Where κ is the number of knots assigned to each Gauss coefficient. Since the number of $\{g_n^m(t), h_n^m(t)\}$ for a given N is $\aleph$, the length of the vector $\mathbf{m}$ will be $\kappa\aleph$. For a time-dependent model, the model parameters are $\{(\alpha_n^m)_k, (\beta_n^m)_k\}$[12], related linearly to $\mathbf{B}$ by the matrix Λ as shown in equations (17). It remains to determine $\mathbf{\Lambda}$ whereupon an inversion can be preformed to obtain the desired model parameters.

Equation (16) shows the product of $\mathbf{\Lambda m}$. The matrices $\mathbf{G}$ and $\mathbf{M}$ are already known, where the former relates the magnetic field to the Gauss expansion coefficients through the spatial geometry of the solution, and the former establishes a time-dependence to each Gauss coefficient through a series expansion. The following choice (starting with the radial component) can easily be shown to fulfill the needed task by multiplying the given matrices:

$$\mathbf{\Lambda}_r = \begin{pmatrix} G^r_{11} & & & \cdots & G^r_{1\aleph} & & \\ & G^r_{21} & & \vdots & & G^r_{2\aleph} & \\ & & \ddots & \vdots & & & \ddots \\ & & G^r_{d1} & \ldots & & & G^r_{d\aleph} \end{pmatrix} \begin{pmatrix} \mathbf{M}(\mathbf{t};\mathbf{t_k}) & & \\ & \mathbf{M}(\mathbf{t};\mathbf{t_k}) & \\ & & \ddots \\ & & \mathbf{M}(\mathbf{t};\mathbf{t_k}) \end{pmatrix} \tag{18}$$

[12]Which from now on will be labeled by $\mathbf{m}$ to retain the notation of equation (7).

Where the empty space in the matrices represent zeros, and the matrix $\mathbf{M}(\mathbf{t};\mathbf{t_k})$ is defined earlier in (15). Additionally we can see that the dimension of $\mathbf{\Lambda}_r$ is as expected:

$$\dim(\mathbf{\Lambda}_r) = \dim\Big\{(d \times d\aleph)(d\aleph \times \kappa\aleph)\Big\} = d \times \kappa\aleph$$

However, we need to use the data from all the available three components of the field, so a similar procedure is done for $\mathbf{G}_\theta$ and $\mathbf{G}_\phi$. With those at disposal, we may write the above construction together with the ϕ and θ part in a single matrix in the following compact way:

$$\mathbf{\Lambda} = \begin{pmatrix} \mathrm{diag}\big(\mathbf{G}^r_{j1}\big) & \dots & \mathrm{diag}\big(\mathbf{G}^r_{j\aleph}\big) \\ \mathrm{diag}\big(\mathbf{G}^\theta_{j1}\big) & \dots & \mathrm{diag}\big(\mathbf{G}^\theta_{j\aleph}\big) \\ \mathrm{diag}\big(\mathbf{G}^\phi_{j1}\big) & \dots & \mathrm{diag}\big(\mathbf{G}^\phi_{j\aleph}\big) \end{pmatrix} \begin{pmatrix} \mathrm{diag}\big[\mathbf{M}(\mathbf{t};\mathbf{t_k}) & \dots & \mathbf{M}(\mathbf{t};\mathbf{t_k})\big] \\ \mathrm{diag}\big[\mathbf{M}(\mathbf{t};\mathbf{t_k}) & \dots & \mathbf{M}(\mathbf{t};\mathbf{t_k})\big] \\ \mathrm{diag}\big[\mathbf{M}(\mathbf{t};\mathbf{t_k}) & \dots & \mathbf{M}(\mathbf{t};\mathbf{t_k})\big] \end{pmatrix} \tag{19}$$

Where the index j here has the values $j = 1, \dots, d$. The matrix Λ will be used in all model computations done in this project.
initially, a solution to (17) can be obtained by the LS estimation described earlier, i.e.

$$\mathbf{m} = \big(\mathbf{\Lambda}^{\mathrm{T}}\mathbf{\Lambda}\big)^{-1}\mathbf{\Lambda}^{\mathrm{T}}\mathbf{d} \tag{20}$$

2.5 Iteratively reweighted LS Estimation

Estimating model parameters is very common in geophysics in order to make reasonable inference from available data. Earlier we have introduced least square estimation as an initial approach to determine the desired model parameters. However, by doing so one is assuming that the errors of the model are Gaussian (normally distributed), an assumption that is not necessarily valid [Constable, 1988]. Some or the majority of the errors may be Gaussian, but even a small fraction of outliers may result in a considerable negative effect on the models. In least square estimation one is minimizing the sum of the squares of the residuals (the loss function). This loss function is however not *robust* to outliers and hence would cause an unwanted effect in the models due to the big contribution from their side on the squares of the residuals. What is needed is a robust regression method which in our case will be the iteratively reweighted least squares (IRLS) with the weights w being the Huber weights. The method and Huber weights are described in detail in [Olsen et al., 2014].
The estimator will be different compared to ordinary LS estimator introduced in equation (10). In this case the minimizer will contain a diagonal weight matrix $\mathbf{W}$:

$$\Theta(\mathbf{m}) = (\mathbf{d} - \mathbf{G}\mathbf{m})^{\mathrm{T}}\mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}) \tag{21}$$

Taking the gradient and equating to zero as done previously with ordinary LS yields the unique solution:

$$\mathbf{m} = \big(\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G}\big)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{d} \tag{22}$$

It is noted that the LS solution introduced earlier in eq. (11) is only a special case of the weighted least square estimator since in ordinary LS all errors are equally weighted, i.e. $w_i = 1$. This method is iterative in the sense that the weights (initially taken to be $w_i = 1$) are iteratively calculated until the matrix $\mathbf{W}$ converges. The iteration number N_{it} can hence be found by saving the weights and choosing the number that achieves convergence, which in our case is determined to $N_{\mathrm{it}} = 6$.

2.6 Regularization

2.6.1 Regularization in Field Modelling

It is now worth pausing for a brief recapitulation of the objectives of our work. The aim is to build models that describe the core (main) field and allow us to investigate problems related to unmodelled contributions. The modelling is done by a spherical harmonic representation of the field which allows solving a linear equation that yields the model parameters. This is done on the assumption that the satellite is measuring the core field (superposed with a crustal and external modeled contribution) and hence the delivered data contains measurements of the core together with errors as independent random variables with some probability density function that initially is assumed to be a normal distribution. In our case this is clearly an invalid assumption as we know that measurements are contaminated by nonmodeled signals (in particular ionospheric and magentospheric currents) and hence models will lack robustness due to the fact that the loss function (sum of the squares of the residuals) has a tendency to be dominated by outliers. The reiteratively weighted least squares partially solves this problem with the use of the Huber loss function.
This, however, does not eliminate the nonmodeled signals and one must impose a further restriction on the estimator. In addition to minimizing the sum of the squares of the residuals, an additional functional is minimized that characterizes the "roughness" of the curve. More specifically, we seek a solution of the class $C^2[.,.]$[13] with a misfit to data measured by the LS loss function and an additional term that can be adjusted by a parameter λ that is denoted the *smoothing parameter*:

$$f(\mathbf{m}) = (\mathbf{d} - \mathbf{Gm})^{\mathrm{T}}\mathbf{W}(\mathbf{d} - \mathbf{Gm}) + \lambda \mathbf{m}^T \mathbf{R}\mathbf{m} \tag{23}$$

Where $\mathbf{R}$ is the regularization matrix. The second term in the equation above is motivated as follows: Focusing on one Gauss coefficient (call it g) from $\{g_n^m(t), h_n^m(t)\}$, we want to represent g by a spline expansion as done in equation (13). The roughness (which shall be minimized for a given λ) will be represented by the functional:

$$F[S(t)] = \lambda \int_{t_1}^{t_N} \left(\partial_t^2 S(t)\right)^2 dt = \lambda \int_{t_1}^{t_N} \left[\partial_t^2 \sum_k \alpha_k M(t, t_k)\right]^2 dt \tag{24}$$

Where we use $S(t) = \sum_k \alpha_k M(t, t_k)$ as a spline expansion. This term can then be written in matrix expression as $\lambda \mathbf{m}^T \mathcal{R} \mathbf{m}$ if we write the elements of $\mathcal{R}$ as:

$$\mathcal{R}_{ij} = \int_{t_1}^{t_N} \partial_t^2 M_i(t) \partial_t^2 M_j(t) dt \tag{25}$$

Where $\mathbf{m}$ will contain the expansion parameters $\{(\alpha_k)\}$. With this in mind we may build $\mathbf{R}$ of equation (23) to include all the Gauss coefficients by diagonally concatenating $\aleph = N(N+2)$ identical versions of $\mathbf{R}$:

$$\mathbf{R} = \mathrm{diag}\left[\overbrace{\mathcal{R}\ ,\ \ldots\ ,\ \mathcal{R}}^{\aleph \text{ times}}\right] \tag{26}$$

The unique minimizer to equation (23) is derived analogously to previous solutions (taking the gradient and equating to zero):

$$\mathbf{m} = \left(\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G} + \lambda \mathbf{R}\right)^{-1} \mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{d} \tag{27}$$

[13] where C^2 denotes functions of the second class of differentiability, i.e. twice continuously differentiable functions on some interval, granting the first two time-derivatives.

This method is called *spline smoothing*, where the smoothing parameter λ determines the degree of "smoothing". In the case $\lambda \to 0$ we recover equation (22), and for increasing λ the term $\lambda\mathbf{R}$ (smoothing term) dominates over $\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G}$ in (27) above. Models built with equation (27) will be denoted **R**-regularized models.

2.6.2 The Regularization Issue

As described earlier the effect of regularization is minimizing the "roughness" (integral of the squared second derivative) of the curve appropriately. Regularization is very useful in geomagnetic modeling, for instance in ionospheric signals in geomagnetic observatory measurements [Constable and Parker, 1988]. It is also used extensively in main field modeling as an approach to deal with unmodeled signals as done in the CHAOS-6 model [Finlay et al., 2016][14]. Most of the current field models are built through imposing regularization as a necessary approach to ensure convergence in spatial and time domain. In figure (1) we see the effects of regularization on example models of one year knot spacing at ground observatories in three different geographical locations. Unregularized models exhibit a high level of oscillations compared to ground observatory data, and they risk to explode when downward-continued towards the core.
However, some negative consequences follow along implementing the kind of regularization discussed above (**R**) in main field modeling. We recall that our model parameters allow to make predictions at any given location and at any given time within a certain domain. In the context of core dynamics one particular region is of great interest, namely the **core mantle boundary** (**CMB**), lying in the region between the silicate mantle and the iron-nickel core and is located approximately 2891 km deep from the Earth's surface. This region is important for inferring physical processes taking place in the fluid core. The smoothing as forced by regularization can however remove some of the important signals of the core as we downward-continue towards core regions such as the CMB.
The noise[15] that is superposed with the core contribution hampers a satisfactory modeling process in which excessive regularization is not needed. Ideally one wishes to mathematically model the unwanted signals to systematically filter them out, but the complexity of the ionospheric and magnetospheric currents makes such an endeavour a rather difficult task.
In what follows we shall attempt to use an alternative method of regularization as proposed by [Gillet et al., 2013].

2.6.3 An Alternative Approach: AR-2 Process

The project shall attempt to utilize an alternative method to the regularization introduced previously as they yield unrealistic covariances. The details of this alternative regularization is rather complicated and shall not be discussed here. We refer to [Gillet et al., 2013] for details and suffice to note that using this method allows for model solutions with more complex time-dependence than models that are derived minimizing smoothing norms, and hence not suppressing rapid changes in the SV that might be caused by core-related events. The aim is then to use this approach to examine whether (and to which degree) it is possible to eliminate unmodelled fields from the ionosphere and magnetosphere. The **R** in equation (27) will be replaced by the new matrix, denoted $\mathbf{R}_{\mathrm{AR2}}$. This will be the equation that yields the models we shall investigate:

$$\mathbf{m} = \left(\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{G} + \lambda\mathbf{R}_{\mathrm{AR2}}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{W}\mathbf{d} \tag{28}$$

[14]The CHAOS-6 model (and previous versions), impose a more complicated version of regularization then the one mentioned in (27). The minimized cost function includes terms involving data error covariance matrix and two regularization matrices penalizing the squared values of the second and third time-derivatives of B_r at CMB. See equation (3) in [Finlay et al., 2016]

[15]More correctly, the unmodelled contributions to the field.

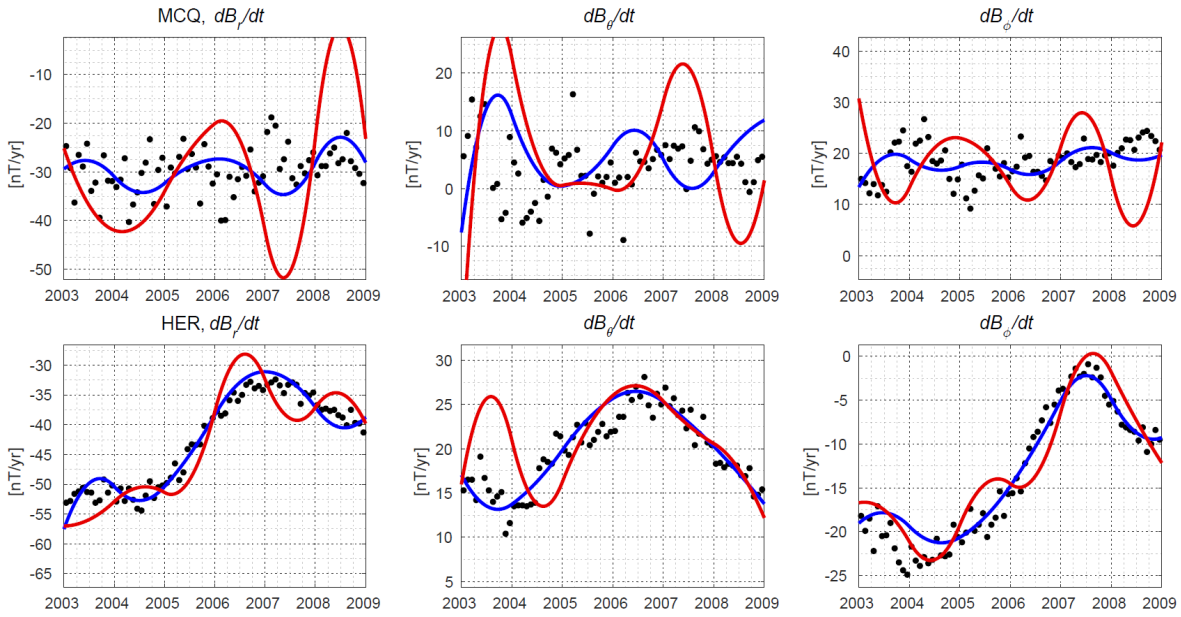


Figure 1: The effect of implementing regularization is seen in the values of SV at MCQ and HER ground observatories (see table 1), here for models with one year knot spacing. Black dots: monthly mean ground observatory; red curves: non-regularized model predictions; blue curves: $\mathbf{R}_{\mathrm{AR2}}$-regularized model prediction (introduced in 2.6.3).

Since $\mathbf{R}_{\mathrm{AR2}}$ will not impose a (as) strict form of smoothing, the models obtained by using this new matrix will be very oscillatory (as we shall see in later sections), in part due to the ionospheric and magnetospheric currents. As mentioned, the reason for using $\mathbf{R}_{\mathrm{AR2}}$ is to avoid potential suppression of weak core signals by the smoothing imposed by matrices such as $\mathbf{R}$ as modellers need to do today.

2.7 Computing Model Parameters

The model parameters are computed using the methods introduced in the modelling subsections above. Here we shall give an outline of the actual implementation of the computation process.
The goal is to determine the set $\left\{(\alpha_1^0)_1, (\alpha_1^0)_2, ..., (\beta_N^N)_{\kappa-1}, (\beta_N^N)_\kappa\right\}$ given the data $\mathbf{d} = (\mathbf{B}_r, \mathbf{B}_\theta, \mathbf{B}_\phi)$, a specified expansion order N and the observation matrix $\mathbf{\Lambda}$. The following list summarizes the procedure of obtaining models:

- The matrices $\mathbf{G} = (\mathbf{G}_r, \mathbf{G}_\theta, \mathbf{G}_\phi)$ given by the expressions in (8) are determined using a script called `design_SH.m` (Olsen, 2003). The script determines the matrices $\mathbf{G}$ given the data and an expansion order N. Unless otherwise stated the models will have $N = 16$, an appropriate choice for core field.

- The matrices $\mathbf{M}$ are the B-spline collocation matrices. The matrix is constructed using MATLAB's spline toolbox command `spcol(knots,k,t)`. This returns a matrix of dimension $(d \times \kappa)$ with (i, j) entries of $D^{m(i)}B_j(t(i))$ [de Boor, 1997]. This is the value of the $m(i)$'th derivative of the j'th spline of order k for the knots κ evaluated at $t(i)$. The spline degree will be set to $k = 4$, i.e. a third degree or cubic (spline) polynomial, and the knots κ will be given by the model span $(2002 : \Delta t : 2010)$ with a *knot spacing* Δt. We shall build models that have knot spacing of 0.5 yr and 1 yr. However, the final tests shall be implemented on 0.5 yr knot spacing. The utility of one year knots models are limited as we wish to resolve core signals on a higher temporal resolution. Figure (2) shows an example of a one year vs. a half year knot spacing model. It

is clearly seen how the one year model suppresses some oscillations seen in some of the Gauss coefficients (to be discussed later).

- With the matrices $\mathbf{G}$ and $\mathbf{M}$ at disposal the observation matrix $\mathbf{\Lambda}$ is straightforwardly computed using equation (19). This matrix is then a function of t, $\mathbf{r}$ and N, i.e. $\mathbf{\Lambda} = \mathbf{\Lambda}(t, \mathbf{r}, N)$.

- Once $\mathbf{\Lambda}$ is given, all the tools needed to build a model for the core field are then available. Using the LS estimation as given by equation (20) one obtains the models parameters $\mathbf{m} = \left\{(\alpha_1^0)_1, , ..., (\beta_N^N)_\kappa\right\}$. However, in the residuals of the models are noticed some outliers, i.e. values that exhibit an extreme departure from all other residuals. They are not numerous but their great residual value affects the quality of the model. Since these are not related to the inversion process they are removed from the data set and a new inversion is made.

- To handle the problem of weighing the residuals equally a reiterative LS estimation is performed. The algorithm is rather simple and is shown in the MATLAB code in the appendix. As mentioned before, the iteration number is shown to be $N_{\text{it}} = 6$, so for each model 6 other inversions must be made before obtaining the solution from equation (22).

- As an additional constraint one needs to make a choice of regularization. In (2.6.2) we discussed the unwanted effects of the regularization methods used in most current field models. We shall mainly construct model solutions that use the covariance matrix R_{AR2} as an alternative approach. With this approach one may investigate the possibility of improvements in field models using data selection based on different criteria.

- From model parameters one may obtain the so-called pp-form. This is the polynomial version of $\mathbf{m}$ that can be used to evaluate the Gauss coefficients $\{g_n^m(t), h_n^m(t)\}$ and their time derivative $\left\{\dot{g}_n^m(t), \dot{h}_n^m(t)\right\}$ at a specific time. Additionally, we mention that in various figures only results in the interval $[2003, 2009]$ are shown, instead of the actual model span of approximately $[2002, 2010]$. This is due to the instability of models at the vicinity of the first and the final knot.

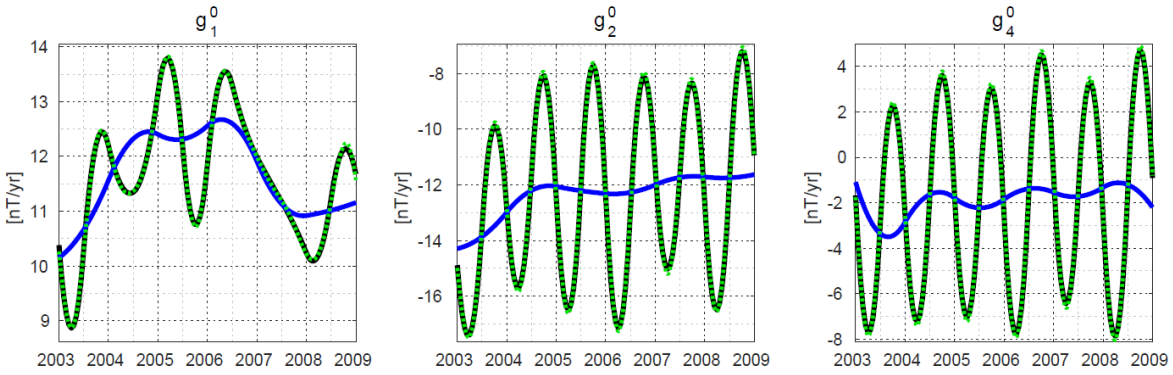


Figure 2: The effect of changing knot-spacing from one year to half year on time-series of the time derivatives of $\left\{g_1^0, g_2^0, g_4^0\right\}$ for $\mathbf{R}_{\text{AR2}}$-regularized models. Blue curves: model prediction of one year knot spacing. Black curves: model predictions of half-year knot spacing. Green circles: a fit of degree 8 Fourier series to the half-year spaced model, used in 4.2.1 for an estimation for the periods of the signals.

3 Data

In the following we shall describe the data used in building the main field models. Raw data cannot be used directly in the computations since they are containing many contributions that are not desired to be included in the model. Hence one needs to perform some selections on the data set itself before applying the modelling machinery.

3.1 Data from CHAMP

The data set provided is taken from the satellite CHAMP. The dataset constitutes 16 arrays **of length 68077**. It includes an array of time in the unit of md2000 (days after 1st January 2000) and spans a period of 8.7 years with a temporal resolution of 10 minutes. The spatial representation is in geocentric spherical coordinates (r, θ, ϕ), given in units of km and degree. The three components (B_r, B_θ, B_ϕ) of the observed field are given in units of nT.
The data set also includes field components of the following contributions:

- CHAOS-6 predictions of the crustal field.
- CHAOS-6 predictions of the external field.

This amounts to removing the modelled part of the field which is not related to the core. Subtracting these from the observed data then yields the corrected field[16]:

$$\mathbf{B}_{\text{corr}} = \mathbf{B}_{\text{obs}} - \mathbf{B}_{\text{lith}} - \mathbf{B}_{\text{ext}} \tag{29}$$

This data set will be the source of the models that shall be built[17]. We remind that the models built in what follows are by no means aiming at replacing existing models, but rather trying to address some specific problems that modellers face today[18]. When modellers build models (such as the CHAOS models) they would use as many available data as possible, but in our case the choice of 68077 would allow to test some hypotheses without the need to perform immensely time-requiring computations.

3.1.1 Initial Selections

Besides the corrections due to the lithospheric and external field there are more selections on the data that are done. They are listed here:

- Data taken from dark areas, i.e. when the Sun is 10° below the horizon. This minimizes disturbances from the ionosphere that intensifies when directly exposed to solar radiation.
- $|\mathrm{d}RC/\mathrm{d}t| < 2$ nT/hr. The RC index has been mentioned earlier in the introduction. This condition is associated with magnetically quiet times [Finlay et al., 2017], i.e. a condition on the field due to the magnetospheric ring current.
- The index Kp of the global geomagnetic activity. The constraint here is Kp$\leq$2° as adopted by the CHAOS models.
- Outliers (greater than 500 nT) from the CHAOS-6 models are removed in any vector component.

These are prior selections with documented effects [Finlay et al., 2017]. The following subsection discusses further selections that will be the center of our work.

[16] Corrected within the framework of the available models. This surely does not correct for other unmodelled signals in the satellite altitude from e.g. atmospheric currents

[17] From now on **data** will mean the corrected data in the sense of equation (29).

[18] The nature of these problems will be investigated in further details in the next section.

3.2 Data Selections

To which extent is it possible to exclude unwanted signals by means of further data selections? This is one of the questions that we shall attempt to address in this project. Data selections are indeed very important in field modelling and is used in most available models today to handle some of the more rapidly-varying current systems in the ionosphere and magnetosphere. The disturbances from these external signals and potentially their induced currents inhibits the quality of probing the physical properties of the Earth's interior, so the success of core field models are significantly dependent on how to handle these unwanted contributions.
As an initial step the available data has been filtered based on some of the solar-driven disturbances and indices for geomagnetic quiet times, as discussed earlier. This has been done beforehand, but this selection (although important) is not adequate as we shall see. In what follows we shall discuss data-selection criteria based on data from the interplanetary magnetic field (IMF) and satellite-based auroral electrojet (AE) indices.

3.2.1 Interplanetary Magnetic Field (IMF)

The Sun forms a strong magnetic field that extends throughout the Solar system. The magnetic field is carried into space by the electrically conducting plasma of the solar wind. As mentioned earlier, the IMF interacts with the Earth's magnetosphere causing disturbances in the data set. The Low Resolution OMNI (LRO) data set[19], however, provides additional options for data selection based on the IMF. The LRO data set includes updated compilation of hourly-averaged, near-Earth components of the IMF and solar wind speed. The data set is obtained from several spacecrafts in geocentric or L1 (Lagrange point) orbits. The components of IMF are given in Geocentric Solar Magnetospheric (GSM) orthogonal coordinates (B_x, B_y, B_z) where B_x and B_y are parallel to the ecliptic, while B_z is perpendicular to the ecliptic. In particular, we shall consider certain criteria on the magnitude and sign of the B_z and the B_y components of the IMF. Additionally, we shall consider a quantity that is derived from the IMF and information on the solar wind speed, namely the merging electric field at the magnetopause, denoted E_m. Although the sunlit data are excluded from our data set, observations from the night-side may also be influenced by the IMF [Friis-Christensen et al., 2017].
The LRO data set is given in a temporal resolution of 1 min in unit of (MD2000). The data set also includes the B_z, B_y components of the IMF and the speed of the solar wind. The measurements are done in the L1 point between the Earth and the Sun. In order to perform selections on the data set from CHAMP, we match the times of the two arrays and for each point in t we take an average for the past two hours. The following is a short description of the criteria based on the IMF:

- $\boldsymbol{B_z}$: This component of the IMF is an important factor for auroral activity as discussed in [Kauristie et al., 2017]. The component B_z amounts to a north-south direction relative to the GSM system, and when this component is pointing southward (i.e. $B_z < 0$) it will connect to the Earth's magnetosphere which in the GSM system is pointing northward. Hence when $B_z < 0$ the solar wind particles have a substantial chance of entering the magnetosphere and guided by the Earth's field lines they continue their path towards the lower part of the atmosphere. We shall set a condition that removes data where $B_z < 0$ for an average of past two hours.

- $\boldsymbol{B_y}$: Recent work (Christensen, et. al 2017) shows that a condition on the B_y component of the field has an effect on the ionospheric currents. It is indicated that (in night-time) B_y has a negative effect in the northern polar region for $B_y > 0$ and correspondingly in the southern polar region for $B_y < 0$. These will equivalently be a condition that shall be tested.

[19] Available at: `ftp://spdf.gsfc.nasa.gov/pub/data/omni/high_res_omni/`

- $\boldsymbol{E_m}$: This represents the merging electric field at the magnetopause which characterizes magnetic reconnection rate at the magnetopasue on dayside. An expression of E_m has been derived in [Kan and Lee, 1979] and revised in [Newell et al., 2007]. In the spirit of CHAOS-6 [Finlay et al., 2016] we shall use the revised expression of E_m, given by:

$$E_m = 0.33 v^{4/3} B_t^{2/3} \sin^{8/3}(|\Theta|/2) \tag{30}$$

where v is the solar wind speed, $B_t = \sqrt{B_y^2 + B_z^2}$ and $\Theta = \arctan(B_y/B_z)$. A condition that removes data where $E_m < 0.8$ mV/m will be implemented[20].

3.2.2 Satellite-based Auroral Electrojet (AE) Indices

Electrojets are current systems travelling in the ionosphere at an altitude of 90-150 km (E-region). The electrojets are divided in equatorial electrojets (EEJ) and the auroral electrojet (AE), placed in the vicinity of the polar circles. There are indices that describe the activity of the AE, and we shall test the application of recent satellite-based auroral indices [Aakjær et al., 2016] to exclude data that are disturbed by auroral electrojet currents as indicated by the. The data includes:

- Time array describing times of the first data point along an orbit, in unit of (md2000).
- maximum and minimum values of the current density along an orbit given by the start time of the orbit, denoted $J_{aur,max}$ and $J_{aur,min}$, respectively. These are measured in A/km.
- Integrated absolute value of the current along the orbit, denoted J_{L1}. This is measured in A.

This additional criterion has been included for testing satellite-based AE indices, and selections based on these indices require the determination of an appropriate threshold. One way of estimating such a threshold is by performing several models as the threshold is varied. One could choose to consider the quantiles of the given indices, e.g. 25%, 50% and 75%, and then choose the quantiles as thresholds, where for each choice some model characteristics are considered and compared. Our choice is to use the median as threshold, as it is shown that a combination of the AE indices and other selections offer an improvement in the statistics of the model (see the discussion in 4.3.1).
There is no physical ground for such a determination of the threshold, and the choice has by no means been determined by an extensive investigation, in part due to the computational cost of the task. However, the availability of the satellite-based AE indices allows us to perform simple tests in our framework, and is indeed motivated by the belief that many of the disturbances that we shall meet are polar/auroral driven.

Nr.	Name of Observatory	Latitude	Longitude	Country
1.	Niemengk (NGK)	52.07°N	12.68°E	Germany
2.	Hermanus (HER)	34.43°N	19.23°E	South Africa
3.	Kakioka (KAK)	36.23°N	140.18°E	Japan
4.	Macquarie Island (MCQ)	54.5°S	158.95°E	Australia

Table 1: Geographical locations of the ground observatories that shall be used for comparison with model predictions.

The thresholds of $J_{L1}, J_{\max}$ and $J_{\min}$ are given by 581140 A, 126.2 A/km and -129.9 A/km, respectively, as determined by the median of the data set.

[20] The values of $E_m^* = E_m/0.33$ are calculated in advance in the given data set on IMF, hence the condition becomes $E_m^* < 0.80/0.33$.

Number of data		Intersections with E_m	
E_m	46148	$\bigcup N$	61203
B_z	34972	$E_m \cap B_z$	33196
B_y	15892	$E_m \cap B_y$	12720
J_{L1}	25903	$E_m \cap J_{L1}$	17881
$J_{\max}$	23943	$E_m \cap J_{\max}$	16718
$J_{\min}$	24329	$E_m \cap J_{\min}$	16383

Table 2: The criteria to be tested on different models together with the number of data removed from each contribution. $\cap$ denotes intersection, while $\bigcup N$ denotes the union of all criteria. The number of our given data set from CHAMP is 68077.

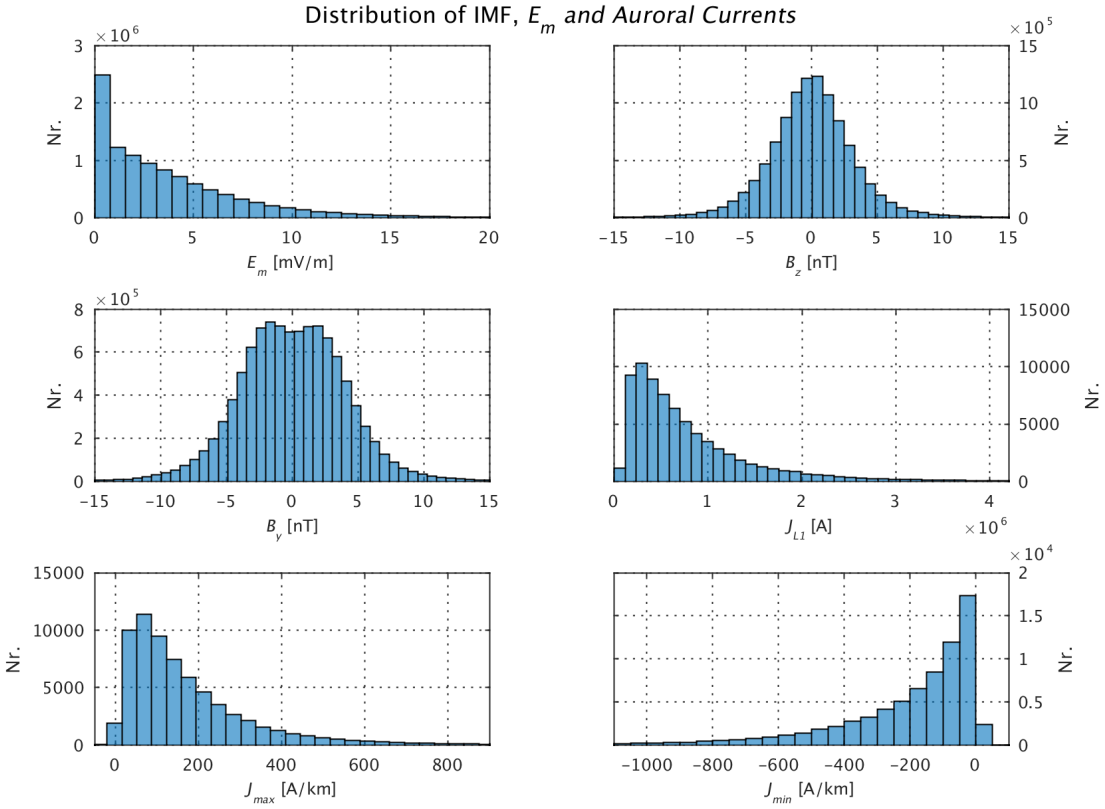


Figure 3: Histograms showing the distribution of the different data selection based on the IMF and the auroral indices.

4 Results and Discussion

The aim of this section is twofold. First we shall attempt to characterize the oscillations in the models regularized with $\mathbf{R}_{\mathrm{AR2}}$. It is of great interest to know the behaviour of these signals with regard to their temporal and geographical distribution. A proper understanding of the signals may allow us to investigate their origin and place modellers in better position to dealing with them. Next we attempt to address the question posed earlier: whether and to which degree it is possible to contain the unwanted signals by means of data selection based on knowledge and available data/indices of current systems in the ionosphere and magnetosphere. The attempt made here is rather modest, both in the selection criteria and in the number of data used to obtain the models. Furthermore, in this section models where all three components (B_r, B_θ, B_ϕ) are used in the inversion are denoted **Type I** models, whereas models where (B_θ, B_ϕ) are excluded from polar latitudes in the inversion are denoted **Type II** models. We shall be using both types of models in the analysis below. Furthermore, some of the figures that will be referred to in this section are placed in the appendix. This is done in order to make the section more readable.
Before proceeding to the results, we give a brief description of the different series of models that have been built during the course of this project.

4.1 Model Series

Although we shall focus on the results of one series of models in the project, the path that was taken to arrive at this series includes several other series of models. Here an outline of these model-series together with some comments on their outcome is given, after which they shall not be considered further in greater detail. The aim here has not been different to our general setting: test how data selection affects signals in model predictions.
In our framework, the compexity of the model time-dependence can be controlled either changing the knot spacing or in the choice of regularization[21]. All four possibilities have been considered and a series of models have been built in order to characterize the model solutions:

- $\mathbf{R}$-regularized models of one year knot spacing.
- $\mathbf{R}$-regularized models of half year knot spacing.
- $\mathbf{R}_{\mathrm{AR2}}$-regularized models of one year knot spacing.
- $\mathbf{R}_{\mathrm{AR2}}$-regularized models of half year knot spacing.

Where $\mathbf{R}$ and $\mathbf{R}_{\mathrm{AR2}}$ denote the two regularization matrices introduced in 2.6.1 and 2.6.3, respectively. As mentioned in the sections on regularization, $\mathbf{R}$-regularized models are subject to smoothing conditions to ensure a realistic model in terms of convergence in spatial and time domain. The first two series of models listed above utilize such a regularization by equation (27). However, $\mathbf{R}$-regularized models are not optimal for testing our objective, i.e. the effect of data selections. One problem in $\mathbf{R}$-regularization is finding an appropriate smoothing parameter λ that serves our purpose. On one extreme, *unregularized* models (i.e. $\lambda \to 0$) exhibit signals with very high amplitudes such that inference based on data selection is very difficult. On the other extreme, smoothed models (i.e. $\lambda \to \infty$) suppress the actual signals that we are interested to investigate.
This leaves $\mathbf{R}_{\mathrm{AR2}}$-regularized models as an alternative, where the "roughness" of the models (see 2.6.1) is not *strongly* penalized. Choosing models of one year knot spacing of has its own disadvantage, as they conceal some of the periodic signals present in the time-series, as seen in figure (2). Furthermore, they are of limited use due to their poor resolution. Hence we have made the choice of considering

[21] Of course, specific data selections, as we shall see, also offer a variety of models, but here we distinguish various models in the *modelling framework* and various models in the context of *which choice of data to be used in the models.*

$\mathbf{R}_{\text{AR2}}$-*regularized models of half year knot spacing* to perform our desired analysis.
In the remainder of this section, the models discussed will be $\mathbf{R}_{\text{AR2}}$-regularized models of half year knot spacing. The models will either be of type I or type II, i.e. based on the field components used at polar latitudes as described in the beginning of this section.

4.2 Characterizing the Oscillations

4.2.1 The Amplitudes

When model solutions are obtained by equation (28), a certain behaviour is seen in the model characteristics, e.g. the time series of the time-derivative of the Gauss coefficients[22]. An example of this behavior is seen in the models of half year knots as the black curves in figure (2) of the coefficients $\{g_1^0, g_2^0, g_4^0\}$.
At first sight one clearly notices the periodic behaviour of these specific signals. In particular, the coefficients g_2^0 and g_4^0 exhibit dominance in a one-year periodicity. To investigate the periodicity property further we may fit Fourier series to the time-series (green marker in figure (2) of the form:

$$f(t) = a_0 + \sum_{n=1}^{8} [a_n \cos(nt\omega) + b_n \sin(nt\omega)] \tag{31}$$

Where a_n, b_n are the Fourier coefficients and ω is the angular frequency, related to the signal's period T by $T = 2\pi/\omega$. The signal is expanded to degree 8, and by considering the magnitude of the coefficients a_n, b_n one can infer which period is "dominant" in the signal. The expansion is determined numerically by `fit(t, g, 'fourier8')`, and in table (3) we find the magnitude of the dominant terms in the Fourier expansion and hence the corresponding dominant period. The period of the signals for g_2^0 and g_4^0 is shown to be approximately 1 yr (as can also be read off in figure (2) above).

$\dot{g}_n^0$	ω	T	dominant term(s)	all other terms	ω^* of dominant term(s)	T^* of dominant term(s)
$\dot{g}_2^0$	0.8933	7.0340 yr	$a_7 = 5.779$, $b_7 = 2.246$	$< \lvert 1.261 \rvert$	$\omega^* = 7\omega$	$T^* = 7/T$ $= 0.995$ yr
$\dot{g}_4^0$	0.8947	7.0227 yr	$b_7 = 7.02$	$< \lvert 0.7118 \rvert$	$\omega^* = 7\omega$	$T^* = 7/T$ $= 0.997$ yr

Table 3: Angular frequency $\omega(\omega*)$ and period $T(T^*)$ of the 8th-degree Fourier series fit to the time series of $\dot{g}_2^0$ and $\dot{g}_4^0$ (green circles in figure (2)). The fit yields the Fourier coefficients in equation (31), where the contribution of the dominant terms is given together with an upper bound to the absolute value of the remaining terms. This gives an estimation of the period of the dominant term, given by T^*.

Figures of the time series show a clear disturbance in temporal dependence as carried by $\{g_n^m(t), h_n^m(t)\}$. However, not all Gauss coefficients exhibit a clear disturbance signal in their time-derivatives, such as the one-year periodicity of $\dot{g}_2^0$ and $\dot{g}_4^0$. In figure (5) a group of $\left\{\dot{g}_n^m(t), \dot{h}_n^m(t)\right\}$ is selected from the first 35 coefficients, together with predictions of CHAOS-6 models. The upper 9 coefficients in the figure exhibit an oscillatory behaviour (including $\dot{g}_2^0$ and $\dot{g}_4^0$) to some degree. The lower 9 coefficients exhibit a less clear pattern, and all the coefficients are generally following the trend of CHAOS-6 predictions. Hence we notice a difference in how the coefficients are affected by the signals.
Furthermore, models of type I produce predictions where the amplitudes of the signals in the time series are higher than ones obtained through models of type II. This agrees with the assertion made

[22] One could make time-series of the coefficients themselves, but the oscillations are more clear in the time derivatives, hence we restrict our results to time-derivatives, both when presenting coefficients and when presenting field components.

earlier that (θ, ϕ)-components are more vulnerable in polar latitudes. This will be discussed in greater detail below, see e.g. figure (A.1) of the appendix.

4.2.2 Observatory Examples

Besides the temporal behaviour, one may also be interested in the geographical distribution of the signals, i.e. how do the signals change in amplitude as we vary the spatial coordinates (r, θ, ϕ). One way of investigating this is by considering $\dot{\mathbf{B}}$ at observatory locations. The advantage of testing models at observatory locations is the availability of well-processed data that can be compared with model predictions. In figure (6) the three-component secular variation of the field is shown at HER, KAK, NGK and MCQ. The geographical locations of the observatories is show in table (1).

An oscillatory behaviour in model predictions is again seen at these observatory locations[23]. At HER and KAK, having a latitude of 34.43°N and 36°N respectively, we see weaker amplitudes in $\partial B_r/\partial t$ compared to those seen in NGK and MCQ that lie at 52.07°N and 54.4°S, respectively. In particular, the MCQ observatory is an interesting case, where observatory measurements themselves show an unclear pattern. This indicates the presence of a strong disturbance source at this observatory location.

4.2.3 Maps of the Signals

By using the CHAOS-6 model as a reference, the signals can be characterized by a more global approach by producing maps of $\dot{\mathbf{B}}$ at a certain altitude. We choose to consider the radial part $\partial B_r/\partial t$ and to produce a map at the Earth's surface. By making a grid of points at the Earth's surface where the values of $\dot{\mathbf{B}}$ are evaluated using the model and CHAOS-6, the two contributions are subtracted from each other such that the leftover signals from the oscillations in model prediction are clearly seen.
These maps where model predictions are subtracted from CHAOS-6 predictions are made both in type I models and in type II models. The former is shown in figure (4) while the latter is shown in figure (A.2) of the appendix. First and foremost we notice the difference in the amplitudes of the two maps. Model of type I have considerably stronger signals, at least in polar regions and mid-latitudes. The signals are also clearly present at both polar regions, where in the type I model there is a clear distinction between polar regions and lower/higher latitudes, while such a distinction is less clear in the type II model. Another interesting observation is how the signals appear as latitudal bands in some of the examples shown in the figures. This shows that the signals possess some latitudal dependence.

4.3 Data Selections

The previous subsection gives an idea of the noise problem in satellite data when core field models are built. The data contains complicated unmodeled signals that manifest themselves as disturbances (in some cases as oscillations) in model predictions, where the amplitudes of the signals are larger in higher latitudes. As mentioned earlier, modellers attempt to carefully exclude data that are believed to be disturbed for improved model solutions. The initial selections discussed in (3.1.1) are an example of such a selection, while using models of type II is another example[24]. In this subsection we shall extend this by testing whether choices on data selections based on criteria on the IMF and satellite-based AE indices (see 3.2) can improve cleaning the data from unmodelled signals.
The procedure of implementing the selections is described in (3.2). Before proceeding to presenting the results, a few comments should be made about the selections. The criteria listed in table (2) shows

[23] In particular, the radial components. Oscillations in (B_θ, B_ϕ)-components are considerably damped due to their exclusion at higher latitudes in this case (i.e. type II).

[24] In fact, the CHAOS-6 models uses the magnitude of the field at higher latitudes [Finlay et al., 2016]. The utilization of type II models is only a simplified way to minimize some of the disturbances at the expense of removing data that might include important information.

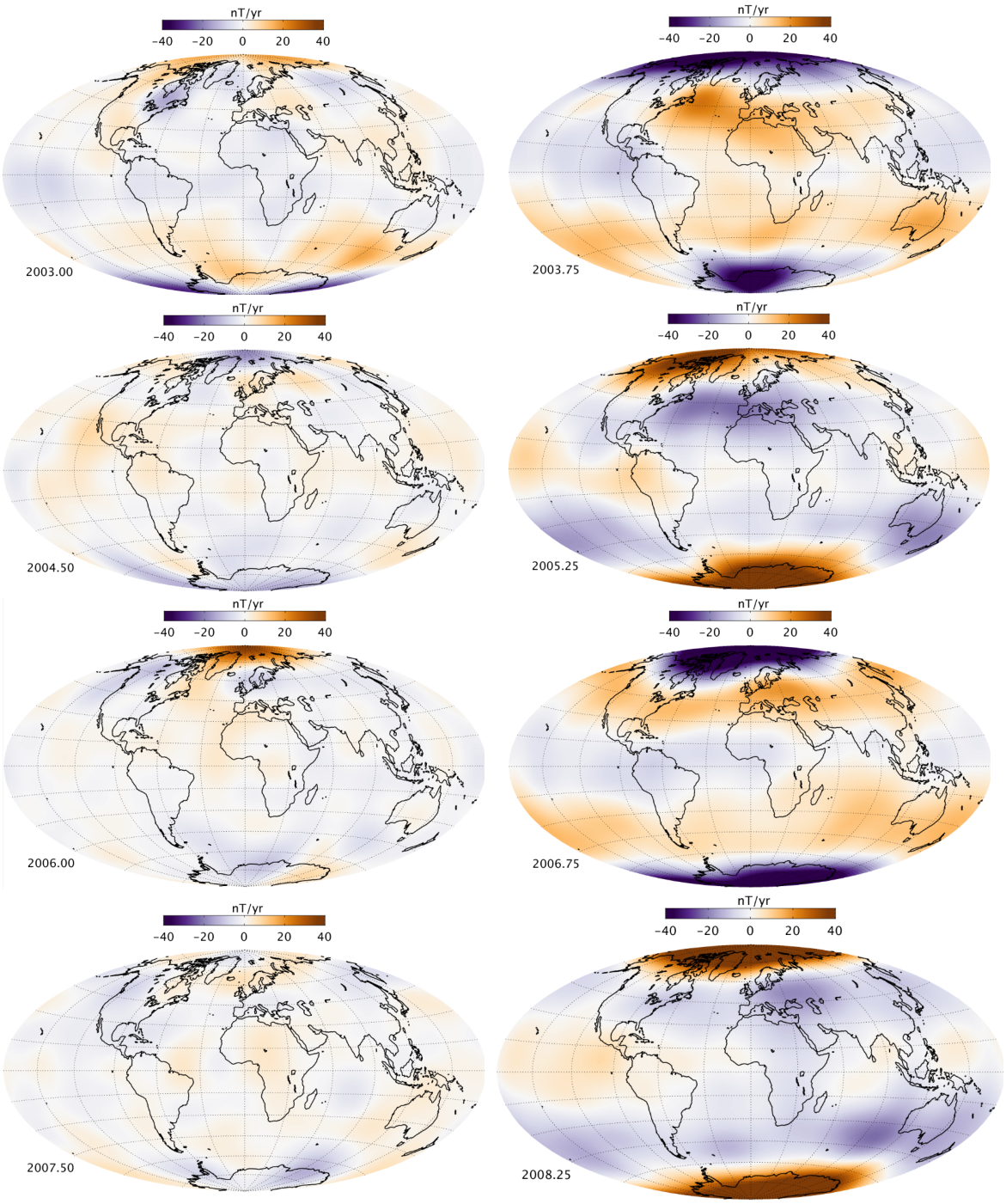


Figure 4: Maps showing the difference between CHAOS-6 and type I model predictions in $\partial B_r/\partial t$ evaluated at different times. The map is evaluated at the Earth's surface $r = 6371.2$ km. Note the difference in the scale of the colorbar in comparison to figure (A.2).

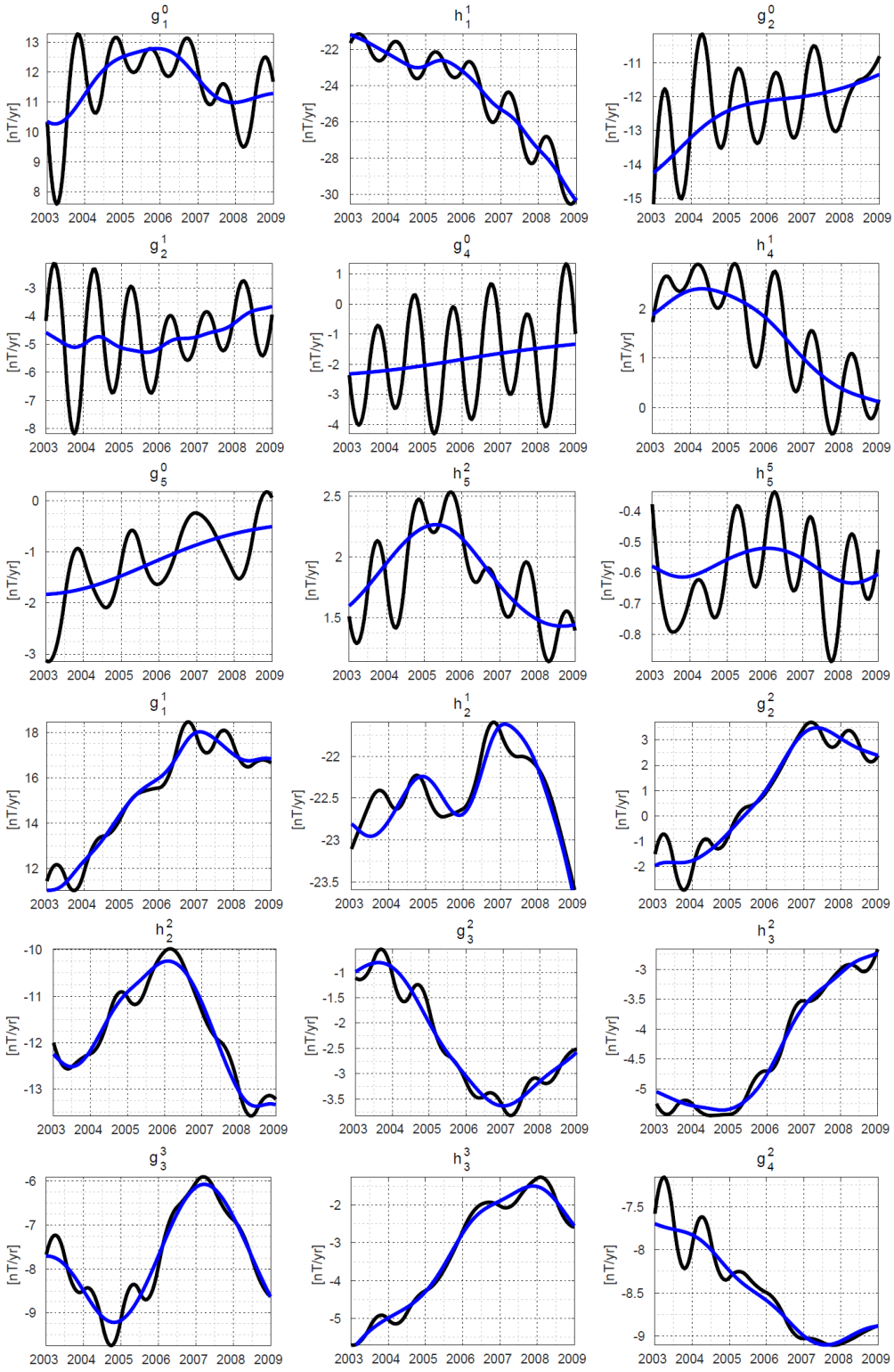


Figure 5: Time derivatives of various coefficients as obtained by a type II model with no additional selections. Blue curves: CHAOS-6 predictions; black curves: model predictions.

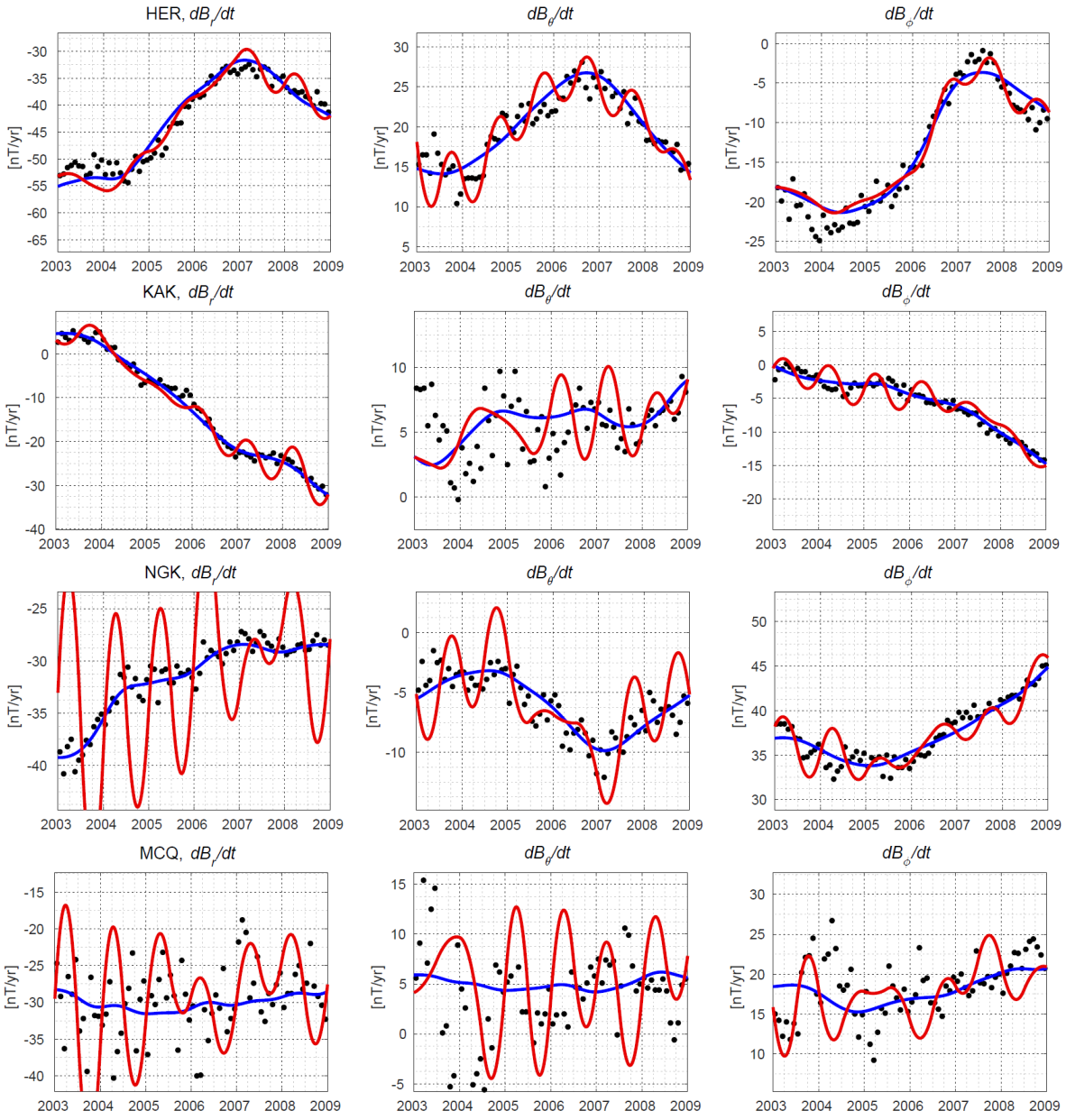


Figure 6: Type II $\mathbf{R}_{\mathrm{AR2}}$-regularized model predictions of the SV components at HER and KAK. Black dots: monthly mean ground observatory values; blue curves: CHAOS-6 predictions; red curves: model predictions. Note the difference in signal amplitudes in $\partial B_r/\partial t$ of HER and KAK compared to NGK and MCQ.

the number of data affected by the respective criterion. By choosing a specific criterion we remove all of the associated data. These selections remove approximately 23% to 91% of the original satellite data set that has a length of 68077, depending on the selection combination. This is a considerable reduction in the number of data to be used in obtaining model solutions. The number of data has an important effect on the quality of the model, so a combination involving a removal of a large percentage of the data may risk yielding a poor model.

The criterion that removes the biggest number of data by itself is E_m, removing about 68% of the data set. Another important detail about E_m is the fact that it is highly intersective with the other criteria as can be seen from table (2). This makes an independent testing almost impossible in the given data set. The focus will hence be aimed towards testing different *combinations* of models rather then testing selections against each other.
The analysis of data-selection effects shall be restricted to the following investigations:

- Utilizing model solutions of type I and type II, models are built where different combinations of criteria are chosen from table (2). The selections are chosen such that the removal does not leave the number of data to be less than approximately 11000 in order to avoid building fragile models.

- In order to make a quantitative assessment of the obtained models, values of the root mean square of the residuals (RMS) are calculated. In our case, the weighted RMS values of the residuals can be calculated by using:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}(\mathbf{d} - \mathbf{\Lambda m})^{\mathrm{T}}\mathbf{W}(\mathbf{d} - \mathbf{\Lambda m})} \tag{32}$$

  Where N is the number of data and $\mathbf{W}$ is the Huber weight matrix. Due to the strong dependence of the residuals on the latitude, the values of the RMS will be separated into polar[25] and non-polar regions.

- Time series of selected $\left\{\dot{g}_n^m(t), \dot{h}_n^m(t)\right\}$ will be considered. A number of coefficients have been chosen from the lowest 35 model coefficients[26] without selections and evaluated at the period 2003-2009. These are shown in figure (5). Some of the coefficients exhibit a clear oscillatory pattern, while others show an unclear disturbance. In the mentioned figure an attempt has been made to include examples from both groups. In particular, the coefficients that are showing clear periodicity (some of the upper 9 examples in the figure) will be chosen in order to be assessed against data selections, while some of the other coefficients (the lower 9 in the figure) that possess a random pattern will not be analyzed against data selection.

- Finally, we would like to investigate how data selections affect SV components of the field, in particular B_r. Instead of restricting to observatory locations (which will yield a very restricted geographical description), we consider maps of differences between model and CHAOS-6 prediction in B_r at polar regions. This is done with and without selections for comparison. The polar regions are chosen due to the strong signals in their vicinity and since the effect of data selections in lower latitudes are not easily detected in our case.

4.3.1 Effects on Statistics

Table (4) shows the models built using type I (upper table) and type II models (lower table). Each model is characterized by a data selection, number of data and RMS of the residual in field components at polar and non-polar regions. We give some remarks on these results:

- *General remarks*: Type I and type II models yield different values of RMS, as one may expect. The RMS misfit in B_r component is smaller in type I than in II in polar regions. It is expected that when only one component (i.e. B_r) is used in the inversion to yield 3-component predictions of the field, the model predictions may be less accurate. The advantage, on the other hand, is to remove unwanted contributions from the (B_θ, B_ϕ)-components. However, in non-polar regions,

[25] $<35°$ and $>125°$ in co-latitude.
[26] i.e. $\left\{g_1^0, g_1^1, h_1^1, \ldots, h_5^5\right\}$.

Field Model	Selection	N	Polar [nT]			Nonpolar [nT]		
			B_r	B_θ	B_ϕ	B_r	B_θ	B_ϕ
FM_type1_01	None	61383	10.097	17.353	18.082	4.291	4.337	3.240
FM_type1_02	E_m	19635	9.974	17.443	18.250	3.962	4.213	3.136
FM_type1_03	B_z	29854	9.869	17.104	17.852	3.880	4.127	3.076
FM_type1_04	B_y	46787	10.112	17.351	17.930	4.311	4.457	3.298
FM_type1_05	E_m, B_z	18089	9.964	17.425	18.181	3.925	4.162	3.119
FM_type1_06	E_m, J_{L1}	11613	10.144	17.740	18.198	4.069	4.312	3.125
FM_type1_07	J_{L1}, $J_{\max}, J_{\min}$	30402	10.097	17.353	18.082	4.291	4.337	3.240

Field Model	Selection	N	Polar [nT]	Nonpolar [nT]		
			B_r	B_r	B_θ	B_ϕ
FM_type2_01	None	61383	12.304	3.152	3.663	2.976
FM_type2_02	E_m	19635	12.247	3.039	3.651	2.981
FM_type2_03	B_z	29854	12.226	3.006	3.582	2.935
FM_type2_04	B_y	46787	12.331	3.199	3.704	2.979
FM_type2_05	E_m, B_z	15223	12.285	3.056	3.680	2.971
FM_type2_06	E_m, J_{L1}	11613	12.508	3.085	3.683	3.015
FM_type2_07	J_{L1}, $J_{\max}, J_{\min}$	14328	13.307	3.130	3.832	3.225

Table 4: Tables presenting different models built with specified selection criteria, where N is the number of data used in obtaining the model solution, together with RMS misfit values of the residuals subdivided into polar and non-polar regions. Upper table: models of type I; lower table: models of type II (as described in the beginning of this section).

the RMS values in all three components are generally less in type II compared to type I. Indeed, it is an interesting fact that the RMS values of the residuals are lower in type II than in type I at non-polar regions, although the type II models are only excluding data from polar latitudes.

- The changes in RMS misfit from model to model are generally on a small scale. Furthermore, some models reduce the value of RMS misfit in some components will increasing the value in other components. The B_z criterion seems to improve the RMS value in both types of models. In the case of E_m and $\{E_m, J_{\mathrm{L1}}\}$ some components seem to improve, while B_y and $\{J_{\mathrm{L1}}, J_{\max}, J_{\min}\}$ does not seem to have an improvement. The $\{E_m, B_z\}$ and $\{E_m, J_{\mathrm{L1}}\}$ conditions in upper table show a small improvements in the non-polar regions, while no clear improvements can be detected in the corresponding lower table. For polar regions, improvements in the radial components are also seen for E_m, B_z and $\{E_m, B_z\}$ for both types of models.

4.3.2 Effects on Coefficients

Besides model misfit statistics, it is of interest to consider the time-dependence in model coefficients. The time series to be presented here are $\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\}$. These are chosen due to their clear periodic signal as can be seen in figure (5), where in particular $\{\dot{g}_2^0, \dot{g}_4^0\}$ exhibit a dominant one year periodicity as seen in table (3).

The effects of the selections based on E_m, B_z, $\{E_m, J_{\mathrm{L1}}\}$ and $\{J_{\mathrm{L1}}, J_{\max}, J_{\min}\}$ on $\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\}$ is seen in figure (7) and (8). An immediate observation is the fact that type II models have lower amplitudes than type I models, as can also be seen in the world maps mentioned earlier. Within each of these two types the tests are performed, where the following points summarize the observations:

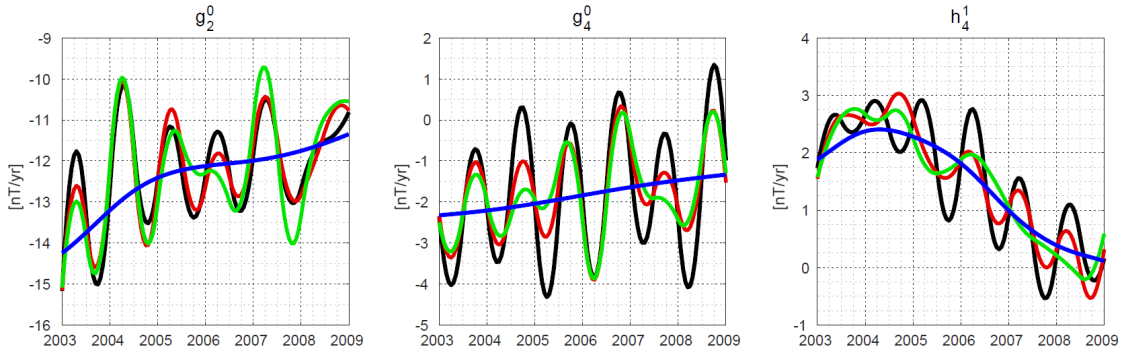
- In type II models, some of the signals get *disturbed* after a selection is chosen. This can be seen in e.g. h_4^1 in figure (7) and (8). This effect is not shown in the case of type I models (with the exception of the period between 2003-2004). The structure of the signal is generally maintained after performing the selections in this case, where the amplitudes are altered (with few exceptions). This is possibly a sign of instability in the models of type II due to the removal of (B_θ, B_ϕ)-components of the data at polar regions.
- Some of the data selections do seem to reduce the amplitudes of the signals. In the case of type II models exceptions seem to be more common, where selections change the signal structure in some time periods. The selections of E_m and B_z show a reduction in amplitudes, similarly with the B_z selection.
- The satellite-based EJ index J_{L1} to have a positive contribution when combined with E_m. This is most clearly seen in the type II models as the green lines in figure (7), while making a choice based on all three auroral selections yield a smaller contribution in comparison as seen in the green lines of figure (8).

Furthermore, there is another clear difference between type I and type II models in the derivatives of the coefficients. Figure (A.1) shows the derivatives of the upper nine coefficients that were introduced in figure (5). The effect of the transition from type I model to type II model is very clear in some of the coefficients, such as g_4^0, while the effect is less clear in others. The coefficient g_2^0 even shows a total reversal in the signal. The difference between the two types have also been observed in the maps, where they show lower amplitudes (as a difference to CHAOS-6) in type II models[27]. There is hence an advantage in excluding the (B_θ, B_ϕ)-components at polar regions[28].
As a last comment in this connection, it is worth noting that the time-variation in some of the mentioned coefficient-derivatives, such as the lowest 9 in figure (5), may in some cases be difficult to separate in core-related variations and external-field-related variation (or even a third effect), as it is well known that there exists rapidly changing core-related events such as geomagnetic jerks. However, the fact that data selections have lowered the amplitudes of some the signals seen in $\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\}$ is arguably an evidence that these signals have an external origin, i.e. related to current systems.

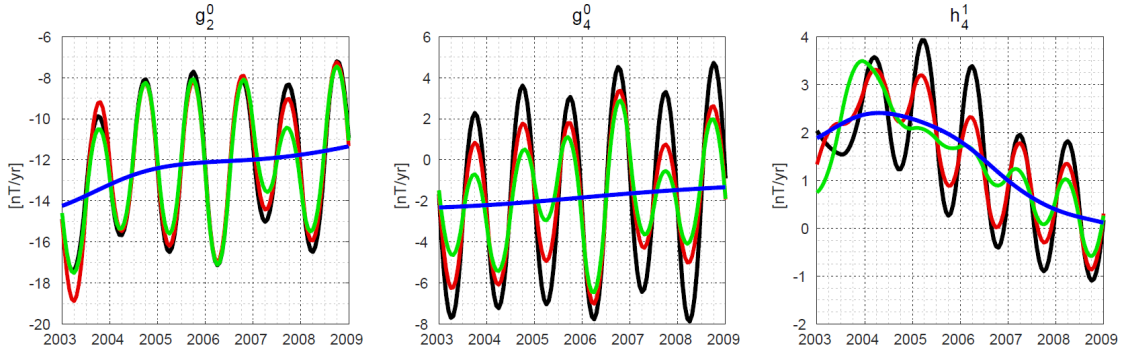
4.3.3 Effects on Polar regions

Finally, we test the effects of data selection based on E_m on the polar region by considering maps of $\partial B_r/\partial t$ restricted to the polar regions and evaluated at the Earth's surface. This choice has been made due to dependence of the disturbance signals on the latitude, with the strongest disturbance localized in polar regions. The effect of data selections on specific locations are hard to investigate, hence we don not consider examples at observatory locations, although we have seen how mid-latitude observatories such as NGK and MCQ exhibit stronger disturbances than HER and KAK which are lying closer to the equator.
The maps, as before, are produced by taking the difference between model predictions and the predictions of CHAOS-6. One important detail when making comparison based on such maps is the fact that one considers a *snap shot* of the SV component at a specific time. If a data selection reduces some of the negative effects of a signal at one given time then one cannot conclude that this is a

[27] Note the difference in the scale of the colormaps when considering figure (A.2) and (4)
[28] Or even better, deal with them differently in order to cope with the disturbances they give rise to, as already done in the CHAOS models.



(a) Time series of $\left\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\right\}$ for type I models based on different selections. Blue curves: CHAOS-6 predictions; black curves: `FM_type1_01`; red curves: `FM_type1_02` ; green curves: `FM_type1_06`.



(b) Time series of $\left\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\right\}$ for type II models based on different selections. Blue curves: CHAOS-6 predictions; black curves: `FM_type2_01`; red curves: `FM_type2_02` ; green curves: `FM_type2_06`.

Figure 7

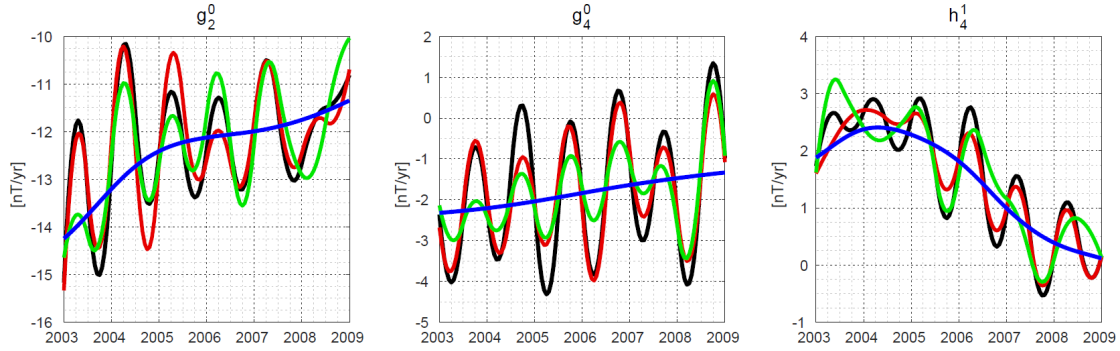
general improvement. We make the following choice: the polar regions are considered where $\partial B_r/\partial t$ is evaluated at times where the signals amplitude is in the vicinity of a maximum or a minimum. By taking the interval time between each image to be a half-year[29], maps spanning from 2003.25-2008.75 are shown in the series of figures (A.3), (A.4) and (A.5)[30]. This is done for type II model where the signals are more clear, and as a selection the E_m criterion is used. It is underlined that the aim here is not to give a quantitative description of the improvements, but rather give a qualitative example of how the disturbance signals are prominent in polar regions, and that a detectable effect can be shown by an example of data selection (in this case E_m).
As observed in the maps, at several moments the signal amplitudes are lowered substantially after the selection. This confirms the assertion that polar regions are quite vulnerable to disturbances.

4.4 Comments on Results

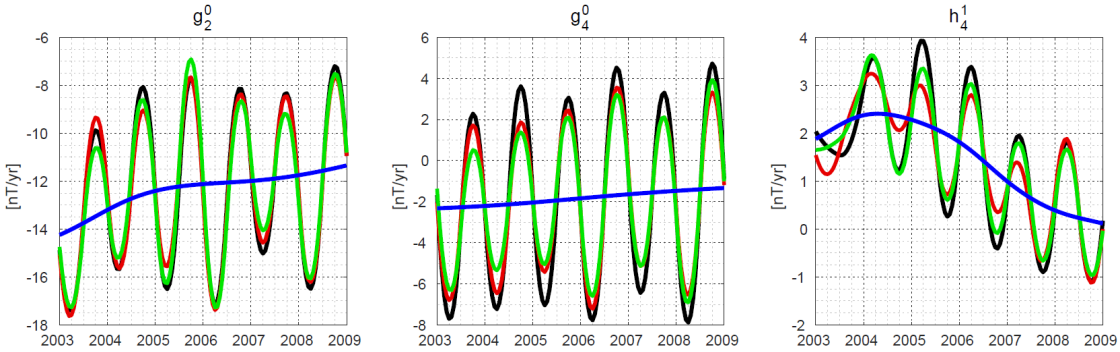
In this section an attempt has been made to investigate some of the disturbance signals identified when building $\mathbf{R}_{\mathrm{AR2}}$-regularized models with a half year knot spacing. As stressed earlier, the motivation for this investigation is to address some of the problems modellers face when building high resolu-

[29] At polar regions a dominant period of one year is shown in the maps. We have also seen a one-year periodicity in the signals that appeared in the time-series of $\dot{g}_2^0$ and $\dot{g}_4^0$.

[30] Note that the scale of the colorbar has been chosen to show the effects of the polar disturbances rather than lower latitudes.



(a) Time series of $\left\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\right\}$ for type I models based on different selections. Blue curves: CHAOS-6 predictions; black curves: `FM_type1_01`; red curves: `FM_type1_03` ; green curves: `FM_type1_07`.



(b) Time series of $\left\{\dot{g}_2^0, \dot{g}_4^0, \dot{h}_4^1\right\}$ for type II models based on different selections. Blue curves: CHAOS-6 predictions; black curves: `FM_type2_01`; red curves: `FM_type2_03` ; green curves: `FM_type2_07`.

Figure 8

tion core field models [Finlay et al., 2017]. Data selection comprises an extremely important part in field modelling based on data and knowledge on the physics of the ionosphere and magnetosphere. Addressing the unmodelled signals can in principle be done either be excluding (i.e. by performing appropriate data selections) the affected data or providing additional constraints (i.e. regularization) on the models based on assumptions made on the core dynamics. The attempt in this project has been to use a weaker temporal regularization and to focus on characterizing the effects of data selections.
This section is finalized with a few remarks on the results. It has been shown how the the disturbances in the model predictions are localized in the polar region. This is no surprise; the polar regions are home to auroral currents in the polar ionosphere that are known to cause problems in field modelling, as documented in [Finlay et al., 2017] and [Friis-Christensen et al., 2017]. We do not wish to delve into a detailed account of the current systems, but merely mention that the results, expectedly, indicate that polar disturbances are quite clear in our internal models.
Indeed, the fact that data selections based on IMF and AE indices reduce some of the disturbances is a strong indication that these disturbances are solar-related. It also indicates that such data selections may have a role in future improvement schemes on field models. In particular, a more sophisticated approach in finding selection-thresholds of the satellite-based AE indices [Aakjær et al., 2016] may constitute a further work in this connection.
Within the framework of this project, data selections of several kinds have been shown to exhibit some effects on the characteristics that we have investigated. But there is also an interest in knowing to *which degree* it is possible to eliminate the unmodelled signals by means of the data selections similar

to the ones that we have considered? This is a rather difficult question to address, at least based on the results above. As stressed earlier, data selections must be in the recipe of any field model, and the results of this section have shown how a removal of selected data can improve reducing the amplitudes of some of the signals. However, the attempt here is arguably incomplete and unsatisfactory for several reasons. The big reduction in the number of data can cause defects in the model, so an increase in data number may be a help. In this connection, SWARM, with its improved precision and resolution, offers a great opportunity for modellers to understand and possibly make better models of the disturbances. Additionally, the higher disturbances in the type I models must be addressed more efficiently before further investigations, such as the method implemented in the CHAOS-6 model that we have mentioned earlier, rather than adopting a type II model. Another method to deal with the polar disturbances could be accomplished by an estimation of identified polar disturbances and incorporating them in model parametrization as done by [Olsen et al., 2016] in the SIFMPlus model.

5 Appendix

A

Figures

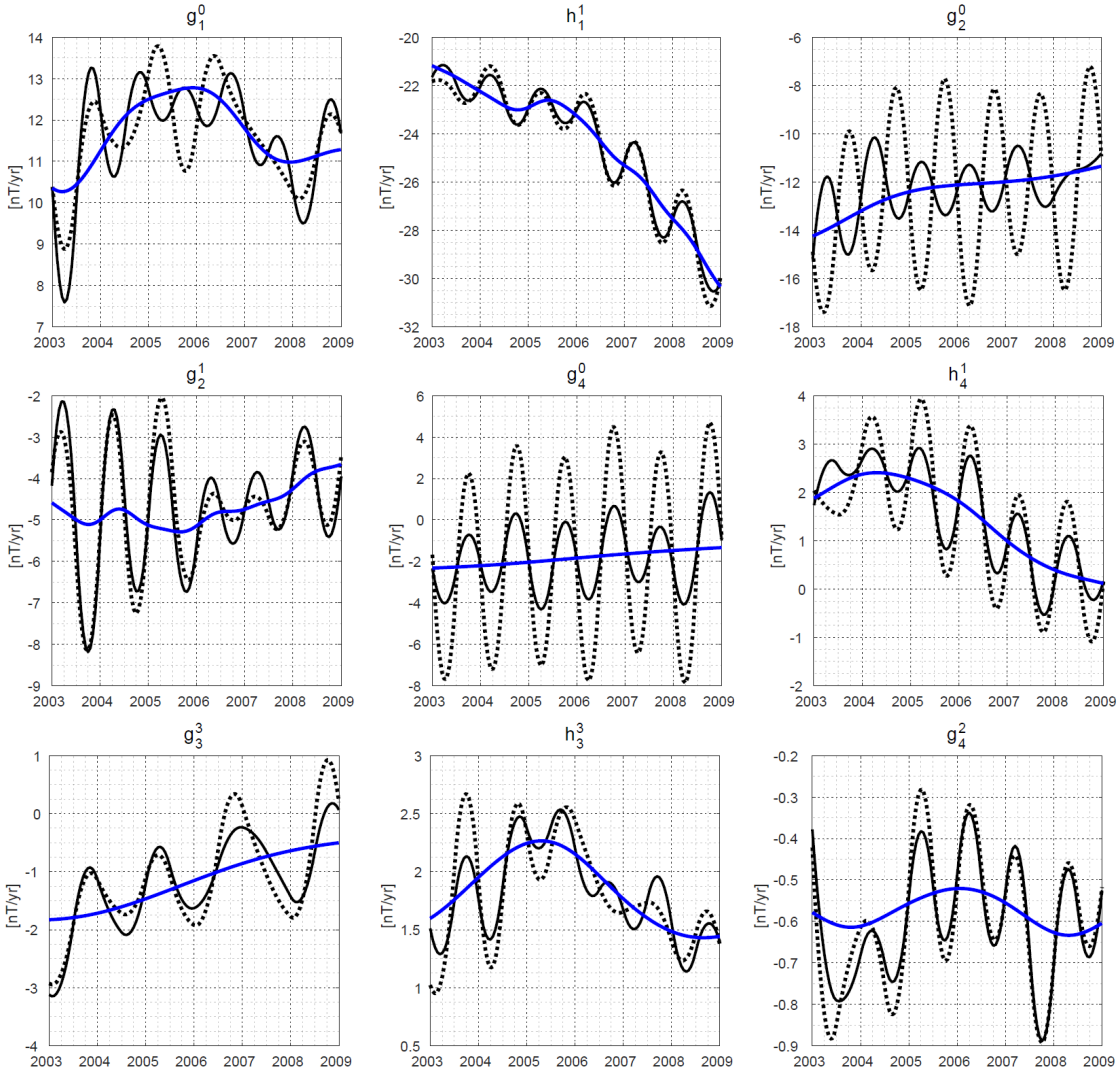


Figure A.1: Time-derivatives of the first 9 coefficients in figure (5). Blue curves: CHAOS-6 model; dashed black curves: type I model with no additional selection; solid black curves: type II model with no additional selection.

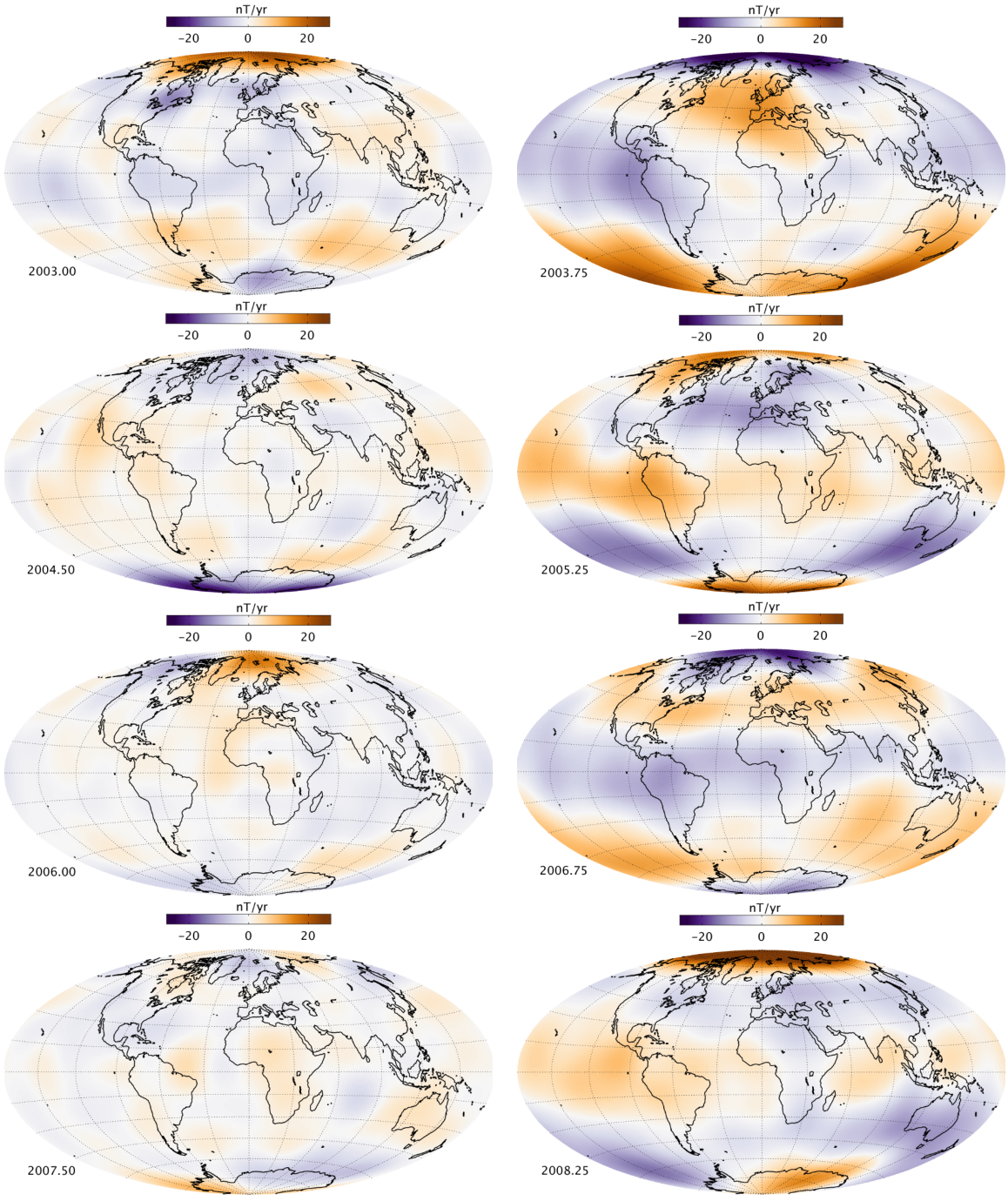


Figure A.2: Maps showing the difference between CHAOS-6 and type II model predictions in $\partial B_r/\partial t$ evaluated at different times. The map is evaluated at the Earth's surface $r = 6371.2$ km. Note the difference in the scale of the colorbar in comparison to figure (4).

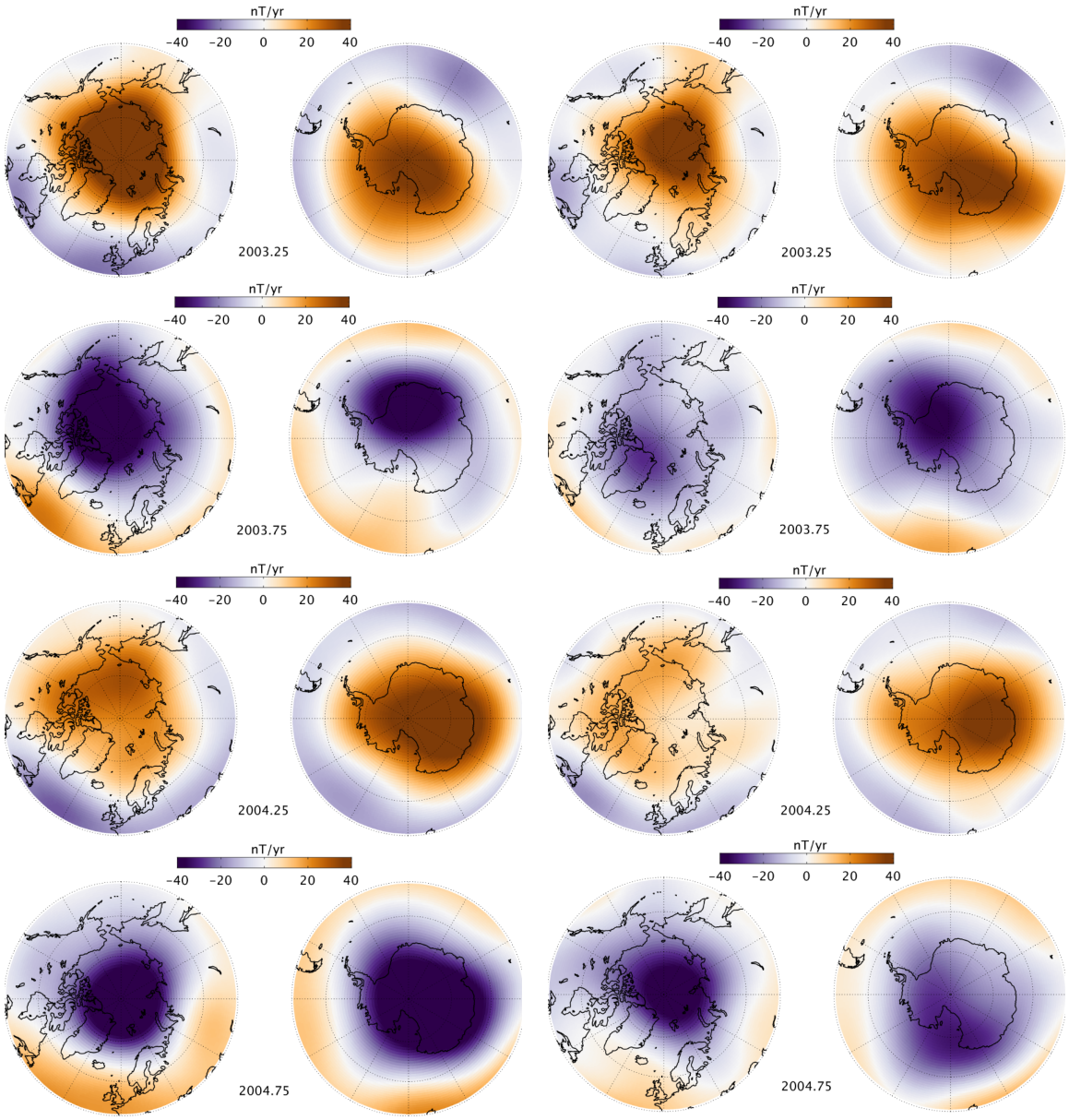


Figure A.3: Maps showing the difference between CHAOS-6 and type II model predictions in $\partial B_r/\partial t$ at polar regions, spanning 2003.25 to 2004.75. The map is evaluated at the Earth's surface $r = 6371.2$ km.

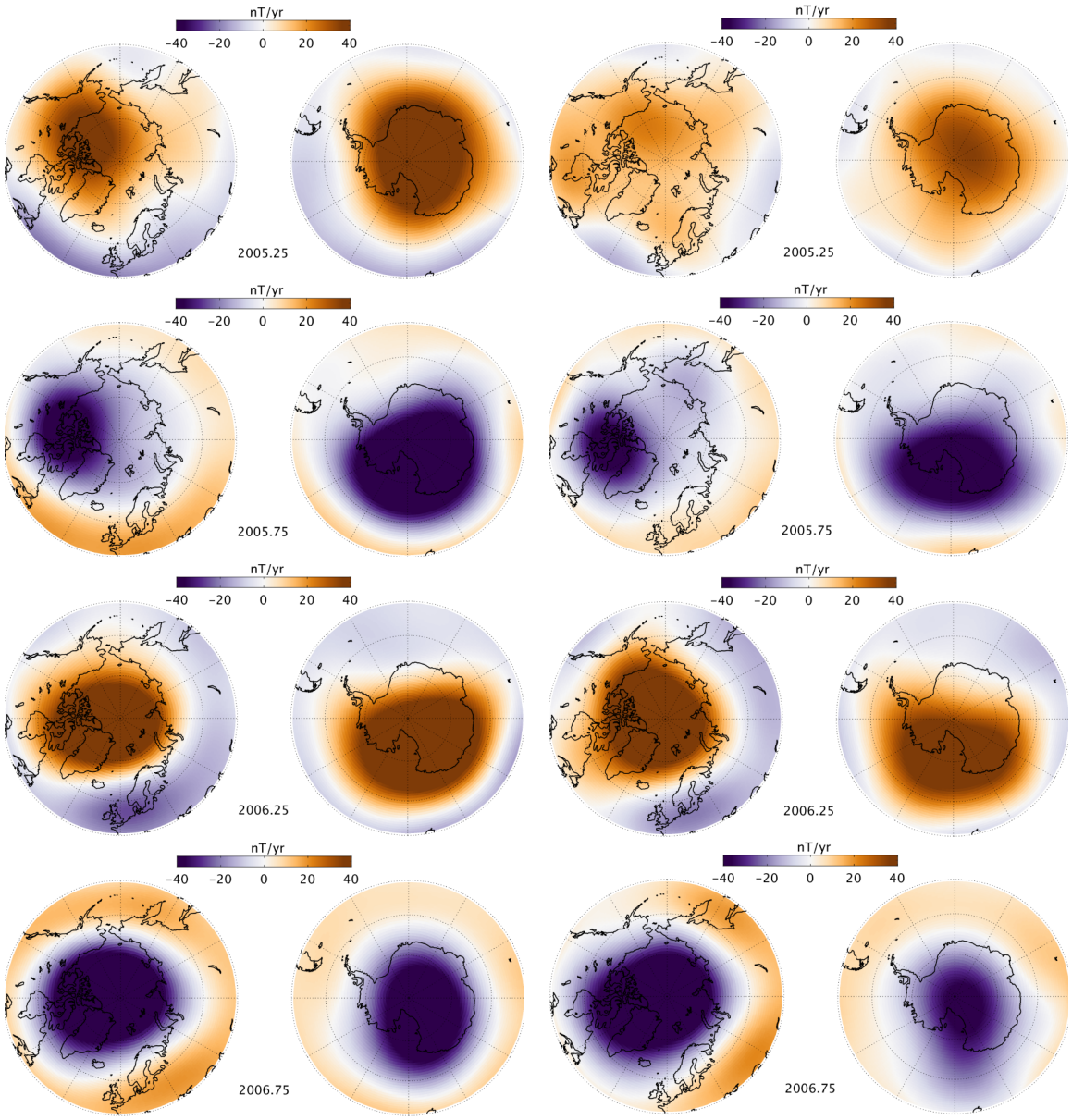


Figure A.4: Maps showing the difference between CHAOS-6 and type II model predictions in $\partial B_r/\partial t$ at polar regions, spanning 2005.25 to 2006.75. The map is evaluated at the Earth's surface $r = 6371.2$ km.

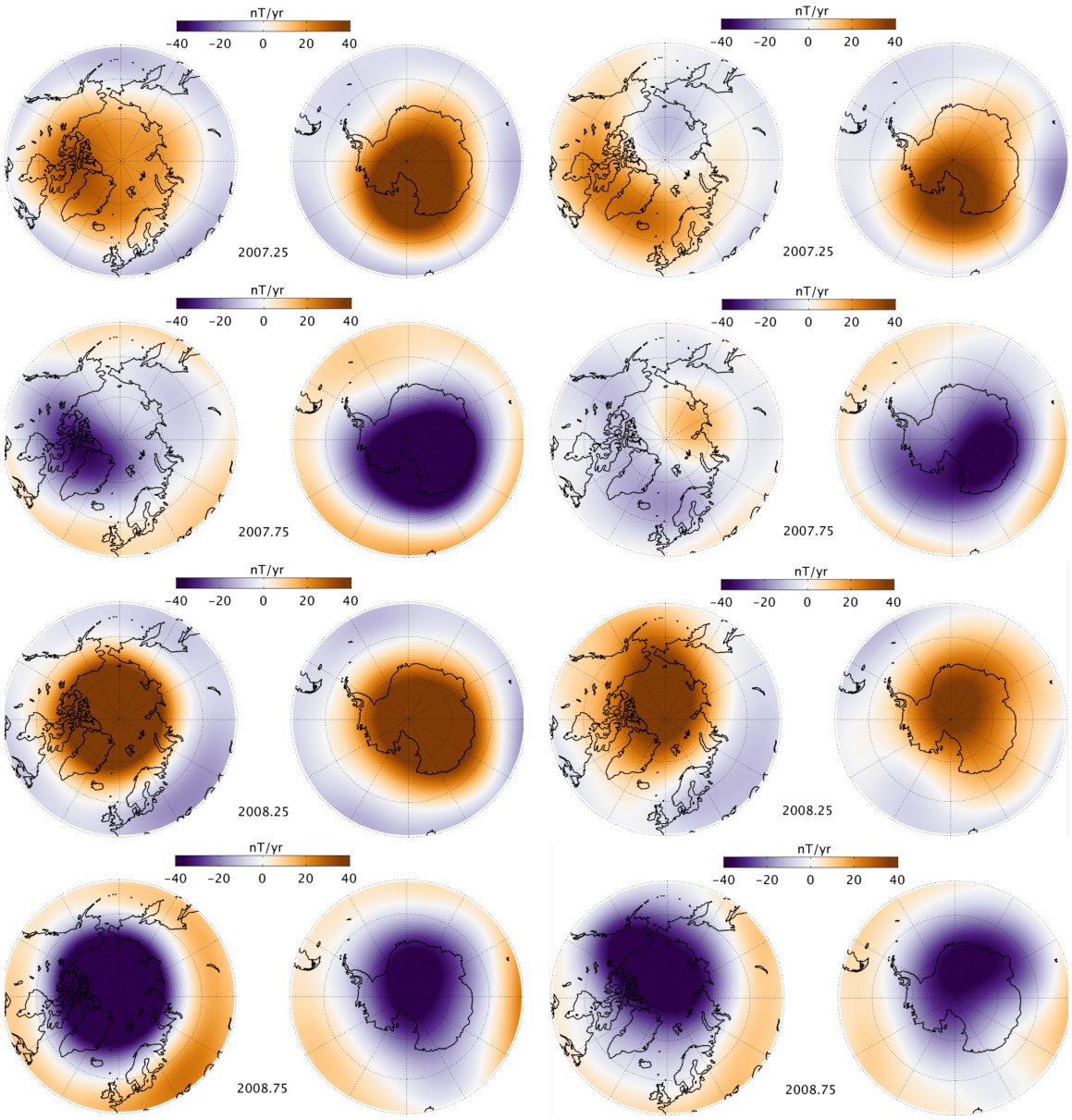


Figure A.5: Maps showing the difference between CHAOS-6 and type II model predictions in $\partial B_r/\partial t$ at polar regions, spanning 2007.25 to 2008.75 for model. The map is evaluated at the Earth's surface $r = 6371.2$ km.

B
MATLAB code

Λ Matrix:

```
% Description:
% Model matrix (lambda)
% input: positions vector, time and expansion degree N
% output: model matrices Lambda_i, i = r, theta, phi,
%         and R, regularization matrix.
% computed as eq. (19)

function [G_1, G_2, G_3, R] = Giant_G(t, r, theta, phi, N)

% angles in radians and distance in units of Earth radius.
theta_rad = theta./180.*pi;
phi_rad = phi./180.*pi;
r_a = r./6371;

% Olsen, design_SHA
[G_r, G_theta, G_phi] = design_SHA(r_a, theta_rad, phi_rad, N);
%

G_gh_1 = []; G_gh_2 = []; G_gh_3 = [];

for j = 1:N*(N+2)
    G_gh_1 = [G_gh_1, spdiags(G_r(:,j), 0, numel(G_r(:,j)),...
        numel(G_r(:,j)))];
    G_gh_2 = [G_gh_2, spdiags(G_theta(:,j), 0, numel(G_theta(:,j)),...
        numel(G_theta(:,j)))];
    G_gh_3 = [G_gh_3, spdiags(G_phi(:,j), 0, numel(G_phi(:,j)),...
        numel(G_phi(:,j)))];
end

G_gh = [G_gh_1, G_gh_2, G_gh_3]; % spatial part of lambda

% unit of time (MD2000)

t = 2000+t/365.25;
sp_order = 4;  % spline order
delta_t = 0.5; % knot spacing
knots = [2002:delta_t:2010]';
knots = augknt(knots, sp_order);
G_ab_0 = spcol(knots, sp_order, t, 'sp', 'no');
ACell = repmat({G_ab_0}, 1, N*(N+2));
G_ab = blkdiag(ACell{:});  % temporal part of lambda

% regularization (not AR2-regularization).
coll_t = spcol(knots, sp_order, [knots(1):delta_t:knots(end)]', 'sp');
coll_2 = diff(coll_t, 2)/(delta_t/365.25)^2;
R_time = coll_2'*coll_2*(delta_t/365.25);
BCell = repmat({R_time}, 1, N*(N+2));
R = blkdiag(BCell{:}); % the regularization matrix

% the final matrices
G_1 = G_gh_1*G_ab;
G_2 = G_gh_2*G_ab;
G_3 = G_gh_3*G_ab;

end
```

Models (type II):

```
%% Description

% Script for generating models that include all data,
% together with selection options.

%% Load data set

clear
data = load('CHAMP_10min_dark_quiet_Ali_Bvect.dat.txt');

%% Define the data vectors from dataset

B_r = data(:,14);
B_theta = data(:,15);
B_phi = data(:,16);
r = data(:,2);
theta = data(:,3);
phi = data(:,4);
t = data(:,1);

tau = t; % this will be used later
%% Sampling data randomly

k = 1;
numb = 0.5*length(data);    % sampling number, max = length(data);
i = zeros(numb,1);
for n = 1:length(i)
    i(k) = n*floor(length(data)/numb); % index for selection
    k = k + 1;
end
% The random selection
B_r = B_r(i);
B_theta = B_theta(i);
B_phi = B_phi(i);
r = r(i);
theta = theta(i);
phi = phi(i);
t = t(i);

%% Selecting data on chosen criteria (e.g. IMF )

load('indices_individual_revised.mat') % IMF selections
load('Auroral.mat')                    % J_L1 selection
load('Auroral_max.mat')                % J_max selection
load('Auroral_min.mat')                % J_min selection

%% The data selection (IMF / AE indices)

J = union(J_max, J_min); % a combination of selections

% removal of data
B_r(J) = [];
B_theta(J) = [];
B_phi(J) = [];
r(J) = [];
theta(J) = [];
phi(J) = [];
t(J) = [];
%% Removal of systematic outliers

```

```
N = 16;
[G_1, G_2, G_3] = Giant_G(t, r, theta, phi, N); % Gamma matrices
G = [G_1; G_2; G_3];
B = [B_r; B_theta; B_phi];
GtG = G'*G;
GtB = G'*B;
m = GtG\GtB; % inversion (model solution)

% defining residuals
res_r = abs(B_r-G_1*m);
res_theta = abs(B_theta-G_2*m);
res_phi = abs(B_phi-G_3*m);

% outliers are defined as abs(res_i)>1000, i = r, theta, phi.
for j=length(B_r):-1:1
    if res_r(j)>1000 || res_theta(j)>1000 || res_phi(j)>1000
        B_r(j)=[]; B_theta(j)=[]; B_phi(j)=[];
        t(j)=[]; r(j)=[]; theta(j)=[]; phi(j)=[];
    end
end

%% Re-building model

N = 16;
[G_1, G_2, G_3, R] = Giant_G(t, r, theta, phi, N);
G = [G_1; G_2; G_3];
B = [B_r; B_theta; B_phi];
GtG = G'*G;
GtB = G'*B;

%% Regularization choice

% R_AR2:
load('R_AR2_0.5.mat')
alphasq2=1; % smoothing parameter
m = (GtG+alphasq2*R_AR2)\GtB; % regularized model

% clearing workspace from definitions made previously
clearvars -except B G m N t tau R_AR2 alphasq2

%% Robust estimation (Huber weights)

e = B-G*m;
sigma = rms(e);
w = ones(length(B),1); % initial weights
c=1.5;
N_it = 6; % iteration number
for i=1:N_it
    wB = w.*B;
    tic
    WG = bsxfun(@times,G,sqrt(w));
    GtWG = WG.'*WG;
    GtWB = G'*wB;
    m = (GtWG + alphasq2*R_AR2)\(GtWB);
    epsilon = (B-G*m)/sigma;
    w = min(c./abs(epsilon), 1);
    toc
end

%% pp-form of model parameters

sp_order = 4;  % spline order
delta_t = 0.5; % knot spacing
```

```
124 knots = [2002:delta_t:2010]';
125 knots = augknt(knots, sp_order);
126
127 N_knots = length(knots) - sp_order; % number of knots
128 N_coeff_SHA_spline = N*(N+2); % number of SH coefficients per knot
129 N_coeff_spline = N_coeff_SHA_spline*N_knots;
130
131 coefs = reshape(m(1:N_coeff_spline), N_knots, N_coeff_SHA_spline);
132 pp = fn2fm(spmak(knots', coefs'), 'pp'); % the model pp-form
133
134 %% Savings
135
136 [ismem, index] = ismember(t,tau); % This is used in model statistics
137 save('FM_final_1.mat','pp','m', 'w', 'index')
```

List of scripts:

1. `attempt_selection.m`
2. `Model_Generator.m`
3. `Model_Polar.m`
4. `design_SHA.m` (*)
5. `Giant_G.m`
6. `Oscillations.m`
7. `plot_CMB_field_edited.m` (**)
8. `predictions.m`
9. `residuals.m`
10. `statistics.m`
11. `statistics_polar.m`
12. `synth_grid.m` (*)
13. `synth_values.m` (*)

(*) denotes files that are not part of the author's work but have been used in the project.
(**) denotes files that are of type (*) but additionally have been modified by the author.

References

[Aakjær et al., 2016] Aakjær, C. D., Olsen, N., and Finlay, C. C. (2016). Determining polar ionospheric electrojet currents from swarm satellite constellation magnetic data. *Earth, Planets and Space*, 68(1):140.

[Constable, 1988] Constable, C. (1988). Parameter estimation in non-gaussian noise. *Geophysical Journal International*, 94(1):131–142.

[Constable and Parker, 1988] Constable, C. and Parker, R. (1988). Smoothing, splines and smoothing splines; their application in geomagnetism. *Journal of Computational Physics*, 78(2):493–508.

[de Boor, 1997] de Boor, C. (1997). Matlab spline toolbox.

[De Boor et al., 1978] De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., and De Boor, C. (1978). *A practical guide to splines*, volume 27. Springer-Verlag New York.

[Finlay et al., 2017] Finlay, C., Lesur, V., Thébault, E., Vervelidou, F., Morschhauser, A., and Shore, R. (2017). Challenges handling magnetospheric and ionospheric signals in internal geomagnetic field modelling. *Space Science Reviews*, 206(1-4):157–189.

[Finlay et al., 2016] Finlay, C. C., Olsen, N., Kotsiaros, S., Gillet, N., and Tøffner-Clausen, L. (2016). Recent geomagnetic secular variation from swarm. *Earth, Planets and Space*, 68(1):1–18.

[Friis-Christensen et al., 2017] Friis-Christensen, E., Finlay, C., Hesse, M., and Laundal, K. (2017). Magnetic field perturbations from currents in the dark polar regions during quiet geomagnetic conditions. *Space Science Reviews*, 206(1-4):281–297.

[Gillet et al., 2013] Gillet, N., Jault, D., Finlay, C., and Olsen, N. (2013). Stochastic modeling of the earth's magnetic field: Inversion for covariances over the observatory era. *Geochemistry, Geophysics, Geosystems*, 14(4):766–786.

[Gubbins and Herrero-Bervera, 2007] Gubbins, D. and Herrero-Bervera, E. (2007). *Encyclopedia of geomagnetism and paleomagnetism*. Springer Science & Business Media.

[Guo et al., 2014] Guo, J., Liu, H., Feng, X., Pulkkinen, T. I., Tanskanen, E., Liu, C., Zhong, D., and Wang, Y. (2014). Mlt and seasonal dependence of auroral electrojets: Image magnetometer network observations. *Journal of Geophysical Research: Space Physics*, 119(4):3179–3188.

[Kan and Lee, 1979] Kan, J. and Lee, L. (1979). Energy coupling function and solar wind-magnetosphere dynamo. *Geophysical Research Letters*, 6(7):577–580.

[Kauristie et al., 2017] Kauristie, K., Morschhauser, A., Olsen, N., Finlay, C., McPherron, R., Gjerloev, J., and Opgenoorth, H. J. (2017). On the usage of geomagnetic indices for data selection in internal field modelling. *Space Science Reviews*, 206(1-4):61–90.

[Newell et al., 2007] Newell, P., Sotirelis, T., Liou, K., Meng, C.-I., and Rich, F. (2007). A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research: Space Physics*, 112(A1).

[Olsen et al., 2016] Olsen, N., Finlay, C. C., Kotsiaros, S., and Tøffner-Clausen, L. (2016). A model of earth's magnetic field derived from 2 years of swarm satellite constellation data. *Earth, Planets and Space*, 68(1):124.

[Olsen et al., 2014] Olsen, N., Lühr, H., Finlay, C. C., Sabaka, T. J., Michaelis, I., Rauberg, J., and Tøffner-Clausen, L. (2014). The chaos-4 geomagnetic field model. *Geophysical Journal International*, 197(2):815–827.

[Olsen and Stolle, 2012] Olsen, N. and Stolle, C. (2012). Satellite geomagnetism. *Annual Review of Earth and Planetary Sciences*, 40:441–465.

DTU Space
National Space Institute
Technical University of Denmark

Elektrovej, building 327
DK - 2800 Kgs. Lyngby
Tel (+45) 4525 9500
Fax (+45) 4525 9575

www.space.dtu.dk